# QUANTUM PRINCIPLES AND PARTICLES

BY

WALTER WILCOX
BAYLOR UNIVERSITY

**Quantum Principles and Particles**

## 1. Perspective and Principles

Failure of classical mechanics
Stern-Gerlach experiment
Idealized Stern-Gerlach results
Classical model attempts
Wave functions for two physical outcome case
Process diagrams, operators, and completeness
Further properties of operators
Operator reformulation
Operator rotation
Bra-ket notation/basis states
Transition amplitudes
Three magnet setup example - coherence
Hermitian conjugation
Unitary operators
Matrix representations
Matrix wave function recovery
Expectation values
Wrap up
Problems

## 2. Free Particles in One Dimension

Photoelectric effect
Compton effect
Uncertainty relation for photons
Stability of ground states
Bohr model
Fourier transform and uncertainty relations
Schrödinger equation
Wave equation example
Dirac delta functions
Wave functions and probability
Probabilty current/momentum equation
Time separable solutions
Completeness for particle states
Particle operator properties
Time evolution and expectation values
Problems

## 3. Some One-Dimensional Solutions to the Schrödinger Equation

Introduction
The infinite square well
The finite potential barrier
The harmonic oscillator
The attractive Kronig-Penney model
Scattering, bound states
Problems

## 4. More About Hilbert Space

Introduction and notation
Inner and outer products
Hermitian conjugation
Operator-matrix analogy
Hermitian operators and eigenkets
Schmidt orthogonalization process
Compatible Hermitian operators
Uncertainty relations and incompatible observables
Simulataneously measurable operators
Unitary transformations and change of basis
Coordinate displacements and unitary transformation
Heisenberg picture of time evolation
Constants of the motion
Free Gaussian wavepacket in Heisenberg picture
Potentials and Ehrenfest theorem
Problems

## 5. Two Static Approximation Methods

Introduction
Time independent perturbation theory
Example of perturbation theory
JWKB semiclassical approximation
Use of JWKB approximation in barrier penetration
Use of JWKB approximation in bound states
Problems

## 6. Generalization to Three Dimensions

Cartesian basis states
Wavefunctions in three dimensions
Position/momentum eigenket generalization
Example: three dimensional infinite square well
Spherical basis states
Orbital angular momentum operator
Effect of angular momentum on basis states
Energy eigenvalue equation and angular momentum
Complete set of observables for radial Schrödinger equation
Specification of angular momentum eigenstates
Angular momentum eigenvector equation
The spherical harmonics as angular momentum eigenvectors
Completeness and other properties of shperical harmonics
Radial eigenfunctions
Problems

## 7. The Three Dimensional Radial Equation

Recap of the situation
The free particle, V(r) = 0
The infinite spherical well

## 8. Addition of Angular Momenta

## 9. Spin and Statistics

**CHAPTER ONE:   Perspective and Principles**


In order to set the stage, let us begin our study of the microscopic laws of nature with some experimental indications that the laws of classical mechanics, as applied to such systems, are inadequate.  Although it seems paradoxical, we turn first to a consideration of matter in bulk to learn about microscopic behavior.  Let us consider the simplest type of bulk matter:  gases.  Indeed, let us consider first the simplest sort of gas:  monotonic.  It was known in the 19th century that for these gases one had the relationship:


Internal energy  $\propto$  Absolute temperature × Number of molecules

This can be systematized as

$$\overline{E} = (\tfrac{1}{2} kT) \times N_a \times 3 \tag{1}$$

where

$\overline{E}$ = average internal energy (per mole)

$k = 1.38 \times 10^{-16} \dfrac{\text{erg}}{\text{deg. Kelvin}}$ ("Boltzmann's constant")

$N_a = 6.022 \times 10^{23}$ ("Avagadro's number")

(Note that a "mole" is simply the number of molecules in <u>amu</u> grams of the material.  There are always $N_a = 6.022 \times 10^{23}$ molecules in a mole.  One can simply view $N_a$ as just a conversion factor from amu's (atomic mass units) to grams.)

The factor of 3 above comes from the classical equipartition theorem.  This law basically says that the average value of each independent quadratic term in the energy of a gas molecule is $\frac{1}{2}$ kT.  This comes from using Maxwell-Boltzmann statistics for a system in thermal equilibrium.  Let us use Maxwell-Boltzmann statistics to calculate the average value of a single independent quadratic energy term:

$$\overline{E}_i = \frac{\int_{-\infty}^{\infty} e^{-\beta E(q_1, q_2, \ldots, p_1, p_2, \ldots)} E_i \, dq_1 dq_2 \ldots dp_1 dp_2 \ldots}{\int_{-\infty}^{\infty} e^{-\beta E(q_1, q_2, \ldots, p_1, p_2, \ldots)} \, dq_1 dq_2 \ldots dp_1 dp_2 \ldots} \qquad (2)$$

where $\beta = \frac{1}{kT}$ .  The factor $e^{-\beta E} dq_1 \ldots dp_1 \ldots$ is proportional to the probability that the system has an energy E with position coordinates taking on values between $q_1$ and $q_1 + dq_1$, $q_2$ and $q_2 + dq_2$, and similarly for the momentum coordinates.  Just like all probabilistic considerations, our probabilities need to add to one; the denominator factor in (2) insures this. Notice that because of the large number of particles involved, Maxwell-Boltzmann statistics does not attempt to predict the motions of individual gas particles, but simply assigns a probability for a certain configuration to exist. Notice also that the "Boltzman factor", $e^{-\beta E}$ discourages exponentially the probability that the system is in an E ≫ kT state.  Now let's say that

$$E_i = ap_i^2 \text{ or } bq_i^2 \, , \qquad (3)$$

where "a" or "b" are just constants, representing a typical

kinetic or potential energy term in the total internal

energy, E.   Then we have

$$\overline{E}_i = \frac{\int_{-\infty}^{\infty} e^{-\beta a p_i^2} a p_i^2 \, dp_i}{\int_{-\infty}^{\infty} e^{-\beta a p_i^2} \, dp_i} = \frac{-a \, \frac{\partial}{\partial \beta'} \left( \int_{-\infty}^{\infty} e^{-\beta' p_i^2} \, dp_i \right)}{\int_{-\infty}^{\infty} e^{-\beta' p_i^2} \, dp_i} \, , \qquad (4)$$

where $\beta' = \beta a$.   Introduce the dimensionless variable

$x = (\beta')^{1/2} \, p_i$.   Then

$$\int_{-\infty}^{\infty} dp_i \, e^{-\beta' p_i^2} = (\beta')^{-1/2} \int_{-\infty}^{\infty} dx e^{-x^2}, \qquad (5)$$

and therefore

$$\overline{E}_i = \frac{-a \, \frac{\partial}{\partial \beta'} \, (\beta')^{-1/2}}{(\beta')^{-1/2}} = \frac{a}{2\beta'} = \frac{1}{2\beta} = \frac{1}{2} \, kT. \qquad (6)$$

If we accept the validity of the equipartition theorem,

we have that

$$\overline{E} = \overline{E}_i \times (\text{total no. of quadratic terms in E} \qquad (7)$$
$$\text{in a mole of gas}),$$

so that (remember that $\vec{p}^{\,2} = p_x^2 + p_y^2 + p_z^2$)

$$\overline{E} = \frac{3}{2} \, kN_a T$$

Let us define the "molar specific heat at constant volume"

(also called "heat capacity at constant volume"), $C_v$:

$$C_v \equiv \left( \frac{\partial \overline{E}}{\partial T} \right)_v \, . \qquad (8)$$

1.4

(The subscript "V" reminds us to keep the variable representing volume a constant during this differentiation.) In our case, for a simple monotonic gas, we get

$$C_v = \frac{3}{2} kN_a \equiv \frac{3}{2} R \left( = 12.5 \ \frac{joules}{mole \cdot deg} \right) . \qquad (9)$$

How does this simple result stack up against experiment? (Carried out at room temperature)

| monotonic gas | $C_v$(experiment) |
|---|---|
| He | 12.5 |
| Ar | 12.5 |

A success!

Well, what about diatomic molecules? To get our theoretical prediction, based on the equipartition theorem, all we need to do is just count degrees of freedom for a single molecule. If we say that the energy of such a molecule is a function of only the relative coordinate, r, separating the two atoms, then we have,

$$E = \frac{\vec{P}^2}{2M} + \frac{\vec{L}^2}{2\mu r^2} + \frac{Pr^2}{2\mu} + U(r), \qquad (10)$$

$$\Rightarrow \text{deg. of freedom* } = \quad 3 \quad + \quad 3 \quad + \quad 1 \quad , \qquad (11)$$
$$\qquad \qquad \downarrow \qquad \quad \downarrow \qquad \quad \downarrow$$
$$\text{translation} \quad \text{rotation} \quad \text{vibration}$$

_____

 * If the diatomic atoms were not *point* particles, one of these degrees of freedom would increase by one. Can you understand which one and why?

($\mu$ = reduced mass and $\vec{L}$ = center of mass angular momentum)

if we say that U(r) ~ $r^2$.  Thus for a diatomic molecule we

would expect

$$\overline{E} = (\frac{1}{2} kT)N_a \times (3 + 3 + 1) = \frac{7}{2} kN_aT. \tag{12}$$

$$\Rightarrow C_v = \frac{7}{2} R \left(= 29.1 \frac{joules}{mole \cdot deg}\right) \tag{13}$$

How does this result stack up against experiments at room

temperature?

| diatomic gas | $C_v$(experiment) |
|---|---|
| $N_2$ | 20.6 |
| $O_2$ | 21.1 |

Something is wrong.  We seem to be "missing" some degree of

freedom.  Notice that

$$\frac{5}{2} R = 20.8 \frac{joule}{mole \cdot deg}$$

seems to be a better approximation to the experimental situation

than does our $\frac{7}{2}$ R prediction.  Later considerations have shown

that the vibrational degrees of freedom are the "missing" ones.

Historically, this was the first experimental indication of a

failure in classical physics applied to atoms, and was known

already in the 1870's.

Another application of these ideas is to solids.  Let us

treat the atoms of a solid as point masses "locked in place"

to a first approximation.  Then we have for the energy, E, of

a single atom

1.6

$$E = \frac{\vec{P}^2}{2M} + ax^2 + by^2 + cz^2, \tag{14}$$

where x,y,z measure the displacement of the ideal atom from its equilibrium position.  There are now 6 quadratic degrees of freedom, which means that

$$\overline{E} = 3kN_aT \tag{15}$$

$$\Rightarrow C_v = 3R \left(= 25 \ \frac{joule}{mole \cdot deg}\right) \tag{16}$$

This law, known before the above theoretical explanation, is called the law of Dulong-Petit.  What happens in experiments, again at room temperature?

| Solid | $C_p \approx C_v$(experimental) |
|---|---|
| Copper | 24.5  (23.3) |
| Silver | 25.5 |
| Carbon (diamond) | 6.1 |

(These data have been taken from Rief, "Fundamentals of Statistical and Thermal Physics." For solids and liquids we have $C_v \approx C_p$, "$C_p$" being the molar specific heat at constant pressure, which is easier to measure than $C_v$.)  Although copper and silver seem to obey the Dulong-Petit rule, diamond obviously does not.  What is even harder to understand is that, for example, the $C_v$ for diamond is temperature dependent.  This is not accounted for by the classical physics behind the Dulong-Petit prediction of the universal value, 3R.

Although copper and silver look rather satisfactory from the point of view of the above law, there is still a paradox associated with them according to classical mechanics. If $N_a$ atoms each give up m valence electrons to conduct electricity, and if the electrons are freely mobile, the heat capacity of a conductor should be

$$C_v \;=\; 3kN_a \;+\; \frac{3}{2}\, mkN_a \qquad\qquad (17)$$

$$\downarrow \qquad\qquad\qquad \downarrow$$

"atomic" piece    "electronic" piece

Thus in these materials the electronic component of specific heat seems not to be present, or is greatly suppressed. Classical mechanics is silent as to the cause of this.

Another place that experimental results have pointed to a breakdown in the application of classical mechanics to atomic systems was in a classic experiment done by H. Geiger (of counter fame) and E. Marsden in the early part of this century. They scattered $\alpha$ particles (Helium nuclei) off of gold foil and found that a larger number of $\alpha$ particles were backscattered by the atoms from the foil than could be accounted for by then-popular atomic models. This led Rutherford to hypothesize that most of the mass of the atom is in a central core or "nucleus." Electrons were supposed to orbit the nucleus like planets around the sun in order to give atoms their known physical sizes. For example, the hydrogen atom was supposed to have a single electron in orbit around a positively charged nucleus. Although Rutherford's

conclusions came via classical reasoning (it turns out that the classical scattering cross section derived by Rutherford is essentially unmodified by the new mechanics we will study here), he could not account for the <u>stability</u> of his proposed model by classical arguments since his orbiting electrons would quickly radiate away their energy caused by their accelerated motion.

All of these experimental shortcomings, the "missing" vibrational degrees of freedom in diatomic molecules, the failure of the law of Dulong and Petit for certain solids, the missing or suppressed electronic component of $C_v$, and the instability of Rutherford's atomic model, pointed to a breakdown in classical mechanics. Thus the time was ripe for a new, more general mechanics to arise.

We will begin our study of <u>quantum</u> mechanics with another experimental finding which was at variance with classical ideas.
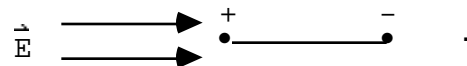
Consider the following simple, static, neutral charge distribution in an external electric field:

$$\vec{E} \quad \longrightarrow \quad \begin{array}{c} \bullet \ +e \\ \uparrow \ \vec{r}_0 \\ \bullet \ -e \end{array}$$
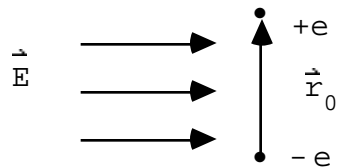
Clearly, this system prefers the orientation

$$\vec{E} \quad \Longrightarrow \quad \overset{-}{\bullet}\!\!\!\rule{2cm}{0.4pt}\!\!\!\overset{+}{\bullet}$$
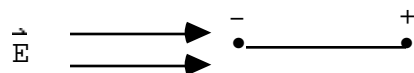
to

This system is called an electric dipole, and there is an energy associated with its orientation. We know that

$$\text{energy} = \text{charge} \times \text{potential}$$

so that ("e" is a positive charge)

$$U = e\phi(+) - e\phi(-) \tag{18}$$

Now we may expand

$$\phi(+) \cong \phi + \frac{\vec{r}_0}{2} \cdot \vec{\nabla}\phi \tag{19}$$

$$\phi(-) \cong \phi - \frac{\vec{r}_0}{2} \cdot \vec{\nabla}\phi \tag{20}$$

where $\phi$ represents the potential of the external field at the midpoint of the dipole. Then

$$\phi(+) - \phi(-) = \vec{r}_0 \cdot \vec{\nabla}\phi \ . \tag{21}$$

But by the definition of the electric field

$$\vec{E} = -\vec{\nabla}\phi \tag{22}$$

so that

$$U = -e\vec{r}_0 \cdot \vec{E} \tag{23}$$

Define $\vec{d} = e\vec{r}_0$, the "electric dipole moment." Then (23) becomes

$$U = -\vec{d} \cdot \vec{E} \ . \tag{24}$$

Eq$^n$. (24) is consistent with the picture that $\vec{r}_0$ prefers to point along $\vec{E}$ since this minimizes the potential energy.

We also know that

$$\text{Force} = \text{charge} \times \text{electric field,}$$

so

$$\vec{F} = e\vec{E}(+) - e\vec{E}(-) = e(\vec{r}_0 \cdot \vec{\nabla})\vec{E} = (\vec{d} \cdot \vec{\nabla})\vec{E} \ . \qquad (25)$$

Since $\vec{E} = -\vec{\nabla}\phi$, then we may also write this as

$$\vec{F} = -(\vec{d} \cdot \vec{\nabla})\vec{\nabla}\phi = -\vec{\nabla}(\vec{d} \cdot \vec{\nabla}\phi) = \vec{\nabla}(\vec{d} \cdot \vec{E}) \ . \qquad (26)$$

This makes sense since we expect that $\vec{F} = -\vec{\nabla}U$ and $U = -\vec{d} \cdot \vec{E}$. Notice that if $\vec{E}$ is uniform, there is no net force on the system.

There is also a torque on the system since

$$\text{Torque} = \text{lever arm} \times \text{force.}$$

Therefore

$$\vec{t} = \frac{\vec{r}_0}{2} \times \left( e\vec{E}(+) \right) + \left( -\frac{\vec{r}_0}{2} \right) \times \left( -e\vec{E}(-) \right), \qquad (27)$$

so

$$\vec{t} = \vec{d} \times \vec{E} \ , \qquad (28)$$

where the $\vec{E}$ is the value of the electric field at the center of the dipole.

In the following we will really be interested in magnetic properties of individual particles.  Rather than deriving similar formulas in the magnetic case (which is

trickier), we will simply depend on an electric-magnetic

analogy to get the formulas we need.  The analogy,

|      Electric      |      Magnetic      |
|:------------------:|:------------------:|
| $\vec{E}$          | $\vec{H}$          |
| $\vec{d}$          | $\vec{\mu}$   ,    |

where $\vec{\mu}$ is the "magnetic moment" then leads to

$$U = -\vec{\mu} \cdot \vec{H} \ , \tag{29}$$

$$\vec{F} = (\vec{\mu} \cdot \vec{\nabla})\vec{H} = \vec{\nabla}(\vec{\mu} \cdot \vec{H}) \ , \tag{30}$$

$$\vec{t} = \vec{\mu} \times \vec{H} \ . \tag{31}$$

These formulas will help us understand the behavior of

magnetic dipoles subjected to external magnetic fields. Remember,

in order to produce a force on a magnetic dipole, we must first

construct an <u>inhomogeneous</u> magnetic field.  Consider therefore

the following schematic experimental arrangement.

1.12
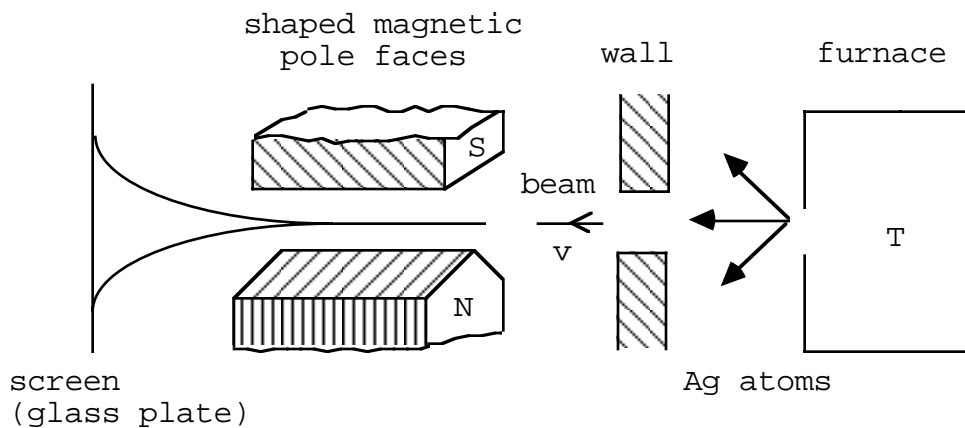
Looking face-on to the magnets, we would see the following:
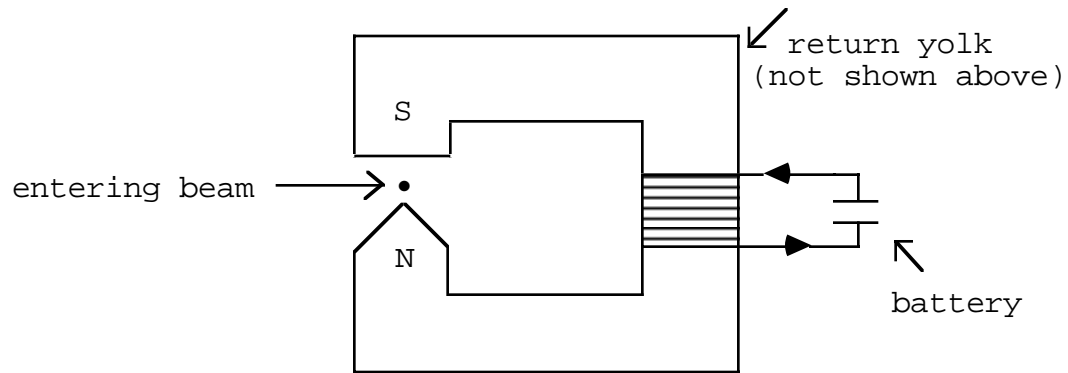


The magnetic field lines near the pole faces are highly non-uniform.  The field looks something like:
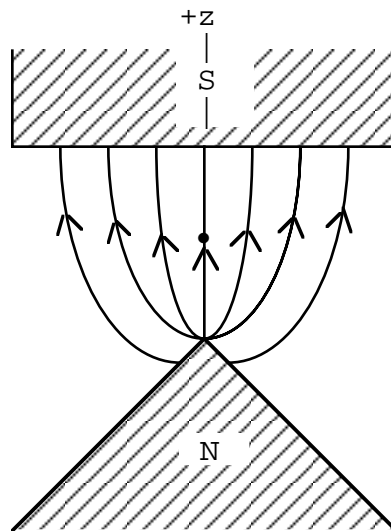


If we take a z-axis centered on the beam and directed upward as in the figure, a non-uniform magnetic field with

$$\frac{\partial H_z}{\partial z} < 0 \ ,$$

will be produced.  The type of experimental setup suggested above was first used by Otto Stern and Walther Gerlach in an

experiment on Ag (silver) atoms in 1922.  The explanation for
their experimental results had to wait until 1925 when Samuel
Goudsmit and George Uhlenbeck, on the basis of some atomic
spectrum considerations, deduced the physical property
responsible.  In the following, we will ignore the
experimental details of this experiment, and will be
considering idealized Stern-Gerlach-like experiments.

From (30), the force on an Ag atom at a single instant
in time is approximately ( $H_z \rightarrow H$ )

$$F_z \approx \frac{\partial}{\partial z} \mu_z H = \mu_z \frac{\partial H}{\partial z} \; . \tag{32}$$

One can imagine measuring the force on a given atom by its
deflection in the magnetic field:

$$\frac{dp_z}{dt} \approx \mu_z \frac{\partial H}{\partial z} \; . \tag{33}$$

Let us assume that the quantity $\frac{\partial H}{\partial z}$ is approximately a

constant in time, fixed by the experimental apparatus.  Then,
we have a situation that looks like:
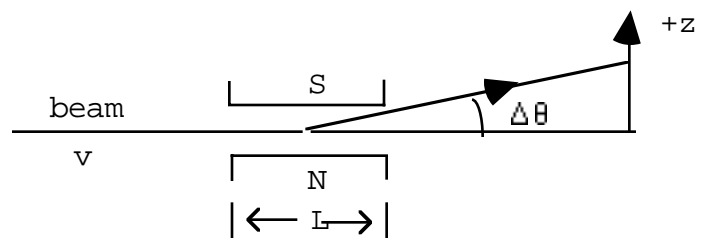


The change in the z-component of the momentum of an Ag atom
is then

1.14

$$\Delta p_z \approx \mu_z \frac{\partial H}{\partial z} t .$$ (34)

But

$$t \approx \frac{L}{v} ,$$ (35)

where v is the velocity of the atoms, so that

$$\Delta p_z \approx \mu_z \frac{\partial H}{\partial z} \frac{L}{v} .$$ (36)

The small angular deflection caused by the magnetic field is

then

$$\Delta\theta \approx \frac{|\Delta p_z|}{|\vec{p}|} \approx \left| \mu_z \frac{\partial H}{\partial z} \right| \frac{L}{mv^2} .$$ (37)

Let us get some numerical feeling for this situation.

The particular values we will take in the following are:

$m = 1.79 \times 10^{-22}$ gm (Ag atom mass)

$T = 10^3 \ ^0K$ (furnace temp.)

$\left| \frac{\partial H}{\partial z} \right| = \frac{10^3 \text{ gauss}}{10^{-1} \text{ cm}} = 10^4 \frac{\text{gauss}}{\text{cm}}$ (field gradient)

$L = 10$ cm (magnet length)

$|\mu_z| \approx 10^{-20} \frac{\text{erg}}{\text{gauss}}$ (Ag z-component magnetic moment)

Using these values, we can estimate the angular derivation $\Delta\theta$

as follows. From the equipartition theorem, we expect the

mean energy of an Ag atom leaving the furnace to be

$$\frac{1}{2} m \overline{v^2} = \frac{3}{2} kT.$$ (38)

which gives

$$\overline{m\,v^2} = 4.14 \times 10^{-13}\,\text{erg}.$$

Then from (37)

$$\Delta\theta \approx \frac{10^{-20}\cdot 10^4 \cdot 10}{4.14 \times 10^{-13}} = 2.4 \times 10^{-3}\ \text{radians},$$

or about .14°.  Naively, we would always expect to be able to back off far enough from the magnets to see this deflection.

Classically, what would we expect to see on the glass screen as a result of the beam of Ag atoms passing through the magnetic field?  Since the atoms will emerge from the furnace with randomly oriented $\mu_z$'s, and since, from (37), we expect the deflection of a given particle to be proportional to $\mu_z$, the classical expectation was to see something like:

|

.

However, our idealized experiment will actually yield only 2 spots:

●

●      .

In a real experiment, the "spots" above would be smeared because of the spread in particle velocities from the furnace and the nonuniformity of the magnetic field.  (We will discuss another source of smearing in just a moment.) Originally, this unexpected two-value-only result was referred to as "space quantization".  However, this is a misleading name since the thing which is quantized here is certainly not space.

1.16

   Now let us catalog some experimental results from other
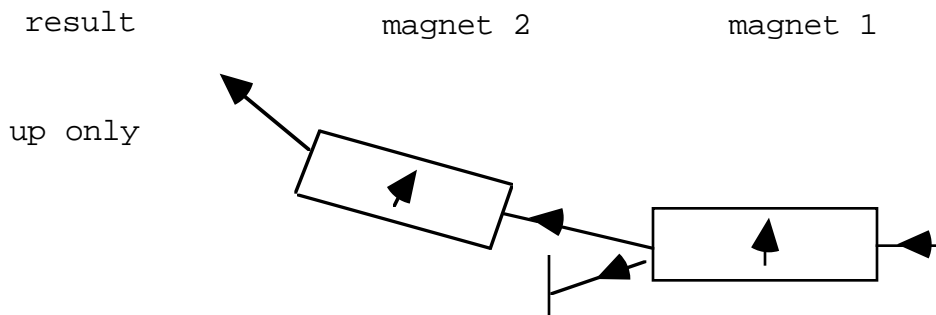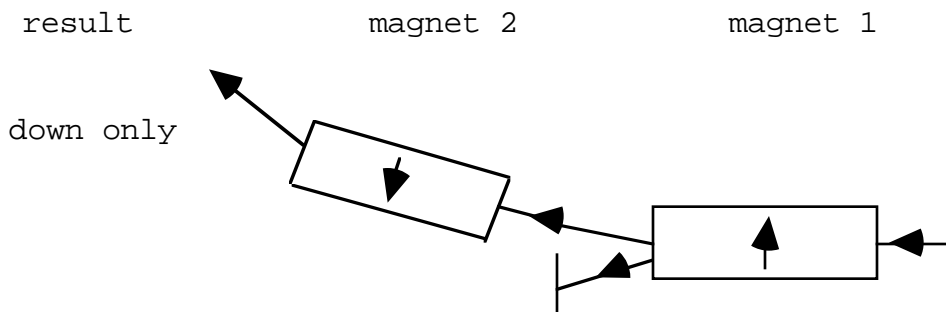setups of Stern-Gerlach apparatuses.

(a)  First, rotate magnet.
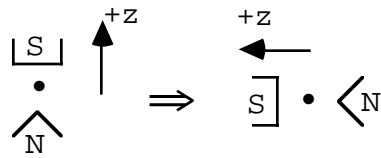


We would see that the beam is now split along the new z-axis.

   Let us now add a second magnet to the system at various
orientations relative to the first.  Let θ represent the
angular orientation of magnet 1 with respect to magnet 2.
For three specific orientations, one finds the following
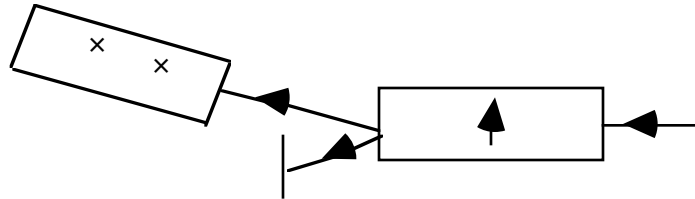experimental results for the intensity of the outgoing beam:


b)  θ = 0°

         result              magnet 2              magnet 1

         up only



c)  θ = 180°

         result              magnet 2              magnet 1

         down only

d) $\theta = 90°$

50% up
50% down



In fact, for an arbitrary orientation $\theta$, the intensity of the "up" orientation is $\cos^2 \frac{\theta}{2}$ and of the "down" is $\sin^2 \frac{\theta}{2}$ .

Here "result" means whether the final beam emerges in an up or down orientation relative to the <u>second</u> magnet.
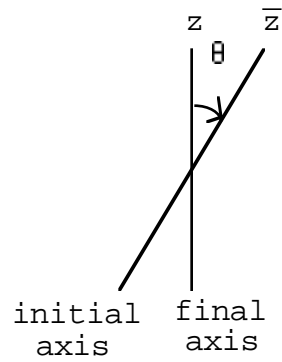
As mentioned above, one would measure the intensity of the outcoming beams to reach these conclusions.  However, let us accept the fact that our description of what is occurring must be based on <u>probabilities</u>. Instead of the <u>intensity</u> of a <u>beam</u> of particles, let's talk about <u>intrinsic</u> <u>probabilities</u> associated with <u>individual, independent particles</u>.
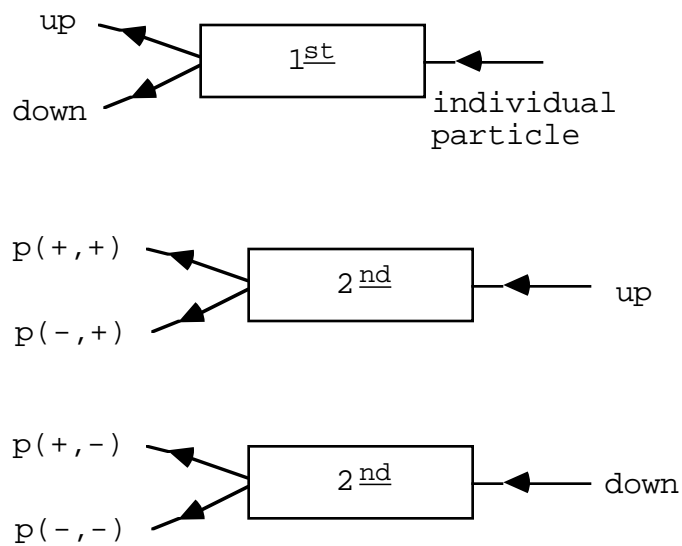
Let's define

$z$ $\bar{z}$

$p(\pm,\pm)$: probability that a particle deflected in the $\pm\bar{z}$ direction from the first S-G gives a particle deflected in the $\pm z$ direction relative to the second S-G. (S-G = Stern-Gerlach experiment)

The axes are related like:

1.18



                    z    z̄
                     θ

            initial  final
             axis     axis

There are 4 probabilities here:



up

down                    individual
                        particle



$p(+,+)$

$p(-,+)$              up



$p(+,-)$

$p(-,-)$              down

From the above we identify $p(+,+) = \cos^2 \dfrac{\theta}{2}$ and $p(-,+) = \sin^2 \dfrac{\theta}{2}$ .

    We must have our probabilities adding to one.

Therefore, we must have

$$p(+,+) + p(+,-) = 1$$

$$\Rightarrow \cos^2 \frac{\theta}{2} + p(+,-) = 1$$

$$\Rightarrow p(+,-) = \sin^2 \frac{\theta}{2} \ .$$

Also

$$p(-,+) + p(-,-) = 1$$

$$\Rightarrow p(-,-) = \cos^2 \frac{\theta}{2} .$$

More abstractly, we have (a' = + -, a" = + - independently)

$$\sum_{a'} p(a",a') = 1, \tag{39}$$

and

$$\sum_{a"} p(a",a') = 1. \tag{40}$$

Notice that

$$p(a',a") = p(a",a'), \tag{41}$$

and that using (41),(40) follows from (39) or vice versa.
Thus (41) may be viewed as a way of ensuring probability
conservation.   Therefore, only one probability is
independent, p(+,+) say; the rest follow from (39) and (41)
(or (40) and (41)).

From (32) we realize that the upward deflected beam is
associated with $\mu_z$ < 0, while the downward beam must have
$\mu_z$ > 0.   We now ask the question:   given the selection of the
up beam along the initial $\bar{z}$ axis ($\mu_{\bar{z}}$ < 0), what is the mean
value expected for $\mu_z$ measured along the final z axis?   The
situation looks like:
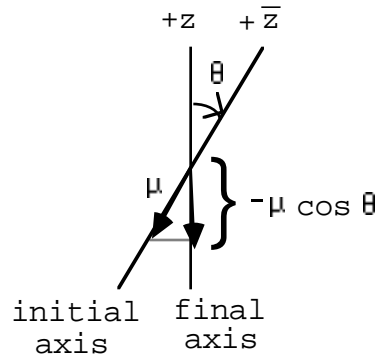
Classically, the answer to this question is given by just picking out the projection of $\vec{\mu}$ along the z-axis.  Thus, the classical answer is $-\mu \cos \theta$.  What do we get from our new, probability point of view?  From this point of view, our mean value of $\mu_z$ is the weighted average of the two probabilities for finding an upward deflected beam from the 2nd S-G ($\mu_z < 0$) and a downward beam from the 2nd S-G ($\mu_z > 0$).  Therefore, we have

$<\mu_z>_+ \equiv$ average value of $\vec{\mu}$ along the z-axis,

given an initial selection of the upward

deflected beam along $\bar{z}$.

$$<\mu_z>_+ \ = \ (-\mu)p(+,+) \ + \ (+\mu)p(-,+)$$

$$= \ -\mu \left( \cos^2 \frac{\theta}{2} \ - \ \sin^2 \frac{\theta}{2} \right)$$

$$= \ -\mu \cos \theta$$

We thus get the same result as expected classically, although the way we have reached our conclusion is not classical at all.

Let us try to build a classical model of the basic S-G experiment.  Magnetic moments classically are produced by the motion of charged particles.  (There are no magnetic monopoles, at least so far.)  A reasonable connection is thus that

$$\vec{\mu} \ = \ \gamma \vec{S} \ . \tag{42}$$

where $\vec{S}$ is a type of angular momentum associated with the Ag
atom.  The symbol "γ" above is just a proportionality
constant, usually called the "gyromagnetic ratio".  What "$\vec{S}$"
represents is not clear yet.  Eq$^{\underline{n}}$ (42) is patterned after a
classical result.  If one has a current loop in a plane,



,
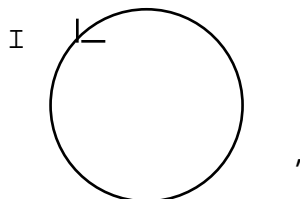
where the elementary charge carriers have a charge -e like an
electron (e > 0 here), the magnetic moment produced by these
moving charges is

$$\vec{\mu} = - \frac{e}{2mc} \vec{L} ,\qquad\qquad (43)$$

where m refers to the charge carrier's mass.  (See Jackson's
Classical Electrodynamics, second edition, p.183).  If (42)
holds for the Ag atom, because the beam is seen to split into
two discrete components, we can associate discrete values of
$S_z$ with the two spots observed.  This behavior of Ag atoms in
a magnetic field is due to its internal structure:  one
unpaired electron outside a closed shell of electrons (which
possess no net magnetic dipole moment).  Thus, the property
of the Ag atoms we are studying is really due to a property
of the electron.  This property, called "spin", sounds very
classical, but is far from being a classically behaving
angular momentum.  Since the magnetic moment we are measuring
in the S-G experiment really refers to a property of

electrons, it is natural to expect that the gyromagnetic ratio in (42) be not too different from the classical one in (43), which also refers to the electron. In fact, the actual gyromagnetic ratio is approximately a factor of two larger than (43):

$$\gamma \approx - \frac{e}{mc} \: .$$ 

(44)

Given this value of $\gamma$ and the experimental determination of the deflection angle $\Delta\theta$ in (37), one can deduce the allowed values of the electron spin along the z-axis:

$$\mu_z = \gamma S_z < 0, \; S_z = \frac{\hbar}{2} \, ,$$

$$\mu_z = \gamma S_z > 0, \; S_z = -\frac{\hbar}{2} \: .$$
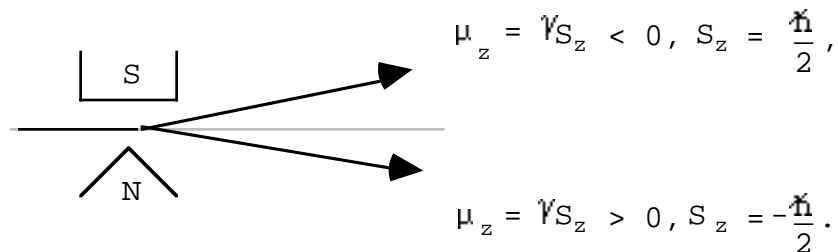
The quantity "h" $\left( \hbar \equiv \frac{h}{2\pi} \right)$ is known as <u>Planck's constant</u>. The z-component of the electron's spin is thus <u>quantized</u>, i.e. limited to the two discrete values $\frac{\hbar}{2}$ and $- \frac{\hbar}{2}$.

Let us continue to develop our classical model. From (31) we have $\vec{t} = \vec{\mu} \times \vec{H}$. From (42) we have $\vec{\mu} = \gamma\vec{S}$, so

$$\vec{t} = \gamma\vec{S} \times \vec{H} \: .$$

(45)

Newton's laws relate $\vec{t}$ to the rate of change of angular momentum,

$$\vec{t} = \frac{d\vec{S}}{dt} \: .$$

(46)

Putting (45) and (46) together gives

$$\frac{d\vec{S}}{dt} = \gamma \vec{S} \times \vec{H} \; . \tag{47}$$

Let us take $\vec{H} = H\hat{e}_z$, where H is a constant.  Then we have

$$\frac{dS_z}{dt} = 0; \quad \frac{dS_x}{dt} = \gamma H S_y; \quad \frac{dS_y}{dt} = -\gamma H S_x. \tag{48}$$

Then for example

$$\frac{d^2 S_x}{dt^2} = \gamma H \frac{dS_y}{dt} = -(\gamma H)^2 S_x. \tag{49}$$

This is a differential equation of the form $\ddot{x} + \omega^2 x = 0$,

where the angular frequency is given by

$$\omega = \gamma H. \tag{50}$$

The picture that emerges is that of a precessing $\vec{S}$

vector:



direction of precession
if $\gamma < 0$.
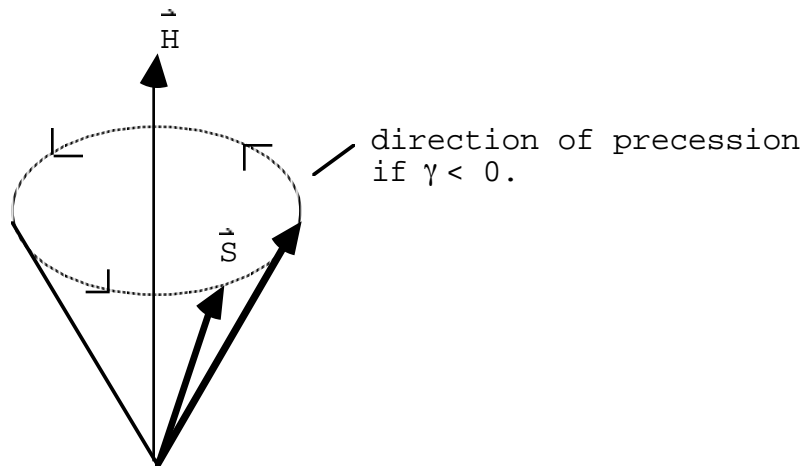
Notice that neither $|\vec{S}|$ nor $S_z$ changes in time.  Since

the time to pass through the magnet poles is given in (35) as
$t \approx \frac{L}{v}$, the <u>total</u> pression angle for an Ag atom is

$$\phi = \omega t \approx \gamma H \frac{L}{v} \; . \tag{51}$$

Again, let's gets some feeling for order of magnitude here. Using our previous result for v (below eq$^n$(38) above), we get

$$\phi = \frac{1.76 \times 10^7 \cdot 10^3 \cdot 10}{4.81 \times 10^4} = 3.7 \times 10^6 \text{ radians!}$$

This is equal to $5.8 \times 10^5$ complete revolutions.

In order to see how far we can push this classical description of spin, we would like to try to "catch" an atom while in the act of rotating. Classically, one should in principle be able to accomplish this by, say, decreasing the value of H and L in (51). Then, the deflection angle (37) will become smaller, but one can always move the screen far enough away to see such a deflection. However, nature makes it impossible to accomplish this goal. To see why, let us examine the experimental arrangement in more detail.

In calculating the deflection angle, $\Delta\theta$, we have assumed we know exactly where the atom is in the magnetic field. In fact, we don't know exactly where an individual atom is since the wall the beam had to pass through actually has a finite width.
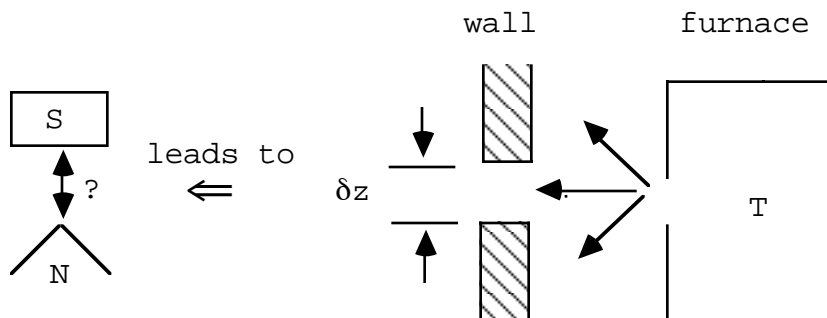
"δz" represents the finite width of the slit.  In our
idealized experiment, up to this point, we have been
imagining two separate operations to be done on the beam:
first, collimation by the wall; second, the measurement done
on the beam by the magnets.  Let us idealize our experiment
even further.  Imagine that the action of the thin wall and
the beginning of the effect of the magnets on the beam both
take place at the same time, or at least approximately
simultaneously.  Then δz represents an uncertainty in the
position of the Ag atoms as they begin their traverse through
the magnetic field.  Because of the gradient in H, this will
cause an uncertainty in the value of the field acting on the
atoms,

$$\delta H \;=\; \frac{\delta H}{\delta z}\;\delta z.$$

This then implies an uncertainty in the precession angle

$$\delta \phi \;=\; \gamma \delta H\;\frac{L}{v} \;=\; \gamma\;\frac{\partial H}{\partial z}\;\frac{L}{v}\;\delta z. \tag{52}$$

Along with the uncertainty in position, δz, there is
also an uncertainty in z-component of momentum, $\delta p_z$, of the Ag
particles after they have emerged from the slit.
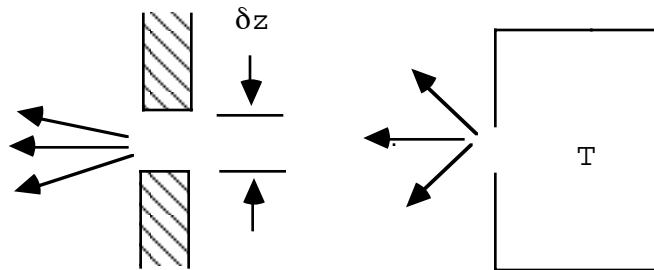
This spread in momentum values will, in fact, wash out our magnetically split beam if it is too large. In order to insure that the experiment works, i.e., that the beam is split so we can tell which way an individual atom is rotating, we need that

$$(\Delta p_z)_+ - (\Delta p_z)_- > \delta p_z, \tag{53}$$

where the $(\Delta p_z)_\pm$ represent the up(+) or down (-) "kick" given to the atoms by the field. From (36) we know that (remember, $\mu_z = \gamma S_z$, with $S_z = \frac{\hbar}{2}$ or $-\frac{\hbar}{2}$ ).

$$(\Delta p_z)_+ = \gamma \, \frac{\hbar}{2} \, \frac{\partial H}{\partial z} \, \frac{L}{v} \, , \tag{54}$$

$$(\Delta p_z)_- = -\gamma \, \frac{\hbar}{2} \, \frac{\partial H}{\partial z} \, \frac{L}{v} \, , \tag{55}$$

From (52) we then have that

$$(\Delta p_z)_+ - (\Delta p_z)_- = \frac{\hbar \delta\phi}{\delta z} \, . \tag{56}$$

Eq$^{\underline{n}}$(53) now says that for the experiment to work, we must have

$$\frac{\hbar \delta\phi}{\delta z} > \delta p_z \, ,$$

$$\text{or} \quad \hbar \delta\phi > \delta p_z \delta z. \tag{57}$$

If nature is such that

$$\delta p_z \delta z \geq \hbar \, , \tag{58}$$

then we must conclude that

$$\delta\phi \geq 1, \tag{59}$$

can not be avoided.  Relations such as (58) or (59) are
called "uncertainty relations," and are an intrinsic part of
quantum theory. Eq$^{\underline{n}}$(58) is Heisenberg's famous
momentum/position uncertainty relation which will be
motivated and discussed extensively in the upcoming chapters.
Given this input, (59) says that the classical picture of a
rotating spin angular momentum, whose precession angle should
be arbitrarily localizable, is untenable.  Ag atoms are <u>not</u>
behaving as just scaled-down classical tops; we cannot
"catch" an Ag in the act of rotating.

As said before, the name "spin", when applied to a
particle like an electron, sounds classical but it is not.
It is, in fact, impossible to construct a classical picture
of an object with the given mass, charge and angular momentum
of an electron. Let's try to.  Consider the following
electron model.
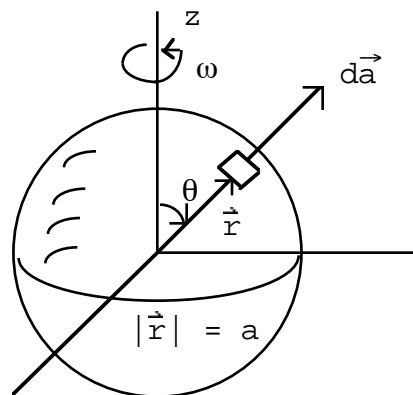
The "electron" consists of an infinitely thin, spherical shell of charge, spinning at the rate that gives it an angular momentum along z of $\frac{\hbar}{2}$ . The moment of inertia of this system is

$$I = \rho_s \int_s (a^2 - z^2) ds \tag{60}$$

where we are doing a surface integral, and

$$ds = a^2 \sin\theta d\theta d\phi,$$

$$z = a \cos\theta,$$

$$\rho_s = \frac{m}{4\pi a^2} .$$

Doing the integral gives $\left( \int_0^\pi \sin^3\theta d\theta = \frac{4}{3} \right)$

$$I = \frac{2}{3} ma^2 \tag{61}$$

Classically, we have (for a principle axis)

$$L = I\omega. \tag{62}$$

Setting $L = \frac{\hbar}{2}$ , we find that the classical electron's angular velocity must be

$$\omega = \frac{3\hbar}{4ma^2} . \tag{63}$$

Now it takes energy in order to assemble this positive shell of charge because of the electrostatic forces of repulsion. This energy must be on the order of

$$E \sim \frac{e^2}{a} . \tag{64}$$

We know from special relatively that mass and energy are equivalent ($E = mc^2$). Thus (64) gives a mass for the

electron, which, if we hypothesize supplies the entire
observed electron mass, implies a radius

$$a \sim \frac{e^2}{mc^2} \,. \tag{65}$$

(This is called the "classical electron radius").  But now
notice that the velocity of the electron's surface at its
"equator" is given by

$$\omega a \sim \frac{3\hbar}{4ma} = \left(\frac{3}{4\alpha}\right) c \sim 103c!$$

The surface is moving much faster than the speed of light,
which is impossible by special relativity.  The constant

$$\alpha = \frac{e^2}{\hbar c}$$

is called the <u>fine structure constant</u> and has the approximate
value $\alpha \approx \frac{1}{137}$ .  The impossible surface speed of the electron
is not the only thing wrong with this model; we are still
stuck with the classical result (43) for the magnetic moment
produced by this spinning charge distribution, which gives
the wrong magnetic moment.

Our conclusion is that electron spin, called "spin $\frac{1}{2}$ "
since $S_z = \pm \frac{\hbar}{2}$ only, is a completely non-classical concept.
Its behavior (as in the S-G setup) and its origin (as above)
are not accounted for by classical ideas.

Now let's go back to the S-G experiment again and look
at it from a more general coordinate system.  Our
experimental results from the two magnet S-G setup are:

$$p(+,+) = \cos^2 \frac{\theta}{2} \qquad\qquad (66)$$

$$p(-,+) = \sin^2 \frac{\theta}{2} \ .$$

The "+" or "-" are labeling whether the particles are deflected "up" ($\mu_z < 0$) or "down" ($\mu_z > 0$) respectively. We now know that the upward deflected particles have $S_z = +\frac{\hbar}{2}$ and the downward ones have $S_z = -\frac{\hbar}{2}$ . Instead of regarding the signs in (66) as labels of being <u>deflected</u> up or down, let us regard them instead as labeling the value of the selected $S_z$ value in units of $\frac{\hbar}{2}$ . (We often call $S_z = +\frac{\hbar}{2}$ spin "up" and $S_z = -\frac{\hbar}{2}$ spin "down"). The result (66) can also be written

$$p(+,+) = \frac{1 + \cos \theta}{2} \ , \ p(-,+) = \frac{1 - \cos \theta}{2} \ . \qquad (67)$$

Remember, "$\theta$ " is the relative rotation angle of magnet 1 with respect to magnet 2. Picking our z-axis arbitrarily compared to the two S-G apparatuses leads to the picture:
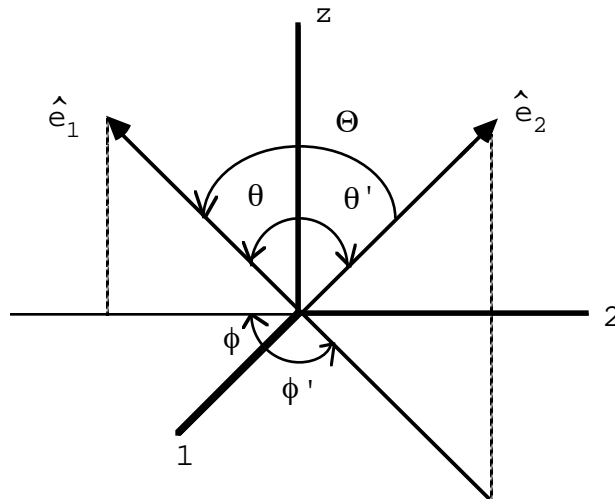
Assuming that the $\theta$'s in (67) retain their meaning as the relative rational angle, the probabilities in this new picture should be written as

$$p(+,+) = \frac{1 + \cos \Theta}{2}, \quad p(-,+) = \frac{1 - \cos \Theta}{2}. \tag{68}$$

An identity relating $\Theta$ to the other angles in the above diagram is

$$\hat{e}_1 \cdot \hat{e}_2 = \cos \Theta = \sin \theta \cos \phi \sin \theta' \cos \phi'$$

$$+ \sin \theta \sin \phi \sin \theta' \sin\phi' + \cos \theta \cos \theta', \tag{69}$$

$$\Rightarrow \cos \Theta = \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos(\phi - \phi'). \tag{70}$$

Let's write $\cos \Theta$ as

$$\cos \Theta = \overbrace{\left(\cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2}\right)}^{(\cos \theta)} \overbrace{\left(\cos^2 \frac{\theta'}{2} - \sin^2 \frac{\theta'}{2}\right)}^{(\cos \theta')}$$

$$+ \underbrace{2 \sin \frac{\theta}{2} \cos \frac{\theta}{2}}_{(\sin \theta)} \cdot \underbrace{2 \sin \frac{\theta'}{2} \cos \frac{\theta'}{2}}_{(\sin \theta')} \cos(\phi - \phi')$$

or

$$\cos \Theta = \cos^2 \frac{\theta}{2} \cos^2 \frac{\theta'}{2} + \sin^2 \frac{\theta}{2} \sin^2 \frac{\theta'}{2}$$

$$- \sin^2 \frac{\theta}{2} \cos^2 \frac{\theta'}{2} - \cos^2 \frac{\theta}{2} \sin^2 \frac{\theta'}{2}$$

$$+ 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \cdot 2 \sin \frac{\theta'}{2} \cos \frac{\theta'}{2} \cos(\phi - \phi')$$

Using $1 = \left(\cos^2 \frac{\theta}{2} + \sin^2 \frac{\theta}{2}\right)\left(\cos^2 \frac{\theta'}{2} + \sin^2 \frac{\theta'}{2}\right)$ , we get

$$p(+,+) = \frac{1 + \cos \Theta}{2} = \cos^2 \frac{\theta}{2}\cos^2 \frac{\theta'}{2} + \sin^2 \frac{\theta}{2}\sin^2 \frac{\theta'}{2}$$

$$+ 2 \cos \frac{\theta}{2} \cos \frac{\theta'}{2} \sin \frac{\theta}{2} \sin \frac{\theta'}{2} \cos(\phi - \phi') \ . \qquad (71)$$

Now suppose that $(\phi - \phi') = 0$. Then we can write $p(+,+)$ in a matrix formulation as

"row matrix"

$$p(+,+) = \left| \left(\cos \frac{\theta'}{2}, \sin \frac{\theta'}{2}\right) \cdot \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} \right|^2 . \qquad (72)$$

"column" matrix

(The absolute value signs are not needed here, yet.) This interesting structure can be repeated when $(\phi - \phi') \neq 0$. Consider the quantity

$$Q = \left| \left(e^{-i\phi'/2} \cos \frac{\theta'}{2}, e^{i\phi'/2} \sin \frac{\theta'}{2}\right)^* \cdot \begin{pmatrix} e^{-i\phi/2} \cos \frac{\theta}{2} \\ e^{i\phi/2} \sin \frac{\theta}{2} \end{pmatrix} \right|^2 \qquad (73)$$

Working backward, we can express this as

$$Q = \left| \cos \frac{\theta}{2} \cos \frac{\theta'}{2} + \sin \frac{\theta}{2} \sin \frac{\theta'}{2} e^{i(\phi - \phi')} \right|^2 . \qquad (74)$$

Now

$$|a + b|^2 = (a + b)^*(a + b)$$

$$= |a|^2 + |b|^2 + \underbrace{a*b + b*a}_{2\,\text{Re}(a*b)} , \qquad (75)$$

$$\Rightarrow Q = \cos^2 \frac{\theta}{2} \cos^2 \frac{\theta'}{2} + \sin^2 \frac{\theta}{2} \sin^2 \frac{\theta'}{2}$$

$$+ 2 \cos \frac{\theta}{2} \cos \frac{\theta'}{2} \sin \frac{\theta}{2} \sin \frac{\theta'}{2} \cos(\phi - \phi'). \qquad (76)$$

This is just the form we want!  In general then

$$p(+,+) = \left| \left( e^{-i\phi'/2} \cos \frac{\theta'}{2}, e^{i\phi'/2} \sin \frac{\theta'}{2} \right)^* \cdot \begin{pmatrix} e^{-i\phi/2} \cos \frac{\theta}{2} \\ e^{i\phi/2} \sin \frac{\theta}{2} \end{pmatrix} \right|^2 . \quad (77)$$

This factored sort of form would not have been possible without using complex numbers.  This is actually a general lesson about quantum mechanics:  complex numbers are a necessity.

Now let's do the same thing for $p(-,+)$:

$$p(-,+) = \frac{1 - \cos \Theta}{2} = \cos^2 \frac{\theta}{2} \sin^2 \frac{\theta'}{2} + \sin^2 \frac{\theta}{2} \cos^2 \frac{\theta'}{2}$$

$$- 2 \cos \frac{\theta}{2} \sin \frac{\theta'}{2} \sin \frac{\theta}{2} \cos \frac{\theta'}{2} \cos(\phi - \phi'),$$

$$\rightarrow p(-,+) = \left| - \sin \frac{\theta'}{2} \cos \frac{\theta}{2} + \cos \frac{\theta'}{2} \sin \frac{\theta}{2} e^{i(\varphi - \varphi')} \right|^2 . \quad (78)$$

With a little hindsight, this can be seen to be equivalent to

$$p(-,+) = \left| \left( -e^{-i\phi'/2} \sin \frac{\theta'}{2}, e^{i\phi'/2} \cos \frac{\theta'}{2} \right)^* \cdot \begin{pmatrix} e^{-i\phi/2} \cos \frac{\theta}{2} \\ e^{i\phi/2} \sin \frac{\theta}{2} \end{pmatrix} \right|^2 . \quad (79)$$

Let's define the column matrices

$$\psi_+(\theta,\phi) = \begin{pmatrix} e^{-i\phi/2} \cos \frac{\theta}{2} \\ e^{i\phi/2} \sin \frac{\theta}{2} \end{pmatrix} , \quad (80)$$

$$\psi_-(\theta,\phi) = \begin{pmatrix} -e^{-i\phi/2} \sin \dfrac{\theta}{2} \\[2ex] e^{i\phi/2} \cos \dfrac{\theta}{2} \end{pmatrix} . \tag{81}$$

Then we may write (the explicit matrix indices are not shown):

$$p(+,+) = \left| \psi_+(\theta',\phi')^+ \, \psi_+(\theta,\phi) \right|^2, \tag{82}$$

$$p(-,+) = \left| \psi_-(\theta',\phi')^+ \, \psi_+(\theta,\phi) \right|^2, \tag{83}$$

where "$^+$" means "complex conjugation + transpose."  (The transpose of a column matrix is a row matrix.)

In general one may show that (a',a" = +- independently)

$$p(a'',a') = \left| \psi_{a''}(\theta',\phi')^+ \, \psi_{a'}(\theta,\phi) \right|^2. \tag{84}$$

In order to make sure we haven't made a mistake, set $\theta' = 0$ in the above expressions.  We should recover our old results, since this means the z-axis is now taken along the $\hat{e}_2$ direction (i.e., along the direction of the field in the final S-G apparatus).  From (77) we get

$$p(+,+) = \left| \left( e^{-i\phi'/2} , 0 \right)^* \begin{pmatrix} e^{-i\phi/2} \cos \dfrac{\theta}{2} \\[2ex] e^{i\phi/2} \sin \dfrac{\theta}{2} \end{pmatrix} \right|^2$$

$$= \left| e^{i(\phi' - \phi)/2} \cos \dfrac{\theta}{2} \right|^2 = \cos^2 \dfrac{\theta}{2} .$$

From (79) we get

$$p(-,+) = \left| \left( 0 \;,\; e^{i\phi'/2} \right)^{*} \begin{pmatrix} e^{-i\phi/2} \; \cos \dfrac{\theta}{2} \\[2mm] e^{i\phi/2} \; \sin \dfrac{\theta}{2} \end{pmatrix} \right|^{2}$$

$$= \left| e^{-i(\phi' - \phi)/2} \; \sin \dfrac{\theta}{2} \right|^{2} = \sin^{2} \dfrac{\theta}{2} \;.$$

No mistakes.

The $\psi_{a'}(\theta,\phi)$ are called "wave functions." In order to find an interpretation for such objects, as well as to learn about other aspects of quantum mechanical systems, we will now try to generalize our S-G type of measurements.

Before, in the S-G case, we were measuring $S_z$ (or $\mu_z$). The physical outcomes were $S_z = \dfrac{\hbar}{2}$ or $S_z = -\dfrac{\hbar}{2}$. The whole measurement can be idealized as:
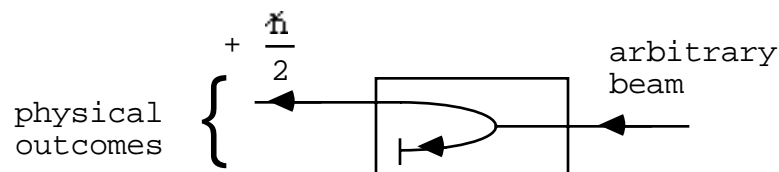


The line entering the box is indicative of a beam of particles entering a S-G apparatus. The separation of the beam suggests the effect of the magnets on the atoms. In addition a selection is being performed whereby only particles with a given physical attribute ($S_z = \dfrac{\hbar}{2}$, say) are permitted to exit.

Let us generalize the above as
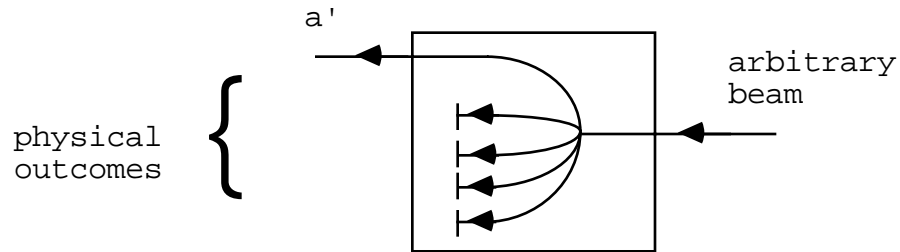
Just as the above S-G apparatus selects the outcome of $\frac{\hbar}{2}$
of the physical property $S_z$, we are imagining the above setup
to select an outcome a' of some more general physical property
A.  We will adopt a symbol which represents the above process,
and call it a "measurement symbol" or an "operator."  The
measurement symbol for the above is:

$$|a'|$$

We will let a',a",... be typical outcomes of such measurements;
we will sometimes explicitly label specific outcomes as
$a_1,a_2,...$ .  For right now think of the outcomes a',a"... as
dimensionless numbers, to keep things simple.

What sort of manipulations are appropriate to these
measurement symbols?  Consider the S-G type of process:
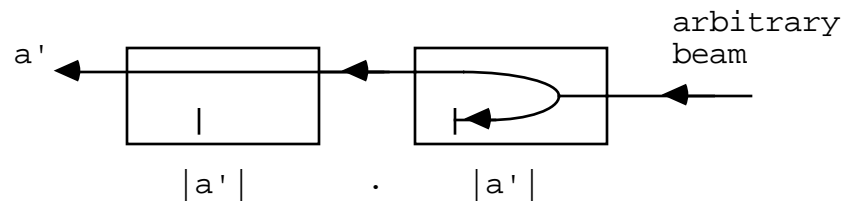


This is clearly the same as just
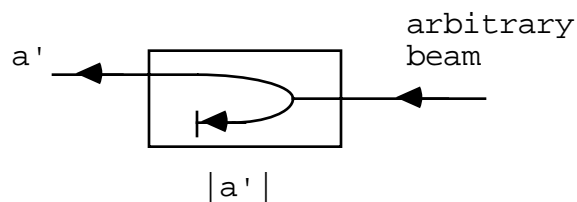
This suggests the rule:

$$|a'||a'| = |a'|. \tag{85}$$

On the other hand, consider (a' ≠ a"):



$$|a''| \qquad \cdot \qquad |a'|$$
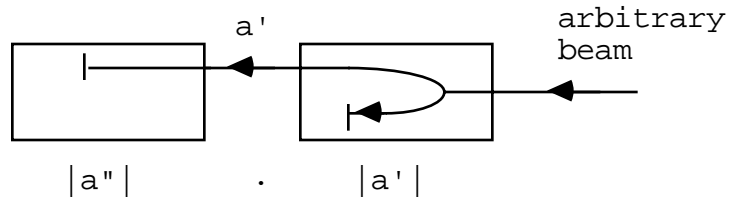
This is equivalent to an apparatus which blocks everything:
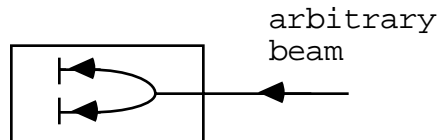


We will call the above a "null measurement" and associate it with the usual null symbol, 0.  Therefore, we adopt

$$|a''||a'| = 0, \ a' \neq a''. \tag{86}$$

Notice

$$|a'||a''| = |a''||a'|, \tag{87}$$

which says that selection experiments are commutative.  This defines multiplication in this context.  What about addition? Let's start at the opposite end to the null measurement in a system with 4 physical outcomes, say
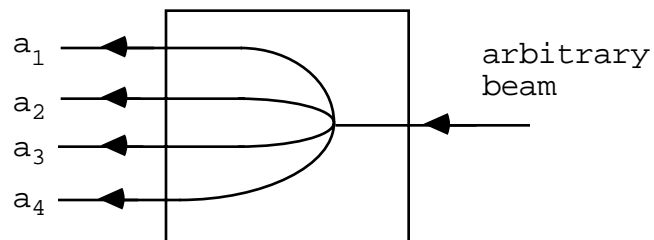
Such a measurement apparatus can perform a separation, but
no selection.  Our symbol for this will be the usual identity
character:  1.  Clearly, we have

$$1 \cdot 1 = 1. \tag{88}$$

Now start blocking out physical outcomes one by one:

$a_1$      $a_2$      $a_3$           a.b.

Symbol:  $1 - |a_4|$

$a_1$      $a_2$           a.b.
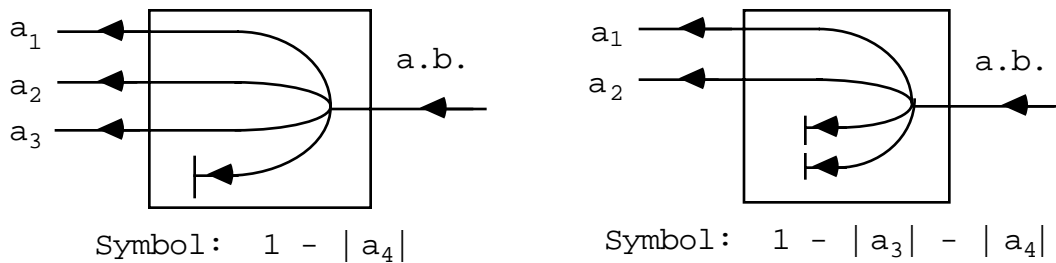
Symbol:  $1 - |a_3| - |a_4|$

("a.b." above means "arbitrary beam".)  Now, block all of the
outcomes:

a.b.
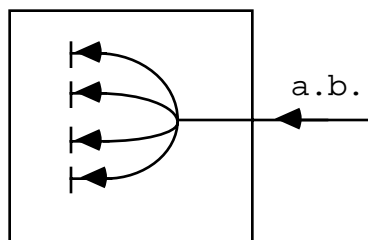
Symbol:  $1 - |a_4| - |a_3| - |a_2| - |a_1|$

This is obviously just the null measurement again.  The
two characterizations must be the same:

$$1 - \sum_i |a_i| = 0. \tag{89}$$

We require that

$$0 + |a'| = |a'|, \tag{90}$$

so (89) can be written

$$\sum_i |a_i| = 1. \tag{91}$$

Eq$^{n}$(91) will be called "completeness." It is no exaggeration to say it is the foundation stone of all of quantum mechanics.

We can write down other mathematical results suggested by the above type of diagrams.  Consider



$$1 \quad \cdot \quad |a_1'|$$
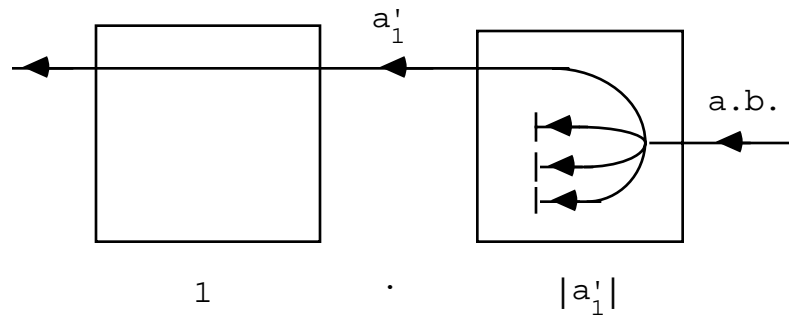
This is clearly equivalent to the opposite order,



$$|a_1'| \quad \cdot \quad 1$$
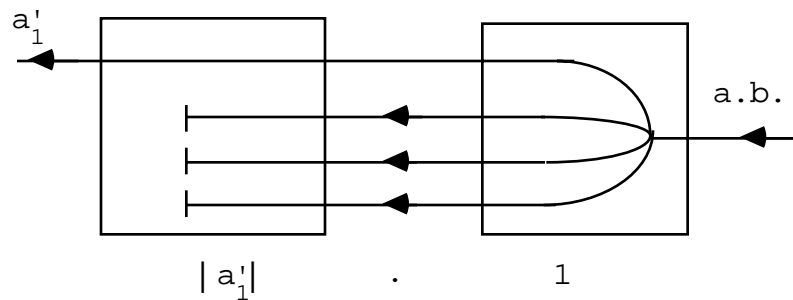
In fact, both are the same as just
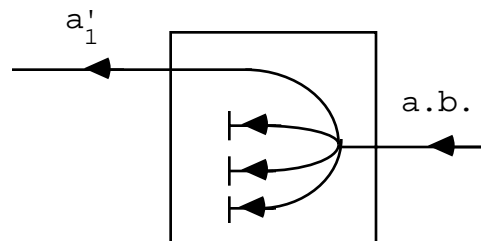


Mathematically, these diagrams tell us that

$$1 \cdot |a'| = |a'| \cdot 1 = |a'|. \qquad (92)$$

One can also show

$$1 \cdot 0 = 0; \ 0 \cdot 1 = 0. \qquad (93)$$

We have to tie this discussion in with real numbers eventually.  All experiments have results and all results are numbers.  There has to be more to an experiment than just accepting or rejecting physical attributes.  There is also the possibility of modulating a signal.  If "C" represents in general a complex number, we adopt the simple rules that

$$\left.\begin{array}{c} C \cdot 0 = 0 \\ C|a'| = |a'|C \\ C1 = 1C \end{array}\right\} \qquad (94)$$

These properties assure that no distinction between 1,0 (measurement symbols) and 1,0 (numbers) is necessary.  For now, let us also regard the numbers "C" as being dimensionless.  We will suggest a modulating device as follows:
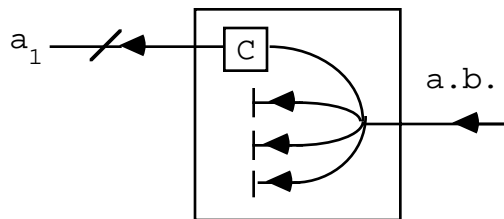


Symbol:   $C|a_1|$

The amplitude of the $a_1$ beam above has been modified by a factor $|C|$, and it's phase has been changed by $\tan^{-1}(\text{Im}(C)/\text{Re}(C))$, just like for an electronic circuit.  A slash through an emerging beam will sometimes be used to denote its modified character, and we can also, if we wish, write the modulating factor in the little box thus - $\boxed{C}$.  Using
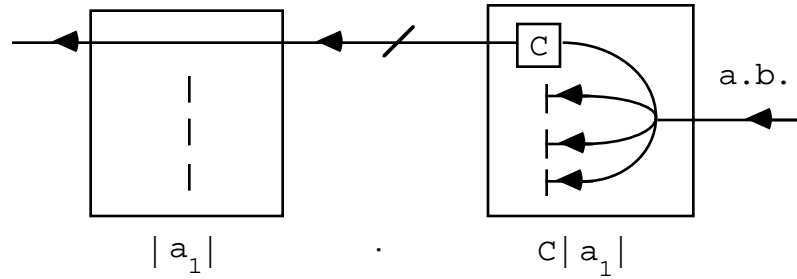
$$|a_1| \quad \cdot \quad C|a_1|$$

we see that, for example,

$$|a_1|(C|a_1|) = C|a_1|, \tag{95}$$

which also follows mathematically from (94) and (85) above.  We
adopt the rule that our beam always travels from right to left,
and will write down our measurement symbols in the same order
as they appear in the diagrams.  An example of a more general
modulated measurement is:
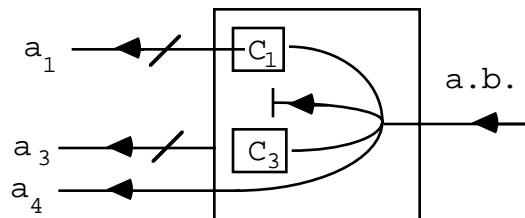


$$\text{Symbol:} \quad C_1|a_1| + C_3|a_3| + |a_4|$$

Now that we know how to associate numbers with measurement
symbols, we may write (85) and (86) together as

$$|a'||a''| = \delta_{a'a''}|a'| \tag{96}$$

where $\delta_{a'a''}$ is the Kroniker delta symbol:

$$\delta_{a'a''} = \begin{cases} 1, & a' = a'' \\ 0, & a' \neq a'' \end{cases}. \tag{97}$$

1.42

In addition, one can show that the distributive law is operative here:

$$(|a'| + |a''|)|a'''| = |a'||a'''| + |a''||a'''|. \qquad (98)$$

Let us now define a very special sort of modulated operator.  If we choose

$$C_i = a_i , \qquad (99)$$

i.e., the amplification factors are chosen as the values of the physical outcomes (which are real), then we have for this measurement

$$A = \sum_i a_i |a_i|. \qquad (100)$$

We have been thinking of the $a_i$ as dimensionless, but we may want to associate physical dimensions with the property A, just as we associate physical dimensions with $S_z$.  We can always supply dimensions by multiplying both sides of (100) by a single dimensionful constant.

$$A' = CA = \sum_i Ca_i |a_i| = \sum_i \bar{a}_i |\bar{a}_i| \qquad (101)$$

$$\uparrow$$

related as $Ca_i = \bar{a}_i$ ,

This mathematical act is somewhat mysterious from the point of view of our diagrams, since it can't be represented in such a manner.  However, every experiment has a readout in units of some kind.  Let us assume the above conversion to physical units represents the machinery's readout of the result in some appropriate units. For now, we will continue to use dimensionless physical outcomes $a_i$; we can always supply a dimensionful constant later.

Let us deduce some properties of the above A.  First,
notice that

$$A|a'| = \left(\sum_i a_i|a_i|\right)|a'| = (a_1|a_1| + a_2|a_2| + ...)|a'|$$
$$= a'|a'||a'| = a'|a'|, \text{ so}$$

$$A|a'| = a'|a'|. \tag{102}$$

Also

$$|a'|A = a'|a'|. \tag{103}$$

"A" has the important property of singling out the value
of the physical outcome a' when it acts in concert with the
selection |a'|.  Pictorically, (102) is saying
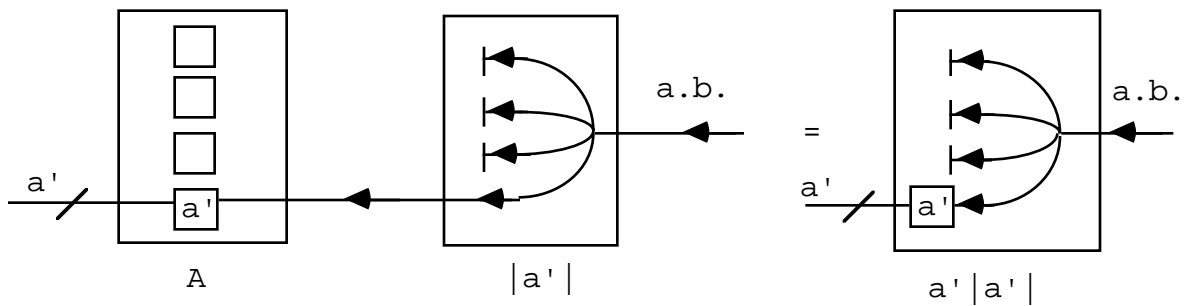


Eq$^n$(103) can be seen as
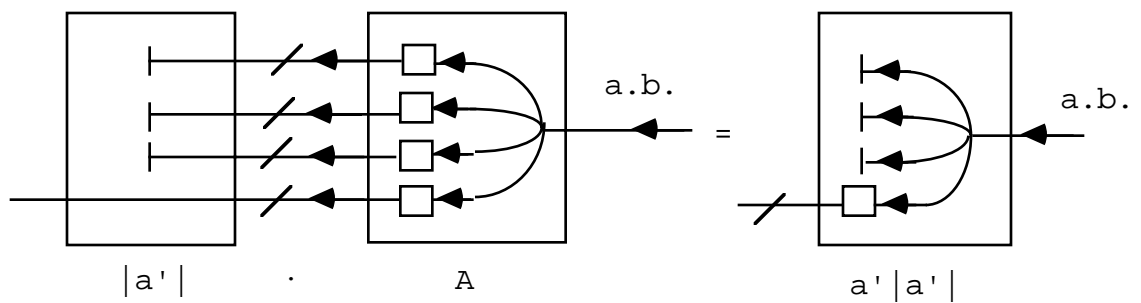
The order of these operations or measurements is not important yet.

It's time to say a little bit more about what the diagrams I have been drawing represent. Although we have used the S-G experimental apparatus to model these idealized measurements after, the above manipulations on the incoming "beam" do not actually represent physical operations carried out in real space. Instead, they represent operations carried out on individual particle characteristics in a mathematical "space" or arena where the concepts "amplitude" and "phase" makes sense. This mathematical space has been given the name of "Hilbert space." Although the above do not represent real-space experimental setups, there is still a correspondence between what happens in a real experiment (involving spin, say) and in our Hilbert space idealizations; this connection will be stated shortly. I will call these ideal manipulations on arbitrary beams ("arbitrary" in the sense of containing nonzero amplitudes for all physical outcomes, a') "Process Diagrams".

Some other properties of the above A are now detailed. Notice that

$$A^2 = A \cdot A = A \sum_i a_i |a_i| = \sum_i a_i A |a_i|.$$

But $A|a_i| = a_i |a_i|$, so

$$A^2 = \sum_i a_i^2 |a_i|. \tag{104}$$

This can be pictured as
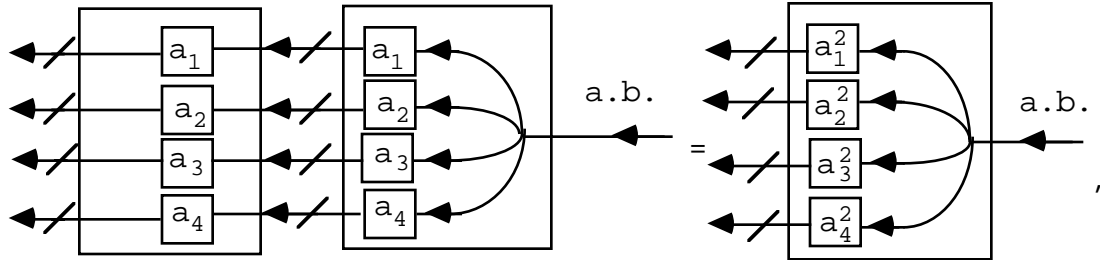
where the amplitude factors $a_1, a_2, \ldots$ associated with each one of the physical outcomes has been written in explicitly.

The generalization of the above rule for $A^2$ is

$$f(A) = \sum_i f(a_i)|a_i|. \tag{105}$$

for some $f(A)$ some power series in A. Let us take some examples to understand (105) better. First, which $f(A)$ results from the choice of $f(a_i) = 1$ for all i?

$$f(A) = \sum_i |a_i| = 1.$$

Next, which $f(A)$ results from $f(a_i) = 0$ for all i?

$$f(A) = \sum_i 0|a_i| = 0.$$

Which $f(A)$ results from the following choice?

$$f(a_j) = 1,$$
$$f(a_i) = 0, \text{ all } a_i \neq a_j?$$

This also is easy:

$$f(A) = |a_j|.$$

However, what is this in terms of A? (This is not so easy.)

Now

$$A|a_j| = a_j|a_j|,$$
$$\Rightarrow (A - a_j)|a_j| = 0,$$

where I have suppressed the unit symbol 1, in writing the second form. Now consider the statement:

$$\left[\prod_{i=1}^{n}\left(A - a_i\right)\right] \, |a_j| = 0 \, . \tag{106}$$

To see that this is true, write this out more explicitly:

$$(A - a_1)(A - a_2)\ldots(A - a_j)\ldots(A - a_n)|a_j|$$
$$= (a_j - a_1)(a_j - a_2)\ldots(a_j - a_j)\ldots(a_j - a_n)|a_j| = 0. \tag{107}$$

Since the above is true for __any__ $|a_j|$, we must have

$$\prod_i (A - a_i) = 0. \tag{108}$$

So this represents a new way of writing the null measurement. (Can you think of a Process Diagram to represent the left hand side of (108)?) Now comparing (108), written in the form

$$\prod_i (A - a_i) = (A - a_j)\prod_{i \neq j}(A - a_i) = 0. \tag{109}$$

with the statement $(A - a_j)|a_j| = 0$ leads to the conclusion that

$$|a_j| = C \prod_{i \neq j}(A - a_i), \tag{110}$$

where "C" is some unknown constant. I will just supply this constant:

$$|a_j| = \prod_{i \neq j}\left(\frac{A - a_i}{a_j - a_i}\right). \tag{111}$$

It is easy to show that (111) works correctly. First, let us show that $|a_j||a_j| = |a_j|$:

$$\left[\prod_{i \neq j}\left(\frac{A - a_i}{a_j - a_i}\right)\right]|a_j| = \left(\frac{A - a_1}{a_j - a_1}\right) \cdots \left(\frac{A - a_n}{a_j - a_n}\right)|a_j|.$$

$$= \left(\frac{a_j - a_1}{a_j - a_1}\right) \cdots \left(\frac{a_j - a_n}{a_j - a_n}\right)|a_j|$$

$$= 1 \ldots 1|a_j| = |a_j|.$$

This tells us we have chosen the constant C correctly. Next, let us check that $|a_j||a_k| = 0$ $(j \neq k)$:

$$\left[\prod_{i \neq j}\left(\frac{A - a_i}{a_j - a_i}\right)\right]|a_k| = \left(\frac{A - a_1}{a_j - a_1}\right) \cdots \left(\frac{A - a_n}{a_j - a_n}\right)|a_k|$$

$$= \left(\frac{a_k - a_1}{a_j - a_1}\right) \cdots \left(\frac{a_k - a_k}{a_j - a_n}\right) \cdots \left(\frac{a_k - a_n}{a_j - a_n}\right)|a_k|$$

$$= 0.$$

Let us study the two-physical-outcome case in some detail. Let $a_1 = 1$, $a_2 = -1$. Then

$$\prod_i (A - a_i) = 0 \Rightarrow (A - 1)(A + 1) = 0 \Rightarrow A^2 = 1. \qquad (112)$$

This is the algebraic equation satisfied by the physical property A. Also (let $|1| = |+|$, $|-1| = |-|$)

$$|+| = \prod_{i \neq 1}\left(\frac{A - a_i}{a_1 - a_i}\right) = \frac{A + 1}{2}, \qquad (113)$$

$$|-| = \prod_{i \neq 2}\left(\frac{A - a_i}{a_2 - a_i}\right) = \frac{1 - A}{2}. \qquad (114)$$

In addition, we have completeness:

$$\sum_i |a_i| = |+| + |-| = \frac{A + 1}{2} + \frac{1 - A}{2} = \frac{1}{2} + \frac{1}{2} = 1.$$

We can explicitly check some properties here:

$$|-||+| = \left(\frac{1 - A}{2}\right)\left(\frac{A + 1}{2}\right) = \frac{1 - A^2}{4} = 0$$

$$|+||+| = \left(\frac{A + 1}{2}\right)^2 = \frac{1 + 2A + A^2}{4} = \frac{2 + 2A}{4} = |+|$$

$$|-||-| = \left(\frac{1 - A}{2}\right)^2 = \frac{1 - 2A + A^2}{4} = \frac{2 - 2A}{4} = |-|.$$

There is another operation we can imagine performing on our "arbitrary beam" of particles that has a quantum mechanical basis. Besides the selection and amplification operations, one can also imagine the following Process Diagram:
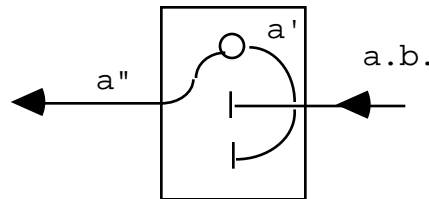


That is, we are imagining an experiment that performs a transition. In the above, the beam with physical outcome a' is transformed into a beam with physical outcome a", keeping the a' beam's amplitude and phase information. In the S-G case, the above would represent

an apparatus which turned spin $S_z = \frac{\hbar}{2}$ into $S_z = -\frac{\hbar}{2}$, say. The

symbol we will adopt for the above is (the "1" is implicit)

$$|a" \quad a'|.$$
$$\uparrow \qquad \uparrow$$
exiting property    entering property

The connection to our earlier measurement symbol is clear:

$$|a'a'| = |a'|. \tag{115}$$

Consider the following:



| a"'a" |  .  |a"a'|

This is the same as



$|a"' a"|$
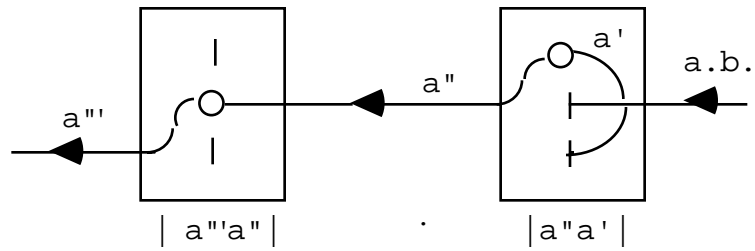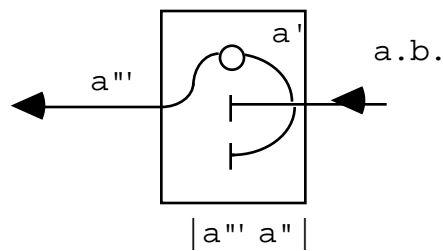
So we adopt the rule that

$$|a"' a"||a"a'| = |a"' a'|. \tag{116}$$

It is also clear that

$$|a^{iv} a"'||a"a'| = 0, \quad a"' \neq a". \tag{117}$$

The algebraic properties of our two types of measurement symbols are summarized in

$$|a'||a"| = \delta_{a'a"}|a'|, \tag{118}$$

$$|a^{iv} a"'||a"a'| = \delta_{a"' a"}|a^{iv} a'|. \tag{119}$$

1.50

As pointed out before, the $|a'|$-type measurements are reversible (i.e., the order of the operations doesn't matter). However, these new types of measurements are not reversible. An example is the following:



$$|a''a'| \qquad \cdot \qquad |a'|$$

$$\ne$$

$$|a'| \qquad \cdot \qquad |a''a'|$$
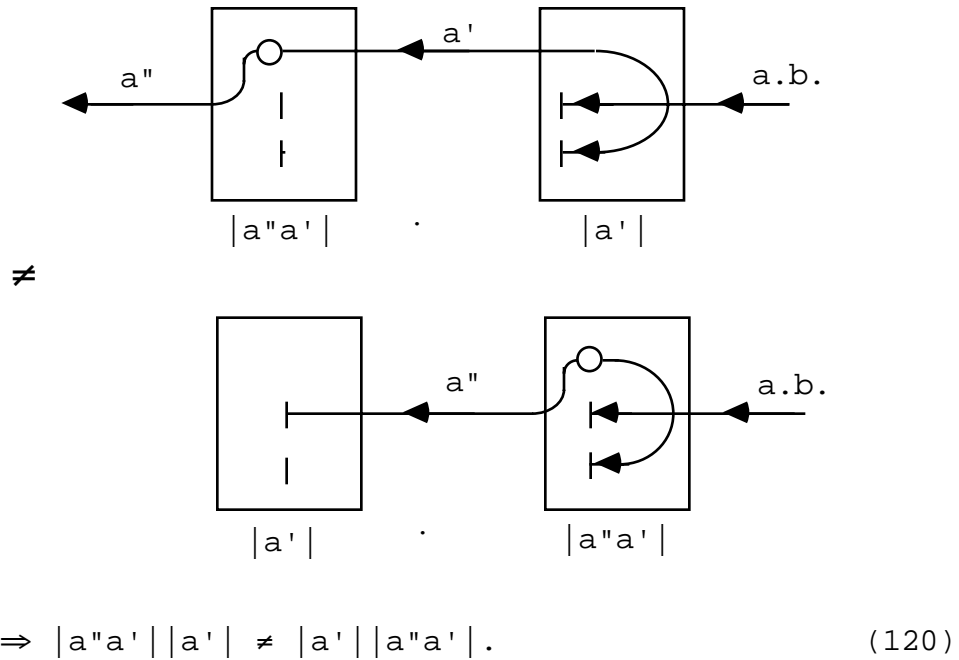
$$\Rightarrow |a''a'||a'| \ne |a'||a''a'|. \tag{120}$$

Eq$^{\text{n}}$(120) follows mathematically from (119) and the fact that $|a'| = |a'a'|$.  Another example is



$$|a'a''| \qquad \cdot \qquad |a''a'|$$

$$\ne$$
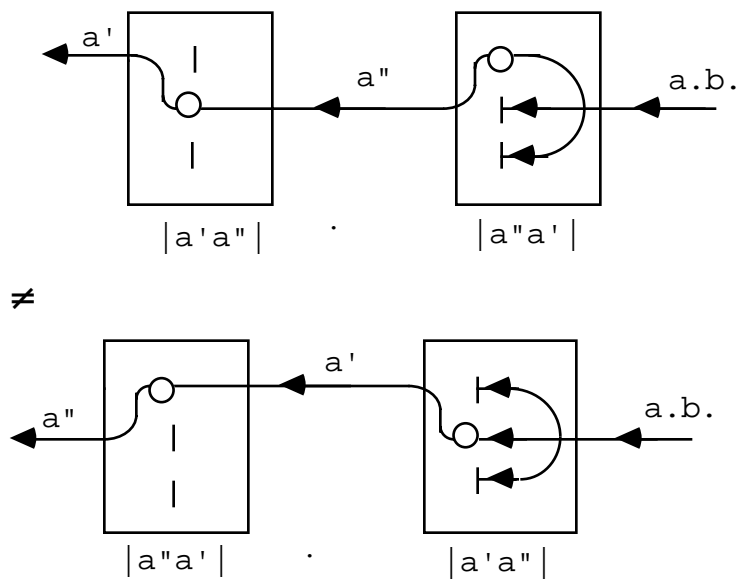
$$|a''a'| \qquad \cdot \qquad |a'a''|$$

$$\Rightarrow \quad |a'a''||a''a'| \neq |a''a'||a'a''|. \qquad (121)$$

The algebra of the $|a'a''|$-type symbols is <u>non-commutative</u>. Each side of (120) and (121) can be reduced further using (119).  We can always label the $\mathrm{O}$-symbols in our transition-type Process Diagrams with the actual transition this device performs in order to remove any ambiguity.  We will do this occasionally in the following.

Let us concentrate on the two-physical outcome case, since this is the simplest situation.  The symbol associated with
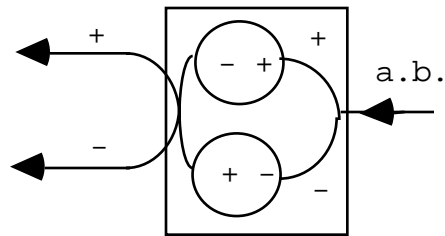


is

$$|-+| \;+\; |+-|.$$

This suggests that the operation,



$$(|-+| \;+\; |+-|) \;\cdot\; (|-+| \;+\; |+-|)$$
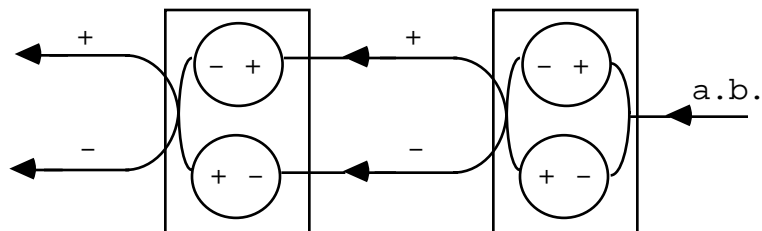
simply reconstitutes the original beam. Let's confirm this mathematically:

$$(|-+| + |+-|)\cdot(|-+| + |+-|)$$

$$= |-+||-+| + |-+||+-| + |+-||-+| + |+-||+-|$$

$$= |-| + |+| = 1.$$

Here are some more examples and the equations that go along with them.

| 0 | = | $|-+|$ | . | $|-+|$ |
|---|---|---|---|---|

a.b.

| $|+-|$ | = | $|+| \cdot (|-+| + |+-|)$ |
|---|---|---|

a.b.

| 0 | = | $|-+| . (|-+| + |-|) . |+|$ |
|---|---|---|

a.b.
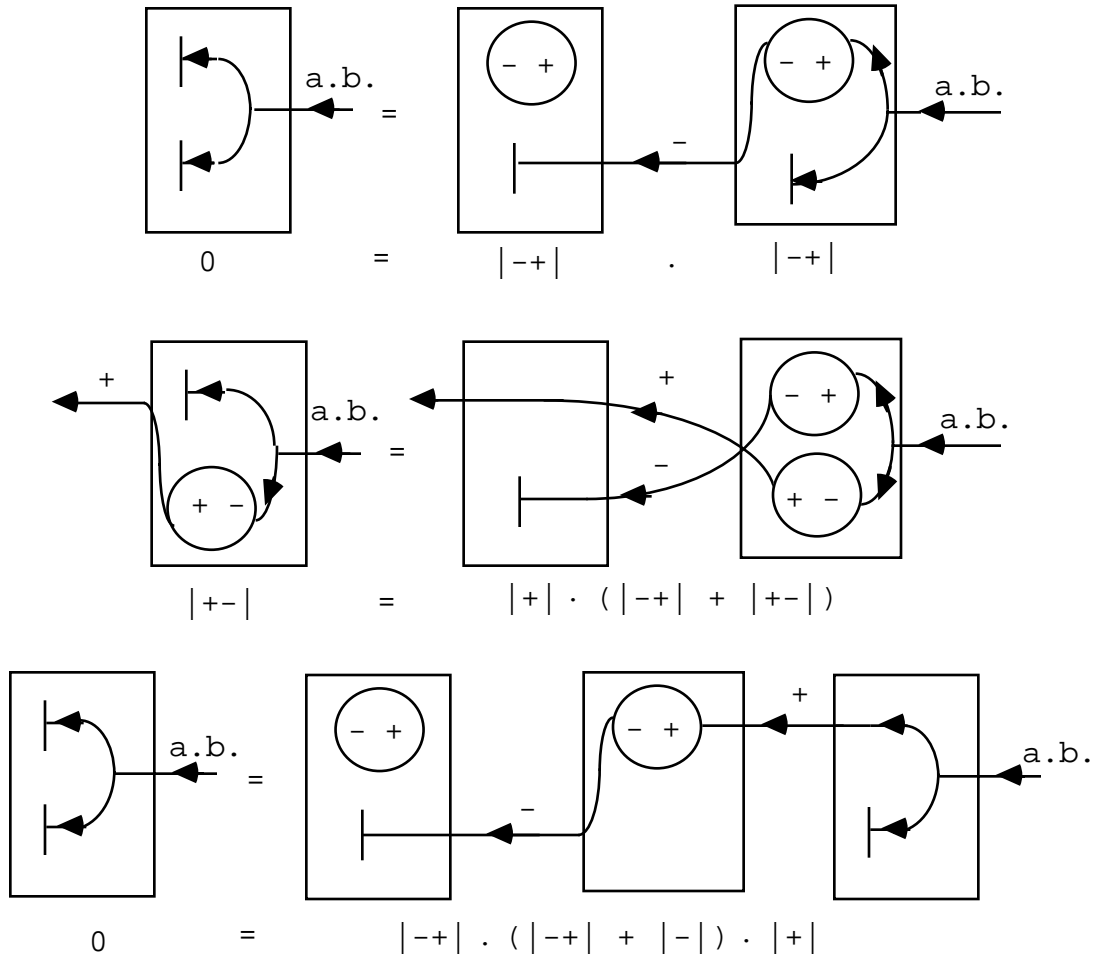
Beams can sometimes combine, as in

$$(|-+| + |-|)$$

a.b.
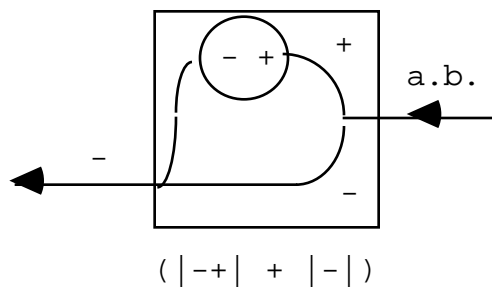
Of course, the "-" beam's amplitude has in general been modified.

Let's continue to investigate the two-physical outcome case. There are four independent measurement symbols:

1. $|{++}| = |{+}|$

2. $|{--}| = |{-}|$

3. $|{+-}|$

4. $|{-+}|$.

We will make a different, more convenient choice of the four independent quantities. We choose the unit symbol

$$1 = |{+}| + |{-}|, \tag{122}$$

as one of them. Another independent choice is

$$\sigma_3 = |{+}| - |{-}|, \tag{123}$$

which we can write as

$$\sigma_3 = \sum_{\sigma_3'} |\sigma_3'| \sigma_3' , \tag{124}$$

where $\sigma_3' = \pm 1$. $\sigma_3$ is just a renaming of the special modulated operator A we investigated on (p. 1.41) above. Let's confirm that (112) holds:

$$\sigma_3^2 = (|{+}| - |{-}|)(|{+}| - |{-}|) = |{+}| + |{-}| = 1.$$

We will take the other two independent combinations to be:

$$\sigma_1 = |{-+}| + |{+-}|,$$

$$\sigma_2 = i|{-+}| - i|{+-}|.$$

In our Process Diagram language, $\sigma_1$ performs a double transition (see some of the examples already drawn) while $\sigma_2$ performs the same transition, but additionally modifies the phases of the signals. It is easy to see that, in addition to $\sigma_3^2 = 1$, we have

$$\sigma_1^2 = 1, \quad \sigma_2^2 = 1. \tag{125}$$

Some other mathematical properties of these new combinations are:

$$\begin{aligned}
\sigma_1 \sigma_2 &= (|-+| + |+-|)(i|-+| - i|+-|) \\
&= -i|-| + i|+| = i\sigma_3,
\end{aligned} \tag{126}$$

$$\begin{aligned}
\sigma_2 \sigma_1 &= (i|-+| - i|+-|)(|-+| + |+-|) \\
&= i|-| - i|+| = -i\sigma_3.
\end{aligned} \tag{127}$$

Therefore, we see that

$$\sigma_1 \sigma_2 = -\sigma_2 \sigma_1. \tag{128}$$

One can also show the following:

$$\sigma_2 \sigma_3 = i\sigma_1, \quad \sigma_3 \sigma_2 = -i\sigma_1, \tag{129}$$

$$\sigma_3 \sigma_1 = i\sigma_2, \quad \sigma_1 \sigma_3 = -i\sigma_2. \tag{130}$$

Summarizing the properties of the $\sigma$'s, we have

$$\sigma_k^2 = 1 \qquad (k = 1,2,3), \tag{131}$$

$$\sigma_k \sigma_\ell = -\sigma_\ell \sigma_k \quad (k \neq \ell), \tag{132}$$

$$\sigma_k \sigma_\ell = i\sigma_m \qquad (k,\ell,m \text{ in } \underline{\text{cyclic}} \text{ order}). \tag{133}$$

Some mathematical phraseology we will use is:

if AB = BA, then A and B <u>commute</u>;

if AB = -BA, then A and B <u>anticommute</u>.

Now our two-physical-outcome case has just been modeled after spin $\frac{1}{2}$. Electron spin, being a type of angular momentum, should involve three components; we hypothesize these spin components are represented by the $\sigma_i$ above, which are seen to satisfy (112), the basic operator equation for the two-physical outcome case. However, in order to supply the correct <u>physical units</u>, we write

$$S_i = \frac{\hbar}{2} \sigma_i, \quad i = 1,2,3. \tag{134}$$

The multiplication by $\frac{\hbar}{2}$ above is the "somewhat mysterious" part of the measurement that cannot be represented by a diagram. (See p. 1.42 above.) Eq$^{\underline{n}}$(134) is the crucial connection that allows us to tie our developing formalism to the real world.

It is important to realize that the above quantities $S_i$ are <u>operators</u>, not numbers. Nevertheless, if they represent an angular momentum they must <u>behave</u> like a vector under rotations in real space. Let us confirm our identification of the $\sigma_i$ as angular momentum components under the following passive rotation ("passive" means the axes, not the vector is rotated) about the third axis:
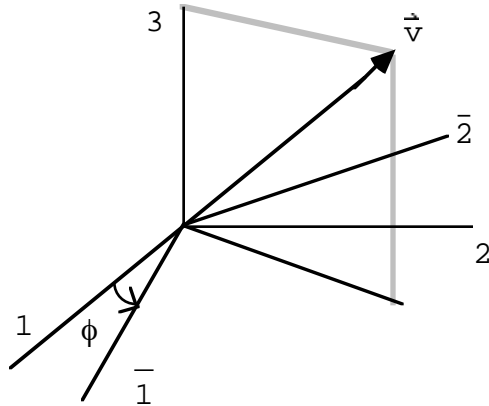
A vector $\vec{v}$ should now have new components $\bar{v}_i$ given by:

$$\left.\begin{array}{l} \bar{v}_3 = v_3 \\ \bar{v}_1 = v_1 \cos \phi + v_2 \sin \phi \\ \bar{v}_2 = -v_1 \sin \phi + v_2 \cos \phi \end{array}\right\} \qquad (135)$$

$v_1, v_2, v_3$ are the components along the old (unbarred) axes.  We will choose the angle $\phi$ in the above figure to be positive by convention.  In analogy to (135), we require that the spin components in the new coordinates be

$$\left.\begin{array}{l} \bar{\sigma}_3 = \sigma_3 \\ \bar{\sigma}_1 = \sigma_1 \cos \phi + \sigma_2 \sin \phi \\ \bar{\sigma}_2 = -\sigma_1 \sin \phi + \sigma_2 \cos \phi \end{array}\right\} \qquad (136)$$

Now, is it still true that $\bar{\sigma}_i^2 = 1$?

$$\bar{\sigma}_3^{\ 2} = \sigma_3^2 = 1$$

$$\overbrace{\phantom{xxxxx}}^{= 0}$$

$$(\bar{\sigma}_1)^2 = \sigma_1^2 \cos^2 \phi + \sigma_2^2 \sin^2 \phi + (\sigma_1\sigma_2 + \sigma_2\sigma_1) \cos \phi \sin \phi$$
$$= \cos^2 \phi + \sin^2 \phi = 1$$

$$\overbrace{\phantom{xxxxx}}^{= 0}$$

$$(\bar{\sigma}_2)^2 = \sigma_1^2 \sin^2 \phi + \sigma_2^2 \cos^2 \phi + (\sigma_1\sigma_2 + \sigma_2\sigma_1) \cos \phi \sin \phi$$
$$= \sin^2 \phi + \cos^2 \phi = 1$$

What about the cyclic property? One can show that

$$\bar{\sigma}_1\bar{\sigma}_2 = i\bar{\sigma}_3, \quad \bar{\sigma}_2\bar{\sigma}_1 = -i\bar{\sigma}_3$$
$$\bar{\sigma}_2\bar{\sigma}_3 = i\bar{\sigma}_1, \quad \bar{\sigma}_3\bar{\sigma}_2 = -i\bar{\sigma}_1 \qquad (137)$$
$$\bar{\sigma}_3\bar{\sigma}_1 = i\bar{\sigma}_2, \quad \bar{\sigma}_1\bar{\sigma}_3 = -i\bar{\sigma}_2.$$

We would like to show now that the properties $\bar{\sigma}_i^2 = 1$ and the cyclic properties (137) above hold for a more general rotation. To start off, let's rewrite $\bar{\sigma}_1$ above as

$$\bar{\sigma}_1 = \sigma_1 \cos\phi + \sigma_2 \sin\phi = \sigma_1 \cos\phi - i\sigma_3\sigma_1 \sin\phi$$
$$\Rightarrow \quad \bar{\sigma}_1 = (\cos\phi - i\sigma_3 \sin\phi)\sigma_1, \qquad (138)$$

or

$$\bar{\sigma}_1 = \sigma_1(\cos\phi + i\sigma_3 \sin\phi). \qquad (139)$$

You will show in a problem that

$$e^{i\lambda\sigma_3} = \cos\lambda + i\sigma_3 \sin\lambda, \qquad (140)$$

where "$\lambda$" is just a real number. In fact, more generally, one can show that

$$e^{i\lambda(\hat{n}\cdot\vec{\sigma})} = \cos\lambda + i(\hat{n}\cdot\vec{\sigma})\sin\lambda. \qquad (141)$$

where $\hat{n}\cdot\vec{\sigma} = n_1\sigma_1 + n_2\sigma_2 + n_3\sigma_3$ and $\hat{n}$ is a unit spatial vector ($\hat{n}^2 = 1$). Using (140) or (141) with $\hat{n} = \hat{e}_3$, we can write (138) and (139) above as

$$\bar{\sigma}_1 = e^{-i\phi\sigma_3} \sigma_1 \qquad (142)$$

or

$$\bar{\sigma}_1 = \sigma_1 e^{i\phi\sigma_3}. \qquad (143)$$

1.58

[There are many ways of seeing the equality of (142) and (143). Since $\sigma_1$ and $\sigma_3$ anticommute, we have for example

$$\sigma_3\sigma_1 = \sigma_1(-\sigma_3)$$

and

$$\sigma_3^2\sigma_1 = \sigma_1(-\sigma_3)^2.$$

Generalizing from these examples, it is easy to see that

$$f(\sigma_3)\sigma_1 = \sigma_1 f(-\sigma_3)$$

for $f(\sigma_3)$ any power series function of $\sigma_3$.] We now wish to prove that

$$e^{i\lambda\sigma_3}\, e^{i\lambda'\sigma_3} = e^{i(\lambda + \lambda')\sigma_3}. \tag{144}$$

We can rewrite

$$e^{i\lambda\sigma_3} = \cos\lambda + i\sigma_3\sin\lambda,$$

$$\Rightarrow \quad e^{i\lambda\sigma_3} = \cos\lambda(|+| + |-|) + i(|+| - |-|)\sin\lambda,$$

$$\Rightarrow \quad e^{i\lambda\sigma_3} = |+|e^{i\lambda} + |-|e^{-i\lambda} \tag{145}$$

Using (145), we now have

$$e^{i\lambda\sigma_3}\, e^{i\lambda'\sigma_3} = (|+|e^{i\lambda} + |-|e^{-i\lambda})\,(|+|e^{i\lambda'} + |-|e^{-i\lambda'})$$

$$= |+|e^{i(\lambda + \lambda')} + |-|e^{-i(\lambda + \lambda')} = e^{i(\lambda + \lambda')\sigma_3},$$

which proves (144) above. Therefore, we can write from (142) for example

$$= \sigma_1 e^{i\phi/2\,\sigma_3}$$

$$\overbrace{\qquad\qquad}$$

$$\bar{\sigma}_1 = e^{-i\phi\sigma_3}\,\sigma_1 = e^{-i\phi/2\,\sigma_3}\,e^{-i\phi/2\,\sigma_3}\,\sigma_1 = e^{-i\phi/2\,\sigma_3}\,\sigma_1 e^{i\phi/2\,\sigma_3} \tag{146}$$

Eq$^{\underline{n}}$(143) leads to the same conclusion. Now, if we call $U = e^{i\phi/2\,\sigma_3}$, then $U^{-1} = e^{-i\phi/2\,\sigma_3}$ since

$$e^{-i\phi/2\ \sigma_3}\ e^{i\phi/2\ \sigma_3} = e^{i(\phi/2\ -\ \phi/2)\sigma_3} = e^0 = 1$$

by (144) above.  Then, we may write

$$\overline{\sigma}_1 = U^{-1}\sigma_1 U. \tag{147}$$

$\overline{\sigma}_2$ and $\overline{\sigma}_3$ above can also be written as in (147):

$$U^{-1}\sigma_2 U = e^{-i\phi/2\ \sigma_3}\ \sigma_2 e^{i\phi/2\ \sigma_3} = \sigma_2 e^{i\phi\,\sigma_3}$$

$$= \sigma_2(\cos\phi + i\sigma_3\sin\phi) = \sigma_2\cos\phi - i\sigma_1\sin\phi$$

$$= \overline{\sigma}_2,$$

$$U^{-1}\sigma_3 U = \overline{\sigma}_3. \quad \text{(trivial)}$$

Thus we have $(S_i = \dfrac{\hbar}{2}\ \sigma_i)$

$$\overline{S}_i = U^{-1}S_i U. \tag{148}$$

These forms shed a new light on why the algebraic properties of the $\sigma_i$ are preserved under a rotation.  Taking a particular example

$$\sigma_1\sigma_2 = i\sigma_3,$$

we can write this as

$$U^{-1}\sigma_1 U U^{-1}\sigma_2 U = iU^{-1}\sigma_3 U$$

or therefore

$$\overline{\sigma}_1\overline{\sigma}_2 = i\overline{\sigma}_3,$$

and the algebra has been preserved.  The entire content of the algebra of the $\sigma_i$ can be summarized as

$$\sigma_i\sigma_j = 1\delta_{ij} + i\sum_k \varepsilon_{ijk}\sigma_k \tag{149}$$

where the permutation symbol $\varepsilon_{ijk}$ has been introduced.  It is defined as

1.60

$$\varepsilon_{ijk} = \begin{cases} 0, & \text{if any index is equal to any other index} \\ +1, & \text{if } i,j,k \text{ form an } \underline{even} \text{ permutation of } 1,2,3 \\ -1, & \text{if } i,j,k \text{ form an } \underline{odd} \text{ permutation of } 1,2,3. \end{cases} \quad (150)$$

Eq$^{\underline{n}}$(149) is, of course, true for the $\bar{\sigma}_i$ also.

Let's take another example of a rotation, as given in the figure below (rotation in the 1,3 plane).
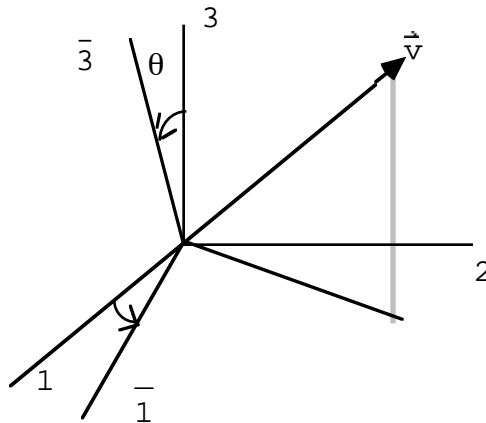


Again, the angle $\theta$ shown above is positive by our conventions. The components of the new $\bar{\sigma}_i$ operators are clearly

$$\left. \begin{aligned} \bar{\sigma}_1 &= \sigma_1 \cos\theta - \sigma_3 \sin\theta \\ \bar{\sigma}_2 &= \sigma_2 \\ \bar{\sigma}_3 &= \sigma_3 \cos\theta + \sigma_1 \sin\theta \end{aligned} \right\} \quad (151)$$

Now let's try to produce the same $\bar{\sigma}_i$ with the following guess for U:

$$U = e^{i\theta/2\ \sigma_2} . \quad (152)$$

We find

$$\begin{aligned} U^{-1}\sigma_1 U &= e^{-i\theta/2\sigma_2}\sigma_1\ e^{i\theta/2\ \sigma_2} = \sigma_1 e^{i\theta\sigma_2} \\ &= \sigma_1(\cos\theta + i\sigma_2 \sin\theta) = \sigma_1 \cos\theta - \sigma_3 \sin\theta \\ &= \bar{\sigma}_1 \end{aligned}$$

Likewise,

$$U^{-1}\sigma_2 U = \bar{\sigma}_2 \text{ (trivial)}$$

$$U^{-1}\sigma_3 U = e^{-i\theta/2\sigma_2} \sigma_3 e^{i\theta/2\sigma_2} = \sigma_3 e^{i\theta\sigma_2}$$

$$= \sigma_3(\cos\theta + i\sigma_2 \sin\theta) = \sigma_3 \cos\theta + \sigma_1 \sin\theta$$

$$= \bar{\sigma}_3.$$

Summing up, we have found that:

$U = e^{i\phi/2\ \sigma_3}$ describes a rotation by $\phi$ about the 3-axis.

$U = e^{i\theta/2\ \sigma_2}$ describes a rotation by $\theta$ about the 2-axis.

Using these operators, we can now describe the more general rotation shown below.
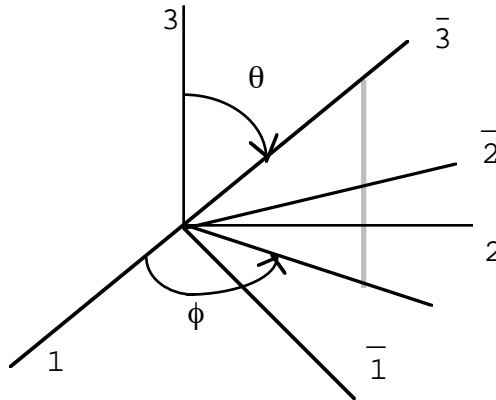


We can generate the new (barred) axes from the old ones by performing the following two steps.

1. Rotate by $\theta$ about the 2-axis

2. Rotate by $\phi$ about the 3-axis

Of course, this specification of how to get the orientation of the new axes from the old is not unique. What is the U that describes this transformation?

At the end of the first step, we have for the spin components $S_i$,

$$e^{-i\theta/2\ \sigma_2}\ S_i e^{i\theta/2\ \sigma_2}\ .$$

At the end of the second step then

$$\overline{S}_i\ =\ e^{-i\phi/2\ \sigma_3}(e^{-i\theta/2\ \sigma_2}\ S_i e^{i\theta/2\ \sigma_2})e^{i\phi/2\ \sigma_3}.$$

By comparison with the usual form $\overline{S}_i\ =\ U^{-1}S_i U$, we then find that

$$\left.\begin{array}{l} U\ =\ e^{i\theta/2\ \sigma_2}\ e^{i\phi/2\ \sigma_3} \\ U^{-1}\ =\ e^{-i\phi/2\ \sigma_3}\ e^{-i\theta/2\ \sigma_2} \end{array}\right\} \qquad (153)$$

describes this more general rotation.  Notice in (153) that the individual exponential factors appear not only with extra minus signs, but also show up in the opposite order.  This is really not so mysterious since if

$$U\ =\ U_1 U_2$$

then

$$(U_1 U_2)(U_1 U_2)^{-1}\ =\ 1$$

$$\Rightarrow\quad U_2(U_1 U_2)^{-1}\ =\ U_1^{-1}$$

$$\Rightarrow\quad U^{-1}\ =\ (U_1 U_2)^{-1}\ =\ U_2^{-1}U_1^{-1}$$

assuming $U_1$ and $U_2$ both possess inverses. Because of the noncommutative algebra, this is not the same as $U_1^{-1}U_2^{-1}$.

Using the U-transformation, we have alternate but equivalent sets of measurement symbols in the two frames of reference which are connected by a rotation. In the "old" description we have the symbols:

$$\left|a'a''\right|.$$

In the "new" description, we have the symbols:

$$|\overline{a'a''}|.$$

The connection is

$$|\overline{a'a''}| = U^{-1}|a'a''|U. \qquad (154)$$

Notice that "completeness" is preserved:

$$\sum_{a'}|\overline{a'a'}| = U^{-1}\left(\sum_{a'}|a'a'|\right)U = U^{-1}U = 1.$$

The new symbols will allow a more complete description of our
original S-G two magnet setup.  In particular, they will allow
us to describe the situation where the two magnets have a
relative rotational angle.  For example we will model the real
space setup
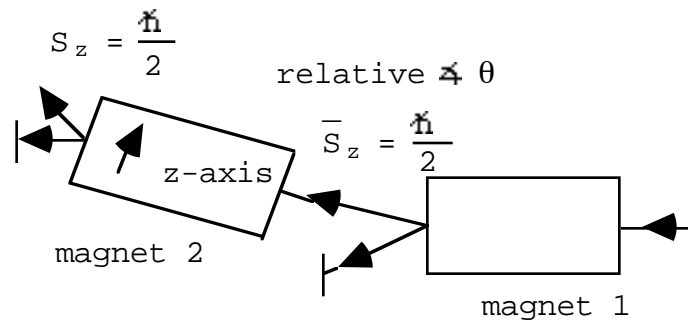


where magnet 1 is rotated an angle $\theta$ with respect to magnet 2,
with the Process Diagram:



$$|+| \qquad \cdot \qquad |\overline{+}|$$
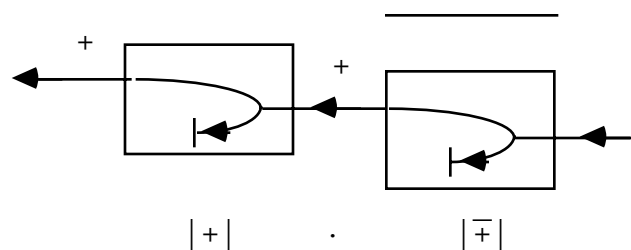
The first measurement symbol, $|\overline{+}|$, selects spin "up" along
the $\overline{z}$-axis.  The resultant beam is then operated on by the
second symbol, $|+|$, in which a further splitting of the beam,

suggestive of what happens in magnet 2 of the real space setup
above, takes place.  Notice in the above that although both
measurements select spin "up", the z-axes of the two
descriptions are different.  This is suggested in the Process
Diagram by barring the first $|\overline{+}|$ selection apparatus.

   We still need to know how to calculate <u>probabilities</u> based
on the above ideas.  First, we have a vital realization to
make.  Consider the transition-type measurement device.
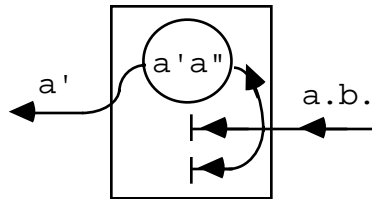


   The realization that we need to make at this point is that
the above can be viewed as a two-step process.  This is
symbolized by a division of the measurement symbol for the
above into two parts, destruction and creation:
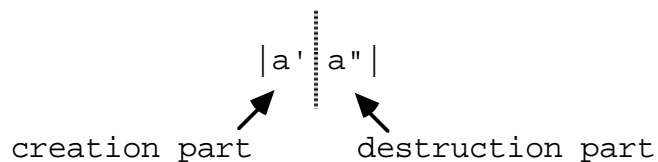


   Viewing these as independent acts we will then write (even
when a' = a")

$$|a'a"| = |a'><a"|. \qquad\qquad (155)$$

We will call

$$\left.\begin{array}{l} |a'>: \text{ a "bra"} \\ <a"|: \text{ a "ket"} \end{array}\right\} \text{ together they make a "bra-ket".}$$

"Completeness" now appears as

$$\sum_i \left| a_i \right> < a_i \right| = 1. \tag{156}$$

The algebraic condition (96) above ($\left| a' \right| \left| a'' \right| = \delta_{a'a''} \left| a' \right|$) now becomes

$$\left| a' \right> < a' \left| a'' \right> < a'' \right| = \delta_{a'a''} \left| a' \right> < a' \right|. \tag{157}$$

Notice, in writing (157) we have shortened $< a' \left| \right| a'' >$ into $< a' \left| a'' \right>$. Now because (157) is true for <u>any</u> $\left| a' \right> < a'' \right|$ combination, we must have that

$$< a' \left| a'' \right> = \delta_{a'a''}. \tag{158}$$

Eq$^{\underline{n}}$(158) is called "orthonormality." Thus we learn that the bra-ket combinations $< a' \left| a'' \right>$ (called the "inner product") are just ordinary numbers. The connection (154) now says

$$\left| \overline{a}' \right> < \overline{a}'' \right| = U^{-1} \left| a' \right> < a'' \right| U \tag{159}$$

We separate the independent pieces:

$$\left| \overline{a}' \right> = U^{-1} \left| a' \right>, \tag{160}$$

$$< \overline{a}'' \right| = < a'' \left| U. \tag{161}$$

Using (160) and (161), we easily see then that

$$< \overline{a}' \left| \overline{a}'' \right> = < a' \left| UU^{-1} \right| a'' \right> = < a' \left| a'' \right> = \delta_{a'a''}. \tag{162}$$

That is, orthonormality is true in the rotated system also. Without making a distinction between bras and kets, I

will often refer to a particular member $|a'>$ or $<a'|$ as a

state.  Also without the same distinction, I will call all

possible states of a system in a particular description a

basis.  The collections $\{|a'>\}$ and $\{|\bar{a}'>\}$ represent different

bases, connected however by a U-transformation.

   We are now ready to model our S-G probabilities.  We model

the general situation (we are not necessarily selecting spin

"up" as my inadequate drawing indicates)

$$S_3 = \pm\frac{\hbar}{2}$$

relative $\sphericalangle\ \theta$

$$\bar{S}_3 = \pm\frac{\hbar}{2}$$

z-axis

magnet 2

magnet 1

with

$$|\,\sigma_3''><\sigma_3''\,| \quad \cdot \quad |\,\bar{\sigma}_3'><\bar{\sigma}_3'\,|$$
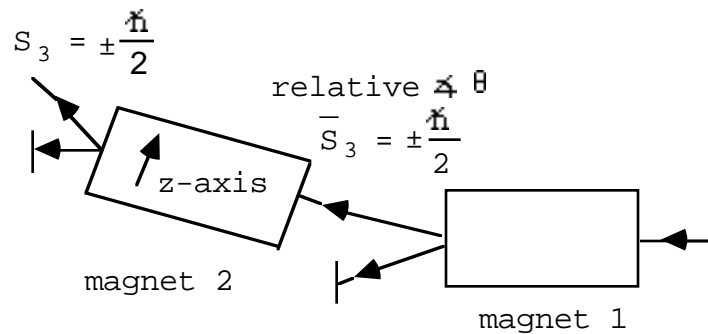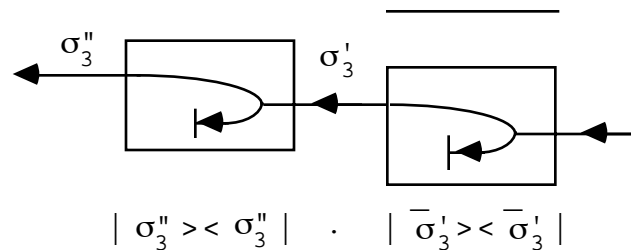
We have performed the separation (155) on the measurement

symbols.  We can now see that this last diagram is equivalent

to (the physical outcomes are given by $S_3' = \frac{\hbar}{2}\,\sigma_3'$, $\sigma_3' = \pm 1$)

$$A(\sigma_3'',\sigma_3')|\,\sigma_3''><\bar{\sigma}_3'\,| \ .$$
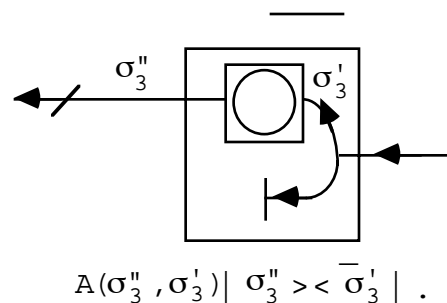
What we have drawn here is a hybrid sort of object that destroys particles with physical property $\sigma_3'$ in the barred basis, transforming them into particles with physical property $\sigma_3''$ in the unbarred basis. In addition, a modulation on the selected beam has been performed.  This happens in general because of the action of the second magnet in the beam produces a further selection and thus a loss of some particles.  From the equivalence of the above last two Process Diagrams, we find that

$$| \sigma_3'' > < \sigma_3'' | \overline{\sigma_3'} > < \overline{\sigma_3'} | \ = A(\sigma_3'', \sigma_3') | \sigma_3'' > < \overline{\sigma_3'} |,$$

so

$$A(\sigma_3'', \sigma_3') \ = \ < \sigma_3'' | \overline{\sigma_3'} >. \tag{163}$$

When the relative angle $\theta = 0$, it is easy to figure out what the value of $A(\sigma_3'', \sigma_3')$ is.  Then the barred and unbarred bases coincide and we have

$$A(\sigma_3'', \sigma_3')_{\theta=0} \ = \ \delta_{\sigma_3' \sigma_3''}, \tag{164}$$

i.e., either 0 or 1.  This suggests in the general case that the modulus (magnitude) of $A(\sigma_3'', \sigma_3')$ lies between these two values.  However, as we will see explicitly later, $A(\sigma_3'', \sigma_3')$ is in general a complex number.  Therefore it is not directly interpretable as a probability.  Now $A(\sigma_3'', \sigma_3')$ is just a transition amplitude for finding a particle with physical outcome $\sigma_3''$ in one basis given a particle with the property $\sigma_3'$ in a different, physically rotated, basis.  If we say that probabilities are like intensities, not amplitudes, one might

guess from an analogy with the relation between intensity and amplitude of waves that

$$P(\sigma_3'', \sigma_3') \overset{?}{=} (A(\sigma_3'', \sigma_3'))^2 ,$$

where $P(\sigma_3'', \sigma_3')$ is the associated probability for this transition.  However, because $A(\sigma_3'', \sigma_3')$ is not real this relation does not yield a real number in general. The next simplest guess is that

$$P(\sigma_3'', \sigma_3') = |A(\sigma_3'', \sigma_3')|^2 \tag{165}$$

where $|\cdots|$ denotes an absolute value.  We will now explicitly calculate $A(\sigma_3'', \sigma_3')$ using the mathematical machinery we have developed to show that (165) gives the correct observed probabilities given in (66).

Let us recall the geometrical situation:
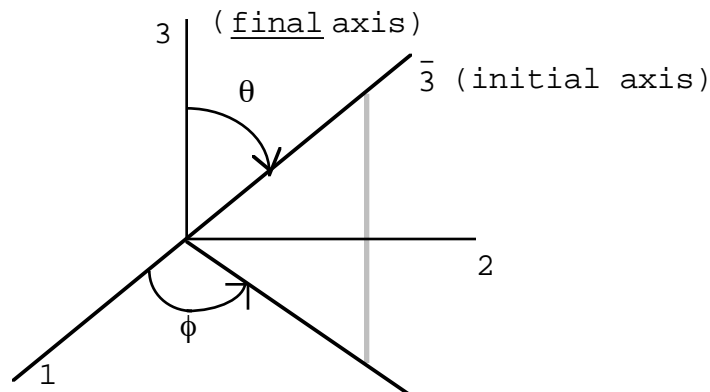


This situation was studied before where we found that

$$U = e^{i\theta/2 \ \sigma_2} e^{i\phi/2 \ \sigma_3} \tag{166}$$

describes the transformation between the barred and unbarred frames.  Although in the situation shown an angle $\phi$ is

necessary to describe the general relationship of the 3 and $\overline{3}$

axes, we expect from our experimentally observed results, (66),

that our answers for the various probabilities will actually be

independent of $\phi$. Let us see if this is so.

Using (160) in (163) gives

$$A(\sigma_3'', \sigma_3') = <\sigma_3''|U^{-1}|\sigma_3'>. \tag{167}$$

[ Eq$^n$(167) allows a more elegant interpretation of $A(\sigma_3'', \sigma_3')$ in

terms of Process Diagrams, but one which is more unlike the

real S-G setup.  Instead of the earlier measurement involving

the $|\sigma_3''><\sigma_3''|$ and $|\overline{\sigma_3'}><\overline{\sigma_3'}|$ symbols, we can also deduce

$A(\sigma_3'', \sigma_3')$ from the diagrammatic equation
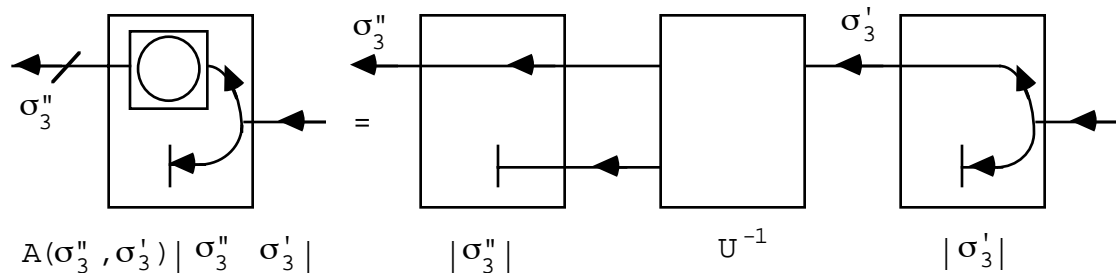


That is, instead of rotating the magnets, we can think of the

$U^{-1}$ above rotating the beam.]

To begin our evaluations, let's choose $\sigma_3' = \sigma_3'' = +$.  Then

$$A(+,+) = <+|e^{-i\phi/2\ \sigma_3}\ e^{-i\theta/2\ \sigma_2}|+>. \tag{168}$$

Some reminders:

$$\sigma_1 = |->< +| + |+>< -|,$$
$$\sigma_2 = i(|->< +| - |+>< -|),$$
$$\sigma_3 = |+>< +| - |->< -|.$$

1.70

Therefore from orthonormality

$$\sigma_1 |+> \ = \ |->, \quad \sigma_1 |-> \ = \ |+>,$$

$$\sigma_2 |+> \ = \ i|->, \quad \sigma_2 |-> \ = \ -i|+>,$$

$$\sigma_3 |+> \ = \ |+>, \quad \sigma_3 |-> \ = \ -|->.$$

(168) now becomes

$$A(+,+) \ = \ <+|e^{-i\phi/2}(\cos \frac{\theta}{2} \ - \ i\sigma_2 \ \sin \frac{\theta}{2} \ |+)$$

$$= \ e^{-i\phi/2} \ <+| \ \cdot \ (|+> \cos \frac{\theta}{2} \ + \ |-> \sin \frac{\theta}{2} \ )$$

$$= \ e^{-i\phi/2} \ \cos \frac{\theta}{2} \ . \tag{169}$$

Thus we get the correct result:

$$p(+,+) \ = \ |A(+,+)|^2 \ = \ \cos^2 \frac{\theta}{2} \ .$$

Likewise, you will show in a problem that

$$A(+,-) \ = \ -e^{-i\phi/2} \ \sin \frac{\theta}{2} \ , \tag{170}$$

$$A(-,+) \ = \ e^{i\phi/2} \ \sin \frac{\theta}{2} \ , \tag{171}$$

$$A(-,-) \ = \ e^{i\phi/2} \ \cos \frac{\theta}{2} \ , \tag{172}$$

which also give correct probabilities.

We have now found the explicit connection between the Process Diagrams and the real-space experimental setup:

**Drawing the Process Diagram of some experimental measurement identifies the transition amplitude of that process; the corresponding probability is the absolute square of the amplitude.**

We know that this works now for the two-magnet S-G setup; let's test this out on a three-magnet experiment. It will yield a valuable lesson. Consider the setup (set $\phi = \phi'=0$ for simplicity):



magnet 3

magnet 2

magnet 1
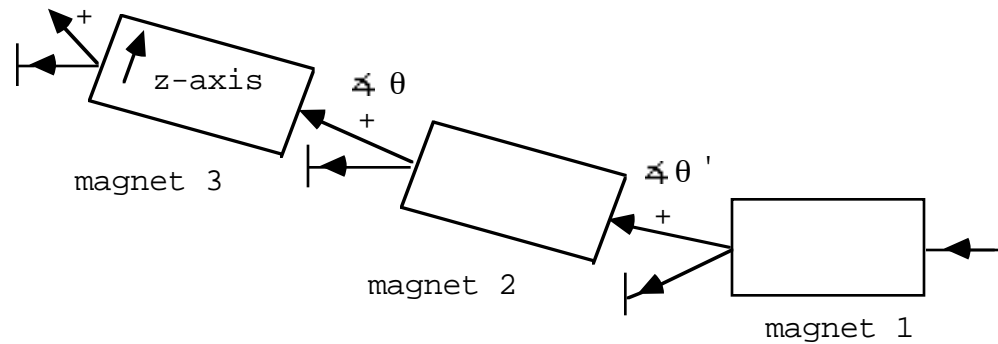
The Process Diagram is



$$|+| \qquad \cdot \qquad |\overline{+}| \qquad \cdot \qquad |\overline{\overline{+}}|$$
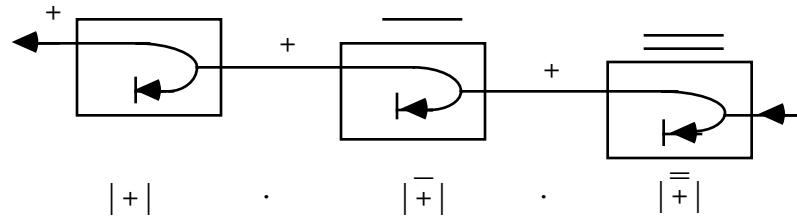
The transition amplitude is identified from

$$|+| \cdot |\overline{+}| \cdot |\overline{\overline{+}}| \; = \; |+><+|\overline{+}><\overline{+}|\overline{\overline{+}}><\overline{\overline{+}}|,$$

as

$$A \; = \; <+|\overline{+}><\overline{+}|\overline{\overline{+}}>.$$

The probability is then

$$P_1 \; = \; |A|^2 \; = \; |<+|\overline{+}>|^2 |<\overline{+}|\overline{\overline{+}}>|^2 \; = \; \cos^2\frac{\theta}{2}\cos^2\frac{\theta'}{2},$$

where we have used (169), with a proper understanding of the role of the angles $\theta$ and $\theta'$. What would we have gotten if we had chosen spin "down" from the second magnet? The appropriate diagram is:
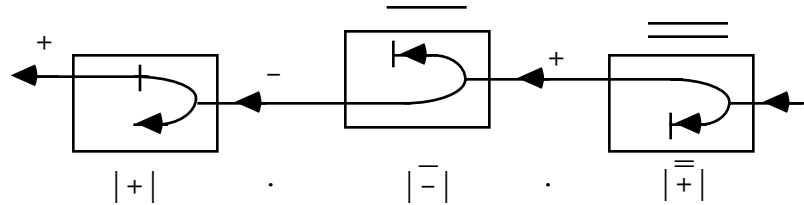
The transition amplitude from

$$| + | \cdot | \overline{-} | \cdot | \overline{\overline{\mp}} | ,$$

is

$$A = < + | \overline{-} > < \overline{-} | \overline{\overline{\mp}} > .$$

The probability is then

$$P_2 = |A|^2 = |< + | \overline{-} >|^2 | < \overline{-} | \overline{\overline{\mp}} >|^2 = \sin^2 \frac{\theta}{2} \sin^2 \frac{\theta'}{2} .$$

Given <u>both</u> choices of the intermediate magnet, the probability that final "up" is selected, given the selection "up" from magnet 1, is

$$P = P_1 + P_2 = \cos^2 \frac{\theta}{2} \cos^2 \frac{\theta'}{2} + \sin^2 \frac{\theta}{2} \sin^2 \frac{\theta'}{2} . \qquad (173)$$

Now let's remove the second magnet entirely, without changing the orientation of magnets one and three.  We know the probability for
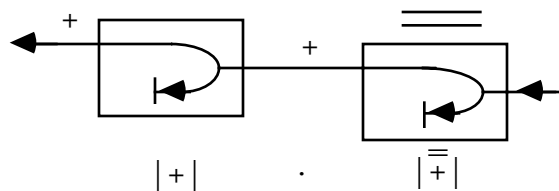


is just

$$P' = | < + | \overline{\overline{\mp}} >|^2 = \cos^2 \left( \frac{\theta + \theta'}{2} \right) . \qquad (174)$$

We get a useful alternate view of the latter probability by inserting completeness (in the single-barred basis),

$$1 = |\overline{+}| + |\overline{-}| = |\overline{+}><\overline{+}| + |\overline{-}><\overline{-}|,$$

as follows:

$$P' = |<\overline{+}|1|\overline{\mp}>|^2 = |<+|\overline{+}><\overline{+}|\overline{\mp}> + <+|\overline{-}><\overline{-}|\overline{\mp}>|^2.$$

Using (169), (170) and (171) here then yields

$$P' = \left|\cos\frac{\theta}{2}\cos\frac{\theta'}{2} - \sin\frac{\theta}{2}\sin\frac{\theta'}{2}\right|^2. \qquad (175)$$

Mathematically, this is the same as (174) of course. Comparing (175) to (173) we see a strong similarity. Eq$^{\underline{n}}$(173) represents the sum of two <u>probabilities</u> whereas (182) arises from the sum of two <u>amplitudes</u>. We say that the two amplitudes that make up (173) add <u>incoherently</u> while the amplitudes in (175) add <u>coherently</u>. Using this terminology, we state an important quantum mechanical principle that is being seen here as an example. That is:

$$\begin{pmatrix}\textbf{Destinguishable}\\ \textbf{Indestinguishable}\end{pmatrix} \quad \textbf{processes add} \quad \begin{pmatrix}\textbf{incoherently}\\ \textbf{coherently}\end{pmatrix}.$$

Thus, there is a difference in the final outcome whether the second magnets are present or not. The act of observing whether an individual particle has spin "up" or "down" in an intermediate stage has altered the experimental outcome. Although the above principle has been stated in the context of the behavior of spin, it is actually a completely general quantum mechanics rule. A paraphrase of the above could be: The fundamental quantum mechanical objects are amplitudes, not probabilities.

We can now recover and get an interpretation of some of our previous results. Let us consider the situation of the two magnet S-G setup described from an arbitrary orientation.
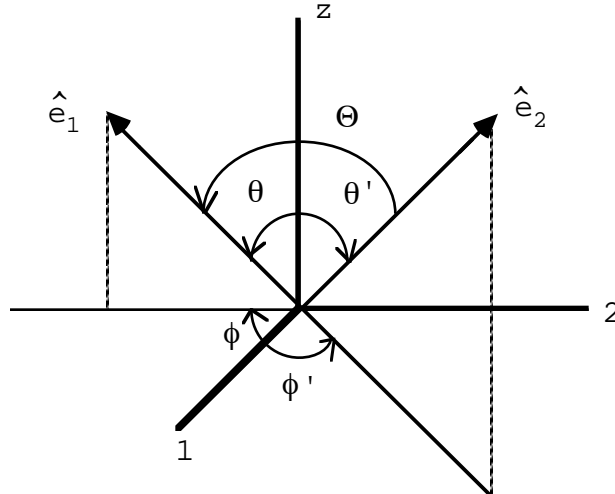


(We already saw the above figure on p. 1.30; $\hat{e}_1$ represents the initial $\bar{z}$-axis and $\hat{e}_2$ represents the final $\bar{\bar{z}}$-axis.)  The spin basis in the single-barred frame we take to be $\{|\bar{\sigma}'_3>\}$ and in the double-barred frame we use $\{|\bar{\bar{\sigma}}'_3>\}$.  Then, from the Process Diagram we identify the transition probability,

$$p(\sigma''_3, \sigma'_3) \; = \; |<\bar{\bar{\sigma}}''_3|\bar{\sigma}'_3>|^2,  \qquad\qquad (176)$$

where $\sigma'_3, \sigma''_3 = \pm 1$ independently as usual.  We are referring both S-G's to a third, independent axis.  This is accomplished by writing

$$p(\sigma''_3, \sigma'_3) \; = \; |<\bar{\bar{\sigma}}''_3| \; 1 \; |\bar{\sigma}'_3>|^2,  \qquad\qquad (177)$$

with

$$\sum_{\sigma3} |\sigma_3><\sigma_3| = 1, \qquad\qquad (178)$$

where the basis $\{|\sigma_3>\}$ refers to the unbarred frame.  Therefore

$$p(\sigma_3'',\sigma_3') \;=\; |\sum_{\sigma_3} <\bar{\bar{\sigma}}_3''|\sigma_3><\sigma_3|\bar{\sigma}_3'>|^2 \tag{179}$$

We have already calculated the following:

$$<+|\bar{+}> \;=\; e^{-i\phi/2}\cos\frac{\theta}{2}\;,$$

$$<+|\bar{-}> \;=\; -e^{-i\phi/2}\sin\frac{\theta}{2}\;,$$

$$<-|\bar{+}> \;=\; e^{i\phi/2}\sin\frac{\theta}{2}\;,$$

$$<-|\bar{-}> \;=\; e^{i\phi/2}\cos\frac{\theta}{2}\;.$$

The new objects we need to calculate are the

$$<\bar{\bar{\sigma}}_3''|\sigma_3> \;=\; <\sigma_3''|U|\sigma_3>, \tag{180}$$

with the same U as on p.1.68 with $\theta,\phi \rightarrow \theta',\phi'$. Using this form, in the same way as before we find

$$<\bar{\bar{+}}|+> \;=\; e^{i\phi'/2}\cos\frac{\theta'}{2}\;, \tag{181}$$

$$<\bar{\bar{+}}|-> \;=\; e^{-i\phi'/2}\sin\frac{\theta'}{2}\;, \tag{182}$$

$$<\bar{\bar{-}}|+> \;=\; -e^{i\phi'/2}\sin\frac{\theta'}{2}\;, \tag{183}$$

$$<\bar{\bar{-}}|-> \;=\; e^{-i\phi'/2}\cos\frac{\theta'}{2}\;. \tag{184}$$

Let's examine the structure of p(+,+):

$$p(+,+) \;=\; |<\bar{\bar{+}}|+><+|\bar{+}> \;+\; <\bar{\bar{+}}|-><-|\bar{+}>|^2$$

$$=\; |(e^{i\phi'/2}\cos\frac{\theta'}{2})(e^{-i\phi/2}\cos\frac{\theta}{2})$$

$$+\; (e^{-i\phi'/2}\sin\frac{\theta'}{2})(e^{i\phi/2}\sin\frac{\theta}{2})|^2$$

$$= \left| (e^{-i\phi'/2} \cos \frac{\theta'}{2}, e^{i\phi'/2} \sin \frac{\theta'}{2})^* \begin{pmatrix} e^{-i\phi/2} \cos \frac{\theta}{2} \\ e^{i\phi/2} \sin \frac{\theta}{2} \end{pmatrix} \right|^2$$

$$= |\psi_+(\theta',\phi')^+ \, \psi_+(\theta,\phi)|^2, \tag{185}$$

where we have used the previous matrix definition of $\psi_+(\theta,\phi)$,

equation (80).  Likewise, we may write

$$p(-,+) = \left| <\bar{\bar{\sigma}}|+> <+|\bar{+}> + <\bar{\bar{\sigma}}|-> <-|\bar{+}> \right|^2$$

$$= \left| (-e^{i\phi'/2} \sin \frac{\theta'}{2} )(e^{-i\phi/2} \cos \frac{\theta'}{2} ) \right.$$

$$\left. + (e^{-i\phi'/2} \cos \frac{\theta'}{2} )(e^{i\phi/2} \sin \frac{\theta}{2} ) \right|^2$$

$$= \left| (-e^{-i\phi'/2} \sin \frac{\theta'}{2}, e^{i\phi'/2} \cos \frac{\theta'}{2})^* \begin{pmatrix} e^{i\phi/2} \cos \frac{\theta}{2} \\ e^{i\phi/2} \sin \frac{\theta}{2} \end{pmatrix} \right|^2$$

$$= |\psi_-(\theta',\phi')^+ \, \psi_+(\theta,\phi)|^2, \tag{186}$$

where we have now used the definition of the matrix $\psi_-(\theta,\phi)$,

equation (81).  Equations (185) and (186) are the exact same as

the forms for $p(+,+)$ and $p(-,+)$ found in (82) and (83).  We

make several realizations from our recovery of the previous

results.  First, we notice that when $\theta' = \theta$ and $\phi' = \phi$ (then,

the distinction between the $|\bar{\sigma}_3' >\}$ and $\{|\bar{\sigma}_3' >\}$ states

disappears), we have from (169) – (172) (listed again above

(180)) and (181) – (184) that

$$< \bar{\sigma}_3' |\sigma_3 >^* = < \sigma_3 |\bar{\sigma}_3' > \tag{187}$$

for any value of $\sigma_3$, $\sigma_3'$.  Notice that (187) is consistent with

(41) and so conserves probabilities. Equation (187) suggests

that the act of complex conjugation, or actually some <u>extension</u>

of its usual meaning, interchanges bras and kets.  Let's define

an operator "+" that does this.  In general for any bra or ket,

$$(<a'|)^+ \equiv |a'>, \qquad (188)$$

$$(|a'>)^+ \equiv <a'|. \qquad (189)$$

We require that the "+" operation not change the <u>character</u> of

the mathematical object it acts on.  That is, it reduces to

complex conjugation when acting on a number, but when acting on

an operator just gives another operator.  (It does not change

an operator into a number or vice versa.)  What's more, we

assume this operation is completely <u>distributive</u>.

Let's test out the effect of "+", which we will call

<u>Hermitian</u> <u>conjugation</u> (or "Hermitian adjoint" or just

"adjoint") on the operators $\sigma_{1,2,3}$:

$$\sigma_3^+ = [\,|+><+|\ -\ |-><-|\,]^+ = \sigma_3 \qquad (190)$$

$$\sigma_2^+ = [\,i|-><+|\ -\ i|+><-|\,]^+ = \sigma_2 \qquad (191)$$

$$\sigma_1^+ = [\,|-><+|\ +\ |+><-|\,]^+ = \sigma_1 \qquad (192)$$

Operators that satisfy $A^+ = A$ are said to be <u>Hermitian</u> or

<u>self-conjugate</u>.  Such operators are of fundamental importance

in quantum mechanics; we will discuss the reason for this later

in this Chapter.  Notice now that we have two operations

designated as "+".  The first time we used this symbol, it was

in a matrix context, and it meant "complex conjugation + transpose." (See p.1.34). In the present context of Dirac bra-ket notation it now means "complex conjugation + bra-ket interchange." Mathematically, it would be better to introduce a distinction between these two uses of the same symbol. In practice, however, physicists let the context tell them which usage is appropriate. We will follow this point of view here. That there is a connection between these sets of mathematical operations will be evident in a moment.

Let us be crystal clear about what we have done in introducing "$+$" in this context. Because the quantity $< \overline{\sigma}_3' | \sigma_3 >$ is just a (complex) number, we have the trivial statement that

$$< \overline{\sigma}_3' | \sigma_3 >^+ = < \overline{\sigma}_3' | \sigma_3 >^* \qquad (193)$$

On the other hand, using (188) and (189) gives

$$< \overline{\sigma}_3' | \sigma_3 >^+ = < \sigma_3 | \overline{\sigma}_3' >. \qquad (194)$$

Comparing (193) and (194) gives (187). Thus, the existence of an operation which interchanges bras and kets, but which reduces to ordinary complex conjugation when applied to complex numbers, renders (187) an identity. We will assume the existence of such an operation.

The existence of the Hermitian conjugation operation has another important consequence. From the left hand side of (194) we must have

$$< \overline{\sigma}_3' | \sigma_3 >^+ = < \sigma_3' | U | \sigma_3 >^+ = < \sigma_3 | U^+ | \sigma_3' >, \qquad (195)$$

whereas the equality stated there demands this is the same as

$$< \sigma_3 \,|\, \overline{\sigma}_3' > \; = \; < \sigma_3 |U^{-1}| \sigma_3' >. \tag{196}$$

Comparing the right hand sides of (195) and (196), for any states $< \sigma_3 |, | \sigma_3' >$, demands that

$$U^{-1} = U^{+} \tag{197}$$

Such an operator as in (197) is called <u>unitary</u>.  We now ask the question:  Are the U-transformation operators we have written down so far unitary?  Let us concentrate on the transformation (166) (on p.1.68).  First, we derive an important rule for Hermitian conjugation.  We have by definition

$$(< a' \,|\, U_1 U_2)^{+} \; = \; (U_1 U_2)^{+} |a' >. \tag{198}$$

However, we may also write

$$(< a' \,|\, U_1 U_2)^{+} \; = \; (< \overline{a}' \,|\, U_2)^{+} \; = \; U_2^{+} |\, \overline{a}' >, \tag{199}$$

where

$$|\, \overline{a}' > \; = \; (< \overline{a}' \,|\,)^{+} \; = \; (< a' \,|\, U_1)^{+} \; = \; U_1^{+} |\, \overline{a}' >. \tag{200}$$

Substituting (200) in (199) gives

$$(< a' \,|\, U_1 U_2)^{+} \; = \; U_2^{+} U_1^{+} |\, \overline{a}' >. \tag{201}$$

Then comparing (198) and (201) for any $|a' >$ then tells us that

$$(U_1 U_2)^{+} \; = \; U_2^{+} U_1^{+}. \tag{202}$$

Although the rule (202) has been stated for unitary operators, it is in fact true of <u>any</u> product of operators.

Now applying (202) to equation (166) gives

$$U^+ = (e^{i\theta/2 \ \sigma_2} e^{i\phi/2 \ \sigma_3})^+ = (e^{i\phi/2 \ \sigma_3})^+ (e^{i\theta/2 \ \sigma_2})^+. \qquad (203)$$

Remembering that (equation (141))

$$e^{i\lambda(\hat{n} \cdot \vec{\sigma})} = \cos\lambda + i(\hat{n} \cdot \vec{\sigma})\sin\lambda,$$

and from (190) – (192) above, we then have

$$U^+ = (e^{i\phi/2 \ \sigma_3})^+ (e^{i\theta/2 \ \sigma_2})^+ = e^{-i\phi/2 \ \sigma_3} e^{-i\theta/2 \ \sigma_2}. \qquad (204)$$

The last form on the right is identically $U^{-1}$, so we have shown that this U is unitary.  We will have a lot more to say about unitary operators as we go along.

Generalizing (185) and (186), we have

$$p(\sigma_3'', \sigma_3') = |\psi_{\sigma_3''}(\theta',\phi')^+ \ \psi_{\sigma_3'}(\theta,\phi)|^2, \qquad (205)$$

which is identical to (84) above if we change the names a', a" to $\sigma_3', \sigma_3''$.  We called the $\psi_{\sigma_3}(\theta,\phi)$ "wavefunctions."  Another realization we make comes from comparing (205) with the earlier expression (179), written in the form

$$p(\sigma_3'', \sigma_3') = |\sum_{\sigma_3} <\sigma_3|\bar{\bar{\sigma}}_3'' >^* <\sigma_3|\bar{\sigma}_3'>|^2, \qquad (206)$$

which now reveals explicitly that

$$\left[ \psi_{\sigma_3'}(\theta,\phi) \right]_{\sigma_3} = <\sigma_3 | \bar{\sigma}_3' >, \qquad (207)$$

where the $\sigma_3$ on the left hand side of this equation is being viewed as the row index of $\psi_{\sigma_3'}(\theta,\phi)$.  Thus the components of $\psi_{\sigma_3'}(\theta,\phi)$ are revealed as just the <u>transition</u> <u>amplitudes</u> in equation (163) above.  This gives us a concrete interpretation of the wavefunction, for which we use the following figure:
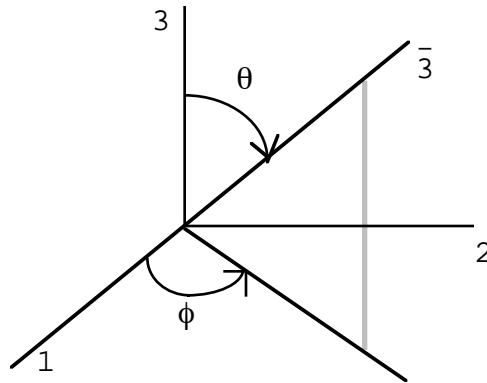
$$\left\{ \begin{array}{l} \left[\, \psi_{\sigma_3'}(\theta,\phi)\,\right]_{\sigma_3} \text{ is the transition amplitude for spin } \frac{1}{2} \\ \text{selected along the } \bar{3} \text{ axis } (\sigma_3' = \pm 1) \text{ having component} \\ \sigma_3 = \pm 1 \text{ along the 3-axis.} \end{array} \right\}$$

Explicitly again

$$\psi_+(\theta,\phi) \; = \; \begin{pmatrix} <+|\,\bar{+}\,> \\ <-|\,\bar{+}\,> \end{pmatrix} = \begin{pmatrix} e^{-i\phi/2}\, \cos\dfrac{\theta}{2} \\ e^{i\phi/2}\, \sin\dfrac{\theta}{2} \end{pmatrix}, \qquad (208)$$

$$\psi_-(\theta,\phi) \; = \; \begin{pmatrix} <+|\,\bar{-}\,> \\ <-|\,\bar{-}\,> \end{pmatrix} = \begin{pmatrix} -e^{-i\phi/2}\, \sin\dfrac{\theta}{2} \\ e^{i\phi/2}\, \cos\dfrac{\theta}{2} \end{pmatrix}. \qquad (209)$$

The fact that the wavefunctions above can be displayed as column matrices suggests that the rest of the algebra we have introduced involving the spin operators $\sigma_{1,2,3}$ can also be viewed as matrix manipulations.  Indeed this is so, and we will discuss how this can be done for spin $\frac{1}{2}$ below, deferring a more general discussion until we come to Chapter 4.  The basic idea is as follows.  Instead of using the spin $\frac{1}{2}$ Hilbert-space operators $\sigma_{1,2,3}$, we may instead use matrix

<u>representations</u> of them.  Given a spin basis $\{|\sigma_3'>\}$ and an operator A, one may produce a two-indexed object $A_{\sigma_3'\sigma_3''}$ as in

$$A_{\sigma_3'\sigma_3''} = <\sigma_3'|A|\sigma_3''>$$

(210)

If we now interpret $\sigma_3'$ as the row index and $\sigma_3''$ as the column index, we may regard this object as a matrix. Explicitly, for $\sigma_{1,2,3}$ we find

$$\sigma_1 = |->< +| + |+><-| \Rightarrow \sigma_1 = \begin{array}{c} \\ + \\ - \end{array} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

with column/row labels $\sigma_3'/\sigma_3''$ : $+$ $-$

$$\sigma_2 = i|-><+| - i|+><-| \Rightarrow \sigma_2 = \begin{array}{c} \\ + \\ - \end{array} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

with column/row labels $\sigma_3'/\sigma_3''$ : $+$ $-$

$$\sigma_3 = |+><+| - |-><-| \Rightarrow \sigma_3 = \begin{array}{c} \\ + \\ - \end{array} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

with column/row labels $\sigma_3'/\sigma_3''$ : $+$ $-$

Notice that I have not attempted to assign a different symbol to the $\sigma$'s when they are regarded as matrices, although to be mathematically clear we probably should.  This again is the common usage; the context will tell us what is meant.  The $\sigma_{1,2,3}$ matrices above are called the <u>Pauli</u> <u>matrices</u>.

Using our matrix language now, let us verify some of the prior algebraic statements involving the $\sigma$'s.  First, let us notice that (208) and (209) become, when $\theta = \phi = 0$,

$$\Psi_+(0,0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \Psi_-(0,0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

(211)

We usually call $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ "spin up" and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ "spin down."  Then in terms of this matrix language, let us write a few examples of how the matrix algebra works.

| Operator statement | Matrix realization |
|---|---|
| $<+\lvert\sigma_1 = <-\rvert$ | $(1 \quad 0) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (0 \quad 1)$ |
| $<-\lvert\sigma_1 = <+\rvert$ | $(0 \quad 1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (1 \quad 0)$ |
| $<+\lvert\sigma_2 = -i<-\rvert$ | $(1 \quad 0) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = (0 \quad -i)$ |
| $<-\lvert\sigma_2 = i<+\rvert$ | $(0 \quad 1) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = (i \quad 0)$ |
| $\sigma_1\lvert +> = \lvert ->$ | $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ |
| $\sigma_1\lvert -> = \lvert +>$ | $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ |
| $\sigma_2\lvert +> = i\lvert ->$ | $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ i \end{pmatrix}$ |
| $\sigma_2\lvert -> = -i\lvert +>$ | $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -i \\ 0 \end{pmatrix}$ |

One may also easily check that the algebraic properties of the $\sigma_i$ stated in (149) above also hold for the Pauli matrices. Although there is an ambiguity in our notation now whether when we write $\sigma_3$, say, we mean the matrix representation or the more abstract operator quantity, let us be clear about the difference of the two in our minds.  The more fundamental of the two is the Hilbert space operator quantity.  The matrix representation is just a realization of the basic operator quantity in a particular basis.  By changing to a rotated basis

$\{|\,\overline{\sigma}_3' >\}$ we could in fact change the representation, although the basic underlying operator structure has not changed. The relationship between operator and representation is summed up very picturesquely in a footnote on p.20 of J.J. Sakurai's <u>Advanced</u> <u>Quantum</u> <u>Mechanics</u>: "The operator is different from a representation of the operator just as the actress is different from a poster of the actress." The fact that there are two languages of spin, operator and matrix, explains why the same symbol "$+$" was introduced in two apparently different contexts. In a matrix context, if you remember, it meant: "complex conjugation + transpose." In an operator context it meant: "complex conjugation + bra-ket interchange." For example, we can now verify the statements (190) – (192) using the first meaning of "$+$":

$$\sigma_3^+ = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^+ = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_3 ,$$

$$\sigma_2^+ = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}^+ = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_2 ,$$

$$\sigma_1^+ = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^+ = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_1 .$$

It should be clear now why I did not introduce a distinction between the two senses of the adjoint symbol. It is because one is carrying out the same basic mathematical operation, but in different contexts. We will go over this point again in a later chapter.

While we are on the subject of the matrix language of spin, let us see if we can recover the $\psi_{\sigma_3'}(\theta, \phi)$ wave functions,

which we have already derived from operator methods, using instead matrix manipulations.

From (102) above, we know that

$$A|a'| = a'|a'|.$$

Since $|a'| = |a'><a'|$, and multiplying on the right by $|a'>$, gives

$$A|a'> = a'|a'> \tag{212}$$

We know that $\sigma_3$ plays the role of "A" above in the two-physical-outcome case, which we know now as spin $\frac{1}{2}$ . Therefore

$$\sigma_3|\sigma_3'> = \sigma_3' |\sigma_3' > \tag{213}$$

The matrix realization of this statement is

$$\sigma_3 \ \psi_{\sigma_3'}(0,0) = \sigma_3' \ \psi_{\sigma_3'}(0,0) \tag{214}$$

where $\sigma_3$ is now a matrix. Likewise, in a rotated basis, the analog of (213) is

$$\overline{\sigma}_3 \ |\overline{\sigma}_3' > = \sigma_3' \ |\overline{\sigma}_3' >. \tag{215}$$

There are of course only two values allowed for $\sigma_3'$ above ($\sigma_3' = \pm 1$ as usual), but $\overline{\sigma}_3$ is the rotated version of $\sigma_3$. We already know how to compute $\overline{\sigma}_3$. It is given by

$$\overline{\sigma}_3 = U^{-1}\sigma_3 U \tag{216}$$

with

$$U = e^{i\theta/2 \ \sigma_2} \ e^{i\phi/2 \ \sigma_3}. \tag{217}$$

(You will carry out the above evaluation of $\overline{\sigma}_3$ in a problem.) Another, easier, way of deriving this quantity is simply to

project out the component along the new $\overline{3}$ axis. With the help of the unit vector

$$\hat{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \qquad (218)$$

pointing in the direction given by the spherical coordinate angles $\theta$ and $\phi$, we find $\overline{\sigma}_3$ as

$$\overline{\sigma}_3 = \vec{\sigma} \cdot \hat{n} = \sigma_1 \sin \theta \cos \phi + \sigma_2 \sin \theta \sin \phi + \sigma_3 \cos \theta. \qquad (219)$$

We therefore have

$$\overline{\sigma}_3 = \sin \theta \cos \phi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \sin \theta \sin \phi \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \cos \theta \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\Rightarrow \overline{\sigma}_3 = \begin{pmatrix} \cos \theta & e^{-i\phi} \sin \theta \\ e^{i\phi} \sin \theta & -\cos \theta \end{pmatrix}. \qquad (220)$$

The matrix realization of (215) is then

$$\begin{pmatrix} \cos \theta & e^{-i\phi} \sin \theta \\ e^{i\phi} \sin \theta & -\cos \theta \end{pmatrix} \begin{pmatrix} \psi_{\sigma_3'}(+) \\ \psi_{\sigma_3'}(-) \end{pmatrix} = \sigma_3' \begin{pmatrix} \psi_{\sigma_3'}(+) \\ \psi_{\sigma_3'}(-) \end{pmatrix}, \qquad (221)$$

where we are calling the upper and lower matrix components of $\psi_{\sigma_3'}(\theta,\phi)$ as $\psi_{\sigma_3'}(+)$ and $\psi_{\sigma_3'}(-)$. Eq$^{\underline{n}}$(221) represents two equations in two unknowns. Written out explicitly, we have

$$\cos \theta \, \psi_{\sigma_3'}(+) + e^{-i\phi} \sin \theta \, \psi_{\sigma_3'}(-) = \sigma_3' \, \psi_{\sigma_3'}(+), \qquad (222)$$

$$e^{i\phi} \sin \theta \, \psi_{\sigma_3'}(+) - \cos \theta \, \psi_{\sigma_3'}(-) = \sigma_3' \, \psi_{\sigma_3'}(-), \qquad (223)$$

or

$$(\cos \theta - \sigma_3') \psi_{\sigma_3'}(+) + e^{-i\phi} \sin \theta \, \psi_{\sigma_3'}(-) = 0, \qquad (224)$$

$$e^{i\phi} \sin \theta \; \psi_{\sigma_3'}(+) \; - \; (\cos \theta + \sigma_3') \psi_{\sigma_3'}(-) = 0. \qquad (225)$$

We know from the theory of linear equations that for the above two (homogeneous) equations to have a nontrivial solution, we must have that the determinant of the coefficients is zero. Therefore:

$$-(\cos \theta - \sigma_3')(\cos \theta + \sigma_3') - \sin^2 \theta = 0$$

$$\Rightarrow \quad \sigma_3'^{\;2} - \cos^2 \theta - \sin^2 \theta = 0$$

$$\Rightarrow \sigma_3'^{\;2} = 1. \qquad (226)$$

We have recovered the known result that $\sigma_3' = \pm 1$ only. Let's look at the $\sigma_3' = 1$ case now. Eq$^{\underline{n}}$ (224) now gives

$$-2 \sin^2 \frac{\theta}{2} \; \psi_+(+) \; + \; 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} e^{-i\phi} \psi_+(-) = 0$$

$$\Rightarrow \frac{\psi_+(+)}{\psi_+(-)} = e^{-i\phi} \frac{\cos \dfrac{\theta}{2}}{\sin \dfrac{\theta}{2}} . \qquad (227)$$

All that is determined is the <u>ratio</u> $\psi_+(+)/\psi_+(-)$. (Eq$^{\underline{n}}$(225) gives the same result.) Therefore, a solution is

$$\psi_+(+) = A \; e^{-i\phi/2} \cos \frac{\theta}{2}, \qquad (228)$$

$$\psi_+(-) = A \; e^{i\phi/2} \; \sin \frac{\theta}{2}, \qquad (229)$$

where "A" is a common factor. It is not completely arbitrary because these transition amplitudes give probabilities that must add to one:

$$\left| \psi_+(+) \right|^2 + \left| \psi_+(-) \right|^2 = 1 \qquad (230)$$

A statement like (230) is called <u>normalization</u>. Eq$\underline{\text{n}}$ (230) shows that we must have $|A|^2 = 1$. However, A is still not completely determined because we can still write $A = e^{i\alpha}$ with $\alpha$ some undetermined phase. Such an overall phase has no effect on probabilities however, since probabilities are the absolute square of amplitudes. The operator analog of the normalization condition (230) is just the statement

$$< \overline{+} | \overline{+} > = 1. \qquad\qquad (231)$$

(Compare (231) with equation (162) above). Question: can you derive (230) from (231)? [Hint: use completeness.]

So, outside of an arbitrary phase, we have recovered the result (215) for $\psi_+(\theta,\phi)$ which we previously derived using operator techniques. Similarly, we can recover the result (209) for $\psi_-(\theta,\phi)$ if we choose to look at either (224) or (225) for $\sigma_3' = -1$.

I would like to return now to our spin mesurements that we began this Chapter with to see how they now look in our emerging formalism. We found that the average value of an electron's magnetic moment along the z-axis, given an initial selection along the $+\overline{z}$-axis was

$$<\mu_z>_+ = -\mu(p(+,+) - p(-,+)). \qquad\qquad (232)$$

We saw that $<\mu_z>_+$ was obtained as the weighted sum of physical outcomes. In our new language of operators (remember $\mu_z = \gamma S_3$), consider the quantity (we have inserted completeness twice):

$$\langle \overline{+} | \mu_z | \overline{+} \rangle = \gamma \sum_{\sigma_3', \sigma_3'' = \pm 1} \langle \overline{+} | \sigma_3' \rangle \langle \sigma_3' | S_3 | \sigma_3'' \rangle \langle \sigma_3'' | \overline{+} \rangle. \qquad (233)$$

But

$$\langle \sigma_3' | S_3 | \sigma_3'' \rangle = \langle \sigma_3' | \frac{\hbar}{2} \sigma_3 | \sigma_3'' \rangle = \frac{\hbar}{2} \delta_{\sigma_3' \sigma_3''} \sigma_3', \qquad (234)$$

so

$$\langle \overline{+} | \mu_z | \overline{+} \rangle = \frac{\gamma \hbar}{2} \sum_{\sigma_3'} \sigma_3' \langle \overline{+} | \sigma_3' \rangle \langle \sigma_3' | \overline{+} \rangle$$

$$= \frac{\gamma \hbar}{2} \sum_{\sigma_3'} \sigma_3' | \langle \sigma_3' | \overline{+} \rangle |^2. \qquad (235)$$

This gives

$$\langle \overline{+} | \mu_z | \overline{+} \rangle = \frac{\gamma \hbar}{2} (p(+,+) - p(-,+)). \qquad (236)$$

Since $\mu = -\frac{\gamma \hbar}{2}$ for the electron, we learn that

$$\langle \mu_z \rangle_+ = \langle \overline{+} | \mu_z | \overline{+} \rangle. \qquad (237)$$

This is in a spin context. In more generality, we say that the <u>expectation</u> <u>value</u> of the operator A in the state $|\psi\rangle$ is

$$\langle A \rangle_\psi \equiv \langle \psi | A | \psi \rangle. \qquad (238)$$

If $A = A^+$, then expectation values are real:

$$\langle A \rangle_\psi^* = \langle \psi | A | \psi \rangle^* = \langle \psi | A | \psi \rangle^+$$

$$= \langle \psi | A^+ | \psi \rangle = \langle A \rangle_\psi. \qquad (239)$$

This is the reason, mentioned on p.1.77, that
Hermiticity of operators is so important.

I have two final points to make before closing this
(already too long) Chapter.  First, you may have noticed that
we have never asked any questions about the results of a S-G
measurement of the original beam of atoms coming directly from
the hot oven.  That is because we have not developed the
necessary formalism to describe such a situation. The formalism
we have developed assumes that any mixture of spin up and down
that we encounter is a <u>coherent</u> mixture, i.e. the beam of
particles is a mixture of the <u>amplitudes</u> of spin up and down.
The furnace beam however is a completely <u>incoherent</u> mixture of
spin up and down, i.e. the <u>probabilities</u> (or intensities) of
the components are adding.  There is a formalism for dealing
with the more general case of incoherent mixtures, but we will
not discuss such situations here.  In fact, our Process
Diagrams are incapable of representing an incoherent beam of
particles.  The "arbitrary beam" entering from the right in
such Diagrams is really not arbitrary at all but only
represents the most general coherent mixture.  Nevertheless, we
have seen the usefulness of these Diagrams in modeling
situations where the behavior of coherent beams is concerned.
If you are interested, you will find a useful discussion of
these matters in Sakurai's book, starting on p.174.  Such
considerations become important in any realistic experimental
situation where one must deal with beams which are only
partially coherent.

The other point I wish to make concerns the Process Diagrams.  They are only meant to be a stepping stone in our efforts to learn the principles of quantum mechanics.  These principles have a structure of their own and are independent of our Diagrams, which are an attempt to simply make <u>manifest</u> some of these principles.  The basic purpose of these Diagrams is to illustrate the existence of quantum mechanical states with certain measurable physical characteristics and to make transparent the means of computing such amplitudes.  You should now be in possession of a rudimentary intuitive and mathematical understanding of the simplest of all quantum mechanical systems, spin $\frac{1}{2}$.  We will now go on to see how the principles of quantum mechanics apply to other particle and system characteristics.

**Problems**

1. The total instantaneous power radiated from a nonrelativistic, accelerated charge is

$$P = \frac{2}{3} \frac{e^2}{c^3} |\vec{a}|^2 \quad ,$$

where $\vec{a}$ is the acceleration and $e$ and $c$ are the particle's charge and the speed of light, respectively. According to this classical result, a hydrogen atom should collapse in a very short amount of time because of energy loss. <u>Estimate</u> the time period for the hydrogen atom's electron to radiate away it's kinetic energy, $E = \frac{e^2}{2a_0}$. Take as a crude model of the atom an electron moving in a circular orbit of radius $a_0 = \frac{\hbar^2}{me^2}$, with a velocity $v^2 = \frac{e^2}{ma_0}$. ( These numerical values for $a_0$ and $v$ come from the simple Bohr model of the hydrogen atom, which we will study next Chapter.)

2. In the notes, I use the estimate $|\mu_z| \approx 10^{-20}$ erg/gauss for the magnitude of the Ag z-component magnetic moment. Reach a more accurate estimate of this value by looking up and plugging in values in

$$|\mu_z| = \frac{e}{mc} |S_z| \quad ,$$

assuming $|S_z| = \frac{\hbar}{2}$. Should one use the Ag-atom or the electron mass for $m$ ?

3. Think of a gas in thermal equilibrium at a temperature $T$, each atom acting as a tiny magnetic dipole, $\vec{\mu}$. Let's not worry what the source of this field is at present.

(Classically, magnetic dipole fields arise from current loops.) Imagine putting such atoms in an external magnetic field, $\vec{H}$. Of course, without $\vec{H}$, we would expect there to be no preferred direction and so no <u>net</u> gas magnetic moment. Let's use Maxwell-Boltzmann statistics to describe the gas:

$$\left. \begin{array}{l} \text{The probability that an atom} \\ \text{has a magnetic moment} \\ \text{between } \vec{\mu} \text{ and } \vec{\mu} + d\vec{\mu}. \end{array} \right\} = \frac{d\Omega e^{-\beta(-\vec{\mu}\cdot\vec{H})}}{\int d\Omega\, e^{-\beta(-\vec{\mu}\cdot\vec{H})}} \; .$$

Therefore

$$< \vec{\mu} >_T \;\; = \;\; \frac{\int d\Omega e^{\beta\vec{\mu}\cdot\vec{H}}\, \vec{\mu}}{\int d\Omega e^{\beta\vec{\mu}\cdot\vec{H}}} \;\; .$$

(a) Consider the situation where $\left| \vec{\mu}\cdot\vec{H} \right| \ll kT$ .

$$< \vec{\mu} >_T \;\; = \;\; \frac{\int d\Omega\vec{\mu}\left(1 + \frac{\vec{\mu}\cdot\vec{H}}{kT}\right)}{\int d\Omega} \;\; = \;\; \frac{0 + \frac{1}{kT}\int d\Omega\vec{\mu}(\vec{\mu}\cdot\vec{H})}{4\pi} \;\; .$$

Since $\vec{H}$ is the only direction in the problem, we hypothesize that

$$\int d\Omega\vec{\mu}\left(\vec{\mu}\cdot\vec{H}\right) \;\; = \;\; C\vec{H},$$

where "C" is just some constant. By taking $\vec{H}$ along the z-direction and using $\mu_z = \left|\vec{\mu}\right|\cos\theta$, find the value of C and show that for weak magnetic fields

$$< \vec{\mu} >_T \;\; \approx \;\; \frac{\vec{\mu}^2}{3kT}\,\vec{H}.$$

This result is called the "Curie law" and was established by Pierre Curie. We see that the collection of gas molecules

has a small net magnetic moment pointed <u>along</u> $\vec{H}$.  Such
behavior is called <u>paramagnetism</u>.

(b) Calculate   $<\vec{\mu}>_T$ *without* the weak-field
approximation. [Take the z-axis along  $\vec{H}$, so that  $\vec{\mu} \cdot \vec{H}$
= $\mu H \cos\theta$ . Then   $< \mu_x > = < \mu_y > = 0$ , but

$$< \mu_z > \;=\; \frac{\mu \int d\Omega \; \cos\theta \; e^{\beta\mu H\cos\theta}}{\int d\Omega e^{\beta\mu H\cos\theta}} \;,$$

where  $d\Omega$  is the element of solid angle. Plot
(qualitatively) your result for $< \mu_z >$ as a function of H.
Find the limits,

$$\lim_{H\to 0} \; < \mu_z >_T \;=\; ?,$$

$$\lim_{H\to\infty} \; < \mu_z >_T \;=\; ?.$$

and <u>give a physical interpretation of these two limits</u>.
[Hint: The top integral can be done by considering the
derivative of the denominator with respect to  $\beta$. The
answer to (b) is

$$< \mu_z >_T \;=\; \mu \left( \coth(\mu\beta H) - \frac{1}{\mu\beta H} \right). ]$$

4. Define (as in the notes)

$< \mu_z >_-$ = average value of $\vec{\mu}$ along the z-axis,
   given an initial selection of the <u>downward</u>
   deflected beam along z'.

Find $< \mu_z >_-$ . Does it agree with what you expect?

5.(a)  The energy of a charge, q, moving with velocity $\vec{v}$ in an external electromagnetic field is (classically)

$$E = q\,\Phi - \frac{q}{c}\,\vec{v}\cdot\vec{A},$$

where $\Phi$ is the scalar potential and $\vec{A}$ is the vector potential of the electromagnetic field.  The relationship between $\vec{A}$ and $\vec{H}$ is

$$\vec{H} = \vec{\nabla} \times \vec{A}.$$

For a constant $\vec{H}$ field in space, verify that

$$\vec{A} = \frac{1}{2}\,\vec{H} \times \vec{r}$$

is a possible form of the vector potential.  [Hint:  The vector identity,

$$\vec{\nabla} \times (\vec{a} \times \vec{b}) = \vec{a}(\vec{\nabla}\cdot\vec{b}) - \vec{b}(\vec{\nabla}\cdot\vec{a}) + (\vec{b}\cdot\vec{\nabla})\vec{a} - (\vec{a}\cdot\vec{\nabla})\vec{b}$$

may be of use.]

(b)  From the $-\dfrac{q}{c}\,\vec{v}\cdot\vec{A}$ term in E, identify a form for $\vec{\mu}$.

(c)  A moving charge will execute circular (or helical) motion in a constant $\vec{H}$ field. Using (b), show that $\vec{\mu} = \dfrac{q}{2mc}\,\vec{L}$, where $\vec{L}$ is the angular momentum of the particle. Assuming the plane of the orbit is perpendicular to $\vec{H}$, and given that $\vec{F} = q(\vec{E} + \dfrac{\vec{v}}{c} \times \vec{H})$ is the force, does $\vec{\mu}$ with q > 0 point along $\vec{H}$ or opposite to $\vec{H}$?  (Extra:  Can you think of a famous law of electromagnetism that explains this direction?)

6.  A well known formula of classical electromagnetism states that (see Jackson, p. 182)

$$|\vec{\mu}| = \frac{I}{c} A$$

where $\vec{\mu}$ is the magnetic moment and I $(= \frac{dq}{dt}$ where "q" is the charge) is the current in a circuit of area A.  Evaluate $|\vec{\mu}|$ for the classical spinning electron model in the notes (thin spherical shell model), and show that

$$|\vec{\mu}| = -\frac{1}{3}\frac{e\omega a^2}{c} = (-\frac{e}{2mc}) L,$$

where (-e) is the electron charge and L = I$\omega$ is the magnitude of the electronic angular momentum. [This gives the classical gyromagnetic ratio seen in Eq$^n$ (43) of the notes.]

7. Replace the thin-shelled spherical electron model with a solid spherical ball of charge, throughout which the charge is uniformly distributed.

    (a)  Find the moment of inertia of this object (replacing (61)).

    (b)  Show that the relation between $|\vec{\mu}|$ and L for this new model is still given by

$$|\vec{\mu}| = -\frac{e}{2mc} L .$$

[Use the technique in Problem 6 above to compute $|\vec{\mu}|$.]

8. (a)  Show that Eq.(79) of the text is equivalent to the
preceeding equation, (78).

(b)  Evaluate p(-,-) and p(+,-) from (84) in the case $\theta'=0$.
Are the results as expected?

9. Illustrate, using Process Diagrams, the product of
measurement symbols,

$$(\,|a'|\ +\ |a''|\,)|a'''|$$

in the cases where

(a) $a'= a'' \mp a''$
(b) $a' \mp a'' = a''$
(c) $a'= a'' = a''$.

(Consider the case with 4 possible physical outcomes, just
to be concrete.)

10. In the two-physical-property case ($a_1=1$, $a_2=-1$) evaluate

$$e^{i\lambda A}$$

as a function of A.   [Hint: Expand $e^{i\lambda A}$ in a power series
and use Eq. (112) in Ch. 1 of the notes, then sum the
resulting infinite series.]

11. In the three-physical-property case [$a_1=1$, $a_2=0$, $a_3=-1$]
find

(a)  The algebraic equation satisfied by $A = \sum\limits_{i} a_i\ |a_i|$.

(b)  $|+|$, $|0|$, $|-|$ as functions of A.
(c)  $e^{i\lambda A}$ as a function of A using part (a).
(d)  Give a Process Diagram interpretation to the
equation found in part (a). (Try to draw this neatly.

1.98

Supplement your diagram with an explanation in words, if you think this is necessary.)

12. Draw Process Diagrams representing the equations:

    (a)   $(|+ -|)^2 = 0$
    (b)   $\sigma_1{}^3 = \sigma_1$
    (c)   $\sigma_1 \sigma_3 = -i \sigma_2$

13. Using

$$\sigma_1 = |- +| + |+ -|, \quad \sigma_2 = i(|- +| + |+ -|),$$
$$\sigma_3 = |+| - |-|,$$

show

    (a)  $\sigma_2 \sigma_3 = i\sigma_1$

    (b)  $\sigma_3 \sigma_1 = i\sigma_2$

14. Using

$$\bar{\sigma}_3 = \sigma_3$$
$$\bar{\sigma}_2 = -\sigma_1 \sin \phi$$
$$\bar{\sigma}_1 = \sigma_1 \cos \phi + \sigma_2 \sin \phi$$

   and

$$\sigma_i{}^2 = 1$$
$$\sigma_i \sigma_j = i\sigma_k \quad (i, j, k \text{ cyclic})$$
$$\sigma_i \sigma_j = -\sigma_j \sigma_i \quad (i \neq j)$$

   show

    (a)  $\bar{\sigma}_1 \bar{\sigma}_2 = i\bar{\sigma}_3$
    (b)  $\bar{\sigma}_2 \bar{\sigma}_3 = i\bar{\sigma}_1$
    (c)  $\bar{\sigma}_3 \bar{\sigma}_1 = i\bar{\sigma}_2$

15. Show that

$$e^{i\phi(\hat{n}\cdot\vec{\sigma})} = \cos\phi + i(\hat{n}\cdot\vec{\sigma})\sin\phi$$

where $\hat{n}$ is an arbitrary unit vector ($\hat{n}^2 = 1$). [Hint: First show that $(\hat{n}\cdot\vec{\sigma})^2 = 1$, then expand the exponential in a power series, remembering problem 10 above.]

16. Show that

$$\frac{1}{2}\vec{\sigma} \times \frac{1}{2}\vec{\sigma} = \frac{i}{2}\vec{\sigma} .$$

Remember that $(\vec{A} \times \vec{B})_i = \sum_{j,k} \varepsilon_{ijk} A_j B_k$. [Hint: It may be useful to recall that

$$\sum_{j,k} \varepsilon_{ijk} \varepsilon_{\ell jk} = 2\delta_{i\ell} . ]$$

17. Finish the equation:

$$\sum_{k=1}^{3} \sigma_k \vec{\sigma} \sigma_k = ?$$

18. Verify the results stated in Eqs. (170), (171), (172) of the notes.

19. Most general rotation:

$$U = e^{\psi/2\sigma_3} e^{i\theta/2\sigma_2} e^{i\phi/2\sigma_3} .$$

Components of the wavefunction:

$$[\Psi_{\sigma_3'}(\psi, \theta, \phi)]_{\sigma_3} = ?$$

1.100

[The change in $\Psi$ is very trivial compared to our earlier form for $[\Psi_{\sigma_3'}(\theta, \phi)]_{\sigma_3}$.]

20. Verify that

$$\bar{\sigma}_3 = \sigma_1 \sin\theta \cos\phi + \sigma_2 \sin\theta \sin\phi + \sigma_3\cos\theta$$

is produced by

$$\bar{\sigma}_3 = U^{-1} \sigma_3 U, \quad U = e^{i\theta/2\sigma_2} e^{i\phi/2\sigma_3}$$

[Hints: Simply use the algebraic properties of the $\sigma$'s, write the exponentials in their factorial forms (e.g., $e^{i\phi/2\sigma_3} = \cos\frac{\phi}{2} + i \sigma_3 \sin\frac{\phi}{2}$) and do the algebra.]

21. Just as the two physical outcome case ($a_1 = 1$, $a_2 = -1$) can be used to represent spin $\frac{1}{2}$, the three physical outcome case ($a_1 = 1$, $a_2 = 0$, $a_3 = -1$) can be used to represent spin one. The components of the spin vector $\vec{S}$ can be taken as

$S_1 = i\hbar(|-0| - |0-|)$,
$S_2 = i\hbar(|+-| - |-+|)$,
$S_3 = i\hbar(|0+| - |+0|)$.

(a) Show that these are Hermitian.
(b) Show that $\vec{S}^2 = S_1^2 + S_2^2 + S_3^2 = 2\hbar^2 \cdot 1$.
(c) Find a matrix representation for $S_1$, $S_2$, and $S_3$. Show the matrices are Hermitian.

22. Using $\vec{S}$ in prob.21, find the answer to

$$(\vec{S} \times \vec{S})_3 = ?$$

Deduce a general statement of which the above is an example.

23. Still using the same $\vec{S}$ as in the last two problems. Solve the equation

$$S_3 \, \psi \;\; = \;\; S_3' \, \psi$$

where $\psi$ is a column matrix and $S_3' = -\hbar$, $0$, or $\hbar$. Use the matrix representation for $S_3$ you found in (c) of problem 3. There are three solutions for $\psi$ corresponding to the three values of $S_3'$. What are their physical interpretation?

**Other Problems**

24. The expected <u>classical</u> result for the single magnet Stern-Gerlach experiment was to find a single continuous line of atoms on the screen positioned beyond the magnet.
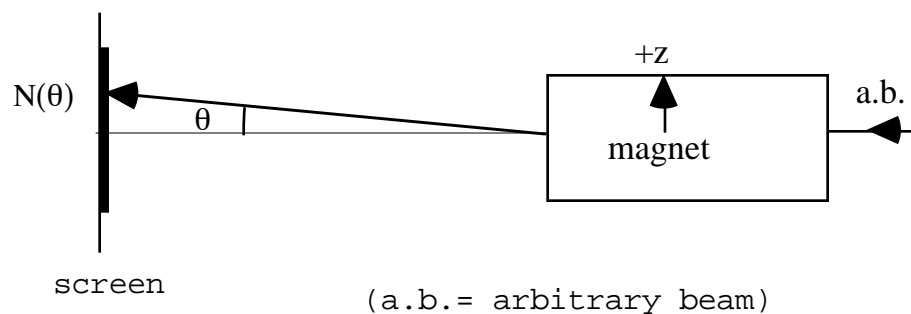


Figure 1

We now want to find an expression for the expected classical number of atoms, $N(\theta)$, as a function of deflection angle, $\theta$. Do this in two parts:

1.102

(a)   Assuming that the magnetic moments, $\vec{\mu}$, of the particles in the arbitrary beam are randomly oriented, and taking the +z axis along the magnet's magnetic field, show that the number of atoms with $\mu_z = \mu \cos\theta'$ relative to those with $\mu_z = \mu$ (i.e. at $\theta' = \frac{\pi}{2}$) is given by

$$\frac{N(\theta)}{N(\theta' = \frac{\pi}{2})} = |\sin\theta'|,$$

[Hint:  Consider thin strips of area dA and dA' in the figure shown and compare the relative number of atoms included:



z axis

dA'

$\theta'$ $\vec{\mu}'$

dA

$\vec{\mu}$

Equitorial plane
$(\ \theta' = \frac{\pi}{2}\ )$
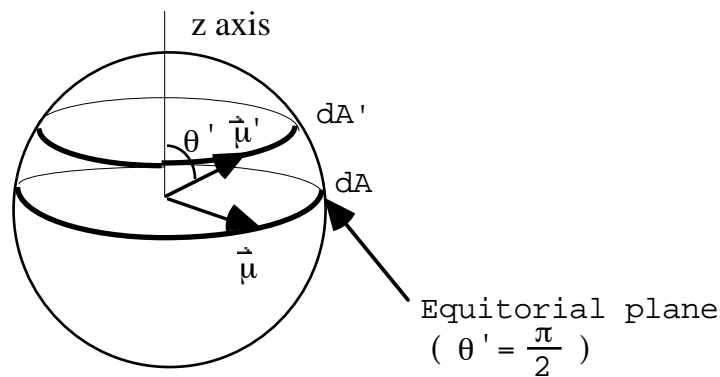
Figure 2

(b)   Now, establish a relationship between the angles $\theta$ and $\theta'$ in Figures 1 and 2.  Use your result in part (a) to show that

$$N(\theta) = N(\theta = 0)\left(1 - \frac{\theta^2}{|\mu\frac{\partial H}{\partial z}|^2 \frac{L^2}{m^2 v^4}}\right)^{1/2}$$

(The symbols L, m, v, $\frac{\partial H}{\partial z}$ are defined in the notes.)  This gives the number of atoms deflected at angle $\theta$ relative to the number with no deflection, $\theta = 0$.

25.   Finish the equality:

$$[\vec{\sigma} \cdot \vec{a}, \vec{\sigma} \cdot \vec{b}] = ?$$

[Remember $[A,B] \equiv AB - BA$. $\vec{a}$ and $\vec{b}$ are vectors whose components are numbers, whereas $\vec{\sigma} = \sigma_1\hat{i} + \sigma_2\hat{j} + \sigma_3\hat{k}$, where the σ's are operators.]

26. A beam of atoms with spin one ($S_3' = -\hbar,0,\hbar$) pass through the two Stern-Gerlach setups shown.  A polar angle θ (relative azimuthal angle φ = 0) gives the direction of the magnetic field in first magnet relative to the second magnet's magnetic field.
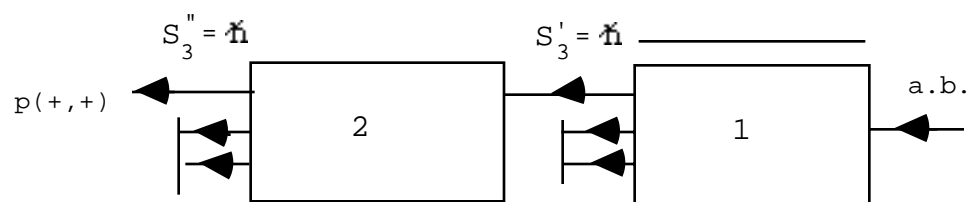


Figure 3

$S_3' = \hbar$ is selected from the first magnet and allowed to pass into the second magnet.

(a) Verify that the state

$$\frac{1}{\sqrt{2}}( |+> +i|0> )$$

is associated with the outcome $S_3' = +\hbar$ from the first magnet. [From prob.21, The components of $\vec{S}$ in this basis are

$$S_1 = i\hbar(|-0| - |0-|),$$
$$S_2 = i\hbar(|+-| - |-+|),$$
$$S_3 = i\hbar(|0+| - |+0|).]$$

(b) Using the result from (a), calculate the probability $p(+,+)$ associated with the selection $S_3'' = +\hbar$ along the second magnet's 3-axis.  [Remember the result of Problem #11 for the three-physical-outcome case:

$$e^{i\lambda A} = A^2 \cos\lambda + iA \sin\lambda + (1 - A^2).$$

What are A and $\lambda$ in this case?]

27.(a) Show that the relation,

$$< \bar{\sigma}_{3'} |\sigma_3 >^* = < \sigma_3|\bar{\sigma}_{3'} >,$$

implies that

$$p(\sigma_{3'},\sigma_3) = p(\sigma_3,\sigma_{3'}).$$

(b) Derive the normalization condition,

$$|\psi_+(+)|^2 + |\psi_+(-)|^2 = 1,$$

where $\psi_{\sigma_3'}(\sigma_3) = < \sigma_3|\bar{\sigma}_{3'} >$ (see eq$^{\underline{n}}$ (209) of Chapter 1) from the orthonormality condition,

$$< \bar{\mp}|\bar{\mp} > = 1.$$

(As usual, the bar denotes a rotated state. This is the spin one-half case.)

28. One representation of $S_{1,2,3}$ (different from that given above) in the spin-one (three-physical outcome case) is

$$S_3 = \hbar(|+| - |-|),$$

$$S_2 = \frac{i\hbar}{\sqrt{2}}(|+0| + |-0| - |0+| - |0-|),$$

$$S_1 = \frac{\hbar}{\sqrt{2}}(|0-| + |-0| - |0+| - |+0|).$$

(a) Find matrix representations of $S_{1,2,3}$.

(b) A beam of atoms with spin one ($S_3'=-\hbar,0,\hbar$) pass through two Stern-Gerlachs which have a relative polar angle $\theta$ between their z-axes.
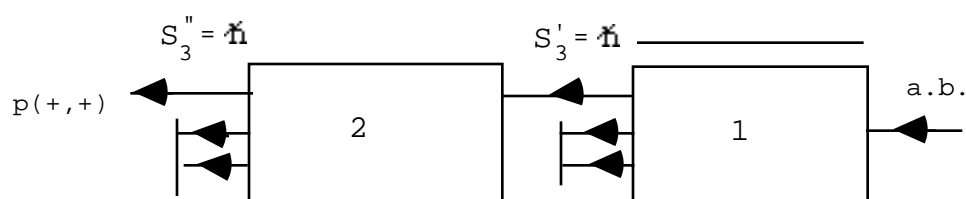


Given that this angle is <u>very small</u> ($\theta \ll 1$) and that the probability $p(+,+)$ is given by

$$p(+,+) = |< +|e^{i\theta(S_2/\hbar)}|+ >|^2,$$

show that
$$p(+,+) \approx 1 - \theta^2/4.$$

[Hint: Expand the exponential and use the matrix representation for $S_2$ you found in (a).]

(c) Verify that the state

$$\frac{1}{\sqrt{2}}( |+ > +i|0 > )$$

is associated with the outcome $S_3' = +\hbar$ from the first magnet. [The components of $\vec{S}$ in this basis are

$$S_1 = i\hbar(|-0| - |0-|),$$

$$S_2 = i\hbar(|+-| - |-+|),$$

1.106

$$S_3 = i\hbar(|0+| - |+0|).]$$

(d) Using the result from (c), calculate the exact
probability p(+,+) associated with the selection $S_3'' = +\hbar$
along the second magnet's 3-axis.  [Remember the result of
for the three-physical-outcome case:

$$e^{i\lambda A} = A^2 \cos\lambda + iA \sin\lambda + (1 - A^2).$$

What are A and $\lambda$ in this case?]

**CHAPTER 2:    Free Particles in One Dimension**


We started the last chapter with experimental
indications of a breakdown in classical mechanics.  We start
this chapter by citing two famous experiments that helped to
begin to construct a new picture.

It had been known since the work of H. Hertz in 1887
that electromagnetic waves incident on a metal surface can
eject electrons from that material.  This was in the context
of an experiment where he noticed a spark could jump a gap
between two metallic electrodes more easily when the gap was
illuminated.  In particular, it was established before
Einstein's explanation of the effect in 1905 that shining
light on metallic surfaces leads to ejected electrons.
Einstein's formulas were not verified until 1916 by R.A.
Millikan.  Einstein received the Nobel Prize in 1921 for this
work.

Consider the following experimental arrangement for
measuring this photoelectric effect:
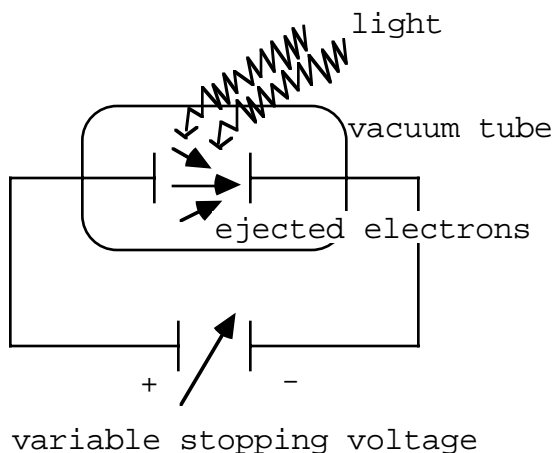
We can shine light of varying intensities and frequencies on the metal surface indicated; a stream of electrons will then be emitted, some of which will strike the other metal plate inside the vacuum tube, thus constituting a current flow. The classical picture of the behavior of the electrons under such circumstances did not explain the known facts. In particular, it was known that there existed a threshold frequency, dependent on the type of metal being studied, below which no photoelectrons were emitted. Once this threshold frequency has been exceeded, one may then adjust the magnitude of the stopping voltage to cancel the photocurrent. (Let us call this voltage $V_{max}$.) In his explanation of the photoelectric effect, Einstein assumed that the incoming beam of light could be viewed as a stream of particles traveling at the speed of light, each of which carried an energy

$$E = h\nu \qquad\qquad\qquad (1)$$

where "h" is Planck's constant. Planck had originally postulated a relation like (1) above to hold for the atoms in the wall of a hot furnace, the so-called black body problem.[*] Einstein then supposed that, based on this particle picture of light, the maximum energy of an electron ejected from the metal's surface could be written as

---

[*] We will deal with black body radiation again when we come to discuss Bose-Einstein statistics in Chapter 9. There we will see that the black body radiation law is simply a consequence of the particle nature of light, independent of assumptions about the atoms in the wall of the furnace.

$$T_{max} = h\nu - W \qquad (2)$$

Eq$^{\underline{n}}$ (2) is based on energy conservation.  The picture Einstein was led to was one in which the energy transferred to an electron by the photon, $h\nu$, is used to overcome the minimum energy necessary to remove an electron from that metal, W, which is called the work function.
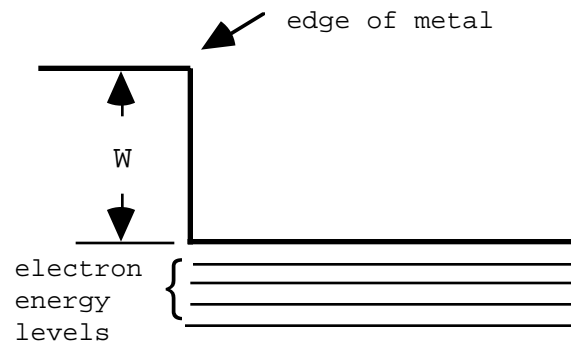


The work function, W, is not accounted for in Einstein's theory, but is assumed to be a constant characteristic of a particular metal.  A clear implication of (2) is that there exists a frequency $\nu_0$, given by

$$\nu_0 = \frac{W}{h} , \qquad (3)$$

below which we would expect to see no ejected electrons since the energy available to any single electron will not be large enough to overcome that material's work function.  Now we know that $V_{max}$, the maximum voltage that allows a photocurrent to flow, is related to $T_{max}$ by

$$eV_{max} = T_{max}. \qquad (4)$$

where e is the magnitude of the charge on the electron. Plugging (4) into (2) then gives

$$V_{max} = \frac{h}{e}\nu - \frac{W}{e} \; . \hspace{3cm} (5)$$

Therefore, another implication of Einstein's theory is that if we plot $V_{max}$ as a function of photon frequency, we should see something like:
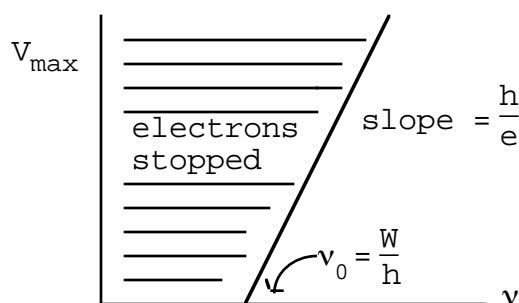


The first implication above explains the threshold frequency behavior observed in metals.  The second implication involving $V_{max}$ was the part that was verified later by Millikan.  Einstein's theory also agreed nicely with another experimental observation; namely, that the photocurrent was directly proportional to the light intensity. This is because light intensity is proportional to the number of light quanta, as is the number of electrons emitted from the surface .

Another famous experiment which has added to our understanding of a new mechanics was done by Arthur Compton in 1923, and has since become known as the Compton effect. Other investigators had actually performed versions of this experiment before Compton, but had not reached his conclusions.  A schematic representation of his setup is shown below.
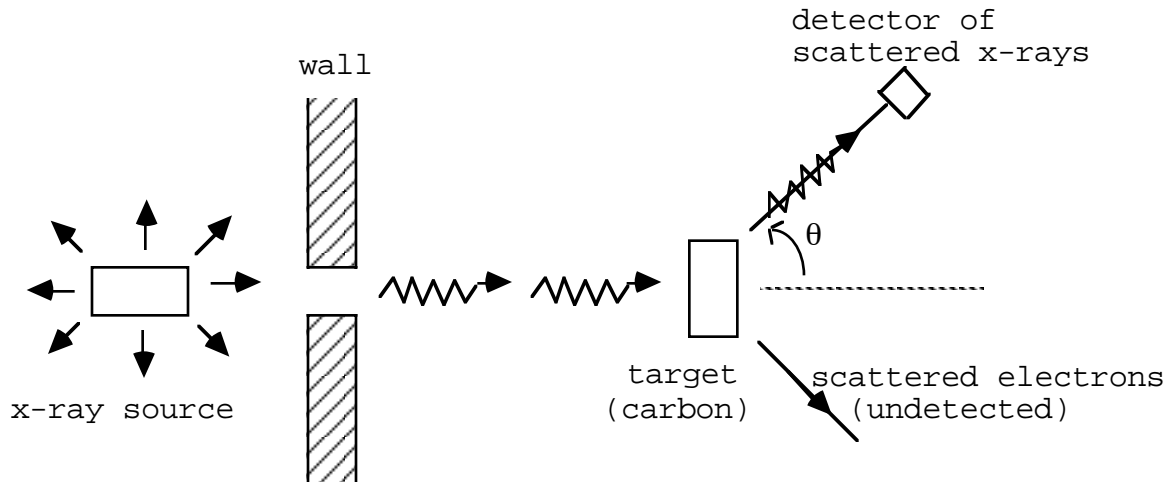
One can measure the <u>intensity</u> of the scattered x-rays as a function of $\theta$. In addition, we can imagine fixing the scattering angle $\theta$ and <u>tuning</u> the detector to measure x-rays of varying wavelengths.  When Compton did this, he found a result that looked like the following:
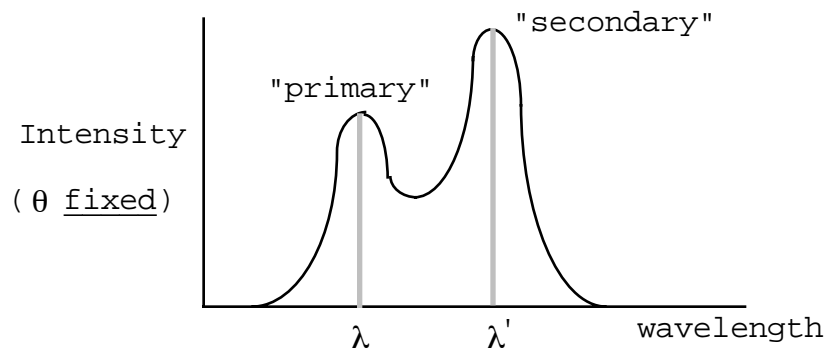


That is, Compton saw two peaks in the intensity spectrum as a function of wavelength.  The peak at $\lambda$, labeled "primary", occurred at the same wavelength as the approximately monochromatic source; the additional peak, labeled "secondary", occurred with $\lambda' > \lambda$.  Compton explained his results by assuming a particle picture for the x-ray beam, as

had Einstein, and by assuming energy and momentum conservation during the collision of the "photon", the particle of light, and the electron.

The kinematics of this collision can be pictured as in the following:
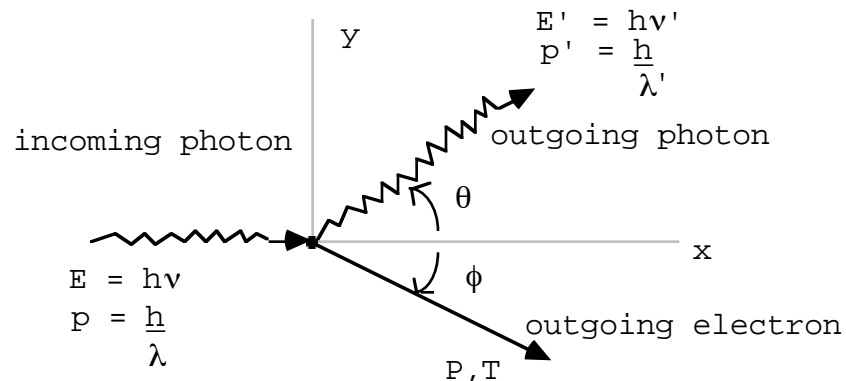


One is assuming that in this frame of reference, the electron is initially at rest.  It then acquires a momentum, P, and kinetic energy, T, from the photon.  We know from Einstein's special theory of relativity that the energy of a particle of mass m and momentum p is

$$E = \sqrt{m^2c^4 + p^2c^2} \qquad (6)$$

We also know from this theory that the speed of light, c, is unattainable for any _material_ particle; however, for a particle with zero mass, the speed of light is the _required_ velocity.  Under these circumstances, which apply to the photon, the relationship between E and p from (6) is

$$E = pc . \qquad (7)$$

Then using (1) we may write

$$p = \frac{h}{\lambda} , \qquad (8)$$

for the photon's momentum.  ($p = |\vec{p}|$ here)

Writing momentum conservation in the x and y directions for the above diagram tells us

$$x: \quad p = p' \cos \theta + P \cos \phi, \qquad (9)$$

$$y: \quad 0 = p' \sin \theta - P \sin \phi. \qquad (10)$$

Therefore, we have

$$P^2 \cos^2 \phi = (p - p' \cos \theta)^2, \qquad (11)$$

$$P^2 \sin^2 \phi = (p' \sin \theta)^2, \qquad (12)$$

Adding (11) and (12) gives

$$P^2 = p^2 + p'^2 - 2pp' \cos \theta. \qquad (13)$$

We now write conservation of energy

$$E - E' = T, \qquad (14)$$

or

$$p - p' = \frac{T}{c} . \qquad (15)$$

Squaring this gives

$$\left( \frac{T}{c} \right)^2 = p'^2 + p^2 - 2pp'. \qquad (16)$$

Let's now subtract (16) from (13):

$$P^2 - \left( \frac{T}{c} \right)^2 = 2pp'(1 - \cos \theta) . \qquad (17)$$

2.8

The <u>relativistic</u> connection between P and T is <u>not</u>

$T = \dfrac{P^2}{2m}$. We write the relation between the electron's energy

and momentum Eq$^n$ (6) above, as

$$E_e^2 = P^2 c^2 + m^2 c^4. \tag{18}$$

The definition of kinetic energy T is

$$E_e = T + mc^2. \tag{19}$$

Substituting (19) in (18) and solving for $P^2$ then gives

$$P^2 = \left(\frac{T}{c}\right)^2 + 2Tm, \tag{20}$$

or

$$P^2 - \left(\frac{T}{c}\right)^2 = 2Tm. \tag{21}$$

Comparing (21) and (17) we now have

$$2Tm = 2pp'(1 - \cos\theta). \tag{22}$$

However, from (15), this now becomes

$$2mc(p - p') = 2pp'(1 - \cos\theta), \tag{23}$$

or

$$\left(\frac{1}{p'} - \frac{1}{p}\right) = \frac{1}{mc}(1 - \cos\theta). \tag{24}$$

Now using (8) we find that

$$\lambda' - \lambda = \frac{h}{mc}(1 - \cos\theta). \tag{25}$$

Numerically,

$$\frac{h}{mc} = .0243\text{Å} = 2.43 \times 10^{-10} \text{ cm}, \qquad (26)$$

a quantity which is called the electron <u>Compton</u> <u>wavelength</u>. The shifted wavelength, $\lambda'$, corresponds to the secondary intensity peak seen by Compton.  Compton explained the primary peak as elastic scattering from the atom as a whole, leading to only a very tiny shift in wavelength due to the Carbon atom's large mass relative to the electron's. Although a more detailed theory would be needed to explain the complete intensity spectrum shown above, the above experiment supplied rather convincing evidence of the particle nature of light.[*]

   We now know that light particles are described by $p = \frac{h}{\lambda}$.  Let's examine the consequence of this statement in view of the fact that the wave nature of light results in <u>diffraction</u>.

   Consider the Fraunhoffer single-slit device below.



_____
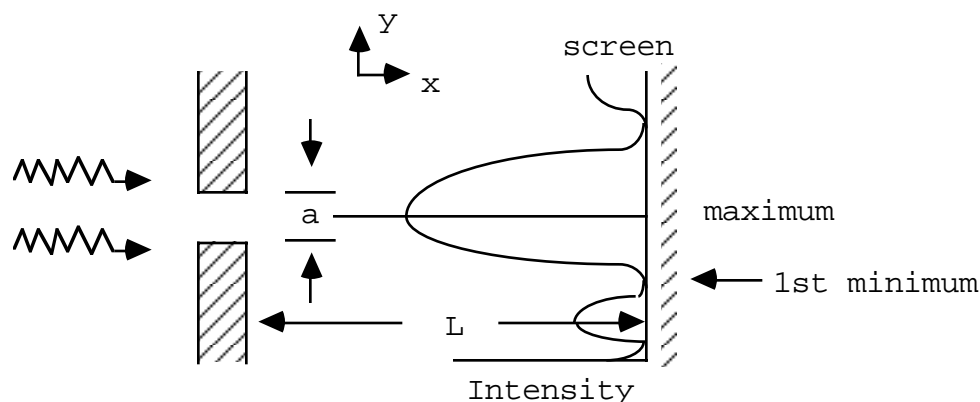
[*] The wavelength shift calculated holds for the scattering of *free* electrons, but of course the electrons in an atom are bound to a nucleus. However, because the X-ray photon energies are so much larger than the electronic binding energy, the above treatment is still very accurate.

Viewing the incoming light rays as a stream of particles, we see that there is an uncertainty in the y-position of an individual photon in passing through the diffraction slit.  We say that

$$\Delta y = a, \tag{27}$$

where "$\Delta$" means "uncertainty in."  Now, we expect to see the first interference minimum, indicated above, when the following conditions obtain:
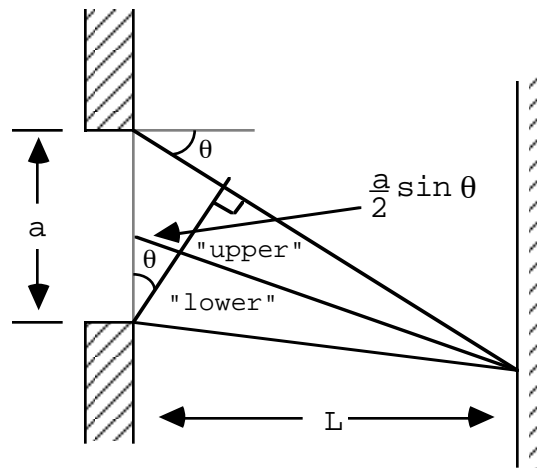


From the above, we see that if L >> a we will have an interference minimum between rays in the "upper" half of the triangle and the "lower" half when

$$\frac{a}{2} \sin \theta = \frac{\lambda}{2} , \tag{28}$$

for then they will be completely out of phase.  Now we know that most of the photons fall within the central maximum of the pattern.  This means, in particle terms, there is also an uncertainty in y-momentum of individual photons passing

through the slit.  If we define this uncertainty so as to roughly correspond to the diffraction minimum, we have

$$\frac{\Delta p_y}{p} \approx \frac{L \sin \theta}{L} = \sin \theta. \qquad (29)$$

Therefore

$$\Delta y \; \Delta p_y \approx a(p \sin \theta) = a(p \; \frac{\lambda}{a}) = p\lambda. \qquad (30)$$

Now using $p = \frac{h}{\lambda}$, this tells us

$$\Delta y \; \Delta p_y \approx h . \qquad (31)$$

The product of the uncertainties in position and momentum of an individual photon is on the order of Planck's constant.  Eq$^{n}$ (31) is an example of the <u>Heisenberg uncertainty principle</u> as applied to photons.  However, relations such as (31) also hold for material particles that can be brought to rest in a laboratory, such as protons and electrons.  Assuming energy and momentum conservation, the Compton effect discussed above shows this to be true, for if we could measure the initial and final position and momentum of the electron, we could determine the position and momentum of the photon to an arbitrary accuracy, contradicting (31).

A more general and rigorous statement of the uncertainty relation for material particles like an electron, involving x and $p_x$, will be seen (in Ch. 4) to be

$$\Delta x \; \Delta p_x \geq \frac{\hbar}{2} . \qquad (32)$$

An illustration of (32) will be given later in the present
Chapter.

If uncertainty relations apply to electrons, then we can
immediately understand the reason for the stability of the
ground state (that is, lowest energy state) of the hydrogen
atom, which, according to classical ideas, should quickly
decay.   The electron's energy is

$$E = \frac{\vec{p}^2}{2m} - \frac{e^2}{r} \, , \tag{33}$$

where $\vec{p}$ and m refer to the electron's momentum and mass.   Now
the hydrogen atom's ground state has zero angular momentum,
so the $\vec{p}$ in (33) can be replaced by $p_r$, and we have
essentially a one-dimensional problem in r, the radial
coordinate.   We hypothesize that

$$rp_r \sim \hbar, \tag{34}$$

for the hydrogen atom ground state; i.e. we suppose the
ground state is also close to being a minimum uncertainty
state.   Then using (34) in (33) gives

$$E = \frac{1}{2m} \left( \frac{\hbar}{r} \right)^2 - \frac{e^2}{r} \, . \tag{35}$$

A plot of (34) looks like the following:
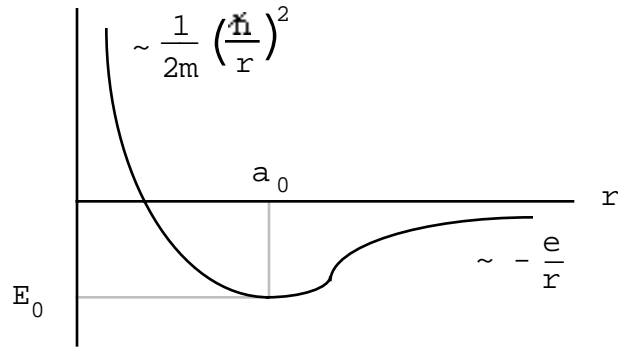
We find $E_0$ by setting

$$\frac{\partial E}{\partial r} = 0,$$

(36)

which gives

$$a_0 = \frac{\hbar^2}{me^2},$$

(37)

and

$$E_0 = -\frac{me^4}{2\hbar^2} \quad (= -13.6 \text{ eV}).$$

(38)

The negative sign above means the same thing as for a classical mechanics system - that the system is _bound_. The value for $E_0$ above turns out to be very close to the actual ground state energy of the real hydrogen atom. Our calculation above is actually a bit of a swindle because Eq[n] (34) is only a rough guess. The quantity $a_0$ in (37) is called the "Bohr radius" and is numerically equal to .53Å ($1$Å $= 10^{-8}$ cm). Thus the uncertainty principle implies the existence of a ground state and gives a rough value of the associated binding energy.

The famous Bohr model from which the above name derives was a significant step forward in our understanding of atomic systems and yields an important insight into the nature of such systems. One way to recover the content of this model is to assume that the relation (8), written above for the photon, holds also for electrons and other material particles. That is, we are assuming a wave-like nature for entities we usually think of as particles, just as we previously were led to assume both wave and particle characteristics for light. If indeed objects such as electrons have wave characteristics, then they should exhibit constructive and destructive interference under appropriate circumstances. Consider therefore a simplified model of a wave-like electron trying to fit in the circular orbit shown.
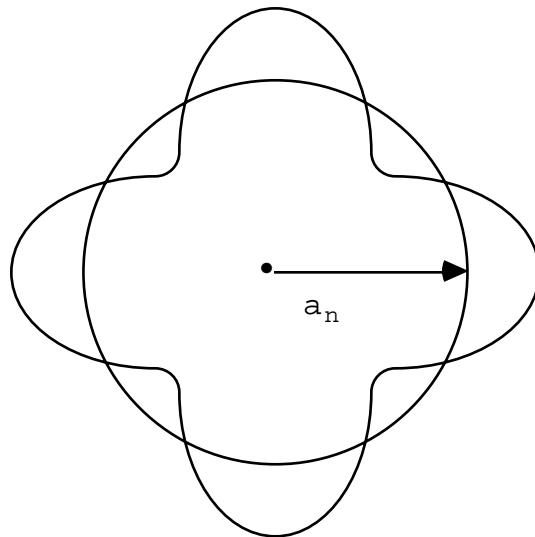


In order for the electron wave to avoid a situation where destructive interference would not permit it to persist in a stable configuration, we should have

$$2\pi a_n = n\lambda, \tag{39}$$

where $n = 1,2,3,...$ We then have

$$a_n = n\,\frac{\lambda}{2\pi} = n\,\frac{\hbar}{p}\,. \tag{40}$$

Multiplying both sides of this by p, we get another interpretation of (39) from

$$a_n p = n\hbar. \tag{41}$$

The left hand side of (39) represents <u>orbital</u> angular momentum, which the right hand side now says is quantized in units of $\hbar$.  Minimizing the energy

$$E = \frac{\vec{p}^2}{2m} - \frac{e^2}{a_n}\,,$$

$$= \frac{1}{2m}\left(\frac{n\hbar}{a_n}\right)^2 - \frac{e^2}{a_n}\,, \tag{42}$$

with respect to $a_n$, we find that

$$\frac{\partial E}{\partial a_n} = 0 = -\frac{n^2\hbar^2}{ma_n^3} + \frac{e^2}{a_n^2}\,, \tag{43}$$

so that

$$a_n = \frac{n^2\hbar^2}{me^2}\,. \tag{44}$$

Plugging this back into (40) now gives

$$E_n = -\frac{e^2}{2a_n} = -\frac{me^4}{2n^2\hbar^2}\,. \tag{45}$$

Setting n = 1 in (45) just gives (38) above.  The energy
levels in (45) agree closely with what is seen
experimentally.  The Bohr model leads us to believe that
material particles can also have wave-like characteristics.
Thus, the stability of the system is guaranteed by the
uncertainty priciple and the disceteness of the energy levels
is due to the assumed wave nature of particles.  The ultimate
verification of such a hypothesis would be to actually
observe the effects of diffraction on a beam of supposedly
particle-like electrons.  That such a phenomenon can exist
was hypothesized by Louis DeBroglie in his doctoral
dissertation (1925), and this was experimentally verified by
Davisson and Germer (1927).  The wavelength associated with
electrons (called their deBroglie wavelength) according to
(8), which we originally applied only to photons, would in
most instances be very tiny.  For this reason Davisson and
Germer needed the close spacing of the atomic planes of a
crystal to see electron diffraction take place.

The model Bohr presented of the behavior of the electron
in a hydrogen atom is correct in its perception of the wave
nature of particles, but is wrong in the assignment of the
planetary-like orbits to electrons.  In fact, as pointed out
earlier, the hydrogen atom ground state actually has zero
angular momentum, as opposed to the single unit of $\hbar$ assigned
to it by (39) above.

Our minds may rebel at the thought of something that
shares both wave and particle characteristics.  This seems

like a paradox.  It remains an experimental fact, however,
that electrons and photons show either aspect under
appropriately designed circumstances.  We must, therefore,
think of particle and wave characteristics as being not
paradoxical, but <u>complementary</u>.  This is the essence of what
is called <u>Bohr's Principle of Complementary</u>.

   We will have to build this dualism into the structure of
the theory we hope to construct.  The next step we take in
this direction is the realization that the appropriate
mathematical tool to use in order to do this is the Fourier
transform.

   Leaving physics behind a minute and specializing to a
single dimension, let us inquire into the Fourier transform
of a function that looks like the following:
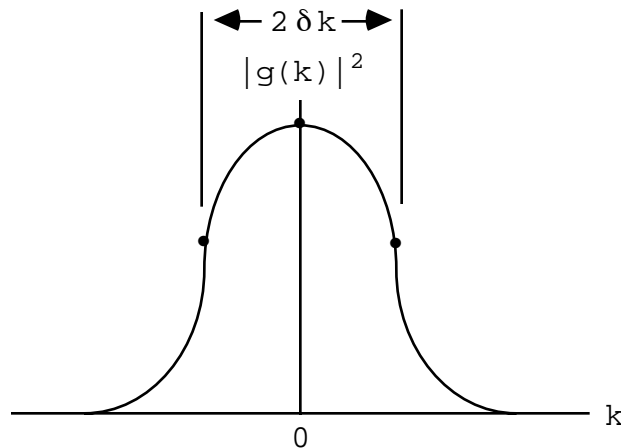


That is, we are Fourier transforming a function g(k) that is
nonzero only when

$$-\delta k \lesssim k \lesssim \delta k. \tag{46}$$

We define the Fourier transform of g(k) to be

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k)e^{ikx}\, dk. \tag{47}$$

We ask:  what will be the magnitude of the resulting function, $|f(x)|^2$, look like?  We answer this question as follows:

| Choose x such that: | $e^{ikx}$ is: | $|f(x)|^2$ value is: |
|---|---|---|
| $x\delta k \ll 1$ | ~1 | <u>maximum</u> |
| $x\delta k \sim 1$ | $e^{-i}$ to $e^{i}$ <br> $(1/\pi$ of a phase rot.$)$ | <u>lowered</u> because of partial destructive interference |
| $x\delta k \gg 1$ | rapidly oscillating phase | <u>small</u> because of strong destructive interference |

Qualitatively then, we would expect $|f(x)|^2$ to look like:
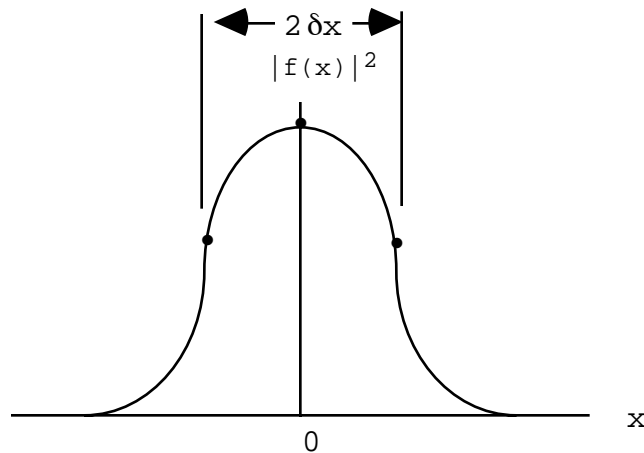


where

$$\delta x \delta k \approx 1. \tag{48}$$

Of course, the detailed appearance of $|f(x)|^2$ depends on the exact mathematical form we assume for $g(k)$. To get some experience, let us do a specific example. Let

$$g(k) = e^{-\alpha k^2} ,\qquad (49)$$

which is a Gaussian centered around $k = 0$. Then $f(x)$ is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\alpha k^2} e^{ikx} dk. \qquad (50)$$

We can write

$$-\alpha k^2 + ikx = -\alpha \left(k - \frac{ix}{2\alpha}\right)^2 - \frac{x^2}{4\alpha} . \qquad (51)$$

This is called "completing the square." Then (50) reads

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/4\alpha} \int_{-\infty}^{\infty} e^{-\alpha(k - ix/2\alpha)^2} dk. \qquad (52)$$

Now it is justified to let $k' = k - \dfrac{ix}{2\alpha}$ and still keep the integral along the real axis. The remaining integral we must do is

$$\int_{-\infty}^{\infty} e^{-\alpha k'^2} dk' = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{\infty} e^{-x^2} dx \equiv \frac{I}{\sqrt{\alpha}} . \qquad (53)$$

$I^2$ can be evaluated as follows:

$$I^2 = \int_{-\infty}^{\infty} dx e^{-x^2} \int_{-\infty}^{\infty} dy e^{-y^2} = \int_{-\infty}^{\infty} dx dy \, e^{-(x^2 + y^2)}$$

$$= \int_{0}^{\infty} r dr \int_{0}^{2\pi} d\theta e^{-r^2} = \pi \int_{0}^{\infty} dr^2 e^{-r^2} = \pi. \qquad (54)$$

Therefore $I = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$ , and we have

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/4\alpha} \sqrt{\frac{\pi}{\alpha}} = \frac{1}{\sqrt{2\alpha}} e^{-x^2/4\alpha} . \qquad (55)$$

(In this special example f(x) turns out to be real, but in general it is complex.)  Eq$^{\underline{n}}$ (55) has the advertised properties.  That is, the width of the Gaussian in k-space is

$\delta k \approx \frac{1}{2\sqrt{\alpha}}$ (this makes $\left| \frac{g(k = \frac{1}{2\sqrt{\alpha}})}{g(k = 0)} \right|^2 = e^{-1/2}$) while the width

of the x-space Gaussian, in the same sense, is $\delta x \approx \sqrt{\alpha}$ .
Therefore, the product of these two widths is

$$\delta x \delta k \approx \sqrt{\alpha} \cdot \frac{1}{2\sqrt{\alpha}} = \frac{1}{2} . \qquad (56)$$

Thus, we can make $\delta x$ or $\delta k$ separately as small as we wish, but then the other distribution will be spread out so as to satisfy (56).

The argument leading to (56) was mathematical.  The argument leading to (31) was a physical one.  We now realize, however, that the uncertainty product in (31) will be guaranteed to hold if we were to identify

$$k = \frac{P_x}{\hbar} , \qquad (57)$$

because then (56) would read (compare with (32) above)

$$\Delta x \Delta p_x \approx \frac{\hbar}{2} , \qquad (58)$$

where we are now interpreting the x and $p_x$ distribution

widths as uncertainties, for which we use the "$\Delta$" symbol.

Thus, the Fourier transform plus the identification (57)

accomplishes the goal of insuring that uncertainty relations

involving position and momenta are built into the theory.

Eq$^n$ (47) says that the f(x) distribution function or

"wave packet" is actually a superposition of functions given

by $e^{ikx}$ with continuous k-values.  Notice that (k can be a

positive or negative quantity)

$$e^{ik(x+2\pi/|k|)} = e^{ikx},\tag{59}$$

which says that $e^{ikx}$ is a periodic function with a wavelength

$\frac{2\pi}{|k|}$ .  The following graphs display the real and imaginary

parts of $e^{ikx}$ as a function of x.
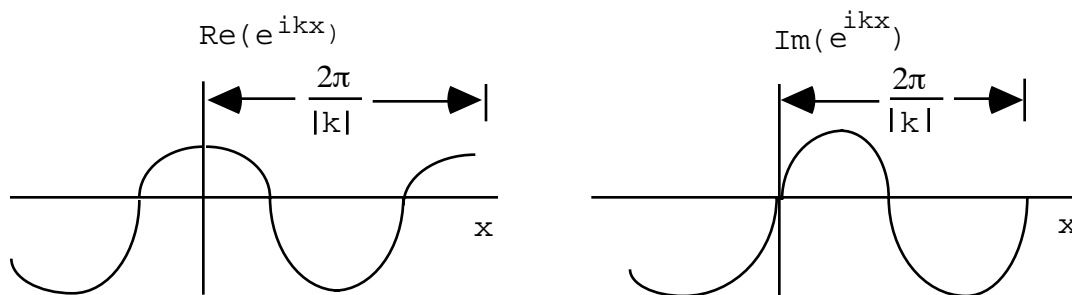


The relation $k = \frac{p_x}{\hbar}$ is consistent with the statement

that $\frac{2\pi}{|k|}$ repesents a particle wavelength since

$$\lambda = \frac{2\pi}{|k|} = \frac{2\pi\hbar}{|p_x|} = \frac{h}{|p_x|},\tag{60}$$

which gives back the previous form of the DeBroglie wavelength. The function $e^{ikx}$ (or $e^{i\vec{k}\cdot\vec{x}}$ in 3 dimensions) is usually called a "plane wave" because points of constant phase of this function form a plane in 3 dimensions.

   Still dealing with only a single spatial dimension, let us now notice that the function (t = time)

$$e^{i(kx - \omega t)} ,$$

represents a <u>traveling</u> plane wave.  We can understand the velocity of the motion by following a point of constant phase in the wave, for which

$$e^{i(kx - \omega t)} = \text{constant}. \tag{61}$$

As a function of time, these positions of constant phase then must satisfy (if x=0 at t=0)

$$kx = \omega t \Rightarrow x = \frac{\omega}{k}\, t. \tag{62}$$

Therefore, the <u>phase velocity</u> of the traveling plane wave is just $\omega/k$.  We then see that these phases move in the +x direction if k > 0 and in the -x direction when k < 0.  The quantity $\omega$ represents the <u>angular frequency</u> of the moving wave.  This is easy to see by fixing the position, x.  Then $e^{-i\omega t}$ tells us how many waves pass our observation point per unit time.  We have

$$e^{-i\omega(t + 2\pi/\omega)} = e^{-i\omega t} , \tag{63}$$

so the period is given by the usual formula

$$T = \frac{2\pi}{\omega} \tag{64}$$

Free particles always have a positive definite angular frequency, $\omega$.

For photons, the deBroglie wavelength is the same as the usual wavelength concept, and the relationship between $\omega$ and $|k|$ is given simply by

$$\omega = \frac{E}{\hbar} = \frac{|p_x|c}{\hbar} = |k|c, \tag{65}$$

using (1),(7), and (57). The above relationship becomes $\omega = |\vec{k}|c$ in three dimensions. (This simple linear relationship is strictly true only for light rays traveling in free space.) The phase speed is then

$$\frac{\omega}{|k|} = c, \tag{66}$$

as it should be. Since $\frac{\omega}{|k|}$ is a constant independent of $|k|$, the phases of all wavelengths travel at the same velocity.

The equation that describes the time propagation of free photons is now easy to find. Based on the above observations, we generalize (47) to account for time dependence.

$$f(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k) e^{i(kx - \omega t)} dk, \tag{67}$$

The differential equation that (67) obeys is easy to construct. Observe that

$$\frac{1}{c^2} \frac{\partial^2 f(x,t)}{\partial t^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k) e^{ik(x - ct)} (-k^2) dk, \qquad (68)$$

and that

$$\frac{\partial^2 f(x,t)}{\partial x^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k) e^{ik(x - ct)} (-k^2) dk, \qquad (69)$$

so that

$$\frac{1}{c^2} \frac{\partial^2 f(x,t)}{\partial t^2} = \frac{\partial^2 f(x,t)}{\partial x^2} \ . \qquad (70)$$

Eq$^{\underline{n}}$(70) is called the "wave equation." In three

dimensions the $\frac{\partial^2}{\partial x^2}$ operator would be changed to $\vec{\nabla}^2$. (We are

ignoring the phenomenon of light polarization in this

discussion.)

Let's now go through a similar discussion for a

nonrelativistic particle in order to get the analog of

Eq$^{\underline{n}}$(70). The crucial step here is the assumption that the

relation $\omega = \frac{E}{\hbar}$, found to hold for photons, also holds for

material particles, where $\omega$ retains its meaning as the

deBroglie angular frequency. This is a very reasonable

supposition since we know that for any wave motion, frequency

is a conserved quantity in time. (Think about Snell's law,

for example). Since energy is also conserved in the motion

of a free particle, it is natural to assume that these

quantities are proportional.  With this hypothesis, we find
for a free nonrelativistic particle

$$\omega = \frac{E}{\hbar} = \frac{p_x^2}{2m\hbar} = \frac{\hbar k^2}{2m}, \tag{71}$$

which would read $\omega = \dfrac{\hbar \vec{k}^2}{2m}$ in three dimensions.

The equation analogous to (67) now becomes (one
dimension again)

$$f(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k)e^{i(kx - \omega(|k|)t)} dk, \tag{72}$$

where the relation between $\omega$ and $|k|$ is given in (71).  The
phase speed of these particles is

$$\frac{\omega}{|k|} = \frac{\hbar |k|}{2m} = \frac{|p_x|}{2m} . \tag{73}$$

This is just half of the value of the mechanical speed of
propagation, $\dfrac{|p_x|}{m}$ .  Therefore, we conclude that the phase
speed of deBroglie waves is not the same thing as the actual
propagation speed of the particle.  Eq$\underline{^n}$(72) also differs from
(66) in that the phases of different deBroglie wavelengths
travel at different velocities. This is called <u>dispersion</u>,
and it's effects in an example will be studied below.  For a
slowly varying function g(k) in (72), (corresponding, say, to
a sufficiently peaked function in position space) most of the
contribution to the rapidly varying exponential integral will
come from the integration domain

$$\frac{\partial(kx - \omega(|k|)t)}{\partial k} \approx 0,$$

(this is called the stationary phase approximation) which identifies the <u>average</u> particle propagation velocity as

$$\frac{\bar{p}_x}{m} = \frac{\partial\omega}{\partial k} \quad . \tag{74}$$

However, remember that a given wave packet contains a continuous range of $\vec{k}$ or $\vec{p}$ values (in three dimensions), so that a particle's velocity can only be defined in an average sense. The vector quantity $\dfrac{\partial\omega}{\partial\vec{k}}$ is called <u>group</u> <u>velocity</u>. The group speed of light in free space is

$$\left| \frac{\partial\omega}{\partial k} \right| = c, \tag{75}$$

the same as its phase speed.

Using (72) above, we find that

$$\frac{\partial f(x,t)}{\partial t} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k)e^{i(kx - \omega(|k|)t)} \left(\frac{-i\hbar k^2}{2m}\right) dk, \tag{76}$$

and

$$\frac{\partial^2 f(x,t)}{\partial x^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k)e^{i(kx - \omega(|k|)t)}(-k^2)dk, \tag{77}$$

so that

$$i\hbar \frac{\partial f(x,t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 f(x,t)}{\partial x^2} \quad . \tag{78}$$

Eq$^{\underline{n}}$ (78) is a special case of the celebrated <u>Schrödinger equation</u> written in one spatial dimension for a free particle.

We will now show that

$$\psi_g(x,t) = $$

$$\frac{1}{(2\pi)^{1/4}\sqrt{\delta x + \frac{i\hbar t}{2m\delta x}}} \cdot \exp\left( \frac{i}{\hbar}\left(\bar{p}x - \frac{\bar{p}^2}{2m}t\right) - \frac{1}{4\delta x}\cdot\frac{\left(x - \frac{\bar{p}}{m}t\right)^2}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)} \right) ,$$

(79)

is a solution to (78).  The left hand side of (78) involves

$$\frac{\partial\psi_g(x,t)}{\partial t}$$ , which can be written as

$$\frac{\partial\psi_g(x,t)}{\partial t} =$$

$$\psi_g(x,t)\left\{ -\frac{1}{2}\frac{\left(\frac{i\hbar}{2m\delta x}\right)}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)} - \frac{i}{\hbar}\frac{\bar{p}^2}{2m} - \frac{1}{4\delta x}\left[ \frac{-2\left(x - \frac{\bar{p}}{m}t\right)\frac{\bar{p}}{m}}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)} - \frac{\frac{i\hbar}{2m\delta x}\left(x - \frac{\bar{p}}{m}t\right)^2}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)^2} \right] \right\}.$$

(80)

Some necessary algebra is

$$-\frac{i}{\hbar}\frac{\bar{p}^2}{2m} + \frac{\bar{p}}{2m\delta x}\frac{\left(x - \frac{\bar{p}}{m}t\right)}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)} + \frac{i\hbar}{8m\delta x^2}\frac{\left(x - \frac{\bar{p}}{m}t\right)^2}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)^2} = -\frac{i}{2m\hbar}\frac{\left(\bar{p}\delta x + \frac{i\hbar}{2\delta x}x\right)^2}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)^2} ,$$

(81)

so we can write

$$i\hbar\frac{\partial\psi_g(x,t)}{\partial t} = \frac{\hbar^2}{2m}\psi_g(x,t)\left\{ -\frac{1}{2\left(\delta x^2 + \frac{i\hbar t}{2m}\right)} - \frac{1}{\hbar^2}\frac{\left(\bar{p}\delta x + \frac{i\hbar}{2\delta x}x\right)^2}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)^2} \right\}.$$

(82)

For the other side of the equation, we find

$$\frac{\partial \psi_g(x,t)}{\partial x} = \psi_g(x,t) \left\{ -\frac{1}{2\delta x} \frac{\left(x - \frac{\bar{p}}{m}t\right)}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)} + \frac{i}{\hbar}\bar{p} \right\} . \qquad (83)$$

and so

$$\frac{\partial^2 \psi_g(x,t)}{\partial x^2} = \psi_g(x,t) \cdot$$

$$\left\{ -\frac{1}{2\delta t} \frac{1}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)} + \frac{1}{4\delta x^2} \frac{\left(x - \frac{\bar{p}}{m}t\right)^2}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)^2} - \frac{\bar{p}^2}{\hbar^2} - \frac{i\bar{p}}{2\hbar\delta x} \frac{\left(x - \frac{\bar{p}}{m}t\right)}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)} \right\} .$$

$$(84)$$

We recognize the last three terms in (84) as the left hand

side of (81), apart from an overall factor. This gives us

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi_g(x,t)}{\partial x^2} = -\frac{\hbar^2}{2m} \psi_g(x,y) \left\{ -\frac{1}{2\left(\delta x^2 + \frac{i\hbar t}{2m}\right)} - \frac{1}{\hbar^2} \frac{\left(\bar{p}\delta x + \frac{i\hbar}{2\delta x}x\right)^2}{\left(\delta x + \frac{i\hbar t}{2m\delta x}\right)^2} \right\} .$$

$$(85)$$

The right hand sides of (82) and (85) are now seen to be the

same, which proves that $\psi_g(x,t)$ is a solution of (78).

Let's now try to interpret this solution. At t = 0 we

have

$$\psi_g(x,0) = \frac{1}{\sqrt{\sqrt{2\pi}\,\delta x}} \exp\left(\frac{i}{\hbar}\bar{p}x - \frac{x^2}{4\delta x^2}\right) , \qquad (86)$$

so

$$|\psi_g(x,0)|^2 = \frac{1}{\sqrt{2\pi}\,|\delta x|} \exp\left(-\frac{x^2}{2\delta x^2}\right) . \qquad (87)$$

Eq$^{\underline{n}}$(87) is a Gaussian in x, as we just finished studying.  At
a time t > 0 from (79) we have

$$|\psi_g(x,t)|^2 = \frac{1}{\sqrt{2\pi}\ |\delta x(t)|}\ \exp\left(-\frac{(x - \frac{\bar{p}}{m}t)^2}{2\delta x(t)^2}\right),\qquad (88)$$

where we have defined

$$\delta x(t)^2 \equiv \delta x^2 + \left(\frac{\hbar t}{2m\delta x}\right)^2.\qquad (89)$$

Comparing (88) with (87) gives us a picture of the time
evolution of a Gaussian <u>wave packet</u>, which we can draw as
follows:
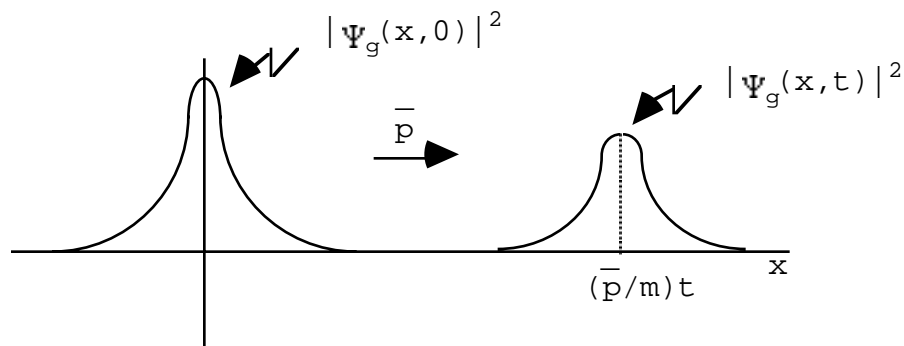


We see that the the peak of the position-space wave
packet $|\psi_g(x,t)|^2$ moves with velocity $\frac{\bar{p}}{m}$, which you will show
in an exercise is the expectation value of the wave packet's
momentum divided by mass, and thus corresponds to the usual
notion of particle velocity. In addition, it does not
maintain it's same shape but <u>spreads</u> in time because of
dispersion in values of momentum.  We shall see momentarily
that the magnitude squared of the momentum-space distribution

does <u>not</u> change with time.  Interpreted as uncertainties, this means that the product $\Delta x \, \Delta p_x$ grows in time.  This is consistent with the $\geq$ sign in the uncertainty relation (32).  The behavior of the wavepacket described by (72) is in contrast with the function $f(x,t)$ in (67) which obeys $f(x,0)=f(x+ct,t)$ (for $g(k)=0$, $k<0$) and therefore propagates undistorted in time with a single velocity, c.

It will behoove us in what is to come to study a little bit about <u>Dirac delta functions</u> now.  As stated in (45), the Fourier transform of $g(k)$ is defined as

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(k)e^{ikx}dk, \tag{90}$$

while the $g(k)$ distribution can be shown to be given as

$$g(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-ikx}dx, \tag{91}$$

Substituting (91) in (90), we get

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk\, e^{ikx} \int_{-\infty}^{\infty} dx'\, f(x')e^{-ik'x'}$$

$$= \int_{-\infty}^{\infty} dx'\, f(x') \left[ \int_{-\infty}^{\infty} \frac{dk}{(2\pi)} e^{ik(x-x')} \right]. \tag{92}$$

We define

$$\delta(x-x') = \int_{-\infty}^{\infty} \frac{dk}{(2\pi)} e^{ik(x-x')}, \tag{93}$$

so

$$f(x) = \int_{-\infty}^{\infty} dx'\, f(x')\delta(x-x'). \tag{94}$$

Let's set x=0 in (94).  Then

$$f(0) = \int_{-\infty}^{\infty} dx' f(x') \delta(x'). \tag{95}$$

Setting f(x) = 1 in (95) tells us also that

$$1 = \int_{-\infty}^{\infty} dx' \delta(x'). \tag{96}$$

Eq$^n$ (96) tells us that the area under the curve $\delta(x')$ is unity, but Eq$^n$ (95) tells us that the only non zero value of $\delta(x')$ is at x' = 0.  Such a function only exists in a limiting sense.  In this limit, $\delta(x')$ is an even function of x' (this can be established from (93)):

$$\delta(x') = \delta(-x'). \tag{97}$$

We can show various properties of the Dirac delta function based on the above.  For example we have

$$\delta(ay) = \delta(-ay), \tag{98}$$

from (97), where "a" is a constant.  Therefore we have

$$\delta(ay) = \delta(|a|y). \tag{99}$$

We can now write (the integral limits are understood to include the point y = 0)

$$\int dy f(y) \delta(ay) = \int dy f(y) \delta(|a|y)$$

$$= \frac{1}{|a|} \int dz f\left(\frac{z}{|a|}\right) \delta(z) = \frac{1}{|a|} f(0). \tag{100}$$

On the other hand

$$\frac{1}{|a|} f(0) = \int dy f(y) \left( \delta(y) \frac{1}{|a|} \right) , \tag{101}$$

so that

$$\delta(ay) = \frac{1}{|a|} \delta(y) . \tag{102}$$

Eq$^{n}$ (102), like all identities satisfied by Dirac delta functions, are really understood to be true in the context of integration. The functions f(x) above must be sufficiently smooth, but are otherwise completely arbitrary.

The form (93) is not the only explicit representation of the Dirac delta function. An easy way of inventing representations of $\delta(x)$ is to set

$$\delta(x) = \int_{-\infty}^{\infty} \frac{dk}{(2\pi)} e^{ikx} K(k), \tag{103}$$

following (93) above, requiring for K(k) that

$$K(k) = \begin{cases} \sim 1, \text{ below a cutoff in } |k| \\ \sim 0, \text{ above a cutoff in } |k| . \end{cases} \tag{104}$$

The Dirac delta function is then defined in the limit of the cutoff going to infinity.

For example, consider

$$K(k) = \lim_{\in \to 0^{+}} e^{-\frac{\in k^{2}}{4}} \tag{105}$$

where $\in \to 0^{+}$ means the limit as $\in$ goes to zero through positive values. We then have

$$K(k) = \begin{cases} \sim 1, \text{ for } k \lesssim \dfrac{1}{\sqrt{\in}} \\[3em] \sim 0, \text{ for } k \geq \dfrac{1}{\sqrt{\in}} \end{cases} \tag{106}$$

Putting (105) into (103) and interchanging the limit and the integral, we then find that $\delta(x)$ can be represented as

$$\delta(x) = \frac{1}{\sqrt{\pi}} \lim_{\in \to 0^+} \frac{1}{\sqrt{\in}} e^{-\frac{x^2}{\in}} \tag{107}$$

Eq$^n$ (107) is a Gaussian peaked at $x = 0$ which becomes increasingly narrow as $\in \to 0^+$, but which continues to have unit area. Using this technique, we can construct many forms of the Dirac delta function.

Let us return to interpreting the Gaussian wave packet solution of the Schrödinger equation. We already found that the magnitude of this solution spreads in time. The k-space (or momentum space since $k = p/\hbar$) transform of $\psi_g(x,0)$ is also a Gaussian, the absolute square of which has a width given by

$$\delta p = \frac{\hbar}{2\delta x} . \tag{108}$$

(The width of the Gaussian is defined the same as on p.2.20 above). We can then write

$$\delta x(t)^2 = \delta x^2 + (\delta v t)^2, \tag{109}$$

where $\delta v \equiv \dfrac{\delta p}{m}$ . Thus, the wave packet is spreading because of the initial velocity distribution at $t = 0$. Our Gaussian

wave packet seems to be spreading out not because the
particle itself is spreading, but because there was an
initial uncertainty in the particle's velocity.  We are being
led to the point of view that the magnitude of the function
$\psi_g(x,t)$ at the point x somehow represents the <u>probability</u>
that the particle is in a given location.  The simplest
possible positive semi-definite quantity we can form out of a
solution $\psi(x,t)$ of the Schrödinger equation is $|\psi(x,t)|^2$,
(Remember, $\psi(x,t)$ is in general complex) so we hypothesize
that:

$$\left.\begin{array}{l}\text{The probability that the single}\\ \text{particle described by the wave}\\ \text{packet } \psi(x,t) \text{ may be found}\\ \text{between positions x and x + dx}\\ \text{at time t}\end{array}\right\} = |\psi(x,t)|^2 dx. \qquad (110)$$

Thus $|\psi(x,t)|^2$ is more properly called a <u>probability density</u>.
Because of the interpretation in (110), we will often require
that

$$\int_{-\infty}^{\infty} dx \, |\psi(x,t)|^2 = 1, \qquad (111)$$

stating that the probability that the particle is somewhere
is unity.  Eq$^{\underline{n}}$ (111) is called a <u>normalization</u> condition.
(We do not always explicitly require (111) for solutions of
Schrödinger's equation, as we will see.)  In order for the
integral in (111) to exist, it is clear that we must have
$\psi(x,t) \to 0$ as $|x| \to \infty$.  In fact, we must have $|\psi(x,t)|$
decreasing faster than $|x|^{-1/2}$ as $|x| \to \infty$.

We now ask the question:  What does (111) imply for the momentum-space distribution of the wave packet?  Referring back to (47) above, we define the momentum-space wave packet $\psi(p_x)$ implicitly through $\left(k = \dfrac{p_x}{\hbar}\right)$

$$\psi(x,0) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp_x \psi(p_x) e^{ip_x x/\hbar} . \tag{112}$$

In fact, we may generalize (112) to

$$\psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp_x \psi(p_x,t) e^{ip_x x/\hbar} , \tag{113}$$

using (67) above as a model, where we have defined ($\omega = \dfrac{E}{\hbar}$ , where $E = \dfrac{p_x^2}{2m}$ for a free particle)

$$\psi(p_x,t) = \psi(p_x) e^{-i\omega t} . \tag{114}$$

From (114) it is clear that the function $|\psi(p_x,t)|^2$ is time-independent for our Gaussian wave packet, which means it does not spread out in time, contrary to the position-space distribution.  Using (112) in (111), we have the statement that at time t

$$\int dx |\psi(xt)|^2 = \int dx \psi(x,t) \psi*(x,t)$$

$$= \frac{1}{2\pi\hbar} \int dx \int dp_x \psi(p_x,t) e^{ipx/\hbar} \int dp_x' \psi*(p_x',t) e^{-ip_x' x/\hbar}$$

$$= \int \int dp_x dp_x' \psi(p_x,t) \psi*(p_x',t) \left[\frac{1}{2\pi\hbar} \int dx e^{-ix(p_x-p_x')/\hbar}\right]. \tag{115}$$

In the above integrals and others following where the limits of integration are not explicitly stated, we are to understand these to be from $-\infty$ to $+\infty$ on both position and momentum. From our recent discussion of Dirac delta functions, we realize that the above quantity in square brackets is just a delta function in momentum,

$$\delta(p_x - p_x') = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dx\, e^{ix(p_x-p_x')/\hbar} \quad . \tag{116}$$

Therefore, we have

$$\int dx\, |\psi(x,t)|^2 = \int dp_x\, |\psi(p_x,t)|^2 \quad . \tag{117}$$

This result is known as <u>Parseval's theorem</u>. Thus, if we normalize using (111), Eq$^n$(117) suggests an interpretation of $|\psi(p_x,t)|^2$ similar to (110) above for $\psi(x,t)|^2$:

$$\left.\begin{array}{l}\text{The probability that the single}\\ \text{particle described by the wave}\\ \text{packet } \psi(p_x,t) \text{ may be found with}\\ \text{momentum values between } p_x \text{ and}\\ p_x + dp_x \text{ at time } t\end{array}\right\} = |\psi(p_x,t)|^2\, dp_x. \tag{118}$$

It is crucial for this discussion that probability be a conserved quantity. That is, once we have imposed (111) at $t = 0$, it must continue to be true that the total probability is one. Our one-dimensional Schrödinger equation is

$$i\hbar\, \frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\, \frac{\partial^2\psi(x,t)}{\partial x^2} \quad . \tag{119}$$

The complex conjugate of this is

$$-i\hbar \ \frac{\partial \psi*(x,t)}{\partial t} \ = \ - \ \frac{\hbar^2}{2m} \ \frac{\partial^2 \psi*(x,t)}{\partial x^2} \ . \tag{120}$$

Using (119) and (120), we find (we suppress the arguments of $\psi(x,t)$)

$$i\hbar \ \frac{\partial}{\partial t}(\psi*\psi) \ = \ i\hbar \ \left( \frac{\partial \psi*}{\partial t} \ \psi \ + \ \psi* \ \frac{\partial \psi}{\partial t} \right)$$

$$= \ \frac{\hbar^2}{2m} \left( \frac{\partial^2 \psi*}{\partial x^2} \ \psi \ - \ \psi* \ \frac{\partial^2 \psi}{\partial x^2} \right). \tag{121}$$

This last statement can be written as (called the "continuity equation")

$$\frac{\partial}{\partial t} \ (\psi*\psi) \ + \ \frac{\partial}{\partial x} \ j(x,t) \ = \ 0 \tag{122}$$

where

$$j(x,t) \ = \ \frac{i\hbar}{2m} \left[ \frac{\partial \psi*}{\partial x} \ \psi \ - \ \psi* \ \frac{\partial \psi}{\partial x} \right] \ , \tag{123}$$

$$= \ \frac{\hbar}{m} \ \text{Im} \ \left[ \ \psi \ \frac{\partial \psi*}{\partial x} \right] \ . \tag{124}$$

$j(x,t)$ is called the "probability flux" and represents the flux or change in probability at a given position and time. Notice from (124) that it is a _real_ quantity, as it should be. Integrating Eqⁿ (122) over all positions, we then come to the conclusion that

$$\frac{\partial}{\partial t} \ \int_{-\infty}^{\infty} dx |\psi(x,t)|^2 \ = \ -j(x,t)\Big|_{-\infty}^{\infty}. \tag{125}$$

Now we argued above that the requirement (111) on $\psi(x,t)$ meant that it had to decrease faster than $|x|^{-1/2}$ as $|x| \to \infty$. This means that $j(x,t)$ must decrease faster than $|x|^{-2}$ as $|x| \to \infty$, from (124). Therefore $j(x,t)$ vanishes at the limits $x = \pm\infty$ and we conclude that

$$\frac{\partial}{\partial t} \int_{-\infty}^{\infty} dx \, |\psi(x,t)|^2 = 0. \tag{126}$$

Thus, the existence of the continuity equation, (122), plus the requirement that the norm of $\psi(x,t)$ be bounded, (111), insures a probabilistic interpretation for $|\psi(x,t)|^2$.

The Schrödinger equation, Eq$^{\underline{n}}$ (119), is written as a differential statement involving the position-space function $\psi(x,t)$. The analogous statement, which can be derived using (119) and (113) above, for the momentum-space distribution $\psi(p_x,t)$ is

$$i\hbar \frac{\partial \psi(p_x,t)}{\partial t} = \frac{p_x^2}{2m} \psi(p_x,t). \tag{127}$$

This is called the momentum-space Schrödinger equation. In the form (127) it is still referring to a free particle in a single dimension. The general solution to (127) is just (114) above,

$$\psi(p_x,t) = \psi(p_x) e^{-i\frac{p_x^2}{2m\hbar} t}. \tag{128}$$

We now try to find solutions of (119) that are separable, i.e. that can be written in the form

$$\psi(x,t) = u(x)T(t). \tag{129}$$

Notice that our <u>general</u> solution for $\psi(x,t)$, Eq$^{\underline{n}}$ (113), is <u>not</u> of this form. We will give a physical interpretation of this in a moment. Substituting (129) into (119), we find that

$$i\hbar u(x) \frac{dT}{dt} = -\frac{\hbar^2}{2m} \frac{d^2u}{dx^2} T(t). \tag{130}$$

Dividing by $u(x)T(t)$, we get

$$i\hbar \frac{\frac{dT}{dt}}{T(t)} = -\frac{\hbar^2}{2m} \frac{\frac{d^2u}{dx^2}}{u(x)} . \tag{131}$$

Since x and t are independent variables, this equation can only be satisfied if both sides are equal to a constant, which with foresight we call E. The solution of

$$i\hbar \frac{dT}{dt} = ET(t), \tag{132}$$

is

$$T(t) = Ce^{-iEt/\hbar} . \tag{133}$$

The equation for $u(x)$ is then (we may label $u(x)$ as $u_E(x)$ if desired)

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dx^2} = Eu(x). \tag{134}$$

The two linearly independent solutions of (134) are of the form

$$Ce^{\pm ik'x} , \tag{135}$$

where $k' = \frac{1}{\hbar}\sqrt{2mE}$ is an undetermined positive quantity.

Thus, separable solutions to (119) are of the form

$$\psi(x,t) \rightarrow Ce^{i(k'x - Et/\hbar)} , Ce^{i(-k'x - Et/\hbar)} . \qquad (136)$$

were E is allowed by (134) to take on all positive values.
(The solutions to (134) for E < 0 are unbounded in x and do
not lend themselves to a probabilistic interpretation.)  We
now realize that the forms (136) for $\psi(x,t)$ result from
taking $\psi(p_x) = C'\delta(p_x - \hbar k')$ or $\psi(p_x) = C'\delta(p_x + \hbar k')$ in our
general solution for $\psi(x,t)$, Eq$^n$ (113) above.  Thus the plane
waves (136) represent <u>specific</u> solutions to the Schrödinger
equation which have a given kinetic energy, E.  We have just
found in fact that there are two such solutions, specified by
$p_x = \pm\hbar k'$ (particle moving to the left or right).  The fact
that our general solution (113) is not separable in x and t
means therefore that it is not a solution with a unique value
of E.  The Gaussian wave packet for example contained a
continuous distribution of $p_x$ (and therefore E) values.  It
makes physical sense that solutions to (134) exist for all
positive values of E since the kinetic energy of a free
particle takes on positive values.  We will see in the
following chapter, however, that the generalization of (134)
that takes a quantum mechanical potential into account in
general <u>does</u> restrict the values of the particle's total
energy, often becoming discrete.  The equation (134) is the

time-independent Schrödinger equation for a free particle in one dimension.

We have succeeded in deducing a probability conserving differential equation for a material particle that has the position-momentum uncertainty relation built into it.  We have done this without reference to our earlier discussion involving spin.  However, I would now like to show that the Schrödinger equation is a natural outgrowth of the formalism we developed in the last chapter.  I also wish to provide the connections between the differential equation point of view in this chapter and the operator formalism of the prior chapter.

The foundation on which we built the operator concept was the existence of what we were calling a basis of measurement symbols.  Such a basis must allow a complete specification of all possible outcomes of a given experimental apparatus.  This is the physical content of the mathematical expression of the completeness concept, Eq$^{\text{n}}$ (91) in Ch.1.  The basis there was discrete.  In the case of spin $\frac{1}{2}$ it consisted of two states $|+>$ and $|->$.  In the case of a free particle (considered spinless for now) an appropriate basis would be a specification of all possible locations or momentum values, which we assume take on continuous values. We can imagine doing all the operations discussed in the last chapter, selection, modulation and transitions, but in this case on a "beam" of particles taking

on various values of position or momenta. (A selection
experiment on particle position might be realized by a
diffraction experiment, for example.)  Therefore, in analogy
to the previous discrete specifications of completeness, we
postulate the continuum statements

$$\int dx' \, |x'> <x'| = 1, \qquad (137)$$

and

$$\int dp_x' \, |p_x'> <p_x'| = 1, \qquad (138)$$

as expressing the completeness of a physical description
based upon continuous positions, Eq$^{\underline{n}}$ (137), or momentum
values Eq$^{\underline{n}}$ (138).  The right hand sides of (137) and (138)
are not the number one, but the unity operator.  In order to
be consistent, we must now have (we again write
$<x'|\cdot|x"> \equiv <x'|x">$ and recognize this product is an
ordinary number)

$$1 = 1 \cdot 1 = \int dx'dx" \, |x'> <x'|x"> <x"|, \qquad (139)$$

from which we realize that

$$<x'|x"> = \delta(x' - x"), \qquad (140)$$

for then

$$\int dx'dx" \, |x'><x'|x"><x"| = \int dx' \, |x'> \int dx" \delta(x'-x")<x"|$$
$$= \int dx' \, |x'><x'| = 1. \qquad (141)$$

Likewise, we have that

$$<p_x'|p_x"> = \delta(p_x' - p_x"). \qquad (142)$$

Eq$^{\text{n}}$s (140) and (142) are the expressions of
<u>orthonormality</u> in position and momentum space.  (Compare with
Eq$^{\text{n}}$ (158), Ch.1, where the Kronecker delta has just been
replaced by the Dirac delta function.)  In addition, we
expect there to be <u>operators</u> for position and momentum just
as before we constructed an operator for $S_z$.  Our model in
such a construction is Eq$^{\text{n}}$ (101), Ch.1, where, however, we
would expect the discrete sum there to be replaced with an
integral over continuous positions or momentums.  Thus, as a
natural outgrowth of our earlier experiences with discrete
systems, we expect a representation for position and momentum
operators by

$$x = \int dx' x' |x'><x'| , \qquad (143)$$

and

$$p_x = \int dp_x' p_x' |p_x'><p_x'| , \qquad (144)$$

where, in this context, the $x'$ and $p_x'$ are numbers and the $x$
and $p_x$ are our more abstract operator quantities.  We now
check that Eq$^{\text{n}}$ (102) of Ch.1 above is holding:

$$x|x'> = (\int dx'' x'' |x''><x''|) \cdot |x'>$$
$$= \int dx'' |x''> \delta(x'' - x') x'' = x'|x'> , \qquad (145)$$

and

$$p_x|p_x'> = (\int dp_x'' p_x'' |p_x''><p_x''|) \cdot |p_x'>$$
$$= \int dp_x'' p_x'' |p_x''> \delta(p_x'' - p_x') = p_x'|p_x'> . \qquad (146)$$

An equation of the form $A|a'> = a'|a'>$ is called an _eigenvalue equation_, the state $|a'>$ is called an _eigenvector_ (or _eigenket_) and the number $a'$ is called the _eigenvalue_. We first saw such forms for spin systems. For spin $\frac{1}{2}$ the eigenvectors are just $|+>$ and $|->$ and the eigenvalues of the operator $S_z$ are just $\pm \frac{\hbar}{2}$ .

Let us also assume the existence of the mathematical adjoint operation, denoted "$+$", which connects our new bra and ket states. That is, we assume that

$$(|x'>)^+ = <x'| , \qquad (147)$$

$$(|p_x'>)^+ = <p_x'| . \qquad (148)$$

Now the physical outcomes, $x'$ and $p_x'$ in (143) and (144), represent the result of position or momentum measurements in a one dimensional space. They are necessarily real and this has the consequence that

$$x^+ = (\int dx' x' |x'><x'|)^+ = \int dx' x' |x'><x'| = x, \qquad (149)$$

$$p_x^+ = (\int dp_x' p_x' |p_x'><p_x'|)^+ = \int dp_x' p_x' |p_x'><p_x'| = p_x. \qquad (150)$$

Such a construction always results in a physical property which is Hermitian. As pointed out at the end of Chapter 1, Hermitian operators have real expectation values. Using (149) and (150), we then find that

$$<x'|x = x'<x'| , \qquad (151)$$

$$\langle p_x'|p_x = p_x'\langle p_x'| \ , \tag{152}$$

where we have used the adjoint in (145) and (146). (151) and (152) may be independently verified using the definition of x and $p_x$, Eq$^{\underline{n}}$s (143) and (144) above. We will also refer to equations in the form (151) and (152) as eigenvalue equations.

We make another important realization in the development of our x,$p_x$ formalism by considering Eq$^{\underline{n}}$ (142). Using the x-representation of the unity operator, we write the left hand side of this equation as (inserting the unit operator in the x-basis, Eq$^{\underline{n}}$ (137))

$$\langle p_x'|p_x''\rangle = \int dx' \langle p_x'|x'\rangle\langle x'|p_x''\rangle \tag{153}$$

which, because

$$\langle p_x'|x'\rangle = (\langle x'|p_x'\rangle)^+ = \langle x'|p_x'\rangle^* \ , \tag{154}$$

we can write as

$$\langle p_x'|p_x''\rangle = \int dx' \langle x'|p_x'\rangle^* \langle x'|p_x''\rangle \ . \tag{155}$$

Therefore Eq$^{\underline{n}}$ (142) reads

$$\int dx' \langle x'|p_x'\rangle^* \langle x'|p_x''\rangle = \delta(p_x' - p_x'') . \tag{156}$$

By comparing the left hand side of (156) with an explicit representation of the delta function (see Eq$^{\underline{n}}$ (116) above; remember, the Dirac delta function is even in it's argument)

$$\frac{1}{2\pi\hbar} \int dx' e^{-ix'(p_x' - p_x'')/\hbar} = \delta(p_x' - p_x'') \,, \tag{157}$$

we see that it is consistent to choose

$$\langle x'|p_x'\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ix'p_x'/\hbar} \,. \tag{158}$$

We will take (158) as the <u>definition</u> of the bra-ket product $\langle x'|p_x'\rangle$. We will test the consistency of this defintion shortly.

We are now in a position to make a crucial realization. We introduce the "Hamiltonian" operator

$$H = \frac{p_x^2}{2m} \,, \tag{159}$$

which is the obvious operator quantity to represent the energy of a free particle.  In analogy to the eigenvalue equations for spin, position and momentum, we postulate a similar equation for H.

$$H|a'\rangle = E_{a'}|a'\rangle. \tag{160}$$

The a' are <u>labels</u> of the allowed energy values $E_{a'}$.  (In the case of the free particle, "a'" represents a <u>continuous</u> label, which, in fact, it may be more convenient simply to choose as E, the actual energy value of the particle.)  We now multiply both sides of (160) on the left by the bra $\langle x'|$:

$$\langle x'|\frac{p_x^2}{2m}|a'\rangle = E_{a'}\langle x'|a'\rangle. \tag{161}$$

In order to show the connection of (161) to our previous results, consider the quantity $<x'|p_x|a'>$. By inserting a complete set of $p'_x$ states, we see that

$$<x'|p_x|a'> = \int dp'_x <x'|p'_x><p'_x|p_x|a'>$$

$$= \frac{1}{\sqrt{2\pi\hbar}} \int dp'_x e^{ix'p'_x/\hbar} p'_x <p'_x|a'>, \qquad (162)$$

where we have used (152) and (158). We may now write

$$\int dp'_x e^{ix'p'_x/\hbar} p'_x <p'_x|a'> = \frac{\hbar}{i} \frac{\partial}{\partial x'} \int dp'_x e^{ix'p'_x/\hbar} <p'_x|a'>. \qquad (163)$$

Working backwards, it is clear that

$$\frac{1}{\sqrt{2\pi\hbar}} \int dp'_x e^{ix'p'_x/\hbar} <p'_x|a'> = \int dp'_x <x'|p'_x><p'_x|a'>$$

$$= <x'|a'>. \qquad (164)$$

Putting (162),(163), and (164) together yields the statement that

$$<x'|p_x|a'> = \frac{\hbar}{i} \frac{\partial}{\partial x'} <x'|a'>. \qquad (165)$$

Since (165) is supposed to be true for <u>any</u> $|a'>$, this then implies ("stripping off the $|a'>$")

$$<x'|p_x = \frac{\hbar}{i} \frac{\partial}{\partial x'} <x'|. \qquad (166)$$

In other words, the Hilbert space operator $p_x$ acting on a position bra is the same as a differential operator, $\frac{\hbar}{i} \frac{\partial}{\partial x'}$

acting on the same bra.  Please do <u>not</u> think this means that $p_x = \frac{\hbar}{i} \frac{\partial}{\partial x'}$ .  It only means that the Hilbert-space operator quantity $p_x$ (defined in (144) above) is replaced by the differential operator $\frac{\hbar}{i} \frac{\partial}{\partial x'}$ when acting in position space. (When $p_x$ acts in momentum space, it gives the number $p_x'$).  By taking the Hermitian adjoint of (166) we then find that

$$p_x |x'> = - \frac{\hbar}{i} \frac{\partial}{\partial x'} |x'>. \qquad (167)$$

(Remember, $p_x$ is Hermitian.)  Likewise, by considering the quantity $<p_x'|x|a'>$, it is possible in the same manner to show that

$$<p_x'|x = - \frac{\hbar}{i} \frac{\partial}{\partial p_x'} <p_x'|, \qquad (168)$$

and

$$x|p_x'> = \frac{\hbar}{i} \frac{\partial}{\partial p_x'} |p_x'>. \qquad (169)$$

Let us test the consistency of these conclusions along with the statement (158) above.  Consider the quantity $<x'|x|p_x'>$.  By allowing x to act first to the left, we find that

$$<x'|x|p_x'> = x'<x'|p_x'>. \qquad (170)$$

On the other hand, by allowing x to act first on $|p_x'>$ we find

$$<x'|x|p_x'> = \frac{\hbar}{i} <x'| \frac{\partial}{\partial p_x'} |p_x'> = \frac{\hbar}{i} \frac{\partial}{\partial p_x'} <x'|p_x'>$$

$$= \frac{\hbar}{i} \, \frac{1}{\sqrt{2\pi\hbar}} \, \frac{\partial}{\partial p_x'} \, e^{ip_x'x'/\hbar} = \frac{x'}{\sqrt{2\pi\hbar}} \, e^{ip_x'x'/\hbar}$$

$$= x' \langle x' | p_x' \rangle. \qquad (171)$$

The right hand sides of (170) and (171) agree, as they should. This confirms (158) above, up to an overall constant.

Let us now go back and apply what we have learned to (161) above. We now recognize that

$$\langle x' | p_x^2 = \frac{\hbar}{i} \, \frac{\partial}{\partial x'} \, \langle x' | p_x = - \hbar^2 \, \frac{\partial^2}{\partial x'^2} \, \langle x' |, \qquad (172)$$

so that (161) reads

$$- \frac{\hbar^2}{2m} \, \frac{\partial^2}{\partial x'^2} \, \langle x' | a' \rangle = E_{a'} \, \langle x' | a' \rangle. \qquad (173)$$

We recognize (173) as just the time-independent Schrödinger equation, Eq$^{\underline{n}}$ (134) above, that we originally motivated from a wave-packet point of view. We now see that our wave equation and operator viewpoints will connect if we take

$$u_{a'}(x') = \langle x' | a' \rangle. \qquad (174)$$

That is, we have come to the realization that the time-independent Schrödinger equation is just a position-space statement of the eigenvalue equation for the Hamiltonian, and the functions $u_{a'}(x')$ are <u>wavefunctions</u> that express the transition amplitude from the energy basis to the position basis. That is, along with a characterization of the unit

2.50

operator in position and momentum space as in (137) and (138),
we also assume that there is also an _energy_ characterization:[*]

$$\left\{ \sum_{\substack{p'_x \geq 0, \\ p'_x < 0}} \int_0^\infty dE_{a'} \text{ or } \sum_{a'} \right\} |a'><a'| = 1. \qquad (175)$$

I am suggesting in (175) that in some occasions we will find a
discrete _spectrum_ of energy values, $E_{a'}$.  You should go back to
the discussion Ch.1 to refresh yourself on the concept of a
wavefunction as a transition amplitude.  Also compare (174)
above with Eq$^{\underline{n}}$ (207) of Chapter 1.

   Interpreting $|u_{a'}(x')|^2$ as a probability density is also
consistent with our earlier finding in the case of spin that
probabilities are the absolute squares of transition
amplitudes.  That is, we may use completeness in position
space to write energy-space orthonormality (here we assume the
energies are discrete), $<a'|a"> = \delta_{a'a"}$, as

$$\int dx' <a'|x'><x'|a"> = \int dx' u_{a'}^*(x')u_{a"}(x') = \delta_{a'a"}. \qquad (176)$$

When a' = a" we get

$$\int_{-\infty}^\infty dx' |u_{a'}(x')|^2 = 1, \qquad (177)$$

---

[*] Comment on Eq$^{\underline{n}}$ (175):  What we are seeing here for the first time, in
the case of the continuous energy values of the free particle, is a
case of energy _degeneracy_ of the energy eigenkets, $|a'>$.  Specifying
their energy, $E_{a'}$ still leaves open the question of whether the
particle has positive or negative momentum.  Thus, in the case of the
continuum statement of completeness in (175), it is necessary to add
the sum $\sum_{p'_x \geq 0, p'_x < 0}$ in order to satisfy completeness. We will talk more
about degeneracies of systems in Chapter 4.

which is the same as (111) above when $\psi(x,t)$ refers to an energy eigenstate, $\psi(x,t) \rightarrow u_{a'}(x)e^{-iE_{a'}t/\hbar}$ .

When (160) above is projected into $\langle p'_x|$, we get

$$\langle p'_x| \frac{p_x^2}{2m} |a'\rangle = E_{a'}\langle p'_x|a'\rangle, \tag{178}$$

$$\Rightarrow E_{a'} = \frac{p'^2_x}{2m} , \tag{179}$$

which confirms that $E_{a'}$ represents the energy of a free particle. The quantities $\langle p'_x|a'\rangle$ are the momentum space energy wavefunctions. The position-space energy wavefunctions $u_{a'}(x') = \langle x'|a'\rangle$ can be related to the $\langle p'_x|a'\rangle$ by the use of momentum-space completeness:

$$u_{a'}(x') = \langle x'|a'\rangle = \int dp'_x \langle x'|p'_x\rangle\langle p'_x|a'\rangle$$

$$= \frac{1}{\sqrt{2\pi\hbar}} \int dp'_x e^{ip'_x x'/\hbar} \langle p'_x|a'\rangle. \tag{180}$$

Comparing (180) with (112) above, when $\psi(x,0) = u_{a'}(x)$, we see that

$$u_{a'}(p'_x) \equiv \langle p'_x|a'\rangle. \tag{181}$$

should be interpreted as the momentum-space energy wavefunction.

Notice that there is as yet no reference to time development in our operator formalism, as opposed to the wave packet discussion, where the Schrödinger equation described the evolution in time of our Gaussian wave packet, for example. However, we receive an important hint on one way to

incorporate time development from Eq$^{\underline{n}}$ (129) which tells us how the energy eigenvalue wavefunctions $u_{a'}(x) = \langle x|a'\rangle$ evolve in time.  In bra-ket notation, (129) can written

$$\langle x'|a',t\rangle = e^{-iE_{a'}t/\hbar} \langle x'|a'\rangle, \qquad (182)$$

where we have defined the time-evolved state $|a',t\rangle$.  But notice that

$$\langle x'|e^{-iHt/\hbar}|a'\rangle = e^{-iE_{a'}t/\hbar} \langle x'|a'\rangle. \qquad (183)$$

Therefore

$$\langle x'|a',t\rangle = \langle x'|e^{-iHt/\hbar}|a'\rangle. \qquad (184)$$

Eq$^{\underline{n}}$ (184), being true for all $\langle x'|$ then implies that

$$|a',t\rangle = e^{-iHt/\hbar}|a'\rangle. \qquad (185)$$

The quantity $e^{-iHt/\hbar}$ is called the <u>time evolution operator</u>.  It provides the key to understanding the time development of particle states.  We notice that this operator, like the operators that describe rotations, is unitary.  That is, given that H is Hermitian (which is certainly true for the free particle where $H = \dfrac{p_x^2}{2m}$ ) we have that

$$(e^{-iHt/\hbar})^{+} = e^{iH^{+}t/\hbar} = e^{iHt/\hbar} . \qquad (186)$$

Now we take the time derivative of (185).  This yields

$$i\hbar \frac{\partial}{\partial t} |a',t\rangle = H\, e^{-iHt/\hbar}|a'\rangle = H|a',t\rangle \qquad (187)$$

By multiplication by a position bra, $\langle x'|$, and by use of (166) above, this now reads

$$i\hbar \frac{\partial}{\partial t} \langle x'|a',t\rangle = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x'^2} \langle x'|a',t\rangle. \qquad (188)$$

Eq$^{\underline{n}}$ (188) says that the time-evolved energy eigenstate, projected into position-space, satisfies the Schrödinger equation.  As a continuation of our earlier notation, we will write $u_{a'}(x',t) = \langle x'|a',t\rangle = e^{-iE_{a'}t/\hbar} u_{a'}(x')$.  (It is easy to check that (187), projected into momentum space, gives the momentum space Schrödinger equation with $\psi(p'_x,t) = \langle p'_x|a',t\rangle = e^{-iE_{a'}t/\hbar} \psi(p'_x)$).  The final connection with the Schrödinger equation, (119) above, becomes complete when we realize that this is a linear differential equation. Therefore, given the time-evolved solutions $u_{a'}(x',t) = \langle x'|a',t\rangle$ of the time independent Schrödinger equation, the most general solution is (again assuming a situation where the energies are discrete)

$$\psi(x',t) = \sum_{a'} u_{a'}(x',t)C_{a'} = \sum_{a'}\langle x'|a',t\rangle C_{a'} \qquad (189)$$

where the $C_{a'}$ are an arbitrary set of constants.  Introducing the notation

$$|\psi,t\rangle = \sum_{a'} |a',t\rangle C_{a'} \qquad (190)$$

for the most general linear combination of ket states, we find that

$$\psi(x',t) = \langle x'|\psi,t\rangle. \qquad (191)$$

In the same way, projecting the general state $|\psi,t>$ into momentum space, the corresponding momentum-space energy wave function is

$$\psi(p_x',t) = \sum_{a'} u_{a'}(p_x',t)C_{a'} = \sum_{a'} <p_x'|a',t> C_{a'} \qquad (192)$$

The quantities $\psi(x',t)$ and $\psi(p_x',t)$ are, of course, Fourier transforms of each other, the general connection being Eq$^{\underline{n}}$ (113) above. An alternative treatment of time development will be presented in Chapter 4.

In the context of our coordinate space discussion, if we take $A = A(x)$, expectation values are given as

$$< A(x) >_{\psi,t} = < \psi,t|A(x)|\psi,t >$$
$$= \int dx'dx"< \psi,t|x'> < x'|A(x)|x" > < x"|\psi,t >. \qquad (193)$$

Now

$$A(x)|x" > = A(x")|x" >, \qquad (194)$$

and so

$$< x'|A(x)|x" > = A(x")< x'|x" > = A(x")\delta(x' - x"), \qquad (195)$$

which results in

$$< A(x) >_{\psi,t} = \int dx'A(x')|\psi(x',t)|^2. \qquad (196)$$

Since $|\psi(x',t)|^2$ is the probability density, we see that $< A(x) >_{\psi,t}$ is obtained as a probability density - weighted integral and is explicitly real.

In the same manner, if $A = A(p_x)$, one can show that

$$< A(p) >_{\psi,t} = \int dp'_x A(p'_x) |\psi(p'_x,t)|^2. \qquad (197)$$

We have now recovered the basic dynamical equation of wave mechanics, the Schrödinger equation, from our earlier, spin-inspired, operator formalism. We have done this by applying the lessons we learned in the simpler spin case by analogy to particles in coordinate space. Our understanding of the mathematics of the underlying operator formalism is still quite incomplete. However, we have reached a point where, using what we have learned, we can solve some simple one-dimensional problems in quantum mechanics. This is what we will do in the next Chapter. Following that, we will try to fill in some of the gaps in our understanding of the operator formalism in Chapter 4.

**Problems**

1. Use the uncertainty relation to show that the potential,

$$V(r) = \frac{-k}{r^{2+\varepsilon}} \ , \ k, \ \varepsilon > 0,$$

is unstable for zero angular momentum states.  [Remember the argument for the stability of the hydrogen atom.  Also remember that setting $\frac{\partial E}{\partial r} = 0$ can pick out either a maximum or a minimum.]

2.  Let's return to the original S-G setup. We found that if $\delta p_z \delta z \geq \hbar$ (Eq$\underline{n}$ (57), p.1.26 of the notes), then $\delta\phi \geq 1$, and thus we have an intrinsic uncertainty in the phase angle of the magnetic moment. We now recognize that the condition $\delta p_z \delta z \geq \hbar$ is responsible for the diffraction or spreading out of the atomic beam as it passes through the slit in the wall. (See the picture on p. 1.25 of the notes.)

   (a)  Find the approximate value of the slit width, $\delta z$, that causes the magnetically split S-G beam to "wash out" due to diffraction.  Evaluate $\delta z$ numerically for our usual values:

   $$M_{Ag} = 1.79 \times 10^{-22} \ gm, \ \frac{\partial H}{\partial z} = 10^4 \ gauss \cdot cm^{-1},$$

   $$l = 10 \ cm, \ |\gamma| \approx 10^7 \ gauss^{-1}sec^{-1}, \ \frac{1}{2} mv^2 = \frac{3}{2} kT,$$

   $$T = 10^3 \ °K.$$

   (b)  Now replace the silver atom's mass, $M_{Ag}$, with the electron's mass in this calculation.  Find the new slit
width, $\delta z$, which causes diffraction washout.

3. Consider a wave packet defined by (47) on p. 2.18 with g(k) given by

$$g(k) = \begin{cases} 0, & k < -K \\ N, & -K < k < K \\ 0, & K < k. \end{cases}$$

(a) Find the form $f(x)$ and plot it.

(b) Show that a reasonable definition of $\delta x$ for (a) yields

$$\delta k \; \delta x \sim 1.$$

(c) Find the value of $N$ (up to a phase factor) for which

$$\int_{-\infty}^{\infty} dx \; |f(x)|^2 = 1.$$

[Hint: Think delta function.]

4. Find the momentum space wavefunction $\Psi_g(p_x,0)$ corresponding to the $t = 0$ coordinate space Gaussian wave function $\Psi_g(x,0)$, given in (86) of the notes. What value of $p_x$ maximizes $|\Psi(p_x,0)|^2$? Show that the width of $|\Psi(p_x,0)|^2$, in the same sense as on p.2.20 of the notes, is

$$\delta p_x = \frac{\hbar}{2\delta x} .$$

[Partial answer: $\Psi_g(p_x,0) = \dfrac{\sqrt{2}\delta x}{(2\pi\hbar^2)^{1/4}} \exp\left(- \dfrac{\delta x^2}{\hbar^2}(\bar{p}-p_x)^2\right)$ ]

5. Use the result of problem 2 to show the statement in the notes, p.2.29 (below the figure). That is, show that

$$\frac{\langle p_x \rangle}{m} \equiv \frac{1}{m} \int dp_x \; p_x |\Psi(p_x,0)|^2 = \frac{\bar{p}}{m},$$

and therefore that $\dfrac{<p_x>}{m}$ "corresponds to the usual notion of particle velocity."

6. Starting with the <u>coordinate space</u> free particle Schrodinger equation (119) of Ch. 2, show that the <u>momentum space</u> Schrodinger equation is given by Eq. (127) of Ch. 2.

7.  Define (the "uncertainty in A")

$$\Delta A = \sqrt{<A^2>_\psi - <A>_\psi^2} \; .$$

For $\Psi_g(x,o)$ as in the notes, Eq. (86), find:

(a)  $\Delta x = ?$
(b)  $\Delta p_x = ?$

A useful integral is

$$\int_{-\infty}^{\infty} dx\; x^2\; e^{-\alpha x^2} = -\frac{d}{d\alpha} \int_{-\infty}^{\infty} dx\; e^{-\alpha x^2}$$

$$\text{where} \int_{-\infty}^{\infty} dx\; e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}} \; .$$

8. Show that, in addition to Eq. (197) of Ch. 2, we may also write

$$<A(p_x)>_{\psi,t} = \int dx'\; \Psi^*(x',t)\; A(\frac{\hbar}{i}\frac{\partial}{\partial x'})\; \Psi(x',t),$$

for the expectation value of $A(p_x)$ if $A(p_x)$ is a power series in $p_x$.

**Other problems**

9. Consider a model of a heavy-light molecular system where the potential energy between the attractive heavy molecule (of infinite mass) and the light orbiting molecule (of mass "m") is given by

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

where "r" is the separation distance between the two molecules and "$\omega$" is a constant. The angular momentum of the electrons in the Bohr atom were quantized in units of $\hbar$. Assuming that the angular momentum of the light molecule is similarly quantized, and that only circular orbits are possible, find:

(a) The radius, $r_n$, of the light molecule's orbit in the $n^{\underline{th}}$ angular momentum state.

(b) The total energy, $E_n$, of the $n^{\underline{th}}$ angular momentum state.

10. Given the momentum space free-particle wavefunction ("$\bar{x}$", "a" are just constants) at t=0,

$$\Psi(p_x) = A \frac{\exp(\frac{i}{\hbar} \bar{x} \, p_x)}{\sqrt{p_x^2 + a^2}},$$

(a) Find the value of the constant "A". (An overall phase does not matter.)

(b) Find the expectation value of the momentum, $<p_x>$.

(c) Find the momentum wavefunction at all later times, $\Psi(p_x,t)$.

11. (a) Starting with the definition, $<p_x>_{\psi,t} = <\Psi,t|p_x|\Psi,t>$, show that

2.60

$$<p_x>_{\psi,t} = \int_{-\infty}^{\infty} dx' \psi^*(x',t)\frac{\hbar}{i} \frac{\partial}{\partial x'} \psi(x',t).$$

(b) Using (a) (or any other means), also establish that

$$\frac{<p_x>_{\psi,t}}{m} = \int_{-\infty}^{\infty} dx' j(x',t),$$

where

$$j(x',t) = \frac{i\hbar}{2m} \left(\frac{\partial \psi^*}{\partial x} \psi - \psi^* \frac{\partial \psi}{\partial x}\right)$$

is the probability current.


12. Use the Heisenberg uncertainty principle to estimate the ground state energy of a one-dimensional harmonic oscillator with energy,

$$E = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2.$$


13. Let's say we tried to use visible light instead of X-rays in a Compton-like scattering experiment (photons scattering from electrons). In this case show that the fractional change in the frequency, $\nu$, of the scattered light is given approximately by

$$\frac{|\Delta\nu|}{\nu} \approx \frac{2h\nu}{mc^2} \sin^2\frac{\theta}{2}.$$


14. Given the position space free-particle wavefunction,

$$\Psi(x) = A \exp(\frac{i}{\hbar} \bar{p}x - \frac{|x|}{2\delta x}),$$


(a) Find the value of the constant "A". (An overall phase does not matter.)

(b) Find the average momentum of the particle (the "expectation value" of $p_x$), $\langle p_x \rangle = \langle \Psi | p_x | \Psi \rangle$.

15. Consider the following experiment done with spin $\frac{1}{2}$ particles (in this case thermal neutrons would probably work the best) on a flat table top:
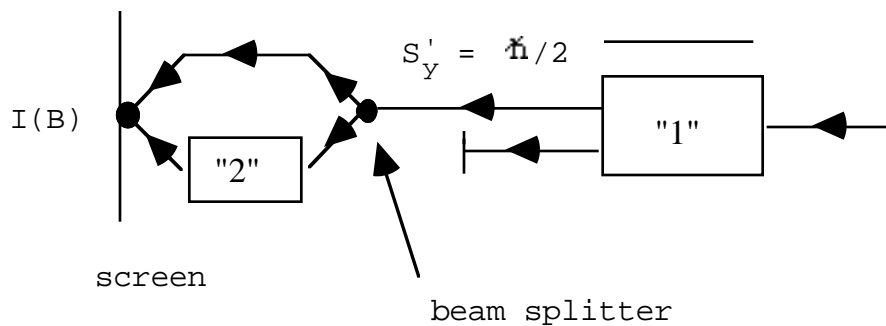


That is $S'_y = \frac{\hbar}{2}$ is first selected by magnet "1", then this beam is split into two parts (with equal amplitudes) that travel along the (equal length) paths shown. Before reaching the beam splitter, the particles have velocity "v" and the beam has intensity $I_0$. A second magnet with a <u>uniform</u> magnetic field pointing along the z-axis is also positioned along one of the beam paths, as shown. Assume that the beams constructively interfere with one another at the screen when magnet "2" is turned off, i.e. $I(B=0) = I_0$.

(a) Show that the state which emerges from magnet "1" $(S'_y = +\frac{\hbar}{2})$ is

$$| \overline{+} \rangle = \frac{1}{\sqrt{2}} (|+\rangle - i|-\rangle).$$

$$(S_z | \pm \rangle = \pm \frac{\hbar}{2} | \pm \rangle \text{ as usual}).$$

(b) Find an expression for the time-evolved state which emerges from magnet "2", given that $H = -\vec{\mu}\cdot\vec{B} = -\gamma BS_z$:

$$|\bar{+},t> = ?$$

(c) Find an expression for the intensity of the beam spot, I(B), as a function of the magnetic field of magnet "2". I'll give you a choice of three methods:

      <u>Method 1</u>:
$$|\Psi> = \frac{1}{2} (|\bar{+},t> + |\bar{+}>),$$
$$I(B) = I_0 |<\Psi|\Psi>|^2.$$

      <u>Method 2</u>:
$$I(B) = I_0 |<\bar{+},t|\bar{+}>|^2.$$

      <u>Method 3</u>:
$$I(B) = I_0 \underline{\text{always}}.$$

## Chapter 3:   Some One-Dimensional Solutions
##                to the Schrödinger Equation

The discussion in the last Chapter centered upon the case of a free particle in a single space dimension.   The Schrödinger equation can be written as

$$i\hbar\,\frac{\partial}{\partial t}\,|\psi,t> = H|\psi,t>, \qquad\qquad (1)$$

where

$$H = \frac{p_x^2}{2m}\,. \qquad\qquad (2)$$

It is natural to assume that more general forms for H are possible.   The form of (2), which is an operator statement, is very classical looking.   We hypothesize that the interaction of a quantum mechanical particle with an external potential V(x) can also be represented by it's classical form:

$$H = \frac{p_x^2}{2m} + V(x). \qquad\qquad (3)$$

The crucial thing that must be checked in writing down (3) is that the probability density interpretation given to $|\psi(x,t)|^2$ in (110) of the last Chapter, which was based on the existence of a conserved probability current, still holds.   You will provide this check in a problem.   In addition, the Schrödinger equation, which now reads

$$i\hbar \frac{\partial}{\partial t} \psi(x,t) = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x,t)}{\partial x^2} + V(x)\psi(x,t), \qquad (4)$$

is still separable in space and time, which implies the time evolution of the energy eigenvalue states is as in (185) of Chapter 2 and that $e^{-iHt/\hbar}$ is still the evolution operator.

Of course, not _every_ V(x) in (3) has a physical significance.   The energies of our system must be real, which implies that

$$<a'|H|a'> = E_{a'}<a'|a'> = E_{a'}, \qquad (5)$$

$$E_a^* = <a'|H|a'>^* = <a'|H|a'>^+ = <a'|H^+|a'>. \qquad (6)$$

Comparing (5) and (6) for any state a' implies that

$$H^+ = H \qquad (7)$$

For (3) this means we must have

$$V(x)^+ = V(x). \qquad (8)$$

That is, the potential operator must be Hermitian.

Since the Schrödinger equation (4) is separable, we can define a time-evolved energy eigenstate

$$|a',t> = e^{-iHt/\hbar}|a'> = |a'> e^{-iE_{a'}t/\hbar}, \qquad (9)$$

as we did in the last Chapter for the free particle. Completeness of the $|a'>$ (we assume a discrete form)

$$\sum_{a'} |a'><a'| = 1, \qquad (10)$$

then implies for any state $|\psi>$ that

$$|\psi> = \sum_{a'} |a'><a'|\psi>. \qquad (11)$$

When projected into position space, this becomes (identical to Eq$\underline{n}$ (189) of Chapter 2 at t=0)

$$\psi(x,0) = \sum_{a'} u_{a'}(x)C_{a'} \qquad (12)$$

where we have learned that $C_{a'} = <a'|\psi>$.  Since we know the time development of the $u_{a'}(x)$, we then have that

$$\psi(x,t) = \sum_{a'} u_{a'}(x)e^{-iE_{a'}t/\hbar}C_{a'} \qquad (13)$$

Eq$\underline{n}$ (13) indicates that the knowledge of the energy eigenfunctions $u_{a'}(x)$ and eigenvalues $E_{a'}$ provide a way of constructing all possible functions $\psi(x,t)$ that solve the Schrödinger equation.  For this reason, the solution to the time-independent Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right] u(x) = Eu(x), \qquad (14)$$

where we have included a potential $V(x)$, is of paramount importance in quantum mechanics.

   We will study the solution to (14) above in this Chapter for a simple set of potentials for which complete analytic solutions are possible.  The four problems we will study here will be:

3.4

    1.    The infinite square well

    2.    The finite potential barrier

    3.    The harmonic oscillator

    4.    The attractive Kronig-Penney model

We will continue to limit the discussion to a single dimension of space for now.

# 1.   The infinite square well

We will take the potential to be as follows.
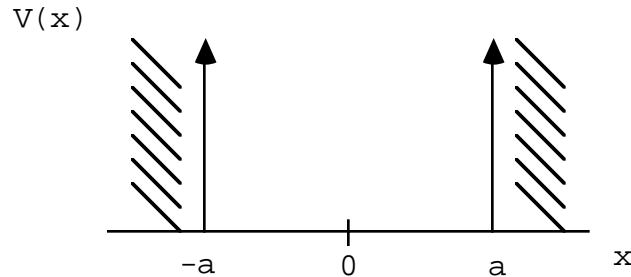


That is, we are assuming $V(x) = 0$ for $-a < x < a$, but $V(x) \rightarrow \infty$ for $|x| > a$.  A consistent way of interpreting this potential is to say that there is zero probability of the particle to escape from the interior region of the well into the shaded region.  This will be insured if we take the boundary conditions

$$u(x)\big|_{x = \pm a} = 0. \tag{15}$$

The easiest way to solve this problem is in the coordinate space representation of the wavefunction.  Thus, we need to solve

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} u(x) = Eu(x). \tag{16}$$

subject to the boundary conditions (15).  Eq$^{\underline{n}}$ (16) can be written as

$$-\frac{d^2}{dx^2} u(x) = k^2 u(x) \tag{17}$$

where (k is now a magnitude only)

$$k \equiv \left( \frac{2mE}{\hbar^2} \right)^{1/2} . \tag{18}$$

We have left the usual subscript off u(x) in anticipation of a labeling scheme for the energy eigenvalues. Of course, the linearly independent solutions to (17) are

$$u(x) = A \sin(kx) \tag{19}$$

or

$$u(x) = A' \cos(kx) . \tag{20}$$

If we apply the boundary conditions (15) to the solutions (19), we find that this means

$$\sin(\pm ka) = 0 \tag{21}$$

which implies that

$$ka = n\pi \tag{22}$$

for n = 1,2,3... .  n = 0 is a trivial solution and n = -1, -2,-3,... are not linearly independent.  Eq$^{\underline{n}}$ (22) tells us the allowed energy levels associated with the odd-space wavefunctions sin(kx) are discrete:

$$E_{n-} = \frac{1}{2m} \left( \frac{\hbar n\pi}{a} \right)^2 \tag{23}$$

(The "n-" notation means the n$^{\underline{th}}$ odd energy level). Likewise, for the even solutions, (20), we have

$$\cos(\pm ka) = 0 , \tag{24}$$

which means

$$ka = \left(n - \frac{1}{2}\right)\pi, \tag{25}$$

for n = 1,2,3,... (n = 0,-1,-2,-3,... are not linearly
independent).  The energies of the even-space wavefunctions
cos(kx) are thus

$$E_{n+} = \frac{1}{2m}\left(\frac{\hbar\,(n-1/2)\pi}{a}\right)^2. \tag{26}$$

Qualitative plots of the lowest few odd and even
wavefunctions are given below.

Odd solutions:



$$E_{2-} = \frac{1}{2m}\left(\frac{2\hbar\pi}{a}\right)^2$$

$$E_{1-} = \frac{1}{2m}\left(\frac{\hbar\pi}{a}\right)^2$$

Even solutions:



$$E_{3+} = \frac{1}{2m}\left(\frac{5\hbar\pi}{2a}\right)^2$$

$$E_{2+} = \frac{1}{2m}\left(\frac{3\hbar\pi}{2a}\right)^2$$

$$E_{1+} = \frac{1}{2m}\left(\frac{\hbar\pi}{2a}\right)^2$$
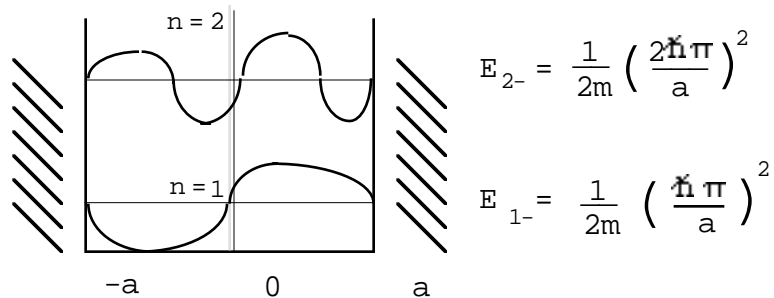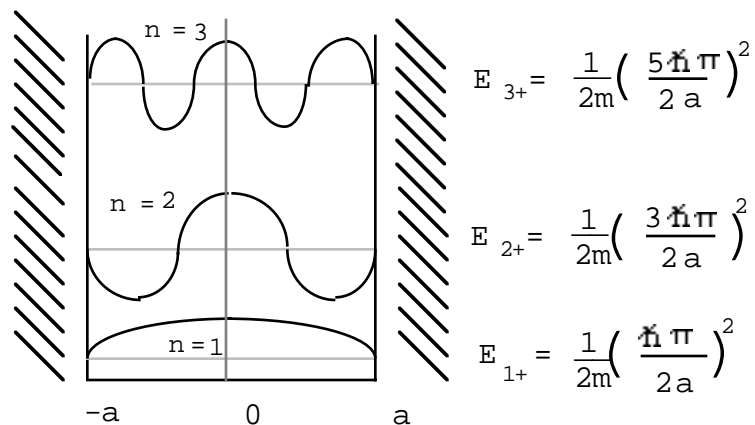
The lowest energy solution is given by $E_{1+}$.  If we didn't
know its exact value, we could guess it approximately from

the uncertainty principle, assuming it is a minimum
uncertainty state.  We have that

$$\Delta p_x \Delta x \approx \hbar , \qquad (27)$$

for such a state.  If we say that $\Delta x \approx 2a$ and that $p_x \approx \Delta p_x$,
then one obtains the estimate

$$E_{lowest} = \frac{p_x^2}{2m} \approx \frac{\hbar^2}{8ma^2} , \qquad (28)$$

which is to be compared with the actual value, $E_{1+} = \frac{\hbar^2 \pi^2}{8ma^2}$.

We now wish to normalize the solutions (19) and (20).
We use the notation

$$u_{n-}(x) = \langle x | n- \rangle = A \sin(k_{n-}x), \qquad (29)$$

$$u_{n+}(x) = \langle x | n+ \rangle = A' \cos(k_{n+}x), \qquad (30)$$

and set

$$1 = \langle n- | n- \rangle = \int_{-a}^{a} dx \, u_{n-}^*(x) u_{n-}(x,) \qquad (31)$$

and

$$1 = \langle n+ | n+ \rangle = \int_{-a}^{a} dx \, u_{n+}^*(x) u_{n+}(x), \qquad (32)$$

Doing the integral in (31) gives

$$\int_{-a}^{a} dx \, u_{n-}^*(x) u_{n-}(x) = |A|^2 a, \qquad (33)$$

and similarly

$$\int_{-a}^{a} dx \; u_{n+}^{*}(x) u_{n+}(x) \; = \; |A'|^{2}a. \tag{34}$$

The overall phase of wavefunctions is arbitrary, so we may choose A,A' as real and positive:

$$A, A' \; = \; \frac{1}{\sqrt{a}} \; . \tag{35}$$

One of the major tenants of quantum mechanics is orthogonality of states.  Verification in the case of the product $<n-|n'+>$ is easy:

$$<n-|n'+> \; = \; \frac{1}{a} \int_{-a}^{a} dx \; \sin(k_{n-}x) \; \cos(k_{n'+}x) \; = \; 0. \tag{36}$$

(The integral of an odd × even = odd function over an even interval is zero.)  The product $<n+|n'+>$ should also be zero.  We can see this as follows.  $(n \neq n')$

$$<n+|n'+> \; = \; \frac{1}{a} \int_{-a}^{a} dx \; \cos \, (k_{n+}x) \; \cos \, (k_{n'+}x)$$

$$= \; \frac{1}{a} \left\{ \left. \frac{\sin[(k_{n+} - k_{n'+})x]}{2(k_{n+} - k_{n'+})} \; \right|_{-a}^{a} + \left. \frac{\sin[(k_{n+} + k_{n'+})x]}{2(k_{n+} + k_{n'+})} \; \right|_{-a}^{a} \right\}$$

$$= \; 0 \; . \tag{37}$$

This comes about since

$$ak_{n+} \; = \; \left(n - \frac{1}{2}\right)\pi \; ,$$

$$\Rightarrow a(k_{n+} - k_{n'+}) \; = \; (n - n')\pi$$

$$\left. \Rightarrow a(k_{n+} + k_{n'+}) \; = \; (n + n' - 1)\pi, \right\} \tag{38}$$

Likewise, one can show that $<n-|n'-> = 0$ for $n \neq n'$. By defining a label P that takes on values ± (labeling

even/odd functions of x), we may summarize the statement of
orthonormality here by

$$\langle nP|n'P'\rangle = \delta_{nn'}\delta_{PP'}. \tag{39}$$

That the solution of the Schrödinger equation forms a
complete orthogonal set of functions is a theorem that can be
proven, so we are seeing here a special case of a very
general situation.  (We will continue to sharpen our
understanding of the mathematical meaning of completeness in
the next Chapter.).

The most general wavefunction consistent with the
boundary conditions can now be written as

$$|\psi\rangle = \sum_{n=1}^{\infty} [C_{n+}|n+\rangle + C_{n-}|n-\rangle], \tag{40}$$

where $C_{n+}$ and $C_{n-}$ are sets of constants. These constants are
not totally arbitrary since we must have

$$1 = \langle\psi|\psi\rangle, \tag{41}$$

which means that

$$\sum_{n=1}^{\infty} [|C_{n+}|^2 + |C_{n-}|^2] = 1. \tag{42}$$

Eq$^{\underline{n}}$ (42) suggest that $|C_{nP}|^2$ be interpreted as the probability
that the general state $|\psi\rangle$ is in the energy eigenstate $|nP\rangle$.
As pointed out below Eq$^{\underline{n}}$ (12) above, these constants are
given by $C_{a'} = \langle a'|\psi\rangle$, which in this specific case means that
(see also (174) of Chapter 2)

$$C_{nP} = <nP|\psi> = \int_{-a}^{a} dx \, u_{nP}^{*}(x)\psi(x). \qquad (43)$$

In an analogy with a 3-dimensional vector space, $C_{nP}$ is like the projection of an arbitrary vector on a given axis or direction. Our interpretation of $|C_{nP}|^2$ as a probability is indeed consistent with our spin discussion of the first Chapter, the main difference being that the number of spin states was finite ("up" and "down" only for spin $\frac{1}{2}$ ), whereas here the number of possible states, np, is infinite. (It is a "countable infinity" in mathematician's jargon.)

Now while the $|nP>$ have sharp energy eigenvalues, we should realize that $|\psi>$ does not. It is a coherent mixture of states with different energies. $|\psi>$ does, however, have a well-defined <u>average</u> energy, given by its expectation value:

$$<H>_{\psi} = <\psi|H|\psi> = \sum_{n'=1}^{\infty} [C_{n'+}^{*} <n'+| + C_{n-}^{*} <n'-|]$$

$$\cdot H \cdot \sum_{n=1}^{\infty} [C_{n+}|n+> + C_{n-}|n->], \qquad (44)$$

$$= \sum_{n=1}^{\infty} [E_{n+}|C_{n+}|^2 + E_{n-}|C_{n-}|^2]. \qquad (45)$$

Eq$^{\underline{n}}$ (45) is consistent with our interpretation of $|C_{nP}|^2$ as the probability that $|\psi>$ is in the state $|nP>$.

Given the energies $E_{nP}$, it is easy to write down the time evolved state $|\psi,t>$:

$$|\psi,t> = \sum_{n=1}^{\infty} [C_{n+}|n+,t> + C_{n-}|n-,t>], \qquad (46)$$

or

$$|\psi,t> = \sum_{n=1}^{\infty} [C_{n+}|n+>e^{-iE_{n+}t/\hbar} + C_{n-}|n->e^{-iE_{n-}t/\hbar}]. \quad (47)$$

Of course, we still have

$$<\psi,t|\psi,t> = \sum_{n'=1}^{\infty} [C_{n'+}e^{iE_{n'+}t/\hbar} <n'+| + C_{n'-}e^{iE_{n'-}t/\hbar} <n'-|]$$

$$\cdot \sum_{n=1}^{\infty} [C_{n+}e^{-iE_n t/\hbar} |n+> + C_{n-}e^{-iE_{n-}t/\hbar} |n'->]$$

$$= \sum_{n=1}^{\infty} [|C_{n+}|^2 + |C_{n-}|^2] = 1, \quad (48)$$

so that probability is conserved.

The quantity P in $|nP>$ is called "parity," It simply categorizes whether a wave function is even (P = +) or odd (P = -) under the substitution $x \rightarrow -x$. Let us define a parity operator by

$$\mathbb{P}|x> = |-x>. \quad (49)$$

This implies that

$$<x|\mathbb{P}^+ = <-x|, \quad (50)$$

We have that

$$<x|(\overset{\leftarrow}{\mathbb{P}^+}-\overset{\rightarrow}{\mathbb{P}})|x'> = <-x|x'> - <x|-x'>$$

$$= \delta(x + x') - \delta(x + x') = 0. \quad (51)$$

Since (51) is true for all $<x|, |x'>$, we have that

$$\mathbb{P}^+ = \mathbb{P} . \tag{52}$$

That is, the parity operator is Hermitian.  We then have that

$$< x | \mathbb{P} | nP > = \underbrace{< -x | nP >}_{u_{nP}(-x)} = P < x | nP > . \tag{53}$$

Since (53) is true for all $< x |$, we may remove it to reveal that

$$\mathbb{P} | nP > = P | nP > , \tag{54}$$

$$\Rightarrow \quad < nP | \mathbb{P} = P < nP | . \tag{55}$$

We thus learn that the $| nP >$ are eigenstates of $\mathbb{P}$ with eigenvalues P.  We also notice that

$$\mathbb{P} H | nP > = E_{nP} \mathbb{P} | nP > = P E_{nP} | nP > , \tag{56}$$

and

$$H \mathbb{P} | nP > = P H | nP > = P E_{nP} | nP > , \tag{57}$$

so that

$$[H, \mathbb{P}] | nP > = 0 . \tag{58}$$

Eqⁿ (58) being true for all states $| nP >$ then it means that

$$[H, \mathbb{P}] = 0 . \tag{59}$$

Thus, the Hamiltonian and the parity operator commute.  The reason that (59) is significant is because we will learn in the next Chapter that any operator which does not explicitly

depend on time and which commutes with the Hamiltonian has

expectation values which are a constant of the motion

(assuming we are working with a conservative system.)  Thus,

the parity, P, of a state $|nP>$ does not change with time.

   We have found the energies of the infinite square well

by solving a differential equation with given boundary

conditions.  As an illustration of techniques which do not

involve the solution of a differential equation, we will now

solve for the energies of this system using operator

techniques.  First, notice that we may write

$$\frac{1}{\sqrt{a}} \sin\left(\frac{n\pi x'}{a}\right) = \frac{1}{\sqrt{a}} \sin\left[\frac{(n+1)\pi x'}{a} - \frac{\pi x'}{a}\right]$$

$$= \frac{1}{\sqrt{a}}\left[\sin\left(\frac{(n+1)\pi x'}{a}\right)\cos\left(\frac{\pi x'}{a}\right) - \sin\left(\frac{\pi x'}{a}\right)\cos\left(\frac{(n+1)\pi x'}{a}\right)\right].$$

$$(60)$$

Also writing

$$\cos\left(\frac{(n+1)\pi x'}{a}\right) = \frac{a}{(n+1)\pi}\frac{d}{dx}\cdot\sin\left(\frac{(n+1)\pi x'}{a}\right), \quad (61)$$

we have that

$$\frac{1}{\sqrt{a}}\sin\left(\frac{n\pi x'}{a}\right) =$$

$$\left[\cos\left(\frac{\pi x'}{a}\right) - \sin\left(\frac{\pi x'}{a}\right)\frac{a}{(n+1)\pi}\frac{d}{dx'}\right]\frac{1}{\sqrt{a}}\sin\left(\frac{(n+1)\pi x'}{a}\right). \quad (62)$$

In terms of the $u_{n-}(x')$, this says that

$$u_{n-}(x') = \left[\cos\left(\frac{\pi x'}{a}\right) - \sin\left(\frac{\pi x'}{a}\right) \frac{a}{(n+1)\pi} \frac{d}{dx'}\right] u_{(n+1)-}(x'). \quad (63)$$

or

$$\langle x'|n-\rangle = \left[\cos\left(\frac{\pi x'}{a}\right) - \sin\left(\frac{\pi x'}{a}\right) \frac{a}{(n+1)\pi} \frac{d}{dx'}\right]\langle x'|(n+1)-\rangle.$$

$$(64)$$

Now using

$$\langle x'|x = x'\langle x'|, \quad (65)$$

and

$$\langle x'|p_x = \frac{\hbar}{i} \frac{\partial}{\partial x'} \langle x'|, \quad (66)$$

Eq$^{\underline{n}}$ (64) may be written as

$$\langle x'|n-\rangle = \langle x'|L_{n+1}|(n+1)-\rangle, \quad (67)$$

where

$$L_{n+1} = \cos\left(\frac{\pi x}{a}\right) - \sin\left(\frac{\pi x}{a}\right) \frac{a}{(n+1)\pi} \frac{ip_x}{\hbar} . \quad (68)$$

Let me emphasize that x and $p_x$ in (68) are operators, not numbers.  Since Eq$^{\underline{n}}$ (67) is true for all $\langle x'|$, we thus have that

$$|n-\rangle = L_{n+1}|(n+1)-\rangle. \quad (69)$$

Similarly, we have that

$$\frac{1}{\sqrt{a}} \sin\left(\frac{n\pi x'}{a}\right) = \frac{1}{\sqrt{a}} \sin\left(\frac{(n-1)\pi x'}{a} + \frac{\pi x'}{a}\right)$$

$$= \frac{1}{\sqrt{a}} \left[ \sin\left(\frac{(n-1)\pi x'}{a}\right) \cos\left(\frac{\pi x'}{a}\right) + \sin\left(\frac{\pi x'}{a}\right) \cos\left(\frac{(n-1)\pi x'}{a}\right) \right]$$

$$= \left[ \cos\frac{\pi x'}{a} + \frac{a}{(n-1)\pi} \sin\frac{\pi x'}{a} \frac{d}{dx'} \right] \frac{1}{\sqrt{a}} \sin\left(\frac{(n-1)\pi x'}{a}\right). \quad (70)$$

which says that

$$< x' |n-> = <x'|L_{n-1}^{+}|(n-1)->, \quad (71)$$

or

$$|n-> = L_{n-1}^{+}|(n-1)->, \quad (72)$$

where the adjoint of $L_{n-1}$ is the effective operator.  In a like manner, for the positive parity states, we can show that

$$|n+> = L_{n-3/2}^{+}|(n-1)+>, \quad (73)$$

$$|n+> = L_{n+1/2}|(n+1)+>. \quad (74)$$

Thus, given $|1+>$ or $|1->$, we can generate all the higher $n$ states by repeated applications of $L_n^{+}$.  For example, we have

$$|3-> = L_2^{+}L_1^{+}|1->, \quad (75)$$

and

$$|4+> = L_{5/2}^{+}L_{3/2}^{+}L_{1/2}^{+}|1+>. \quad (76)$$

Because we can raise or lower the $n$ value of these states by

repeated applications of the $L_n^+$ or $L_n$ (for various n values),
these are called "ladder" operators.

Let us say we know the ladder operators and their effect
on states, but suppose we didn't know the energy eigenvalues
of the system. It is possible to find the energies of the
system as follows. Let us evaluate the quantity

$$[H, L_n^+]|n->.$$

This will give us the energies of the negative parity states.
To see this, it will necessary to do some operator algebra,
which should provide us good practice. We shall have to deal
with the commutator

$$[H, L_n^+] = \left[ \frac{p_x^2}{2m}, \ \cos\left(\frac{\pi x}{a}\right) + \frac{a}{n\pi} \sin\left(\frac{\pi x}{a}\right) \frac{ip_x}{\hbar} \right]. \qquad (77)$$

Before considering (77), let us work out some simpler things.
First, consider the quantity

$$<x'|[p_x, f(x)] = <x'|(p_x f(x) - f(x)p_x). \qquad (78)$$

We then find that

$$<x'|[p_x, f(x)] = \frac{\hbar}{i} \frac{\partial}{\partial x'} (<x'|)f(x) - f(x')<x'|p_x$$

$$= \frac{\hbar}{i} f(x') \frac{d}{dx'} <x'| + \frac{\hbar}{i} \frac{df(x')}{dx'} <x'| - \frac{\hbar}{i} f(x') \frac{d}{dx'} <x'|$$

$$= \frac{\hbar}{i} \frac{df(x')}{dx'} <x'| = \frac{\hbar}{i} <x'| \frac{df(x)}{dx}. \qquad (79)$$

3.18

Being true for all $<x'|$, Eq$^{\underline{n}}$ (79) implies that

$$[p_x, f(x)] = \frac{\hbar}{i} \frac{df(x)}{dx} . \qquad (80)$$

Likewise, one may show that

$$[x, f(p_x)] = -\frac{\hbar}{i} \frac{df(p_x)}{dp_x} . \qquad (81)$$

Using (80) we have

$$\left[ p_x, \cos\left(\frac{\pi x}{a}\right) \right] = i\hbar \frac{\pi}{a} \sin\left(\frac{\pi x}{a}\right) , \qquad (82)$$

and

$$\left[ p_x, \sin\left(\frac{\pi x}{a}\right) \right] = -i\hbar \frac{\pi}{a} \cos\left(\frac{\pi x}{a}\right) , \qquad (83)$$

for example. Also useful are the following commutator identities:

$$[A,B] = -[B,A], \qquad (84a)$$

$$[A + B,C] = [A,C] + [B,C], \qquad (84b)$$

$$[AB,C] = A[B,C] + [A,C]B. \qquad (84c)$$

Using (84c) we can now write

$$[H, L_n^+] = \frac{p_x}{2m} \left[ p_x, \cos\left(\frac{\pi x}{a}\right) + \sin\left(\frac{\pi x}{a}\right) \frac{a}{n\pi} \frac{ip_x}{\hbar} \right]$$

$$+ \left[ p_x, \cos\left(\frac{\pi x}{a}\right) + \sin\left(\frac{\pi x}{a}\right) \frac{a}{n\pi} \frac{ip_x}{\hbar} \right] \frac{p_x}{2m}. \qquad (85)$$

Using (82) and (83) this becomes

$$[H,L_n^+] = \frac{p_x}{2m}\left(i\hbar\,\frac{\pi}{a}\,\sin\left(\frac{\pi x}{a}\right) + \frac{1}{n}\,\cos\left(\frac{\pi x}{a}\right)p_x\right)$$

$$+ \left(i\hbar\,\frac{\pi}{a}\,\sin\left(\frac{\pi x}{a}\right) + \frac{1}{n}\,\cos\left(\frac{\pi x}{a}\right)p_x\right)\frac{p_x}{2m}. \quad (86)$$

We want to try to combine various types of terms together. In order to do this, let us try to move all the $p_x$ operators in (86) to the right of the factors involving x. We must be careful in doing this because x and $p_x$ do not commute. We have that

$$p_x\,\sin\left(\frac{\pi x}{a}\right) = \sin\left(\frac{\pi x}{a}\right)p_x + \left[p_x,\sin\left(\frac{\pi x}{a}\right)\right]$$

$$= \sin\left(\frac{\pi x}{a}\right)p_x - i\hbar\,\frac{\pi}{a}\,\cos\left(\frac{\pi x}{a}\right), \quad (87)$$

$$p_x\,\cos\left(\frac{\pi x}{a}\right) = \cos\left(\frac{\pi x}{a}\right)p_x + \left[p_x,\cos\left(\frac{\pi x}{a}\right)\right]$$

$$= \cos\left(\frac{\pi x}{a}\right)p_x + i\hbar\,\frac{\pi}{a}\,\sin\left(\frac{\pi x}{a}\right), \quad (88)$$

Therefore, (86) can be written

$$[H,L_n^+] = 2\cdot\frac{i\hbar}{2m}\,\frac{\pi}{a}\,\sin\left(\frac{\pi x}{a}\right)p_x + \frac{\hbar^2}{2m}\left(\frac{\pi}{a}\right)^2\cos\left(\frac{\pi x}{a}\right)$$

$$+ 2\cdot\frac{1}{2mn}\,\cos\left(\frac{\pi x}{a}\right)p_x^2 + \frac{i\hbar}{2mn}\left(\frac{\pi}{a}\right)\sin\left(\frac{\pi x}{a}\right)p_x. \quad (89)$$

This gives

$$[H, L_n^+]|n-> = \left\{ \frac{\hbar^2}{2m} \left(\frac{\pi}{a}\right)^2 \cos \left(\frac{\pi x}{a}\right) \right.$$

$$\left. + \frac{\hbar^2}{m} \left(\frac{\pi}{a}\right) \left(1 + \frac{1}{2n}\right) \sin \left(\frac{\pi x}{a}\right) \frac{ip_x}{\hbar} + \frac{1}{mn} \cos \left(\frac{\pi x}{a}\right) p_x^2 \right\}|n->. \quad (90)$$

In the last term in Eqⁿ (90), we may make the replacement

$$p_x^2 |n-> = 2mE_{n-}|n-> . \quad (91)$$

As for the other terms in (90), the following consideration will be of use.  We know that

$$|(n+1)-> = L_n^+|n-> , \quad (92)$$

and

$$|(n-1)-> = L_n|n-> , \quad (93)$$

or more explicitly

$$|(n+1)-> = \left[\cos \left(\frac{\pi x}{a}\right) + \frac{a}{n\pi} \sin \left(\frac{\pi x}{a}\right) \frac{ip_x}{\hbar}\right]|n->, \quad (94)$$

$$|(n-1)-> = \left[\cos \left(\frac{\pi x}{a}\right) - \frac{a}{n\pi} \sin \left(\frac{\pi x}{a}\right) \frac{ip_x}{\hbar}\right]|n->. \quad (95)$$

Subtracting (95) from (94) and multiplying by $\left(\frac{n\pi}{2a}\right)$ gives

$$\left(\frac{n\pi}{2a}\right) [|(n+1)-> - |(n-1)->] = \sin \left(\frac{\pi x}{a}\right) \frac{ip_x}{\hbar} |n->. \quad (96)$$

Adding (94) and (95) gives

$$\frac{1}{2} [|(n+1)-> + |(n-1)->] = \cos \left(\frac{\pi x}{a}\right) |n-> . \quad (97)$$

Substituting (96) and (97) in (90) now gives

$$
[H, L_n^+]|n-> = \left\{ \frac{\hbar^2}{2m} \left(\frac{\pi}{a}\right)^2 \frac{1}{2} [|(n+1)-> + |(n-1)->] \right.
$$

$$
+ \frac{\hbar^2}{2m} \left(\frac{\pi}{a}\right)^2 \left(n + \frac{1}{2}\right) [|(n+1)-> - |(n-1)->]
$$

$$
\left. + \frac{1}{n} E_{n-}[|(n+1)-> + |(n-1)->] \right\}. \quad (98)
$$

Now, multiply on the left by $<(n-1)-|$. On the left hand side
of (98) we have

$$
<(n-1)-|[H, L_n^+]|n-> = <(n-1)-|HL_n^+ - L_n^+H)|n->
$$

$$
= E_{(n+1)-} <(n-1)-|(n+1)-> - E_{n-} <(n-1)-|(n+1)->
$$

$$
= 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad (99)
$$

Explicitly evaluating the right hand side now tells us that

$$
0 = \frac{\hbar^2}{2m} \left(\frac{\pi}{a}\right)^2 \frac{1}{2} - \frac{\hbar^2}{2m} \left(\frac{\pi}{a}\right)^2 \left(n + \frac{1}{2}\right) + \frac{1}{n} E_{n-}
$$

$$
\Rightarrow E_{n-} = \frac{1}{2m} \left(\frac{\hbar\pi n}{a}\right)^2 \qquad\qquad\qquad (100)
$$

Thus, we have evaluated the energies of the negative parity
states simply from a knowledge of the properties of the
ladder operators. We may similarly find the energies of the
positive parity states from the ladder operators.

Our evaluation of the negative parity energies above has
mainly been an exercise in the use of operator methods. The

initial solution for the energies using the coordinate space
Schrödinger equation was much easier in fact.  We will soon
look at a problem, the harmonic oscillator, where the
opposite is true.  That is, the solution of the coordinate
space differential equation is much more difficult to do than
the solution of the problem using operator methods.  We will
only use operator methods there because of this.

## 2. **The finite potential barrier**

The next potential we consider is shown below.
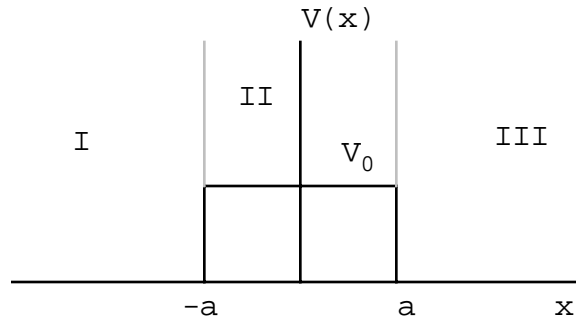


We have broken the x space into three regions in which we will separately solve the Schrödinger equation.

In regions I and III, we simply have the free space time independent Schrödinger equation to solve. The general solution to

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} = Eu(x), \tag{101}$$

can be written as

$$u_I(x) = Ae^{ik_1x} + Be^{-ik_1x}, \tag{102}$$

$$u_{III}(x) = Ee^{ik_1x} + Fe^{-ik_1x}, \tag{103}$$

where $k_1 = \frac{\sqrt{2mE}}{\hbar}$. We saw in Chapter 2 that $e^{i(kx - wt)}$ represents a wave with momentum $p_x = k\hbar$ traveling in the +x direction. Therefore, we interpret the coefficient A in (102) as the amplitude of a plane wave incident from the left. (There is no time dependence in (102) and (103)

because we are interested in time independent or stationary

solutions in this Chapter.)  We will take as a boundary

condition the fact that the incident waves come from the

left. However, it would be wrong to conclude from this that

we could choose B = 0 here because we expect that there will

be _reflected_ waves from the potential steps at x = ±a.

Because we are choosing only waves incident from the left we

_must_ choose F = 0 in region III.

In region II we must solve

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} = (E - V_0)u(x) \qquad (104)$$

The solution of (104) depends on whether $E > V_0$ or $E < V_0$.

When $E > V_0$, we have

$$u_{II}(x) = Ce^{ik_2x} + De^{-ik_2x} , \qquad (105)$$

where $k_2 = \dfrac{\sqrt{2m(E - V_0)}}{\hbar}$ .  Our solution so far consists of

(102), (103) (with F = 0), and (105). How are we to find the

five coefficients A,B,C,D, and E?  First, it should be clear

that one condition on these coefficients is an overall

normalization condition.  In the case of a continuum

distribution of energy values, as it should be evident is the

cases here (since there is no possibility of discrete bound

states), we can use the condition

$$< E' | E'' > = \delta(E' - E''), \qquad (106)$$

as applied to our right-moving waves.  (For another example
of delta-function normalization, see Eq$^n$ (142) of Chapter 2.)
Notice that (106) leads to

$$< E' | E" > = \int_{-\infty}^{\infty} dx u_{E'}^{*}(x) u_{E"}(x) = \delta(E' - E"),$$

in contradistinction to the condition of Eq$^n$ (177) of Chapter
2, which is appropriate to <u>discrete</u> energy levels.  Thus, we
really have only to determine 4 out of these 5 unknown
coefficients.  It will be convenient therefore to solve only
for the 4 ratios

$$\frac{B}{A}, \frac{C}{A}, \frac{D}{A}, \text{ and } \frac{E}{A}.$$

In order to do this, we must bring out an underlying
requirement of solutions of the Schrödinger equation.  To see
this requirement, let us integrate the time independent
Schrödinger equation over an infinitesimal region surrounding
a point, x.

$$\int_{x_0-\varepsilon}^{x_0+\varepsilon} dx \left[ -\frac{\hbar^2}{2m}\frac{d^2u}{dx^2} + V(x)u(x) = Eu(x) \right]. \quad (107)$$

Assuming a piecewise continuous potential, V(x), this says
that, in the limit that $\varepsilon \to 0$

$$\frac{du}{dx}\bigg|_{+} = \frac{du}{dx}\bigg|_{-}, \quad (108)$$

or, in words, that the first derivative of u(x) evaluated in
a limiting sense from points on the left or right of $x_0$, is
continuous.  This eliminates something that looks like
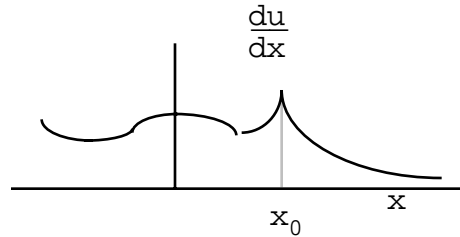
Integrating (108) again tells us that u(x) is also
continuous. If a discontinuity in u'(x) <u>were</u> allowed, one
can show that we would loose our interpretation of $|u(x)|^2$ as
a probability density. The only exceptions to having a
continuous u'(x) come when non-stepwise continuous
potentials are considered. We have already seen an example
of this in the infinite square well where, in fact, we have a
discontinuity in u'(x) at the edges of the well, x = ta.
Another example of a non-stepwise continuous potential where
a discontinuity in u'(x) is allowed is a delta-function
potential. We will study such a situation shortly in the
attractive Kronig-Penny model.

For our problem, we must require the continuity of our
wavefunctions at <u>all</u> positions. In particular, this means
our wavefunctions and their first derivatives must be
continuous in the neighborhood of the joining positions of
regions I, II and III. This gives us four conditions on our
coefficients, which is just enough to determine the four
unknown ratios written down above. At x = -a, we find that

$$Ae^{-ik_1a} + Be^{ik_1a} = Ce^{-ik_2a} + De^{ik_2a} \, , \qquad (109)$$

from continuity of u(x), and

$$Ak_1e^{-ik_1a} - Bk_1e^{ik_1a} = k_2Ce^{-ik_2a} - k_2De^{ik_2a} \ , \tag{110}$$

from continuity of u'(x).  At x = a, we get the equations

$$Ce^{ik_2a} + De^{-ik_2a} = Ee^{ik_1a}, \tag{111}$$

and

$$k_2Ce^{ik_2a} - k_2De^{-ik_2a} = k_1Ee^{ik_1a}, \tag{112}$$

from continuity of u(x) and u'(x).  We now define the
transmission and reflection coefficients

$$T = \left| \frac{E}{A} \right|^2 \ . \tag{113}$$

$$R = \left| \frac{B}{A} \right|^2 \ . \tag{114}$$

We use the absolute squares of the ratios of amplitudes in
order that the results be interpretable as probabilities.
Note that

$$T + R = 1, \tag{115}$$

as one must have if probabilities are conserved.  Solving
(109) – (112) gives us

$$\frac{B}{A} = \frac{\dfrac{i}{2}\left(\dfrac{k_2}{k_1} - \dfrac{k_1}{k_2}\right) e^{-2ik_1a} \sin(2k_2a)}{\cos(2k_2a) - \dfrac{i}{2}\left(\dfrac{k_1}{k_2} + \dfrac{k_2}{k_1}\right) \sin(2k_2a)} \ , \tag{116}$$

$$\frac{C}{A} = \frac{\frac{1}{2}\left(1 + \frac{k_1}{k_2}\right) e^{-ik_2 a} \, e^{-ik_1 a}}{\cos(2k_2 a) - \frac{i}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) \sin(2k_2 a)} \quad , \tag{117}$$

$$\frac{D}{A} = \frac{\frac{1}{2}\left(1 - \frac{k_1}{k_2}\right) e^{-ik_1 a} \, e^{-ik_2 a}}{\cos(2k_2 a) - \frac{i}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) \sin(2k_2 a)} \quad , \tag{118}$$

$$\frac{E}{A} = \frac{e^{-2ik_1 a}}{\cos(2k_2 a) - \frac{i}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) \sin(2k_2 a)} \quad . \tag{119}$$

The transmission coefficients T and R, from (119) and (116) respectively, are then found to be

$$T = \frac{1}{\cos^2(2k_2 a) + \frac{1}{4}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right)^2 \sin^2(2k_2 a)} \quad . \tag{120}$$

and

$$R = \frac{\frac{1}{4}\left(\frac{k_2}{k_1} - \frac{k_1}{k_2}\right)^2 \sin^2(2k_2 a)}{\cos^2(2k_2 a) + \frac{1}{4}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right)^2 \sin^2(2k_2 a)} \quad . \tag{121}$$

Thus, even though in the case we are studying the plane wave has an energy in excess of $V_0$, the classical energy necessary to overcome the potential barrier, there is in general a

nonzero probability that a particle will reflect from the barrier.

Let us examine the solution in the case $E < V_0$ now. The wave function in regions I and III are as before, but now the general solution in region II is

$$u_{II}(x) = Ce^{-Kx} + De^{Kx} \quad , \qquad (122)$$

where $K = \dfrac{\sqrt{2m(V_0 - E)}}{\hbar}$. We notice that the only difference now is the fact that we are working with real exponentials. Therefore, rather than re-working out T and R from the start, it is only necessary to make the substitution $k_2 \rightarrow + iK$ everywhere. We find in the $E < V_0$ case that

$$T = \cfrac{1}{\cosh^2(2Ka) - \dfrac{1}{4}\left(\dfrac{K}{k_1} - \dfrac{k_1}{K}\right)^2 \sinh^2(2Ka)} \quad , \qquad (123)$$

and

$$R = \cfrac{\dfrac{1}{4}\left(\dfrac{K}{k_1} + \dfrac{k_1}{K}\right) \sinh^2(2Ka)}{\cosh^2(2Ka) + \dfrac{1}{4}\left(\dfrac{K}{k_1} - \dfrac{k_1}{K}\right)^2 \sinh^2(2Ka)} \quad . \qquad (124)$$

Putting (120) and (123) together, we find the following qualitative result for T(E):
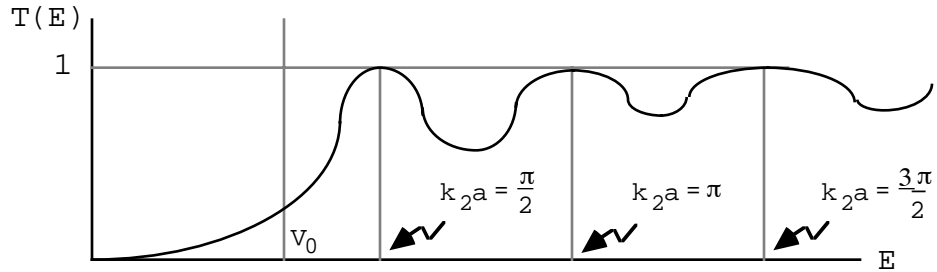
There are several aspects to remark upon on this graph.
The most obvious thing to observe is that T ≠ 0 even when
E < $V_0$.  Classically, such a thing could never happen.  That
is, if we had a classical particle with an energy E < $V_0$,
there would be zero probability that the particle would be
able to overcome the potential barrier.  It would be
prevented by energy conservation.  However, in quantum
mechanics there is an uncertainty relation for energy and
time similar to that for momentum and position.  In order to
understand its interpretation, let us go back for a moment
and consider the Gaussian free wave function in the case
$<p_x> = \bar{p} = 0$ (This will not limit the generality of our
conclusions.)  We know that this wave function spreads in
time according to Eqⁿ (88) of Chapter 2:

$$\delta x(t)^2 = \delta x^2 + \left(\frac{\hbar t}{2m\delta x}\right)^2 \; .$$

Therefore, the time it takes for the wavefunction to evolve
into a considerably wider form is when

$$\frac{\hbar t_c}{2m\delta x} \approx \delta x$$

or

$$t_c \approx \frac{2m\delta x^2}{\hbar} \tag{125}$$

It is easy to show for the Gaussian that its uncertainty in energy (using the definition in one of the problems) is just

$$\Delta E = \frac{\hbar^2}{4\sqrt{2}\,m\delta x^2} \quad , \tag{126}$$

which, with (125), can be written as

$$\Delta E t_c \approx \hbar \quad . \tag{127}$$

Thus, if $t_c$ is large, the uncertainty in energy of the Gaussian is quite small, and vice versa. Eq$\underline{n}$ (127) is very similar to the Heisenberg uncertainty principle $\Delta p_x \Delta x \geq \frac{\hbar}{2}$ . However, whereas we will see in Chapter 4 that the Heisenberg relation can be <u>derived</u> since $p_x$ and x are both operator quantities, we will not be able to do the same for (127) since the time, t, is a parameter, not an operator, in nonrelativistic quantum mechanics. (In relativistic theories, both x and t are just parameters.) Therefore, the energy-time relation (127) is inferred rather than derived. This does not mean it is any less applicable to the real world, however. Its meaning is completely general if $t_c$ is interpreted as a <u>correlation time</u> between wavefunctions during some transition that takes place. In the case of a particle traversing a potential barrier, this implies that there will be an uncertainty in the energy of the particle during the time it takes to complete its traversal. Thus, some components of the wavefunction <u>will</u> have sufficient

energy to overcome the barrier.  This phenomenon is known as
tunneling.

The other remarkable thing about the graph of T(E) is
the fact that there is complete transmission (T = 1) when the
condition

$$2k_2a = n\pi \ , \ n = 1,2,3... \tag{128}$$

is fulfilled.  That (128) leads to maxima in T(E) vs. E is
quite easy to understand.  The path difference between the
incident wave and an internally reflected wave which has
transversed the barrier and back is 4a.  These two waves will
interfere constructively whenever the path difference is an
integral multiple of the wavelength in region II.
Alternatively, the path difference of 4a leads to destructive
interference between the wave reflected at x = -a and a
reflected wave from x = a, given that the reflected wave at
x = a undergoes a phase shift of $\pi$.  However, this simple
argument does not tell us that T = 1 at these positions,
simply that we should expect local maxima there.

## 3.   The harmonic oscillator

One of the most important problems in classical mechanics is the simple harmonic oscillator, described by the equation of motion,

$$m\ddot{x} + kx = 0. \tag{129}$$

If we integrate this equation, we get an equation for the energy of the system:

$$E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2 \ (= \text{constant}). \tag{130}$$

Let us study the same problem in quantum mechanics. That is, we will take

$$H = \frac{p_x^2}{2m} + \frac{1}{2}kx^2. \tag{131}$$

where this is now an operator equation.  In order to simplify life, let us introduce some new variables that will eliminate the quantities m and k from appearing explicitly in our equations.  First, define

$$\omega = \sqrt{\frac{k}{m}}, \tag{132}$$

and rewrite (131) as (we have divided by $\frac{1}{\hbar\omega}$ )

$$\frac{H}{\hbar\omega} = \frac{p_x^2}{2m\hbar\omega} + \frac{m\omega}{2\hbar}x^2. \tag{133}$$

Now define the dimensionless variables

$$H = \frac{H}{\hbar\omega} \quad , \quad p = \frac{p_x}{\sqrt{m\omega\hbar}} \quad , \quad q = \sqrt{\frac{m\omega}{\hbar}} \; x \; . \qquad (134)$$

Then the eigenvalue problem we wish to solve may be written as

$$H \;|n> \; = \; \mathcal{E}_n|n> \; , \qquad (135)$$

$$H = \frac{1}{2} \; (p^2 + q^2) \; , \qquad (136)$$

where $n = 0,1,2,3,...$ is (initially) just a labeling of the expected discrete energy states of this system and $\mathcal{E}_n = E_n/\hbar\omega$. Thus, we let $n = 0$ label the ground state, $n = 1$ picks out the first excited state, and so on.

We can easily verify that

$$[q,p] = i, \qquad (137)$$

for these new variables, and that

$$\left( \frac{q + ip}{\sqrt{2}} \right) \left( \frac{q - ip}{\sqrt{2}} \right) = \frac{q^2 + p^2}{2} + \frac{i}{2} \; [p,q], \qquad (138)$$

$$\Rightarrow \; H = \left( \frac{q + ip}{\sqrt{2}} \right) \left( \frac{q - ip}{\sqrt{2}} \right) - \frac{1}{2} \; . \qquad (139)$$

Let us define

$$A = \frac{q + ip}{\sqrt{2}} \; , \qquad (140)$$

which, since q and p are Hermitian, means that

$$A^{+} = \frac{q - ip}{\sqrt{2}} , \qquad (141)$$

Therefore, we may write

$$H = AA^{+} - \frac{1}{2} . \qquad (142)$$

Now let us consider the commutator $[H,A]$. We have that

$$[H,A] = \frac{1}{2\sqrt{2}} [q^2, q + ip] + \frac{1}{2\sqrt{2}} [p^2, q + ip]$$

$$= \frac{i}{2\sqrt{2}} [q^2, p] + \frac{1}{2\sqrt{2}} [p^2, q]$$

$$= \frac{i}{2\sqrt{2}} \{q[q,p] + [q,p]q\} + \frac{1}{2\sqrt{2}} \{p[p,q] + [p,q]p\}$$

$$= - \frac{(q + ip)}{\sqrt{2}} . \qquad (143)$$

or

$$[H,A] = -A. \qquad (144)$$

Now notice

$$[H,A]|n> = -A|n> . \qquad (145)$$

so

$$H A|n> - A H |n> = -A|n> \qquad (146)$$
$$\underbrace{\qquad} $$
$$\mathcal{E}_n|n>$$

and

$$H (A|n>) = (\mathcal{E}_n - 1)(A|n>). \qquad (147)$$

Therefore $A|n>$ is also an eigenstate of $H$, but with a lower value of energy than the state $|n>$. Let us assume this is just the next lowest state, $|n-1>$, outside of an unknown multiplicative constant

$$A|n> = C_n|n-1>, \tag{148}$$

for $n = 1,2,3,\ldots$ This implies that the energy levels of the system are all equally spaced. We can always choose the $C_n$ in (148) to be real and positive by associating any phase factor which arises with the state $|n - 1>$. Eq$^{\underline{n}}$ (148) then implies

$$<n|A^+ = C_n<n-1|, \tag{149}$$

from taking the adjoint of both sides.

The adjoint of (144) is

$$[\mathsf{H},A^+] = A^+ . \tag{150}$$

Eq$^{\underline{n}}$ (150) then implies, the the same way we derived (147), that

$$\mathsf{H}(A^+|n>) = (\mathcal{E}_n + 1)(A^+|n>). \tag{151}$$

Therefore, $A^+|n>$ is an eigenstate of $\mathsf{H}$ with the next highest energy to $|n>$ since we know the energy levels are equally spaced:

$$A^+|n> = C_n'|n+1> . \tag{152}$$

There is no immediate reason why the $C_n'$ should be real since in picking $C_n$ real, we defined the phase of the states. The operators $A$ and $A^+$ are seen to be ladder operators for the states, but unlike the ladder operators for the infinite square well, there is no dependence of the $A$ or $A^+$ on the state label, n.

Now notice that

$$H \, |n> \; = \; (AA^{+} - \tfrac{1}{2} \, ) \, |n> \; = \; C_n' A \, |n + 1> \; - \; \tfrac{1}{2} \, |n>$$

$$= \; (C_n' C_{n+1} - \tfrac{1}{2} \, ) \, |n> \; = \; \mathcal{E}_n \, |n>. \tag{153}$$

Since $H^{+} = H$, we have that the energies of the system must be real (Remember the argument above). Therefore, since we chose the $C_n$ real, the $C_n'$ must also be real from the above.

We are assuming in this discussion that this system has a lowest energy state, which, because of the Heisenberg uncertainty principle has a positive value. (We estimated its value in a problem.) That is, we can only consistently maintain that $|0>$ is the state of lowest energy if we take

$$A \, |0> \; = \; 0, \tag{154}$$

which means, from (148), that we must take

$$C_0 \; = \; 0. \tag{155}$$

Eq$^{\underline{n}}$ (155) will be useful in a moment.

To complete this argument let us consider the quantity

$$Q \; \equiv \; <0 \, | A^n (A^{+})^n | \, 0>. \tag{156}$$

By allowing the $A^n (A^{+})^n$ operators to act one at a time to the right, we learn that

$$Q \; = \; \underbrace{<0 \, | 0>}_{1} \; C_1 \; \cdots \; C_n C_{n-1}' \; \cdots \; C_0' \, , \tag{157}$$

or

$$Q = (C_1 C_0')(C_2 C_1') \ldots (C_n C_{n-1}') . \tag{158}$$

Now, by allowing $A^n$ to act to the left while $(A^+)^n$ is still acting to the right, reveals that $Q$ may also be written as

$$Q = C_0' \ldots C_{n-1}' \underbrace{<n|n>}_{1} C_{n-1}' \ldots C_0' . \tag{159}$$

In writing down (159), we are using the fact that the $C_n'$ are real, plus the adjoint of Eq$\underline{^n}$ (152), which tells us that

$$<n|A = C_n' <n + 1| . \tag{160}$$

Thus, (159) gives us the alternate evaluation:

$$Q = (C_0')^2 (C_1')^2 \ldots (C_{n-1}')^2 . \tag{161}$$

Comparing (158) and (161) in the case of $n = 1$ tells us that

$$C_1 = C_0' . \tag{162}$$

Therefore, in the case $n = 2$, we conclude that

$$C_2 = C_1' . \tag{163}$$

By induction we may show then in general that

$$C_n = C_{n-1}' . \tag{164}$$

From the equal spacing of the energy levels

$$\mathcal{E}_{n-1} = \mathcal{E}_n - 1, \tag{165}$$

we then have, from (153) and (164) that

$$C_{n+1}^2 = C_n^2 + 1. \tag{166}$$

Using (155) in (166) now tells us that

$$C_n^2 = n . \tag{167}$$

Remember, we choose $C_n$ to be positive so that

$$C_n = \sqrt{n} . \tag{168}$$

The dimensionless energies of the system are now given as

$$\mathcal{E}_n = C_{n+1}^2 - \frac{1}{2} = n + \frac{1}{2} . \tag{169}$$

Putting the dimensions back in, we have

$$E_n = \hbar\omega\mathcal{E}_n = \hbar\omega(n + \frac{1}{2}). \tag{170}$$

Notice that the lowest energy is nonzero, $E_0 = \frac{1}{2} \hbar\omega$. As already pointed out, this is a consequence of the Heisenberg uncertainty principle for momentum and position. This lowest energy of the simple harmonic oscillator is called its <u>zero point energy</u>. Actually, the zero point energy is not observable, since the energy of a system is arbitrary up to an additive constant. However, changes in the zero-point energy are uniquely defined and can be observed in laboratory experiments.

The next thing we do will be to get explicit expressions for the coordinate space wave functions, $<q'|n>$. First of all, we know that

$$A^+|n-1> = \sqrt{n} \ |n>, \tag{171}$$

so that we may write

$$|n> = \frac{1}{\sqrt{n}} A^+|n-1> = \frac{1}{\sqrt{n}} \frac{1}{\sqrt{n-1}} (A^+)^2|n-2>$$

$$= \ldots = \frac{1}{\sqrt{n(n-1)\ldots 1}} (A^+)^n|0>. \qquad (172)$$

Therefore, we may write the energy eigenkets as

$$|n> = \frac{(A^+)^n}{\sqrt{n!}} |0> , \qquad (173)$$

Since

$$(A^+)^n = \left( \frac{q-ip}{\sqrt{2}} \right)^n , \qquad (174)$$

Eq$^{\underline{n}}$ (173) becomes

$$|n> = \frac{1}{\sqrt{2^n n!}} (q-ip)^n|0> . \qquad (175)$$

Our coordinate space wavefunctions are then

$$u_n(q') = <q'|n> = \frac{1}{\sqrt{2^n n!}} <q'|(q-ip)^n|0>, \qquad (176)$$

where the $<q'|$ state is just a relabeling of the state $<x'|$. Using our previous results from Chapter 2, we have that

$$<q'|q = q'<q'|, \qquad (177)$$

and

$$<q'|p = \frac{1}{i} \frac{\partial}{\partial q'} <q'| . \qquad (178)$$

We may now use

$$<q'|(q-ip)^n|0> = <q'|(q-ip)(q-ip)^{n-1}|0>$$

$$= \left(q' - \frac{d}{dq'}\right) < q' | (q - ip)^{n-1} | 0 >$$

$$= \left(q' - \frac{\partial}{\partial q'}\right)^2 < q' | (q - ip)^{n-2} | 0 >$$

$$= \ldots \left(q' - \frac{d}{dq'}\right)^n < q' | 0 >, \tag{179}$$

in Eq$\underline{n}$ (176). Therefore

$$u_n(q') = \frac{1}{\sqrt{2^n n!}} \left(q' - \frac{d}{dq'}\right)^n < q' | 0 >. \tag{180}$$

The question now is: What is the ground state wavefunction $< q' | 0 >$? If we can find this, then (180) gives all the rest of the wavefunctions. We may find this wavefunction by solving a differential equation. We know that $A | 0 > = 0$, but

$$< q' | A | 0 > = \frac{1}{\sqrt{2}} \left(q' + \frac{d}{dq'}\right) < q' | 0 >, \tag{181}$$

so

$$\left(q' + \frac{d}{dq'}\right) < q' | 0 > = 0. \tag{182}$$

The solution to (182)

$$< q' | 0 > = C e^{-q'^2/2}, \tag{183}$$

where C is an unknown constant. Normalizing this to unity, we find

$$|C|^2 \int_{-\infty}^{\infty} dq' \, e^{-q'^2} = |C|^2 \sqrt{\pi} = 1, \tag{184}$$

so we may choose

$$C = \frac{1}{\pi^{1/4}} . \tag{185}$$

Therefore

$$u_0(q') = <q'|0> = \frac{1}{\pi^{1/4}} e^{-q'^2/2}, \tag{186}$$

and we have that

$$u_n(q') = \frac{1}{\sqrt{\sqrt{\pi} \; 2^n n!}} \left(q' - \frac{d}{dq'}\right)^n e^{-q'^2/2}. \tag{187}$$

Let us simplify the result (187) a bit in order to connect up with some standard results. It is possible to show that

$$\left(\frac{d}{dq'}\right)^n e^{-q'^2/2} f(q') = e^{-q'^2/2} \left(\frac{d}{dq'} - q'\right)^n f(q'), \tag{188}$$

for an arbitrary function $f(q')$ by the use of mathematical induction. Now by choosing $f(q') = e^{-q'^2/2}$, this means that

$$\left(\frac{d}{dq'}\right)^n e^{-q'^2} = e^{-q'^2/2}(-1)^n \left(q' - \frac{d}{dq'}\right)^n e^{-q'^2/2}. \tag{189}$$

Using (189) in (180) gives us the alternate form

$$u_n(q') = \frac{(-1)^n}{\sqrt{\sqrt{\pi} \; 2^n n!}} e^{q'^2/2} \left(\frac{d}{dq'}\right)^n e^{-q'^2}. \tag{190}$$

Eq$^{\underline{n}}$ (190) is more familiar when the definition of Hermite polynomials,

$$H_n(q') = (-1)^n e^{q'^2} \left(\frac{d}{dq'}\right)^n e^{-q'^2}. \tag{191}$$

is introduced.  Eq$^{\underline{n}}$ (191) in (190) gives us

$$u_n(q') = \frac{1}{\sqrt{\sqrt{\pi}\ 2^n n!}}\ e^{-q'^2/2}\ H_n(q').\qquad(192)$$

The first few Hermite polynomials are

$$H_0 = 1,\ H_1 = 2q',\ H_2 = 4q'^2 - 2.\qquad(193)$$

The order of the $n^{\underline{th}}$ Hermite polynomial is n.  Also, the
number of zeros in the wavefunction $u_n(q')$ is also n
(excluding the points at infinity).  The first three $u_n(q')$
look like
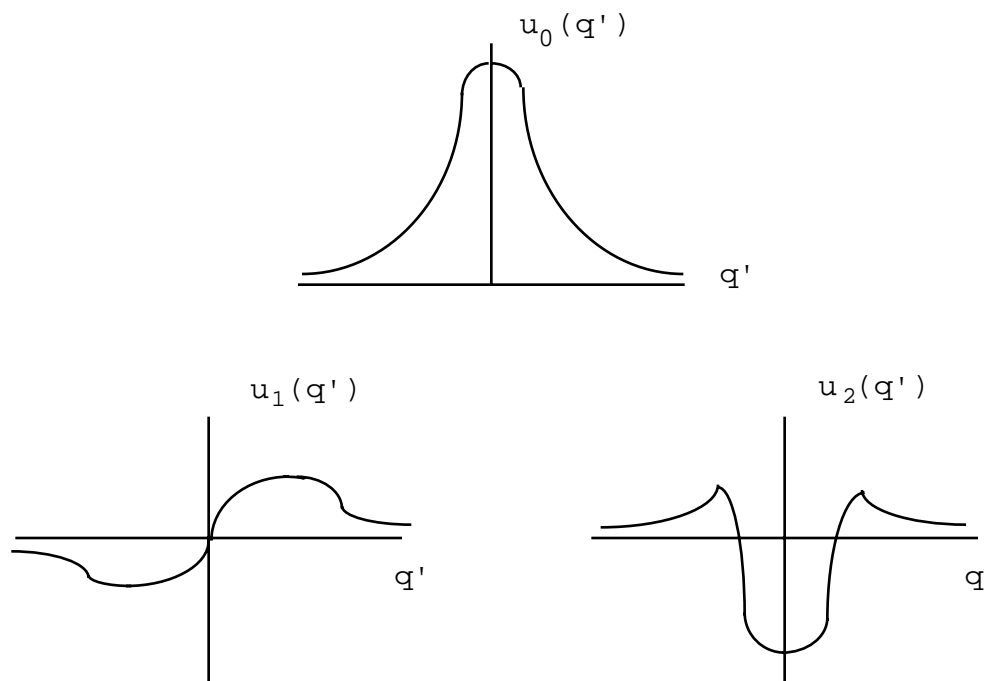


The orthonormality of the states $|n>$ may be demonstrated
as follows.  We know that

$$<n|n'> = <0|\ \frac{A^n}{\sqrt{n!}}\ \frac{(A^+)^{n'}}{\sqrt{n'!}}\ |0>.\qquad(194)$$

3.44

Now we may write

$$A(A^+)^{n'} = (A^+)^{n'}A + [A,(A^+)^{n'}]. \qquad (195)$$

In a problem, you will evaluate the commutator in (196) as

$$[A,(A^+)^{n'}] = n'(A^+)^{n'-1}. \qquad (196)$$

Using (195) in (194), and using $A|0> = 0$, we now find

$$<n|n'> = \frac{n'}{\sqrt{nn'}} <0| \frac{A^{n-1}}{\sqrt{(n-1)!}} \frac{(A^+)^{n'-1}}{\sqrt{(n'-1)!}} |0>$$

$$= \frac{n'}{\sqrt{nn'}} <n-1|n'-1> . \qquad (197)$$

We have three possible cases. If $n = n'$, then

$$<n|n> = <n-1|n-1> = \ldots = <0|0> = 1. \qquad (198)$$

However, if $n > n'$

$$<n|n'> = \frac{n'}{\sqrt{nn'}} <n-1|n'-1> = \ldots = \text{constants} <(n-n')|0>,$$

$$\qquad (199)$$

but for $n > n'$

$$<(n-n')|0> = <0| \frac{A^{n-n'}}{\sqrt{(n-n')!}} |0> = 0, \qquad (200)$$

so $<n|n'> = 0$ for $n > n'$. Similarly, $<n|n'> = 0$ when $n < n'$. Therefore, we have shown that

$$<n|n'> = \delta_{nn'} . \qquad (201)$$

In this discrete context, completeness of the energy eigenfunctions reads

$$\sum_n |n><n| = 1. \tag{202}$$

You will find the $<p'|n>$ wavefunctions in a problem.

We notice that here, as in the square well problem, that the parity operator commutes with H. We have that

$$<q'|\mathbb{P}|n> = <-q'|n> = u_n(-q'). \tag{203}$$

But since, from (191)

$$H_n(-q') = (-1)^n H_n(q'), \tag{204}$$

we have from (192) that

$$u_n(-q') = (-1)^n u_n(q'). \tag{205}$$

Therefore, (203) reads

$$<q'|\mathbb{P}|n> = (-1)^n <q'|n>. \tag{206}$$

Being true for all $<q'|$ tells us that

$$\mathbb{P}|n> = (-1)^n|n>. \tag{207}$$

It is now easy to show that

$$\mathbb{P}H|n> = \mathbb{P}E_n|n> = E_n(-1)^n|n>, \tag{208}$$

$$H\mathbb{P}|n> = (-1)^n|n> = (-1)^n E_n|n>, \tag{209}$$

and thus that

$$[H,\mathbb{P}]|n> = 0, \tag{210}$$

or, again, since this is true for all $|n>$, that

$$[H, \mathbb{P}] = 0 . \tag{211}$$

Thus, the parity of the states $u_n(q')$ does not change when they are time-evolved.

The wavefunctions $u_n(q')$ are dimensionless and normalized so that

$$\int_{-\infty}^{\infty} dq' u_n(q') u_n(q') = 1 \tag{212}$$

If we wish to work with the physically dimensionful quantity $x'$, then we should use the wavefunction

$$\mathbf{u}_n(x') = \left(\frac{m\omega}{\hbar}\right)^{1/4} u_n\left(q' \rightarrow \sqrt{\frac{m\omega}{\hbar}} x'\right) \tag{213}$$

which is normalized so that

$$\int_{-\infty}^{\infty} dx' \mathbf{u}_n(x') \mathbf{u}_n(x') = 1. \tag{214}$$

# 4. The Attractive Kronig-Penny Model

Next, we will study some of the physics of electrons in conductors and insulators. Of course, we will have to simplify the situation a great deal in order to be able to reduce the complexity of this problem. The first idealization is the reduction of the real situation to one spatial dimension. Then we might expect, qualitatively, the potential experienced by a single electron in the material to look somewhat like the following:
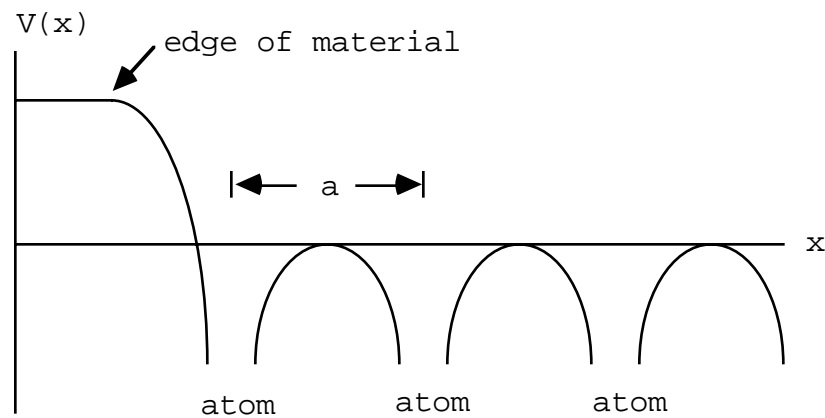


Notice that the potentials are attractive, as should be the case for electrons in the vicinity of atoms with a positively charged core. Also notice that the potential at the edge of the material rises to a constant level as one is getting further away from the attractive potentials of the interior atoms.

Even the potential shown above is too difficult for us to consider here. We will make two additional simplifications. First, we totally ignore all surface effects by assuming we are dealing with an infinite

collection of atoms arrayed in one dimension.  Second, we model the attractive Coulomb potentials by Dirac delta function spikes.  The result is the following potential:
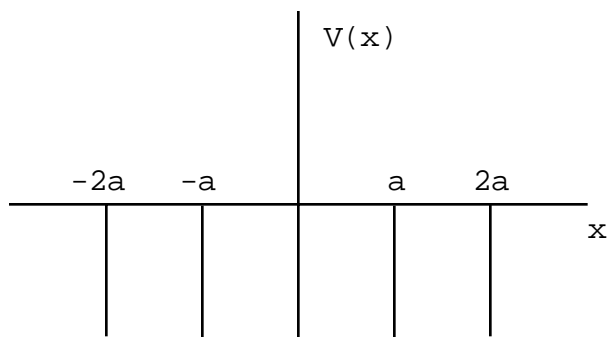


We now have a well-defined potential for which we can solve for the allowed energies.  Notice we have chosen our zero of potential in the above figure to correspond to the constant potential felt by the electrons in the "interior" of the metal, away from the positions of the "atoms."  In the following, we will restrict our attention to solving for the allowed positive energy levels, although states with $E < 0$ also exist.  We will call the $E > 0$ states that <u>conductance</u> electrons and the $E < 0$ states the <u>valence</u> electrons.

The potential in Eq$^{\underline{n}}$ (14) above is now determined as

$$V(x) = -\frac{\hbar^2}{2m}\frac{\lambda}{a}\sum_{n=-\infty}^{\infty}\delta(x - na).\qquad(215)$$

The positive constant $\lambda$ is dimensionless.  This follows since the dimensions of the delta function in (215) from

$$\int dx\ \delta(x - x') = 1,\qquad(216)$$

are $\left[\dfrac{1}{\text{length}}\right]$ and $\dfrac{\hbar^2}{2ma}$ has dimensions of $[\text{Energy} \cdot \text{length}]$.

The equation we want to solve is

$$-\left[\frac{\partial^2}{\partial x^2} + \frac{\lambda}{a}\sum_n \delta(x - na)\right] u(x) = \frac{2mE}{\hbar^2} u(x). \qquad (217)$$

Let us simplify this equation a bit by introducing the dimensionless variable $y = \dfrac{x}{a}$ so that, by using Eq$^{\text{n}}$ (101) of Chapter 2, we have

$$\delta(x - na) = \frac{1}{a}\delta(y - n). \qquad (218)$$

Let's also set $k = \dfrac{\sqrt{2mE}}{\hbar}$ as usual.  Then we have that

$$\left[\frac{\partial^2}{\partial y^2} + \lambda\sum_n \delta(y - n)\right] u(y) = -(ka)^2 u(y) \qquad (219)$$

for the positive energy solutions.  At any positions $y \neq n$, the solutions to

$$\frac{\partial^2}{\partial y^2} u(y) = -(ka)^2 u(y) \qquad (220)$$

are simply $\sin(kay)$ and $\cos(kay)$.  The general solution to $u(y)$ in the region $(n - 1) < y < n$ is then given by

$$u(y) = A_n \sin(ka(y - n) + B_n \cos(ka(y - n)) \qquad (221)$$

where $A_n$ and $B_n$ are (complex) coefficients, and the phase factors, $-(ka)n$, in the sine and cosine are chosen for convenience.  We have learned that for piecewise continuous potentials our probability interpretation for wavefunctions

only holds if u(x) and $\frac{du}{dx}$ are continuous functions.  Of

course here our potential contains delta functions and so

does not satisfy the condition of piecewise continuity.  What

continuity conditions must we therefore impose on the

wavefunctions for this problem?  It is clear that the

troublesome positions are the locations of the delta

functions at y = n.  In order to find the requirements on the

wavefunctions at these locations, let us repeat the argument

based on Eqⁿ (107) above.  Here we have that

$$\int_{n-\varepsilon}^{n+\varepsilon} dy \left\{ \left[ \frac{\partial^2}{\partial y^2} + \lambda \sum \delta(y - n) \right] u(y) = -(ka)^2 u(y) \right\}, \quad (222)$$

where we have integrated over a small neighborhood around

y = n.  Then, in the limit $\varepsilon \to 0^+$ we have

$$\left. \frac{\partial u}{\partial y} \right|_{y=n+} - \left. \frac{\partial u}{\partial y} \right|_{y=n-} = -\lambda u(n). \quad (223)$$

Eqⁿ (223), in contrast to Eqⁿ (108), says that a

discontinuity

in $\frac{\partial u}{\partial y}$ is <u>required</u>.  By integrating again, we find however

that

$$u(n)|_+ = u(n)|_- \quad (224)$$

so that the wavefunctions <u>are</u> continuous across y = n.

Let us now apply (223) and (224) to (221).  The

condition (224) applied at y = n (between the wavefunctions

defined in the regions $(n - 1) \leq y \leq n$ and $n \leq y \leq n + 1$)
says that

$$B_n = -A_{n+1} \sin(ka) + B_{n+1} \cos(ka), \qquad (225)$$

while (223) implies

$$ka(A_{n+1} \cos(ka) - A_n) + kaB_{n+1} \sin(ka) = -\lambda B_n. \qquad (226)$$

These last two equations require that

$$A_{n+1} = A_n \cos(ka) - \left( \frac{\lambda}{ka} \cos(ka) + \sin(ka) \right) B_n, \qquad (227)$$

$$B_{n+1} = \left( \cos(ka) - \frac{\lambda}{ka} \sin(ka) \right) B_n + \sin(ka)A_n, \qquad (228)$$

which are recursion relations for the $A_{n+1}$ and $B_{n+1}$ given $A_n$
and $B_n$.

Eq$^{\text{n}}$s (227) and (228) are not sufficient to complete the
description of this system.  There must be further relations
between wavefunctions in different spatial regions. Let us
notice that under the substitution $x \rightarrow x + a$ (equivalent to
$y \rightarrow y + 1$) we have that

$$V(x) \rightarrow V(x + a) = V(x), \qquad (229)$$

and

$$\frac{d}{dx} \rightarrow \frac{d}{d(x + a)} = \frac{d}{dx} . \qquad (230)$$

This means that the differential equation we are solving,
(217), is <u>invariant</u> or unchanged under this substitution.  We
will therefore demand that all observables are also invariant

under this change.  This means that the probability densities

in neighboring regions must be identical.

$$|u(y + 1)|^2 = |u(y)|^2 . \tag{231}$$

This says in general that

$$u(y + 1) = e^{i\phi}u(y), \tag{232}$$

where $\phi$ is some unknown phase. (From the general form

(221),this phase can not be be y-dependent.)

    We have gone as far as we can without specifying

boundary conditions.  We will use the condition

$$u(y + N) = u(y), \tag{233}$$

where N = 1,2,3,... . (If there are are $\sim 10^{23}$ atoms present,

there will be $10^8 \approx (10^{23})^{1/3}$ atomic planes in any one

direction.)  These are called periodic boundary conditions.

This condition couples the electrons on one side of the

material to the other, resulting in what can be thought of as

a ring of atoms. The phase angle $\phi$ in (232) is now determined

as

$$e^{i\phi N} = 1 \tag{234}$$

$$\Rightarrow \quad \phi = \frac{2\pi}{N} m , \quad m = 0,\pm 1,\pm 2,... \tag{235}$$

Eq$^{\underline{n}}$(235) gives the dimensionless quasi-momentum values in our

model.  In terms of the coefficients in (221), this says that

$$A_{n+1} = e^{i\phi}A_n \ , \tag{236}$$

$$B_{n+1} = e^{i\phi}B_n \ , \tag{237}$$

Let's now substitute (236) and (237) into (227) and (228) above.  We get

$$e^{i\phi}A_n = A_n \cos(ka) - \left( \frac{\lambda}{ka} \cos(ka) + \sin(ka) \right) B_n \ , \tag{238}$$

$$e^{i\phi}B_n = \left( - \frac{\lambda}{ka} \sin(ka) + \cos(ka) \right) B_n + \sin(ka)A_n. \tag{239}$$

These may be put into the form

$$(e^{i\phi} - \cos(ka))A_n = - \left( \frac{\lambda}{ka} \cos(ka) + \sin(ka) \right) B_n \ , \tag{240}$$

$$\sin(ka)A_n = - \left( - \frac{\lambda}{ka} \sin(ka) + \cos(ka) - e^{i\phi} \right) B_n \ . \tag{241}$$

In order for (240) and (241) to be consistent, we must have that

$$\frac{e^{i\phi} - \cos(ka)}{\sin(ka)} = \frac{\dfrac{\lambda}{ka} \cos(ka) + \sin(ka)}{- \dfrac{\lambda}{ka} \sin(ka) + \cos(ka) - e^{i\phi}} \ . \tag{242}$$

This equation can be reduced to

$$\cos \phi = \cos(ka) - \frac{\lambda}{2} \frac{\sin(ka)}{ka} \ . \tag{243}$$

Imagine letting $N \rightarrow \infty$ in (235).  Then $\phi$ essentially becomes a continuous variable and $\cos \phi$ can take on any value between 1 and -1.  This is a type of <u>dispersion relation</u>. This equation determines the energies of the system, given by

$E = \dfrac{(\hbar k)^2}{2m}$, in terms of the allowed quasi-momentum values of

the system, $\phi$.

Plots of the right hand side of (243) follow for the

cases $\lambda < 4$ and $\lambda > 4$:



$\lambda < 4$ case



$\lambda > 4$ case
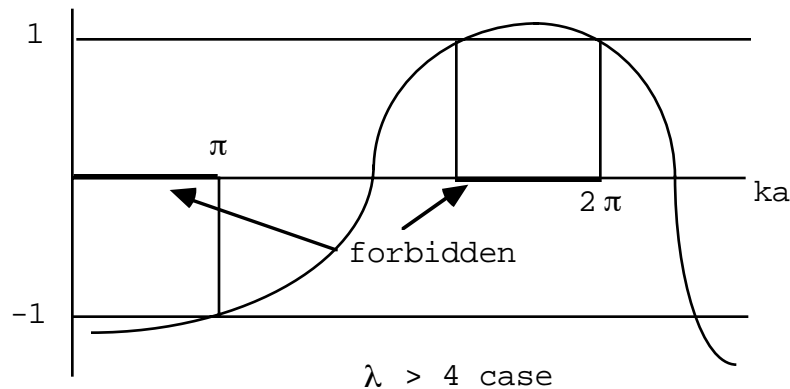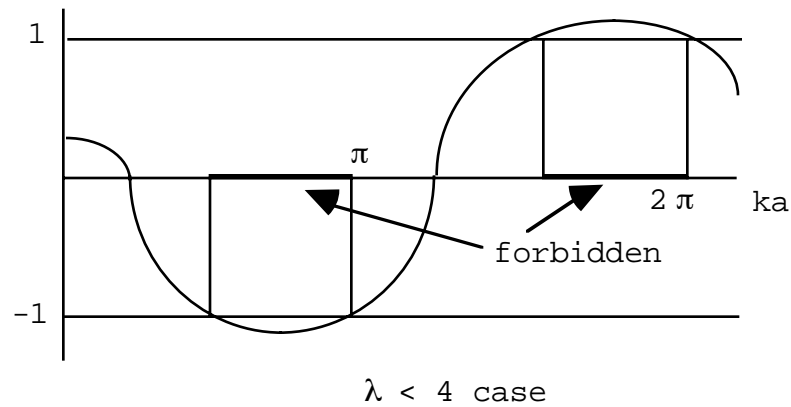
Now notice that since the left hand side of Eq$^{\underline{n}}$ (243) is

bounded by 1 and -1, there are no solutions to (243) in the

regions marked "forbidden" above.  Otherwise we have a

continuum of solutions (at least in the $N \to \infty$ limit).  These

continuum solutions are called energy _bands_.  The forbidden

zones, or energy _gaps_, are a result of destructive

interference between waves reflected off the various delta

function potentials.  We saw such a phenomenon before in our

discussion of the finite potential barrier, where the transmission of the waves through the potential barrier was reduced for some values of the momentum because of destructive interference.  Now, the positive constant $\lambda$ in some sense represents the strength of the attractive delta function potentials. By mapping all the quasi-momentum values in the above figures into the interval $-\pi \lesssim \phi \lesssim \pi$, we get the so-called <u>reduced-zone</u> description of energy eigenstates:
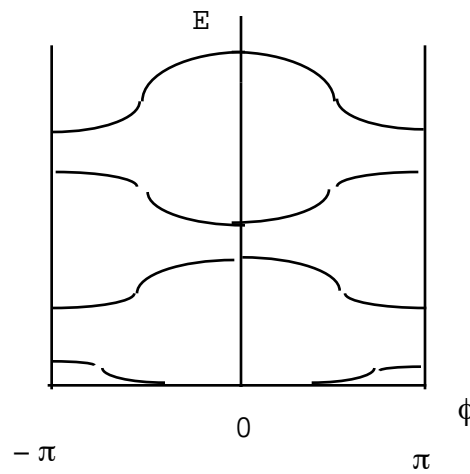


$\lambda < 4$ case
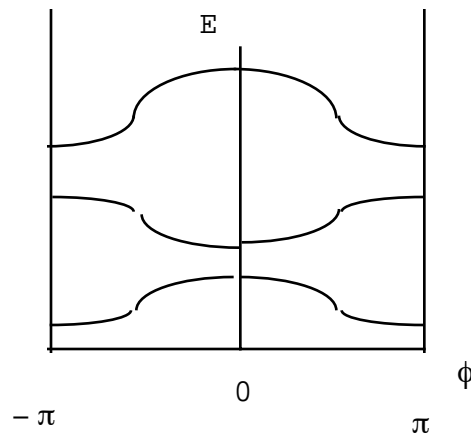


$\lambda > 4$ case

What we see in the above figures is that when the attraction between the electrons and atoms is weak, the conductance energy bands come all the way down to E = 0, that is, to the top of the valence band.  We see that, however, when λ becomes large enough that an energy gap forms between (ka) = 0 and (ka) = π.  This is a very rough model of the band structures in conductors (λ small) and insulators (λ large). In conductors, where the interaction between the electrons and atoms is relatively weak, there is no energy gap and valence electrons can easily occupy states in the conductance band where they are able to transport charge through the material.  On the other hand, insulators hold on tightly to their electrons and they generally have an energy gap, as in this simplified model.  An external electric field is then not effective in moving electrons into the conductance band and no flow of electricity results.  Temperature can play an important role in exciting some electrons upward in energy into the conductance band in some materials.  Actually, conductors usually have a partially filled conductance band at room temperatures and therefore conduct electricity readily.

Of course, there are many simplifications inherent in this one-dimensional model of the interior of conductors and insulators.  In reality, because one is working with a finite system instead of the infinitely long system considered above, the continuous energy conductance bands we have found above are in really a collection of extremely closely spaced

discrete levels.  In addition, there are effects having specifically to do with the surface of such a material, which we have not considered here.  The Kronig-Penny model does correspond to reality in the illustration of the formation of energy gaps, and it was for this reason that I have presented it here.

We have solved the time-independent Schrödinger equation in one spatial dimension for a variety of potentials.  We have seen two basic types of solutions, <u>bound state</u> solutions (as in the infinite square well and the simple harmonic oscillator) and <u>scattering</u> solutions (as for the finite barrier and the Kronig-Penny model).  Bound state solutions have discrete energies and their wavefunctions can be normalized in space.  In scattering solutions, energy values are contiguous and the wave functions cannot be normalized in space (there would be infinite) but instead are normalized to a delta function in energy, say.

Generalizing from the experience we have gained in this Chapter, we would expect to find the following classification of energy states for the following two Examples.
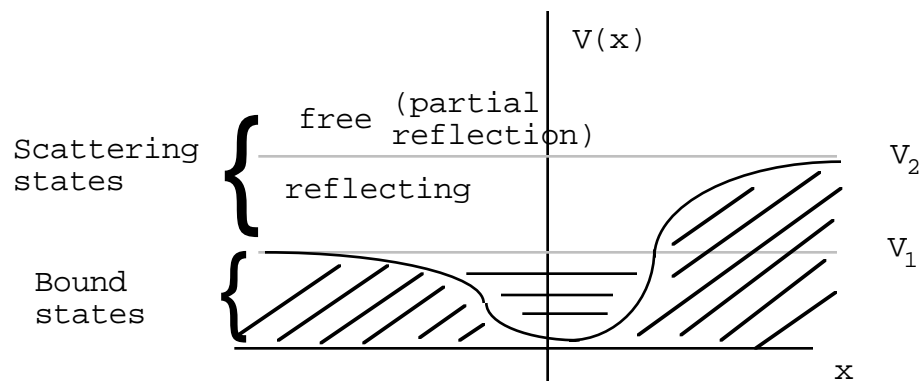


Example 1
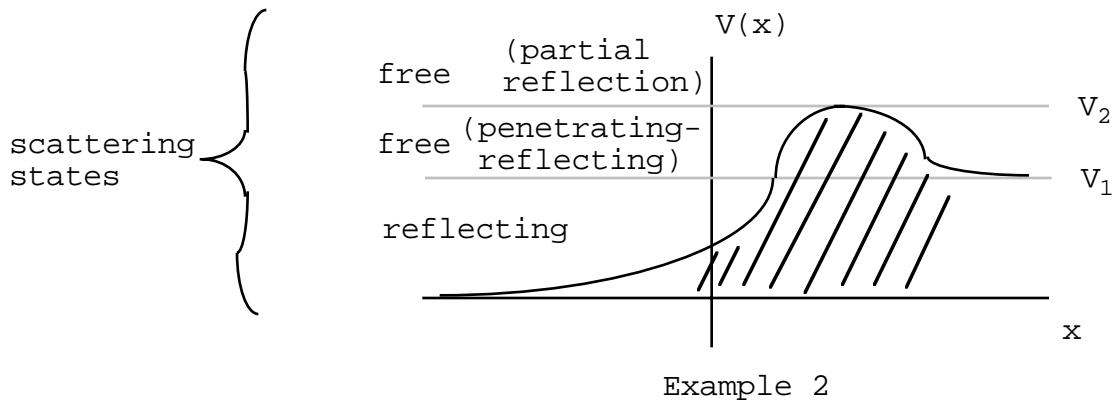
Example 2

In Examples 1 or 2 for $E > V_2$, we would expect the
particles to be "free" in the sense of having oscillatory
wave functions whose wavelengths, however, would be a
function of position, just like we saw for energies higher
than the top of the potential in the finite barrier problem.
In Example 1 when $V_1 < E < V_2$ we expect states that will
describe waves that completely reflect off of the potential
barrier which extends to $x = \infty$.  This does not mean, however,
that the particle will not penetrate into the right hand side
region (the shaded region) where, classically, the particle
would not be allowed to go. We saw in the finite potential
barrier problem that this indeed can happen.  However,
because the particle's wavefunction will be damped in this
region, one has a smaller probability of detecting a particle
with such an energy as we move further to the right in this
figure.  All of the particles in this case must eventually
reflect because in a steady state situation the flux of
particles at $x = \infty$ is zero.  In Example 2 for $V_1 < E < V_2$,

the particle has only a finite potential barrier to overcome
and it can "tunnel" from one side of the barrier to the other
resulting in both penetration and reflection from the
barrier.  In Example 1 for $E < V_1$ we see that the particles

are trapped in a potential well.  Then, because of

destructive interference between waves, we would expect only

a discrete set energy values to be allowed in a steady state.

In Example 2 for $E < V_1$ we again expect complete reflection

from, but partial penetration of, the barrier.  The key to

understanding these qualitative behaviors is the wave picture

of particles as deBroglie waves that we developed in the last

Chapter.

## Problems

1. Show for the Schrodinger equation with an arbitrary potential, V(x), that the continuity equation, Eq. (122) of Chapter 2, still holds true (as asserted in the text on p. 3.1).  Assume that V(x) is Hermitian.

2. Show equations (84 a,b,c) of the text in Chapter 3 are true.

3. Show for the one-dimensional square well that

$$\langle x'|n+\rangle = \langle x'|L_{n+1/2}|(n+1)+\rangle,$$

and

$$\langle x'|n+\rangle = \langle x'|L_{n-3/2}^+|(n-1)+\rangle.$$

[Since the above are true for all $< x'|$, they imply

$$|n+\rangle = L_{n+1/2}|(n+1)+\rangle,$$
$$|n+\rangle = L_{n-3/2}^+|(n-1)+\rangle.]$$

4. A square well wave function at t = 0 is given by

$$|\Psi,0\rangle = \frac{1}{\sqrt{13}} [2|1+\rangle + 3i|2-\rangle].$$

Find $\langle \mathbb{P} \rangle$ ($\equiv \langle \Psi,t|\mathbb{P}|\Psi,t\rangle$) as a function of time.

5. Derive the result, Eq.(119) of Ch.3, for the $\frac{E}{A}$ coefficient for the finite potential barrier problem described there.

6. Find the momentum space harmonic oscillator wave functions, $\Psi_n(p') = \langle p'|n\rangle$, normalized such that

$$\int_{-\infty}^{\infty} dp' \, |\Psi_n(p')|^2 = 1.$$

[p' is the dimensionless position variable related to $p_x$ by

$$p = \frac{p_x}{\sqrt{mw\hbar}} \; .$$

Hint: One way is to take Eq.(175) of Ch.3 and project it into $\langle p'|$ instead of $\langle q'|$.]

7.(a)  From
$$A|n\rangle = \sqrt{n}\,|n-1\rangle,$$
derive
$$\frac{d}{dq'} \, H_n(q') = 2n \, H_{n-1}(q').$$

(b)  From

$$A^+|n\rangle = \sqrt{n+1}\,|n+1\rangle,$$
derive
$$\left(2q' - \frac{d}{dq'}\right) H_n(q') = H_{n+1}(q').$$

(c)  Use (a) and (b) to show that $H_n(q')$ satisfies the differential equation:

$$\left(\frac{d^2}{dq'^2} - 2q'\frac{d}{dq'} + 2n\right) H_n(q') = 0.$$

8. Find some way of showing Eq.(196) of Ch.3 holds true.

9. Consider the Kronig-Penny model with periodic <u>repulsive</u> delta-function potentials, i.e., $\lambda < 0$.

  (a) Given the dispersion relation,

$$\cos\phi = \cos(ka) - \frac{\lambda}{2}\frac{\sin(ka)}{ka},$$

3.62

draw (qualitatively) the appropriate cosφ vs. ka graph (be
sure to indicate any energy gaps):

cos φ

1

ka

-1
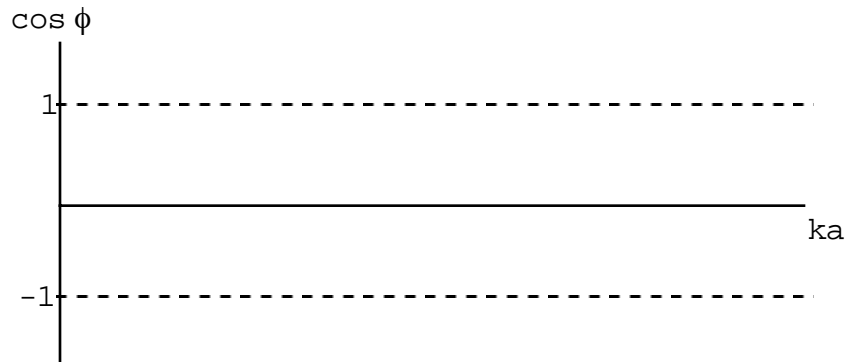
   (b) Give the <u>reduced-zone</u> energy graph corresponding to
your answer to part (a). Show at least the lowest two
energy bands:

E

ka

−π          0          π
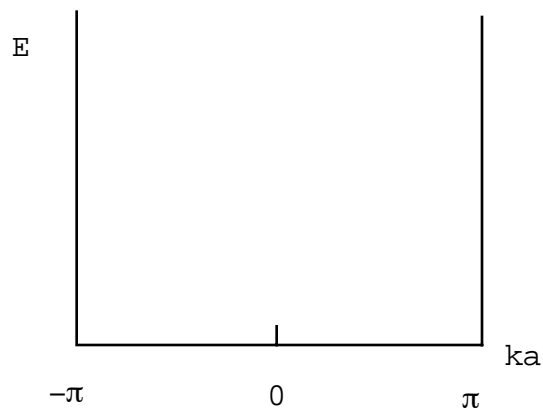
   (c) Show that there are no energy eigenvalue solutions to
the part (a) dispersion relation when both λ < 0 and E < 0.


**Other  Problems**

10. Given the harmonic oscillator eigenvalue equation,
$H|n\rangle = E_n|n\rangle$, where $E_n = \hbar\omega(n+\frac{1}{2})$ and

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

(a) Derive the position space energy eigenvalue equation for $u_n(x') \equiv \langle x'|n\rangle$ ("the Schrödinger equation in position space").

(b) Derive the momentum space energy eigenvalue equation for $v_n(p'_x) \equiv \langle p'_x|n\rangle$ ("the Schrödinger equation in momentum space").

(c) Find a simple substitution of variables which changes the differential equation in (a) into the equation in (b). (This therefore gives $v_n(p'_x)$, outside of an overall normalization, if $u_n(x')$ is known.)

11. We are given a wavefunction, $\psi(x)$, which solves the coordinate space Schrödinger equation (with a piecewise continuous potential $V(x)$). Assume that this $\psi(x)$ has a kink in it at $x = x_0$:
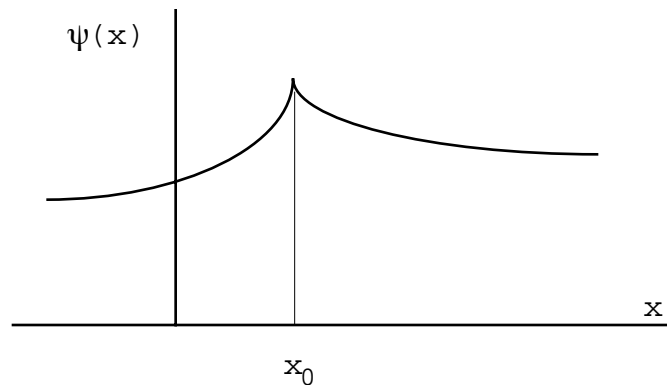


Give a convincing mathematical argument that such a discontinuity would imply that probability is not conserved. (See p.3.26 of the class notes for some relevant comments.)

12. Consider the scattering problem:

3.64



x=0
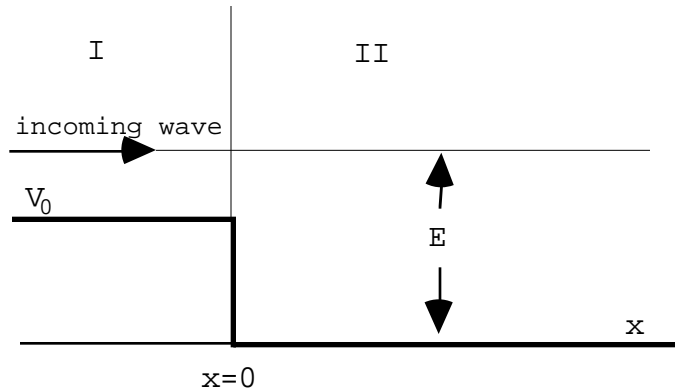
The energy, E, of the incoming particle is greater than the top of the step potential: $E > V_0$. The incoming particle originates from the left (i.e., from x < 0 ). Define

$$k_1 = \frac{\sqrt{2m(E-V_0)}}{\hbar}, \quad k_2 = \frac{\sqrt{2mE}}{\hbar}.$$

   (a) Write down the general solution to the energy eigenvalue equation in Region I.

   (b) Write down the general solution to the energy eigenvalue equation in Region II.

   (c) Show that the reflection coefficient, R, is given by

$$R = \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2.$$


13. The Kronig-Penny model dispersion relation for <u>negative</u> energies is given by:

$$\cos\phi = \cosh(\kappa a) - \frac{\lambda}{2}\frac{\sinh(\kappa a)}{\kappa a},$$

where $\kappa = \sqrt{-2mE}/\hbar$.

   (a)  Draw (qualitatively) the appropriate $\cos\phi$ vs. $\kappa a$ graph (be sure to indicate any energy gaps; assume $\lambda > 4$):

$$\cos \phi$$

$$1 \quad \text{-------------------------------}$$

$$\kappa a$$

$$-1 \quad \text{-------------------------------}$$
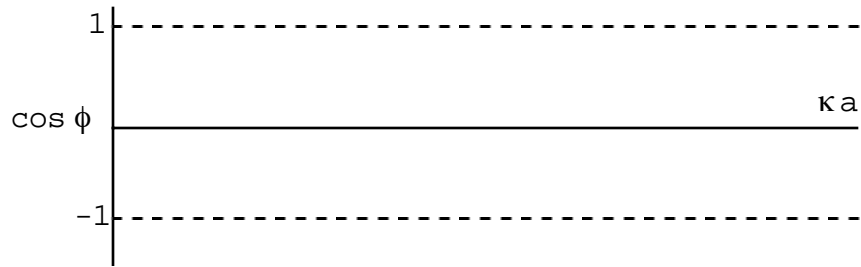
(b) Give the <u>reduced-zone</u> energy graph corresponding to your answer to part (a) (again assume $\lambda > 4$):

$$- E$$

$$-\pi \qquad 0 \qquad \pi \qquad \phi$$
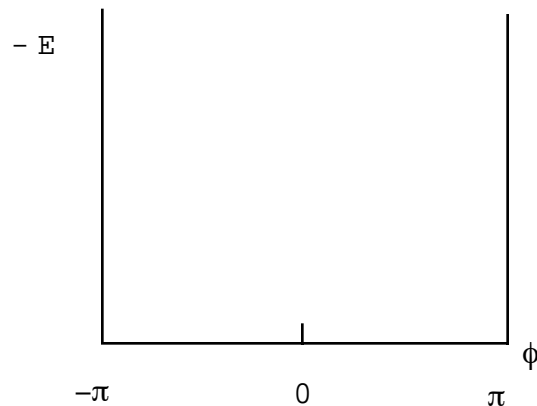
What value of $\phi$ is associated with the ground state (i.e., most negative) energy? <u>Explain how to solve for this energy</u>. (You don't actually have to do it.)

14. Evaluate the double commutator:

$$[p_x,[p_x,f(x)]] = ?$$

15. Find the expectation value of the parity operator, $\mathbb{P}$, as a function of time, given the infinite square well initial state:

$$|\psi,t=0> = \frac{1}{\sqrt{5}}(2i|1+> + |1->).$$

16. Refers to the harmonic oscillator. Evaluate the quantities:

3.66

   (a) $\langle n+2|q^2|n\rangle$ = ?

   (b) $\langle n|p^2|n\rangle$ = ?


17. Work out the transmission coefficient, T, for a one-dimensional Dirac delta-function potential ($\lambda > 0$), $V(x) = \frac{\lambda\hbar^2}{2m}\delta(x)$, for a right-moving incoming wave of energy E:
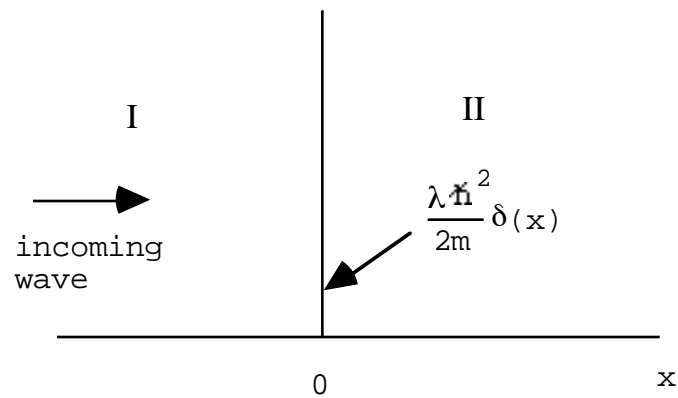
## Chapter 4:  More About Hilbert Space

Although we have been progressing steadily in our understanding of the laws of nature in the microscopic world, I have not been very systematic in the development of the mathematics behind the physics.  I hope I have convinced you of the utility of the bra-ket notation of Dirac, but now I owe you a deeper discussion of the mathematical properties of bras, kets and operators.  In other words, I think it would be best at this stage to consolidate our knowledge of formalism at bit before moving on to more complicated problems.  For this reason, we will continue our limitation to one dimensional problems here, moving on to three dimensions in Chapter 6.

First, a technicality before we move on.  We will formally deal here with systems that have a finite number of degrees of freedom, like spin, but unlike the position or momentum characterizations of a free particle.  This is done in order to simplify the discussion.  However, we will concern ourselves with results which are general in character and straightforwardly carry over to position and space measurements with obvious changes (like changing Kronecker deltas into Dirac deltas, etc.) Technically, only systems with an infinite number of degrees of freedom constitute a Hilbert space. We will use the term more loosely here to include finite systems.

Next, some words about notation and terminology.  We will continue to let

$$|a>, \quad |a'>, \quad |a_i> \qquad \text{(Case 1)}$$

or

$$|i>, \quad |n> \qquad \text{(Case 2)}$$

represent the eigenkets of some operator A. In Case 1
above, we are labeling the eigenket by the actual
eigenvalue; in Case 2 we are labeling these in some
unspecified order by an integer.  Because we are limiting
ourselves here to a finite number of states or physical
outcomes, the upper limit on all such labels is finite.

   Let me also remind you of the meaning of a "basis."
When we were discussing spin $\frac{1}{2}$ , the basis consisted of the
states $|\sigma_3'>$, $\sigma_3'$  = ± 1.  A basis is a set of linearly
independent vectors in Hilbert space that completely
characterize the space in the same way that the usual unit
vectors, $\hat{e}_1, \hat{e}_2, \hat{e}_3$ completely characterize 3-D space.  (When
I refer to a "vector" in Hilbert space, I mean either a bra
or a ket.)  Just as any 3-D vector $\vec{v}$ may be expanded as

$$\vec{v} = \sum_{i=1}^{n} v_i \hat{e}_i , \qquad (1)$$

any state $|\psi>$ may be expanded as

$$|\psi> = \sum_{i=1}^{n} C_i |x_i >, \qquad (2)$$

where I will be using the notation $|x_i>$ to denote a basis
ket.  The $C_i$ in (2) are some (in general) complex constants.

We say that a basis "spans" the Hilbert space.  According to the technical limitation we have imposed, the number of basis states in (2) is a finite number, n.

What does linear independence of basis kets mean mathematically?  In 3-D vector language, we say that two nonzero vectors, $\vec{v}_1$ and $\vec{v}_2$, are linearly independent if there is no constant C such that

$$\vec{v}_1 = C\vec{v}_2. \tag{3}$$

In the same way, we say that the nonnull kets $|x_1\rangle$ and $|x_2\rangle$ are linearly independent if there is no constant C such that

$$|x_1\rangle = C|x_2\rangle. \tag{4}$$

The $|x_i\rangle$ therefore are like vectors "pointing" in various independent directions in Hilbert space.  (This last sentence is a good summary of the entire content of this Chapter.)

We will be assuming the existence of the Hermitian conjugation operation that takes bras and kets and vice versa.  For some arbitrary state $|\psi\rangle$ we have

$$(|\psi\rangle)^+ = \langle\psi|. \tag{5}$$

Bras and kets are independent objects, but the existence of a state $|\psi\rangle$ implies the existence of $\langle\psi|$, and the other way around.  The relationship between these two objects is

called "dual."  We are assuming linearity of this
operation, so that

$$( |\psi_1 > + |\psi_2 >)^+ = <\psi_1| + <\psi_2|. \tag{6}$$

When C is a complex number, we also have

$$(C|\psi >)^+ = C^*<\psi| = <\psi|C^*. \tag{7}$$

Of course, given the state $|\psi>$ in (2), the state $<\psi|$ is
implied to be

$$<\psi| = \sum_{i=1}^{n} C_i^*<x_i|. \tag{8}$$

Given the bra $<x_i|$ and the ket $|x_j>$, we may form two
types of quantities, the <u>inner product</u>, $<x_i|x_j>$ (introduced
previously in Ch. 1), and the <u>outer product</u>, $|x_j><x_i|$
(previously $|x_j x_i|$). Just as the elements of $\hat{e}_i$ are
orthogonal in 3-D, as expressed by

$$\hat{e}_i \cdot \hat{e}_j = 0, \quad i \neq j, \tag{9}$$

we choose a basis such that the inner product must satisfy

$$<x_i|x_j> = 0, \quad i \neq j. \tag{10}$$

For convenience, we often normalize the basis so that (called
an orthonormal basis)

$$<x_i|x_j> = \delta_{ij} , \qquad \text{(all i,j)} \tag{11}$$

where $\delta_{ij}$ is the Kronecker delta.  A vector $|\psi>$ is said to
have a squared "length", $<\psi|\psi>$ in Hilbert space given by

$$< \psi | \psi > \; = \; \sum_{i,j} C_i^* C_j < x_i | x_j > \; = \; \sum_i |C_i|^2 \; \geq \; 0. \qquad (12)$$

It is in this sense that a bra or ket is said to be "nonzero" (as below Eq$\underline{n}$ (3) above).

The outer product $|x_j > < x_i|$ is another name for what we have been calling an operator. The most general expression for an operator A in our finite Hilbert space is

$$A \; = \; \sum_{i,j} A_{ij} |x_i > < x_j|. \qquad (13)$$

The coefficients $A_{ij}$ are just the elements of an n × n matrix representation of A. This is easy to see since

$$< x_k | A | x_\ell > \; = \; \sum_{i,j} A_{ij} < x_k | x_i > < x_j | x_\ell >$$

$$= \; \sum A_{ij} \delta_{ki} \delta_{j\ell} \; = \; A_{k\ell}, \qquad (14)$$

which is the definition of a matrix element of the operator A.

When acting on A in Eq$\underline{n}$ (13), Hermitian conjugation has the effect

$$A^+ \; = \; \sum_{i,j} A_{ij}^* |x_j > < x_i|$$

$$= \; \sum_{i,j} (A^*)_{ij}^T |x_i > < x_j|. \qquad (15)$$

Thus we see that the operator statement

$$A \; = \; A^+, \qquad (16)$$

is equivalent to the matrix statement

$$A = (A^*)^T \ .$$

(17)

(The transpose and complex conjugation operations commute.)
As remarked on before in Chapter 1, we often denote the
complex transpose of a matrix as "$+$".  The meaning of "$+$" is
determined by its context.

Consider the operator product AB.  This can be written
as

$$AB = \sum_{i,j} A_{ij} |x_i> <x_j| \cdot \sum_{k,\ell} B_{k\ell} |x_k> <x_\ell|$$

$$= \sum_{i,j,\ell} A_{ij}(B_{j\ell}) |x_i> <x_\ell| \ .$$

(18)

The Hermitian conjugate is given by

$$(AB)^+ = \sum_{i,j,\ell} A^*_{ij} B^*_{j\ell} |x_\ell> <x_i|$$

$$= \sum_{i,j,\ell} A^*_{\ell j} B^*_{ji} |x_i> <x_\ell| \ ,$$

(19)

whereas we also have that

$$B^+A^+ = \sum_{i,j} B^*_{ji} |x_i> <x_j| \cdot \sum_{k,\ell} A^*_{\ell k} |x_k> <x_\ell|$$

$$= \sum_{i,j,\ell} A^*_{\ell j} B^*_{ji} |x_i> <x_\ell| \ ,$$

(20)

which proves in general that

$$(AB)^+ = B^+A^+ \ .$$

(21)

This is the same rule encountered in Ch. 1.

The operator/matrix relationship may be elaborated more fully as follows.  We know that the product $A|\psi>$ where A and $|\psi>$ are arbitrary, is another ket:

$$|\psi'> = A|\psi>. \tag{22}$$

Using the forms (13) and (2) for A and $|\psi>$, we get

$$A|\psi> = \sum_{i,j} A_{ij}|x_i><x_j| \cdot \sum_k C_k|x_k>$$

$$= \sum_{i,j} A_{ij}C_j|x_i>. \tag{23}$$

On the other hand we know that $|\psi'>$ is another ket and so may be expanded in basis kets:

$$|\psi'> = \sum_i B_i|x_i>. \tag{24}$$

Thus, the operator-ket statement (22) is equivalent to the matrix-vector statement

$$B_i = \sum_j A_{ij}C_j, \tag{25}$$

which is often written with understood indices as

$$B = AC, \tag{26}$$

where B and C are interpreted as <u>column</u> matrices.  Likewise, given the statement

$$<\psi'| = <\psi|A, \tag{27}$$

and the expansions

$$<\psi'| = \sum_i B_i'<x_i| , \tag{28}$$

$$< \psi | \ = \ \sum_i C_i' < x_i | \ , \tag{29}$$

the analogous matrix statement to (27) is

$$B' = C'A, \tag{30}$$

where B' and C' are now <u>row</u> matrices.

Our operator-matrix analogy stands as follows:

$$\text{operator} \longleftrightarrow \text{matrix}$$

$$\text{ket} \longleftrightarrow \text{column matrix}$$

$$\text{bra} \longleftrightarrow \text{row matrix}$$

$$\text{complex conjugation} \oplus \longleftrightarrow \text{complex conjugation} \oplus$$
$$\text{bra,ket interchange} \qquad\qquad \text{transpose}$$

The last line gives the meaning of Hermitian conjugation in the two cases.

The completeness relation, first seen in Eq$^{\underline{n}}$ (91) of Chapter 1, may be deduced in this context as follows. A general ket $|\psi>$ may be expanded as in (2) above. Multiplying on the left of (2) by $<x_j|$ and assuming orthonormality, this tells us that the expansion coefficients $C_i$ in (2) are just

$$C_i \ = \ < x_i | \psi >. \tag{31}$$

Plugging this back into (2) then gives

$$| \psi > \ = \ \sum_i | x_i > < x_i | \psi > \ = \ \left( \sum_i | x_i > < x_i | \right) | \psi > \ . \tag{32}$$

Since $|\psi>$ in (32) is arbitrary, this tells us that

$$\sum_i |x_i > < x_i| \; = 1 \; .\tag{33}$$

This sheds a different, more mathematical, light on the meaning of completeness.

A distinguishing characteristic of a Hermitian operator is that its physical outcomes, which mathematically speaking are its eigenvalues, are real.   (We already showed at the end of Chapter 1 that the expectation value of a Hermitian operator is real.)   We can show this as follows.   Let $|a_i>$ be any nonzero eigenket of A,

$$A|a_i > \; = \; a_i|a_i > \; .\tag{34}$$

Then

$$< a_i|A|a_i > \; = \; a_i < a_i|a_i > \; .\tag{35}$$

The adjoint of (34) is

$$< a_i|A^+ \; = \; a_i^* < a_i| ,\tag{36}$$

which then gives

$$< a_i|A^+|a_i > \; = \; a_i^* < a_i|a_i> \; .\tag{37}$$

If $A = A^+$, then (35) and (37) demand that

$$a_i \; = \; a_i^* \; ,\tag{38}$$

and the eigenvalues are real. (A quicker proof is to recall that at the end of Ch. 1 we proved that the expectation value

of Hermitian operators are real. Since the left hand side of
(35) is just such a quantity, and because $\langle a_i | a_i \rangle \geq 0$, we
must have that the $a_i$ are real.)

It is easy to prove that eigenkets of an Hermitian
operator corresponding to distinct eigenvalues are
orthogonal.  Eq$^{\underline{n}}$ (34) implies that

$$\langle a_i | A | a_j \rangle = a_j \langle a_i | a_j \rangle \, , \tag{39}$$

and (36) (with $A = A^+$, $a_i = a_i^*$) implies

$$\langle a_i | A | a_j \rangle = a_i \langle a_i | a_j \rangle \, . \tag{40}$$

Therefore from (39) and (40) we have

$$(a_i - a_j) \langle a_i | a_j \rangle = 0, \tag{41}$$

which says that $\langle a_i | a_j \rangle = 0$ if $a_i \neq a_j$. (The operator A in
(39) can be visualized as operating to the right while the A
in (40) acts to the left.)

Now there are n independent basis kets $| x_i \rangle$.  The
eigenkets of a Hermitian operator can, of course, be expanded
in such a basis.  The number of linearly independent
eigenkets $| a_i \rangle$ cannot exceed n, the number of basis kets,
for if they did then the $| x_i \rangle$ would not span the space
contrary to our definition.  Therefore, the maximum number of
distinct eigenvalues is always less than or equal to n.  If
all n of them are distinct, then it is clear that we may
choose the $| a_i \rangle$ as a (not necessarily normalized) basis for
our description.  This is a most useful result since it

provides a means of finding a basis. However, it sometimes occurs that the number of distinct eigenvalues is less than n. In this case it is not clear if we can find enough linearly independent eigenkets of the Hermitian operator to span the space so that we can choose them as a basis. What I want to explain is that even when the number of distinct eigenvalues is less than n, the number of linearly independent eigenkets of a Hermitian operator is still n, and then therefore can be chosen as a basis. I will be content to just lay out the bare bones of this explanation since the details involve some mathematical technicalities. In doing so, the connection we have found between operators and matrices will be extremely useful.

Let us say we have a Hermitian operator A and we want to find its eigenvalues and eigenkets. Let A be represented in terms of an arbitrary basis as in Eq$^{\underline{n}}$ (13) above. We want to find all possible $|\psi>$ such that

$$A|\psi> = a|\psi>, \qquad (42)$$

where "a" is an eigenvalue of A. Expanding $|\psi>$ in our basis as

$$|\psi> = \sum_i y_i |x_i>, \qquad (43)$$

we know that (42) is equivalent to the matrix statement

$$Ay = ay. \qquad (44)$$

Showing the explicit matrix indices, (44) may be written as

$$\sum_k (A_{jk} - a\delta_{jk})y_k = 0. \tag{45}$$

It is a well-known result out of linear algebra that the necessary and sufficient condition that (45) have a nontrivial solution is that

$$\det(A - aI) = 0, \tag{46}$$

where "I" is the unit matrix. Eq$^{\underline{n}}$ (46) is called the characteristic equation of the matrix A. It is easily shown that the left hand side of (46) is just a n$^{\underline{th}}$ order polynomial in "a", and therefore has n solutions or roots. These roots constitute the totality of eigenvalues, $a_i$. After these have been found by solving (46), we find the associated eigenvectors $y^{(i)}$ by solving

$$Ay^{(i)} = a_i y^{(i)}. \tag{47}$$

for each value $a_i$. (We actually went through this process for the spin matrix $\bar{\sigma}_3$ on pgs. 1.85 - 1.87 of these notes.) If the n solutions to (46) are all distinct, then the corresponding eigenkets span the space. However, the solution of (46) may involve repeated roots. A root that is repeated k times in the solution of the characteristic equation is called a k$^{\underline{th}}$ order degeneracy. In this case, we cannot use the previous argument to establish that the nonzero eigenkets or eigenvectors of these repeated roots are all orthogonal. However, it may be shown that the number of

linearly independent eigenvectors corresponding to a k-fold root of the characteristic equation of a Hermitian matrix is exactly k (See Cushing, <u>Applied Analytical Mathematics for Physical Scientists</u>, 1st edition, pp. 108-111.).  Given this fact, this means that there will be exactly n linearly independent eigenkets or eigenvectors, and these can now be used as a basis.

Although we are guaranteed that the number of linearly independent eigenkets of a Hermitian operator is n, this does not mean that any such set will satisfy orthogonality.  Those eigenvectors or eigenkets that correspond to <u>distinct</u> eigenvalues are orthogonal from the previous argument.  We only have to worry about the eigenvectors or eigenkets corresponding to each $k\underline{th}$ order eigenvalue degeneracy.  However, there is a procedure, called the <u>Schmidt orthogonalization process</u>, which allows us to construct an orthogonal set of vectors or kets from any linearly independent set.  Let us say that we have k linearly independent vectors $|y_i>$ corresponding to the single eigenvalue a:

$$A|y_i> = a|y_i> \quad , \; i = 1,\ldots,k \qquad (48)$$

(The order of labeling of these objects is arbitrary.)  We can construct an orthogonal set by the following procedure.

First, choose

$$|x_1> = |y_1>, \qquad (49)$$

as the first of our orthogonal set. $|y_1>$ is actually <u>any</u> of the original set of non-orthogonal kets.

Now form

$$|x_2> = |y_2> - \frac{<x_1|y_2>}{<x_1|x_1>} |x_1>. \qquad (50)$$

Notice that

$$<x_1|x_2> = <x_1|y_2> - \frac{<x_1|y_2>}{<x_1|x_1>} <x_1|x_1> = 0. \qquad (51)$$

Thus, the second term in (50) has been chosen in such a way as to remove the overlap between $|x_1>$ and $|x_2>$. The next vector is

$$|x_3> = |y_3> - \frac{<x_1|y_3>}{<x_1|x_1>} |x_1> - \frac{<x_2|y_3>}{<x_2|x_2>} |x_2>. \qquad (52)$$

Again notice

$$<x_1|x_3> = <x_1|y_3> - \frac{<x_1|y_3>}{<x_1|x_1>} <x_1|x_1> - \frac{<x_2|y_3>}{<x_2|x_2>} <x_1|x_2>$$

$$= 0. \qquad (53)$$

$$<x_2|x_3> = <x_2|y_3> - \frac{<x_1|y_3>}{<x_1|x_1>} <x_2|x_1> - \frac{<x_2|y_3>}{<x_2|x_2>} <x_2|x_2>$$

$$= 0. \qquad (54)$$

In general, the procedure is to pick $|x_1> = |y_1>$, after which we have

$$|x_i> = |y_i> - \sum_{j=1}^{i-1} \frac{<x_j|y_i>}{<x_j|x_j>} |x_j> , \quad i \geq 2. \qquad (55)$$

We can now normalize these, so that they will be orthonormal.

The end result of the above considerations is this:  The eigenkets of $A = A^+$ can all be chosen as orthonormal and therefore represents a possible basis for the space.

The above is a very useful result.  However, the situation is not really satisfactory yet.  We have seen that the eigenvalues of a Hermitian operator do not necessarily specify or classify all the eigenkets of the system, and therefore the states of a particle, because of eigenvalue degeneracy.  In our analogy with vectors in 3-D, this is like labeling the unit vectors identically although we know they are linearly independent.  The process just described which produces an orthogonal set of kets or vectors corresponding to the same eigenvalue has an element of arbitrariness in it because of the random choice $|x_1> = |y_1>$.  It represents a formal mathematical way of producing an orthogonal set of basis kets, but there must be a more physical way of doing the same thing so that known physical properties are associated with each and every ket.  A way of doing this is contained in the following theorem.

**THEOREM**: If $A = A^+$ and $B = B^+$, a necessary and sufficient condition that $[A,B] = 0$ is that A and B posses a common complete set of orthonormal eigenkets.

The proof of sufficiency is as follows.

We are assuming that A and B possess a common complete set of orthonormal basis vectors, which we will label as

$|a_i,b_i>$ (some of the eigenvalues $a_i$ or $b_i$ or both may be degenerate). Thus

$$A|a_i,b_i> = a_i|a_i,b_i>, \tag{56}$$

$$B|a_i,b_i> = b_i|a_i,b_i>. \tag{57}$$

Then

$$AB|a_i,b_i> = b_iA|a_i,b_i> = b_ia_i|a_i,b_i>, \tag{58}$$

$$BA|a_i,b_i> = a_iB|a_i,b_i> = a_ib_i|a_i,b_i>. \tag{59}$$

Therefore

$$[A,B]|a_i,b_i> = 0. \tag{60}$$

This statement holds for all i, which then implies that[1]

$$[A,B] = 0. \tag{61}$$

The proof of necessity in this theorem is a little more tricky, and will not be presented here. (I will ask you, however, to prove a restricted version of the necessary condition in a problem).

Hence, if we have a situation where several of the eigenkets have the same eigenvalue of some Hermitian operator A, these can be chosen as eigenvectors of another Hermitian operator B as long as [A,B] = 0. Then, the eigenvalues of B

---

[1] At a number of points here, and before, we have assumed that $O_1|x> = O_2|x>$ for all $|x>$ implies that $O_1 = O_2$. Can you show this?

may possibly serve to distinguish between them.  If B does
not completely distinguish between them, there may be another
Hermitian operator C with [C,A] = [C,B] = 0 that <u>will</u> do so,
and so on if the eigenkets are still not physically
distinguished from one another.  This is in fact what we
effectively did in characterizing the states of the free
particle in Chapter 2.  Just labeling the states by their
energy value did not completely distinguish between them.
However, by labeling in addition the momentum state we are
able to resolve the double degeneracy of the energy kets.
(We actually only found it necessary to label the sign of the
momentum state).  Of course, we have $[H,p_x] = 0$ so that this
is possible.  We could therefore label the states as $|a',p_x'\rangle$
where

$$H|a',p_x'\rangle = E_{a'}|a',p_x'\rangle \ , \qquad\qquad (62)$$

$$p_x|a',p_x'\rangle = p_x'|a',p_x'\rangle \ . \qquad\qquad (63)$$

A minimal set of Hermitian operators A,B,C,... whose
common complete eigenkets can be characterized uniquely by
their eigenvalues is said to constitute a <u>complete set of
observables</u> (or operators).  Such a minimal set, however, is
not unique.  For example, in distinguishing the energy states
of the free particle, we could also have chosen $\mathbb{P}$, the parity
operator to resolve the 2 fold degeneracy since we can show
that $[H,\mathbb{P}] = 0$.  These would be labeled $|a',P\rangle$, say, with
$P = \pm 1$:

$$H|a',P\rangle = E_{a'}|a',P\rangle \ , \qquad\qquad (64)$$

$$\mathbb{P} \, |a',P> = P|a',P> . \tag{65}$$

The $|a',p_x'>$ and $|a',P>$ are just linear combinations of each other. Notice that we can <u>not</u> choose our eigenkets to be common to $H, p_x$ <u>and</u> $\mathbb{P}$ since $\mathbb{P}$ and $p_x$ do not commute. [In fact we can show that $\{\mathbb{P},p_x\} = 0$, that is, the anti-commutator of $\mathbb{P}$ and $p_x$ is zero.] In this case we have that either $\{H,p_x\}$ or $\{H,\mathbb{P}\}$ constitute a complete set of operators. A common terminology is to call Hermitian operators for which $[A,B] = 0$ "compatible observables." In terms of Process Diagrams, this says that the order in which the operations A and B are carried out is immaterial.

The other side of the coin are Hermitian operators for which $[A,B] \neq 0$. I will now prove an extremely important theorem that will make more concrete some of my prior statements concerning uncertainty relations. Let us say that $|\psi_1>$ and $|\psi_2>$ represent arbitrary kets in some finite dimensional Hilbert space. Let us set

$$|\psi> = |\psi_1> + \lambda|\psi_2>, \tag{66}$$

where $\lambda$ is some arbitrary complex number. By (12) we have that

$$<\psi|\psi> \geq 0, \tag{67}$$

so that

$$(<\psi_1| + \lambda^* <\psi_2|) \cdot (|\psi_1> + \lambda|\psi_2>) \geq 0. \tag{68}$$

Writing this out in full we get

$$\langle \psi_1 | \psi_1 \rangle + |\lambda|^2 \langle \psi_2 | \psi_2 \rangle + 2\text{Re}[\lambda \langle \psi_1 | \psi_2 \rangle] \geq 0, \tag{69}$$

where "Re" means the real part of the argument. Since (69) must be true for any $\lambda$, we may choose

$$\lambda = - \frac{\langle \psi_2 | \psi_1 \rangle}{\langle \psi_2 | \psi_2 \rangle}, \tag{70}$$

assuming that $|\psi_2\rangle$ is a non zero ket. Substitution of (70) into (69) yields

$$\langle \psi_1 | \psi_1 \rangle + \frac{|\langle \psi_2 | \psi_1 \rangle|^2}{\langle \psi_2 | \psi_2 \rangle} + 2\text{Re}\left[ - \frac{|\langle \psi_2 | \psi_1 \rangle|^2}{\langle \psi_2 | \psi_2 \rangle} \right] \geq 0. \tag{71}$$

But the argument of Re <u>is</u> purely real, so (71) implies that

$$\langle \psi_1 | \psi_1 \rangle \geq \frac{|\langle \psi_2 | \psi_1 \rangle|^2}{\langle \psi_2 | \psi_2 \rangle}, \tag{72}$$

or

$$\langle \psi_1 | \psi_1 \rangle \langle \psi_2 | \psi_2 \rangle \geq |\langle \psi_2 | \psi_1 \rangle|^2. \tag{73}$$

Eq$^{\underline{n}}$ (73) is called the <u>Schwartz inequality</u>. Let us let $|\psi_1\rangle$ and $|\psi_2\rangle$ in (73) be given by

$$(A - \langle A \rangle_\psi)|\psi\rangle = |\psi_1\rangle, \tag{74}$$

$$(B - \langle B \rangle_\psi)|\psi\rangle = |\psi_2\rangle. \tag{75}$$

where A and B are Hermitian and the expectation values $<A>_\psi$ and $<B>_\psi$ are as usual given by

$$<A>_\psi = <\psi|A|\psi>, \tag{76}$$

$$<B>_\psi = <\psi|B|\psi>. \tag{77}$$

It is understood in (74) and (75) that in an operator context $<A>_\psi = <A>_\psi \cdot 1$ where "1" is the unit operator. We choose $|\psi>$ to be normalized, i.e. $<\psi|\psi> = 1$. Notice that

$$<\psi_1|\psi_1> = (<\psi|(A - <A>_\psi))((A - <A>_\psi)|\psi>)$$

$$= <\psi|(A - <A_\psi>)^2|\psi>. \tag{78}$$

Similarly for $<\psi_2|\psi_2>$. We now get that

$$<\psi|(A - <A>_\psi)^2|\psi><\psi|(B - <B>_\psi)^2|\psi> \geq$$

$$|<\psi|(A - <A>_\psi)(B - <B>_\psi)|\psi>|^2. \tag{79}$$

The quantities $<\psi|(A - <A>_\psi)^2|\psi> = <\psi_1|\psi_1>$ and $<\psi|(B - <B>_\psi)^2|\psi> = <\psi_2|\psi_2>$ are intrinsically positive or zero. We now define the uncertainties in the operators A and B in the state $|\psi>$ (same as the definition in a problem from Chapter 2)

$$\Delta A \equiv \sqrt{<\psi|(A - <A>_\psi)^2|\psi>} \geq 0. \tag{80}$$

$$\Delta B \equiv \sqrt{<\psi|(B - <B_\psi>)^2|\psi>} \geq 0. \tag{81}$$

Since

$$\langle \psi | \langle A \rangle_\psi | \psi \rangle \; = \; \langle A \rangle_\psi \; \langle \psi | \psi \rangle \; = \; \langle A \rangle_\psi , \qquad (82)$$

we may also write

$$\left.\begin{array}{l} \Delta A \; \equiv \; \sqrt{\langle A^2 \rangle_\psi \; - \; \langle A \rangle_\psi^2} \; , \\[12pt] \Delta B \; \equiv \; \sqrt{\langle B^2 \rangle_\psi \; - \; \langle B \rangle_\psi^2} \; . \end{array}\right\} \qquad (83)$$

Now let us deal with the right hand side of Eq$^{\underline{n}}$ (79).  We have

$$\langle \psi | (A \; - \; \langle A \rangle_\psi)(B \; - \; \langle B \rangle_\psi) | \psi \rangle$$

$$= \; \langle \psi | (AB \; - \; \langle A \rangle_\psi B \; - \; A \langle B \rangle_\psi \; + \; \langle A \rangle_\psi \langle B \rangle_\psi) | \psi \rangle$$

$$= \; \langle \psi | (AB \; - \; \langle A \rangle_\psi \langle B \rangle_\psi) | \psi \rangle \; . \qquad (84)$$

By adding and subtracting the quantity $\frac{1}{2} BA$, we may also write

$$\langle \psi | (AB \; - \; \langle A \rangle_\psi \langle B \rangle_\psi) | \psi \rangle$$

$$= \; \langle \psi | \left( \frac{1}{2} (AB + BA) \; + \; \frac{1}{2} (AB \; - \; BA) \; - \; \langle A \rangle_\psi \langle B \rangle_\psi \right) | \psi \rangle$$

$$= \; \langle \psi | \left( \frac{1}{2} \{A,B\} \; - \; i \left( \frac{i}{2} [A,B] \right) \; - \; \langle A \rangle_\psi \langle B \rangle_\psi \right) | \psi \rangle , \qquad (85)$$

where we have introduced both the commutator, $[A,B]$, and the anticommutator, $\{A,B\}$, of A and B.  Let us define the new operators

$$x \; = \; \frac{1}{2} \{A,B\} \; - \; \langle A \rangle_\psi \langle B \rangle_\psi \; , \qquad (86)$$

$$y = \frac{i}{2} [A,B]. \tag{87}$$

Since A and B are Hermitian, it is easy to show that x and y are also Hermitian.  Therefore, our inequality, Eq$^{n}$ (79), now reads

$$(\Delta A)^2 (\Delta B)^2 \geq \left| <\psi|x|\psi> - i<\psi|y|\psi> \right|^2. \tag{88}$$

Since x and y are Hermitian, we have that the expectation values

$$<\psi|x|\psi> = <x>_\psi , \tag{89}$$

$$<\psi|y|\psi> = <y>_\psi , \tag{90}$$

are real numbers.  Therefore we must have that

$$\left| <x>_\psi - i<y>_\psi \right|^2 = <x>_\psi^2 + <y>_\psi^2 \geq <y>_\psi^2 , \tag{91}$$

the last inequality only holding if $<x>_\psi = 0$.  Using (91) in (88) now gives us that

$$(\Delta A)^2 (\Delta B)^2 \geq <y>_\psi^2 = <\frac{i}{2} [A,B]>_\psi^2 . \tag{92}$$

Since both sides of (92) are intrinsically positive or zero, this means that

$$\Delta A \Delta B \geq \left| <\frac{i}{2} [A,B]>_\psi \right|, \tag{93}$$

where the absolute value sign is used on the right hand side

because while $<y>_\psi = <\frac{i}{2} [A,B]>_\psi$ is guaranteed real, it may

not be positive. Eq$^{\underline{n}}$ (93) says that if the Hermitian

operators A and B are not compatible, there will in general

be an uncertainty relation connecting them. Eq$^{\underline{n}}$s (73) and

(93) were derived in a situation where the dimensionality of

the Hilbert space is finite, but they also hold where the

dimensionality increases without limit. In that case, we can

apply (93) to the incompatible observables x and $p_x$, which

tells us that

$$\Delta x \Delta p_x \geq \frac{\hbar}{2} \; , \tag{94}$$

since $[x,p_x] = i\hbar$. Eq$^{\underline{n}}$ (94) is the promised relation Eq$^{\underline{n}}$

(32) of Chapter 2. It is clear that we cannot use this type

of derivation to establish the energy-time uncertainty

relation since the time is a parameter, not an operator, as I

also pointed out in Ch.3 of these notes.

One may try to turn this argument around. That is, given

operators for which

$$\Delta A \Delta B = 0 \; ( \Rightarrow <\psi|[A,B]|\psi> = 0) \; , \tag{95}$$

(these are called simultaneous observables) for all states

$|\psi>$, does this imply that A and B are compatible, i.e., that

$[A,B] = 0$? The answer to this question is affirmative. We

can show this as follows. Let us let $\mathcal{O} = AB - BA$ and let

$|x_i>$, $|x_j>$ be any elements of a basis which spans the space.

Now we know that (95) holds for any state $|\psi>$. Now consider

$$|\psi> = |x_i> + |x_j> . \tag{96}$$

Substituting (96) into (95) tells us that

$$<x_i|\mathbb{O}|x_i> + <x_j|\mathbb{O}|x_j> + <x_i|\mathbb{O}|x_j> + <x_j|\mathbb{O}|x_i> = 0. \tag{97}$$

But the first two terms of (97) are zero because of (95).
Thus, since we may write

$$<x_j|\mathbb{O}|x_i> = <x_i|\mathbb{O}^+|x_j>^*, \tag{98}$$

and

$$\mathbb{O}^+ = (AB - BA)^+ = (BA - AB) = -\mathbb{O} , \tag{99}$$

we have from (97) that

$$<x_i|\mathbb{O}|x_j> - <x_i|\mathbb{O}|x_j>^* = 0, \tag{100}$$

which is the same as saying

$$Im(<x_i|\mathbb{O}|x_j>) = 0. \tag{101}$$

Likewise, consider

$$|\psi> = |x_i> + i|x_j>. \tag{102}$$

Substituting (102) into (95) tells us

$$i<x_i|\mathbb{O}|x_j> - i<x_j|\mathbb{O}|x_i> = 0, \tag{103}$$

which, with the use again of (98) and (99), says

$$i(<x_i|\mathbb{O}|x_j> + <x_i|\mathbb{O}|x_j>^*) = 0, \tag{104}$$

or

$$Re(< x_i | \emptyset | x_j >) = 0. \tag{105}$$

Eq$^n$s. (101) and (105) together imply that

$$< x_i | \emptyset | x_j > = 0, \tag{106}$$

for all i,j, so that

$$\emptyset = [A,B] = 0. \tag{107}$$

Thus, the statement that $\Delta A \Delta B = 0$ for all $|\psi>$, given that A and B are Hermitian, is the same as saying $[A,B] = 0$. Another way of stating this is: two Hermitian operators are simultaneously measurable for any state if and only if they commute. However, if (95) holds just for some particular states $|\psi>$, then $AB \neq BA$ in general. Given that $[A,B] \neq 0$, the $|\psi>$ for which the equality holds in (93) are minimum uncertainty states called <u>coherent states</u>. We saw an example of such a state in the 1-D Gaussian wavepacket of Chapter 2.

I have pointed out an analogy between Hilbert space and our ordinary 3-D world. I have said that a normalized basis, $\{|x_i>\}$, is like a set of unit vectors in ordinary space. Now the description of a general vector in terms of orthogonal unit vectors is not unique. There is always the freedom of a <u>choice</u> of basis. A different choice description cannot, of course, change the length of a vector. If we let $x_i$ represent the projections of an arbitrary vector upon three orthogonal directions and $\bar{x}_i$ represent the projections of the

same vector upon another set of mutually orthogonal 3-D unit vectors, then we must have

$$\sum_i \bar{x}_i^2 = \sum_i x_i^2 . \tag{108}$$

The transformation equations relating $\bar{x}$ to $x$ can be written as

$$\left.\begin{aligned}
\bar{x}_1 &= x_1\lambda_{11} + x_2\lambda_{21} + x_3\lambda_{31} , \\
\bar{x}_2 &= x_1\lambda_{12} + x_2\lambda_{22} + x_3\lambda_{32} , \\
\bar{x}_3 &= x_1\lambda_{13} + x_2\lambda_{23} + x_3\lambda_{33} ,
\end{aligned}\right\} \tag{109}$$

or more compactly as

$$\bar{x}_i = \sum_j x_j\lambda_{ji} . \tag{110}$$

The requirement (108) says that

$$\sum_{i,j,k} \lambda_{ji}\lambda_{ki}x_jx_k = \sum_i x_i^2 , \tag{111}$$

which is only satisfied if

$$\sum_i \lambda_{ji}\lambda_{ki} = \delta_{jk} , \tag{112}$$

where $\delta_{jk}$ is the Kronecker delta symbol. In matrix notation (110) and (112) read

$$\bar{x} = x\lambda, \tag{113}$$

and

$$\lambda\lambda^{T} = 1, \tag{114}$$

respectively. Comparing (114) with the definition of $\lambda^{-1}$,

$$\lambda\lambda^{-1} = 1 \tag{115}$$

means that

$$\lambda^{T} = \lambda^{-1}, \tag{116}$$

and (114) may also be written as

$$\lambda^{T}\lambda = 1. \tag{117}$$

Any nonsingular transformation that satisfies (116) preserves
the length of vectors. This includes both rotations and
inversions of the coordinates.

Now let us attempt to do the same thing for state
vectors in Hilbert space. We have seen that the quantity
$<\psi|\psi>$ is a real, positive quantity for non null vectors $|\psi>$
(Eq[n] (12) above). Interpreting this quantity as the square of
the "length", we require that

$$<\bar{\psi}|\bar{\psi}> = <\psi|\psi> \tag{118}$$

under a change of basis. The analog of the linear
transformation (113) is

$$<\bar{\psi}| = <\psi|U \tag{119}$$

where U is an operator in the Hilbert space. Eq[n] (118) now
gives us that

$$\langle\psi|UU^{+}|\psi\rangle = \langle\psi|\psi\rangle \qquad (120)$$

similar to (111) above. The requirement that (120) hold for all states $|\psi\rangle$ then results in

$$UU^{+} = 1 \qquad (121)$$

(The reasoning that yields (121) from (120) is essentially the same as the way we showed $[A,B] = 0$ follows from $\langle\psi|[A,B]|\psi\rangle = 0$.) Comparison of (121) with

$$UU^{-1} = 1 \qquad (122)$$

tells us that

$$U^{+} = U^{-1} \, , \qquad (123)$$

similar to (116). Therefore, we may also write (121) as

$$U^{+}U = 1. \qquad (124)$$

We recognize (123) as the definition of a <u>unitary transformation</u>, first seen in the discussion of spin $\frac{1}{2}$ in Chapter 1. Thus, what was seen there as a special case is revealed as being general. A unitary transformation describes a change of basis that preserves the length of state vectors in Hilbert space. Because of the strong analogy between real space vectors and Hilbert space vectors, I will sometimes refer to (119) as a rotation in Hilbert space.

We have been regarding (113) as describing the point of view where the $\bar{x}_i$ are the components of a fixed vector in a rotated coordinate system.  This is called a <u>passive</u> rotation of the vector.  However, an equally valid interpretation of (113) is that the $\bar{x}_i$ represent the components of a rotated vector in a fixed coordinate system, provided this rotation is taken in the opposite direction to the passive one.  This is called an <u>active</u> rotation of the vector.  We have also been regarding the analogous Hilbert space statement, Eq$^{\underline{n}}$ (119), in a passive sense.  That is, we have taken (119) as describing a situation where the bra vector is fixed but the basis is rotated.  Just as for real space vectors, however, we could just as well view (119) as an active rotation on the bra in the opposite direction to the passive one.  I am using the passive terminology here mainly to connect smoothly with the discussion of unitary transformations in Chapter 1.  You should be aware that either point of view is equally valid. We will use the point of view that is most convenient at the time.  Once we have chosen an interpretation, however, we must strive for consistency.

Of course, completeness and orthonormality are preserved under a unitary transformation.  That is, given

$$\sum_i |x_i> <x_i| = 1, \tag{125}$$

and

$$<x_i|x_j> = \delta_{ij}, \tag{126}$$

4.30

we have that

$$\sum_i |\bar{x}_i><\bar{x}_i| = U^+ \left( \sum_i |x_i><x_i| \right) U = 1, \qquad (127)$$

and

$$<\bar{x}_i|\bar{x}_j> = <x_i|UU^+|x_j> = \delta_{ij}. \qquad (128)$$

Hermiticity of operators is also preserved under unitary transformations. Given the rotation, either passive or active, on the bra vectors as in (119), the same rotation applied to an arbitrary operator A is

$$\bar{A} = U^+AU. \qquad (129)$$

(The rotation in the opposite direction is given by $UAU^+$.) If $A = A^+$, then

$$(\bar{A})^+ = (U^+AU)^+ = U^+AU = \bar{A}. \qquad (130)$$

Also, given $A|a'> = a'|a'>$, we have

$$\bar{A}|\bar{a}'> = U^+AU(U^+|a'>) = a'|\bar{a}'>, \qquad (131)$$

and

$$<\bar{\psi}|\bar{A}|\bar{\psi}> = <\psi|UU^+AUU^+|\psi> = <\psi|A|\psi>, \qquad (132)$$

so that the eigenvalues and expectation values are also unchanged.

It is a theorem that can be proven that any unitary operator U can be written as

$$U = e^{iA}, \tag{133}$$

where $A = A^{+}$. (See Merzbacher, 2$\underline{nd}$ ed., p.323).  The operator $e^{x}$ is defined as

$$e^{x} = 1 + x + \frac{x^2}{2} + \ldots = \sum_{n=0}^{\infty} \frac{x^n}{n!} \,. \tag{134}$$

From (134) it is easy to see that

$$(e^{x})^{+} = e^{x^{+}}, \tag{135}$$

from which (133) gives us

$$UU^{+} = e^{iA}(e^{-iA}) = 1, \tag{136}$$

as it should.  [By the way, in general we have

$$e^{A}e^{B} \neq e^{A+B}, \tag{137}$$

when A and B do not commute.  To convince yourself of this, just expand both sides in powers of the arguments of the exponents.  There is an equality sign in (137) only in general if A and B commute.]

    We can use unitary transformations to represent a change in basis due to coordinate displacements, rotations (in 2 or more spatial dimensions) and momentum boosts.  In addition, we will see that these transformations also provide an alternate means of viewing the time development of a quantum system.

    As the simplest of these possibilities, let us consider the unitary representation of coordinate displacements.  (We

will deal with the equally simple case of velocity boosts in a problem.)  The appropriate operator is

$$U = e^{-ix'p_x/\hbar} \, , \tag{138}$$

where x' is a number (with dimensions of length) and $x$ and $p_x$ are the usual position and momentum operators.  This unitary transformation cannot be represented by a finite matrix since the number of eigenvalues x' of $x|x'> = x'|x'>$ is infinite. We will not worry about this subtlety and will treat it as if the space were finite.  Let us first consider the quantity

$$\bar{x} = U^{+}xU, \tag{139}$$

with U given by (138).  In order to find the effect of U on $x$, let us construct a differential equation for $\bar{x}$.  We have that

$$\frac{\hbar}{i} \frac{d\bar{x}}{dx'} = e^{ix'p_x/\hbar} [p_x x - x p_x] e^{-ix'p_x/\hbar} \, , \tag{140}$$

where we see that the commutator $[p_x, x] = \dfrac{\hbar}{i}$ has arisen.  Then we have

$$\frac{\hbar}{i} \frac{d\bar{x}}{dx'} = \frac{\hbar}{i} e^{ix'p_x/\hbar} e^{-ix'p_x/\hbar} = \frac{\hbar}{i} \, , \tag{141}$$

$$\Rightarrow \frac{d\bar{x}}{dx'} = 1. \tag{142}$$

The solution to (142) is simply

$$\bar{x} = x' + \mathcal{O} \, , \tag{143}$$

where $\emptyset$ is an unknown "constant" independent of x'. Both sides of (143) actually have an operator character. We should understand x' in (143) to mean x'1, where 1 is the unit operator. The value of the operator $\emptyset$ is specified by letting x' = 0 in (139), which means that

$$\bar{x}(x' = 0) = x, \tag{144}$$

which implies that

$$\emptyset = x. \tag{145}$$

Therefore, we have that

$$x + x' = e^{ix'p_x/\hbar}\, xe^{-ix'p_x/\hbar}\ . \tag{146}$$

The effect of this unitary transformation on coordinate states can be found as follows. We know from (146) that

$$e^{ix''p_x/\hbar}\, x = (x + x'')e^{ix''p_x/\hbar}\ . \tag{147}$$

Therefore

$$(<x'|e^{ix''p_x/\hbar}\ )x = <x'|(x + x'')e^{ix''p_x/\hbar}$$
$$= (<x'|e^{ix''p_x/\hbar})(x' + x''). \tag{148}$$

This means that

$$<x'|e^{ix''p_x/\hbar} = C<x' + x''|, \tag{149}$$

where C is some proportionality factor.  Because unitary

transformations preserve length in Hilbert space, we must

have $|C|^2 = 1$.  We may then <u>choose</u>

$$\langle x' | e^{ix''p_x/\hbar} = \langle x' + x'' | . \tag{150}$$

Therefore, the unitary operator $e^{ix''p_x/\hbar}$ generates a spatial

translation when acting on bra coordinate states.

Using (150) we can recover our previous result for

$\langle x' | p_x$.  To first order in $\delta x''$ we have

$$\langle x' | \left( 1 + \frac{\delta x'' p_x}{\hbar} \right) = \langle x' + \delta x'' | . \tag{151}$$

This gives

$$\langle x' | p_x = \frac{\langle x' + \delta x'' | - \langle x' |}{\delta x''} . \tag{152}$$

Taking the limit of both sides of (152) as $\delta x'' \rightarrow 0$ now yields

$$\langle x' | p_x = \frac{\hbar}{i} \frac{\partial}{\partial x'} \langle x' | , \tag{153}$$

the same as Eq$^{\underline{n}}$ (166) of Chapter 2.

As our second example of a unitary transformation,

consider

$$U(t) = e^{iHt/\hbar} . \tag{154}$$

We saw this operator (or rather, its adjoint) in Chapter 2 where it was termed the time evolution operator.  We saw there that

$$|a',t> = e^{-iHt/\hbar}|a'>$$
(155)

where $|a'>$ is an energy eigenstate of H at t = 0 and $|a',t>$ is the time evolved state.  The original discussion was for the case of a free particle.  However, the 1-D Schrödinger equation is also separable in space and time for Hamiltonians including a potential as in Eq$^{\underline{n}}$ (3) of Chapter 3.  This means that (155) describes the time evolution of energy eigenstates in this case as well.

Let me take this opportunity to point out another way of dealing with time evolution using the unitary operator U(t) above.  Let us consider the expectation value of some physical property A at time t.  This is given by

$$<A>_{\psi,t} = <\psi,t|A|\psi,t> = <\psi|U(t)AU^{+}(t)|\psi>,$$
(156)

with U(t) given above.  This point of view assigns the time evolution to the states.  However, from (156) we see that it is equally valid to assign all time dependence to the operator A.  That is, we may set

$$<A>_{\psi,t} = <\psi|A(t)|\psi>,$$
(157)

where

$$A(t) \equiv U(t)AU^{+}(t).$$
(158)

The problem now is to find the dynamical equation satisfied

by A(t) so that its time behavior can be determined.

Consider therefore the time derivative of A(t).  One finds

that

$$\frac{\hbar}{i} \frac{dA(t)}{dt} = U(t)[HA - AH]U^+(t) + U(t) \frac{\partial A}{\partial t} U^+(t). \qquad (159)$$

The three terms in (159) come from the time dependence in

$U(t), U^+(t)$ and a possible explicit time dependence in the

operator A.  Now we may write

$$U(t)[HA - AH]U^+(t) = [HA(t) - A(t)H] = [H,A(t)], \qquad (160)$$

since in the simple case we are studying, U(t) and H commute.

The most common case is where $\frac{\partial A}{\partial t} = 0$.  Then (159) becomes

$$\frac{\hbar}{i} \frac{dA(t)}{dt} = [H,A(t)]. \qquad (161)$$

This is the operator equation of motion satisfied by A(t).

Eq$^n$ (161) is called the Heisenberg equation of motion.

Taking the expectation value of both sides of (161), we learn

that

$$\frac{\hbar}{i} \frac{d< A(t) >_\psi}{dt} = < [H,A(t)] >_\psi, \qquad (162)$$

in an arbitrary state $|\psi>$.

Notice that if [H,A] = 0, this implies [H,A(t)] = 0

because U(t) and H commute.  Therefore, when H and A commute,

A is a <u>constant of the motion</u>.  In this case (162) gives rise
to

$$\frac{d< A(t) >_\psi}{dt} = 0.$$ 
<div align="right">(163)</div>

This says that expectation values of A are a constant in
time.  This makes sense because if A is assumed to commute
with H then it is clear that

$$< A(t) >_\psi = < \psi|U(t)AU^+(t)|\psi > = < \psi|A|\psi > = < A >_\psi,$$

for all t.  In particular, if at t = 0 the wavefunction is an
eigenvector of A with eigenvalue a', this will continue to
hold true at a later time t. a' is called a <u>good</u> <u>quantum</u>
<u>number</u> and A can be chosen as one of the complete set of
observables that characterize eigenvectors.  We previously
saw an example of a constant of the motion in the parity
operator, $\mathbb{P}$ , in the case of the infinite square well and the
simple harmonic oscillator.  A state with a given $< \mathbb{P} >_\psi$
value will keep this quantity fixed in time.

The representation for U(t) in (154) is in general only
true when the Hamiltonian is not an explicit function of the
time. Other forms for the evolution operator hold when
H = H(t).  (See Sakurai's discussion in "Modern Quantum
Mechanics" on pgs. 72 and 73).  We will not deal with time
dependent Hamiltonians here.

We thus have an alternate and equivalent way of viewing
the time dynamics of quantum systems.  Previously, we were

taking the operators as static and viewing the states $|\psi>$ as evolving either actively or passively in time.  This point of view leads to the Schrödinger equation and is called the Schrödinger picture.  We have now learned that we may instead time-evolve the operators actively or passively and let the states $|\psi>$ be static.  This point of view leads to the Heisenberg equations of motion and is called the Heisenberg picture.

As an example of the use of the Heisenberg picture, let us re-examine the free Gaussian wavepacket of Chapter 2.  The peak of the Gaussian distribution $|\psi_g(x,t)|^2$ spread with time but maintained its shape.  It's easy to check that the expectation values of position and momentum in the Schrödinger picture are

$$< x >_{\psi,t} \; = \; \frac{\bar{p}}{m} \; t, \tag{164}$$

$$< p_x >_{\psi,t} \; = \; \bar{p}, \tag{165}$$

from which we have that

$$< x >_{\psi,t} \; = \; \frac{< p_x >_{\psi,t}}{m} \; t. \tag{166}$$

Let us try to recover (166) from the Heisenberg picture.  We have that

$$\frac{\hbar}{i} \; \frac{dx(t)}{dt} \; = \; [H,x(t)] \; = \; U(t) \left[ \frac{p_x^2}{2m} \; , x \right] U^+(t). \tag{167}$$

The commutator in (167) is

$$\left[\frac{p_x^2}{2m},x\right] = \frac{1}{2m}\{p_x[p_x,x] + [p_x,x]p_x\} = \frac{\hbar}{i}\frac{p_x}{m} . \qquad (168)$$

We now find

$$\frac{\hbar}{i}\frac{dx(t)}{dt} = \frac{\hbar}{i}U(t)\frac{p_x}{m}U^+(t) = \frac{\hbar}{i}\frac{p_x(t)}{m} , \qquad (169)$$

so we have the <u>operator</u> statement

$$\frac{dx(t)}{dt} = \frac{p_x(t)}{m} . \qquad (170)$$

Of course, we also have that

$$\frac{\hbar}{i}\frac{dp_x(t)}{dt} = U(t)\left[\frac{p_x^2}{2m},p_x\right]U^+(t) = 0, \qquad (171)$$

so that

$$\frac{dp_x(t)}{dt} = 0, \qquad (172)$$

and $p_x(t)$ is a constant operator in time $(p_x(t)=p_x)$.

Therefore from (170) we get by integration

$$x(t) = \frac{p_x}{m}t + C, \qquad (173)$$

where C is a constant operator.  We know from

$x(t) = U(t)xU^+(t)$ that $x(0) = x$, so $C = x$.  Taking the

expectation value of both sides of (173) finally gives us

$$<x(t)>_\psi = \frac{<p_x>_\psi}{m}t, \qquad (174)$$

if $<x>_\psi = 0$.  In (166) the time dependence of the

expectation values is in the state, whereas in (174) the time

dependence is in the operator.  We get identical results

either way.  Notice that the relation (166) was derived for a

particular wave packet, whereas (176) shows that this

relation holds for any particle wave packet as long as $<x>_\psi$

$= 0$.

The commutation properties of operators are preserved in

the Heisenberg picture if these are interpreted as <u>equal time</u>

relations.  For any unitary transformation U(t) we have for

example

$$U(t)[x,p_x]U^+(t) = i\hbar,$$

$$\Rightarrow \quad U(t)xp_xU^+(t) - U(t)p_xxU^+(t) = i\hbar,$$

$$\Rightarrow \quad x(t)p_x(t) - p_x(t)x(t) = = i\hbar,$$

$$\Rightarrow \quad [x(t),p_x(t)] = i\hbar. \tag{175}$$

The above application of the Heisenberg picture was to

the case of free particles.  We can go a step beyond this in

applying this formalism in the case where an unspecified

potential is present.  Let us let

$$H = \frac{p_x^2}{2m} + V(x), \tag{176}$$

Then we have

$$\frac{\hbar}{i} \frac{dp_x(t)}{dt} = U(t)[H,p_x]U^+(t), \qquad (177)$$

$$[H,p_x] = [V(x),p_x]. \qquad (178)$$

From Eq$^{\underline{n}}$ (80) of Chapter 2 recall that

$$[p_x,f(x)] = \frac{\hbar}{i} \frac{df(x)}{dx} , \qquad (179)$$

which is an operator statement. For the commutator in (178) we have therefore

$$[V(x),p_x] = - \frac{\hbar}{i} \frac{dV(x)}{dx} , \qquad (180)$$

from which we have

$$\frac{dp_x(t)}{dt} = -U(t) \frac{dV(x)}{dx} U^+(x). \qquad (181)$$

We also have that

$$\frac{\hbar}{i} \frac{dx(t)}{dt} = U(t)[H,x]U^+(t) = \frac{1}{2m} U(t)[p_x^2,x]U^+(t), \qquad (182)$$

where $[p_x^2,x] = \frac{2\hbar}{i} p_x$. Therefore

$$\frac{dx(t)}{dt} = \frac{p_x(t)}{m} , \qquad (183)$$

the same as (170) above. Taking another derivative in (183) now gives us

$$\frac{d^2x(t)}{dt^2} = \frac{1}{m} \frac{dp_x(t)}{dt} = - \frac{1}{m} U(t) \frac{dV(x)}{dx} U^+(t), \qquad (184)$$

where we have used (181).  If V(x) is a power series in x, then

$$U(t) \frac{dV(x)}{dx} U^+(t) = \frac{dV(x(t))}{dx(t)} \tag{185}$$

and we have

$$m \frac{d^2x(t)}{dt^2} = - \frac{dV(x(t))}{dx(t)}, \tag{186}$$

which has the appearance of Newton's second law, but written for operators.  We may, if we wish, take the expectation value of both sides of (186) in an arbitrary state.  This yields

$$m \frac{d^2 <x(t)>_\psi}{dt^2} = - \left< \frac{dV(x(t))}{dx(t)} \right>_\psi. \tag{187}$$

The analogous statement in the Schrödinger language writes this as

$$m \frac{d^2 <x>_{\psi,t}}{dt^2} = - \left< \frac{dV(x)}{dx} \right>_{\psi,t}. \tag{188}$$

Eq$^{\underline{n}}$ (187) or (188) is called <u>Ehrentest's theorem</u>.  In words, it says that the expectation value of the position operator moves like a classical particle subjected to a "force" given by $- \left< \frac{dV}{dx} \right>$ .

# Problems

1. <u>Prove</u>:  In a finite Hilbert space of dimensionality N, if the N eigenvectors of A can all be chosen orthonormal and if the eigenvectors are all real, then $A = A^+$ (it is Hermitian).

2. <u>Prove</u>:  (Restricted form of the necessary condition for the theorem on p.4.16, p.4.17 of the notes.)  If A and B are Hermitian and the eigenvalues of A (or B) are all distinct, and if $[A,B] = 0$, then A and B possess a common set of orthonormal eigenkets.

3. Show for the parity operator, $\mathbb{P}$, that ("{}" is the anti-commutator)

   (a)  $\{\mathbb{P},p_x\} = 0$.

   (b)  $\{\mathbb{P},x\} = 0$.

   (c)  $\mathbb{P}\mathbb{P}^+ = 1$.

[Hint:  $\mathbb{P}$ is defined by

$$\mathbb{P}|x'> = |-x'>.$$

An explicit representation of $\mathbb{P}$ is therefore

$$\mathbb{P} = \int dx' \; |-x'><x'|.]$$

4. <u>Prove</u>:  If $<\Psi|A|\Psi>$ is real for all $|\Psi>$, then $A = A^+$. (A is Hermitian.)

5. <u>Prove</u>:

$$|<\Psi_1|A|\Psi_2>|^2 \leq <\Psi_1|A|\Psi_1> <\Psi_2|A|\Psi_2>$$

4.44

for any $A = A^+$ such that $\langle\Psi|A|\Psi\rangle \geq 0$ for all $|\Psi\rangle$.


**Other Problems**

6.(a) For a free particle ($V(x)=0$) in one dimension, show that there is an uncertainty relation connecting x (the position operator) and x(t) (the Heisenberg time-evolved position operator):

$$\Delta x \; \Delta x(t) \; \geq \; \left|\frac{\hbar t}{2m}\right|.$$

(b) Express in words what this uncertainty relation means in terms of experiments which measure a particle's position.

7.(a) Show that (set $\hbar=1$ here for simplicity)

$$e^{-ip'_x x} \; p_x \; e^{ip'_x x} \; = \; p_x + p'_x,$$

where x and $p_x$ are the usual position, momentum operators, and $p'_x$ is a (real) number. [Hint: Construct a differential equation in $p'_x$.]

(b) Use part (a) to argue that

$$\langle p'_x|x \; = \; i\frac{\partial}{\partial p'_x} \; \langle p'_x|.$$


8. One thousand neutrons are in a one-dimensional box, with walls at x=0, x=L. At t=0, the particle's wavefunction is

$$\psi(x,0) \; = \; A \left(\sin(\frac{\pi x}{L}) \; - \; \frac{1}{10} \sin(\frac{2\pi x}{L})\right).$$

(a) Normalize $\psi(x,0)$ and find the value of the constant A.

(b) How many particles are in the interval $(0, \frac{1}{2} L)$ at t=0?

(c) How many particles have energy $E_2$ $(= \frac{4\hbar^2\pi^2}{2mL^2})$ at t=0?

(d) Find $\psi(x,t)$ for t>0.

9. (a) Show that

$$e^A e^B = e^{A+B} e^{1/2[A,B]},$$

given that A and B each commutes with [A,B]. [Hint: In general, we have

$$e^A B e^{-A} = B + [A,B] + \frac{1}{2!} [A,[A,B]] + \frac{1}{3!} [A,[A,[A,B]]] + \cdots .]$$

(b) Use (a) to show that we may write ($\hbar = 1$ again)

$$[e^{iax}, e^{ibp_x}] = -2ie^{i(ax+bp_x)}\sin(\frac{ab}{2}),$$

where $p_x$ and x are operators; x', a, and b are numbers.

# Chapter 5:   Two Static Approximation Methods

It is in the nature of physical systems that the closer they are observed, the more detail there is to see;  the closer we come to mathematically describing these additional details, the more intricate our mathematical considerations often then become; the more intricate the mathematical description is, the more difficult it usually becomes to solve our system of equations in some <u>exact</u> analytical way. Sometimes, the additional physical details we wish to incorporate into the theory are sufficiently small compared to other quantities already in the theory that their incorporation does not involve an entirely new solution or an entirely new starting point but can be treated as a perturbation of the old solution or starting point.  What I want to talk about here are two methods for carrying out this program.  The first technique, called <u>time independent</u> <u>perturbation</u> <u>theory</u> or <u>Rayleigh-Schrödinger</u> <u>perturbation</u> <u>theory</u> is mainly useful when additional small time-independent interactions are added to a system for which an exact analytical solution is already available.  This is often the case in the intricate interactions that occur in atomic and nuclear physics.  The results that it gives are very general and are not necessarily limited to a nonrelativistic domain. The other technique, called the <u>WKB</u> or <u>JWKB semiclassical approximation</u>, is based upon the use of a "classical" version of the Schrödinger equation as a new

starting point for the description of situations where particles are subjected to slowly varying potentials in space.  The treatment of time independent perturbation theory is independent of any assumptions concerning our space dimensionality. However, we will maintain our limitation to a single spatial dimension for purposes of simplicity in the WKB method discussion, although that single dimension could be, for example, the radial coordinate in a three dimensional space, say.

Time dependent perturbation theory is also an important subject in quantum mechanics. It comes into play mainly when transitions from one state to another or scattering are considered.  Unfortunately, we will not have time to cover this important topic here.

To begin with, let us consider a Hamiltonian that is dependent on some real parameter, $\lambda$. (It could, for example, represent an interaction with an electric or magnetic field.) The basic eigenvalue-eigenvector statement is given by

$$(H(\lambda) - E(\lambda)|E\lambda> = 0, \tag{1}$$

where in general the energies and the states also have a $\lambda$ dependence. ($\lambda$ is simply a parameter in $|E\lambda>$, not an eigenvalue.) Taking the derivative of (1) with respect to $\lambda$ gives

$$(H - E) \frac{\partial}{\partial \lambda} |E\lambda> + \left( \frac{\partial H}{\partial \lambda} - \frac{\partial E}{\partial \lambda} \right)|E\lambda> = 0. \tag{2}$$

Now project both terms in (2) into the state $< E\lambda|$. Since we know that $< E\lambda|(H - E) = 0$, we get (Feynman–Hellman theorem)

$$< E\lambda| \ \frac{\partial H}{\partial \lambda} \ |E\lambda> \ = \ \frac{\partial E}{\partial \lambda} \ , \tag{3}$$

which is an exact statement. This equation is useful on occasions when $E(\lambda)$ is known and we wish to evaluate certain operator expectation values.

We now wish to solve (2) for $\frac{\partial}{\partial \lambda} |E\lambda>$. Assuming the inverse of the operator $(H - E)$ exists, the general solution is

$$\frac{\partial}{\partial \lambda} \ |E\lambda> \ = \ iC(\lambda)|E\lambda> \ + \ \frac{1}{E \ - \ H} \left( \frac{\partial H}{\partial \lambda} \ - \ \frac{\partial E}{\partial \lambda} \right)|E\lambda> \ , \tag{4}$$

where $C(\lambda)$ is an arbitrary real constant. We can see why this term is allowed because if we try to reproduce (2) from (4) by operating on both sides by $(E - H)$, we see that the term proportional to $C(\lambda)$ will project to zero. It actually arises because of the freedom of choice of a $\lambda$–dependent phase in the definition of the state $|E\lambda>$. We will put $C(\lambda) = 0$ in the following, but this will not limit the generality of the results.

We will now specialize to problems that have discrete, non-degenerate energy eigenvalues. Completeness can be written as

$$\sum_{E'} \ |E'\lambda> < E'\lambda| \ = 1. \tag{5}$$

5.4

Using (5) we may write

$$\left( \frac{\partial H}{\partial \lambda} - \frac{\partial E}{\partial \lambda} \right) |E\lambda> = \sum_{E'} |E'\lambda> <E'\lambda| \left( \frac{\partial H}{\partial \lambda} - \frac{\partial E}{\partial \lambda} \right) |E\lambda>, \quad (6)$$

$$= \sum_{E' \neq E} |E'\lambda> <E'\lambda| \frac{\partial H}{\partial \lambda} |E\lambda> . \quad (7)$$

We know by (3) above that when E' = E in the sum in (6) that

the matrix element $<E\lambda| \frac{\partial H}{\partial \lambda} - \frac{\partial E}{\partial \lambda} |E\lambda>$ vanishes. That is why

the sum in (7) leaves out this term. Once this single term

is eliminated, we know that the $\frac{\partial E}{\partial \lambda}$ term in (6) does not

contribute because of the orthogonality of the states $<E'\lambda|$

and $|E\lambda>$. This term is zero and the result is then Eq$^{\underline{n}}$ (7).

Replacing (7) in (4) gives us

$$\frac{\partial}{\partial \lambda} |E\lambda> = \frac{1}{E - H} \sum_{E' \neq E} |E'\lambda> <E'\lambda| \frac{\partial H}{\partial \lambda} |E\lambda>. \quad (8)$$

Now we postulate on the basis of Eq$^{\underline{n}}$ (105) of Chapter 1

that

$$f(H)|E'\lambda> = f(E')|E'\lambda>, \quad (9)$$

so that

$$(E - H)^{-1}|E'\lambda> = (E - E')^{-1}|E'\lambda>. \quad (10)$$

Eq$^{\underline{n}}$ (8) now becomes

$$\frac{\partial}{\partial\lambda}\ |E\lambda> \ = \ \sum_{E'\neq E}\ |E'\lambda> \ \frac{<E'\lambda|\ \dfrac{\partial H}{\partial\lambda}\ |E\lambda>}{(E \ - \ E')} \quad . \tag{11}$$

Let us go back to Eq$^{\underline{n}}$ (3) and work out the second derivative of $E(\lambda)$:

$$\frac{\partial^2 E}{\partial\lambda^2} \ =$$

$$<E\lambda|\ \frac{\partial^2 H}{\partial\lambda^2}\ |E\lambda> \ + \ <E\lambda|\ \frac{\partial H}{\partial\lambda}\left(\frac{\partial}{\partial\lambda}\ |E\lambda>\right) \ + \ \left(\frac{\partial}{\partial\lambda}\ <E\lambda|\right)\frac{\partial H}{\partial\lambda}\ |E\lambda>. \tag{12}$$

The last two terms in (12) are in fact just complex conjugates of each other, so that

$$\frac{\partial^2 E}{\partial\lambda^2} \ = \ <E\lambda|\ \frac{\partial^2 H}{\partial\lambda^2}\ |E\lambda> \ + \ 2\text{Re}\left[<E\lambda|\ \frac{\partial H}{\partial\lambda}\left(\underbrace{\frac{\partial}{\partial\lambda}\ |E\lambda>}\right)\right] \quad . \tag{13}$$

<p style="text-align:center">Replace</p>

Now let us use (11) in (13). We get

$$\frac{\partial^2 E}{\partial\lambda^2} \ = \ <E\lambda|\ \frac{\partial^2 H}{\partial\lambda^2}\ |E\lambda> \ + \ 2\ \sum_{E'\neq E}\ \frac{\left|<E\lambda|\ \dfrac{\partial H}{\partial\lambda}\ |E'\lambda>\right|^2}{(E \ - \ E')} \quad , \tag{14}$$

where we have dropped the real part restriction because the quantity in brackets in (13) <u>is</u> real. Eq$^{\underline{n}}$ (14) is also an exact formula.

Now let's do a Taylor series for $E(\lambda)$. We have (assuming the series exists)

$$E(\lambda) \ = \ E(0) \ + \ \lambda\ \frac{\partial E(0)}{\partial\lambda} \ + \ \frac{\lambda^2}{2}\ \frac{\partial^2 E(0)}{\partial\lambda^2} \ + \ \ldots \tag{15}$$

where the partials with respect to $\lambda$ are evaluated at $\lambda = 0$.

Using (3) and (14) in (15) now reveals that

$$E(\lambda) = E(0) + \lambda < E| \frac{\partial H(0)}{\partial \lambda} |E> + \frac{\lambda^2}{2} < E| \frac{\partial^2 H(0)}{\partial \lambda^2} |E>$$

$$+ \lambda^2 \sum_{E' \neq E} \frac{|< E| \frac{\partial H(0)}{\partial \lambda} |E'>|^2}{(E - E')}) + \ldots . \qquad (16)$$

The result (16) is usually applied to the situation where the Hamiltonian is given by

$$H = H_0 + H_1 . \qquad (17)$$

$H_0$ represents a Hamiltonian for which an exact solution is known and $H_1$ represents the "perturbation." Instead of (17) we may formally write

$$H = H_0 + \lambda H_1 \qquad (18)$$

and then evaluate the Taylor series (16) when $\lambda = 1$ to get the effect of the perturbation $H_1$ on the energy levels. The result is

$$E = E_0 + < E_0|H_1|E_0> + \sum_{E_0' \neq E_0} \frac{|< E_0|H_1|E_0'>|^2}{(E_0 - E_0')} + \ldots \qquad (19)$$

where I have labeled the unperturbed energies as $E_0$. Eq$^n$ (19) says the leading correction to the $E_0$ energy level is just

the diagonal element of the pertubation matrix. Because the perturbation $H_1$ appears linearly in the diagonal term, this

is the first order correction to the energy.  The next term, where the $H_1$ matrix element appears squared, is the second order correction, and so on.  Corresponding to these corrections in the energies are corrections to the energy wavefunctions.  However, these new wavefunctions get complicated quite quickly and will not be examined here. (See for example Ch.9 of Anderson, Modern Physics and Quantum Mechanics, 1st ed.)

It sometimes happens that the leading first order correction in (19) vanishes for certain perturbations, but the second order term does not.  Notice that if $E_0$ represents the ground state energy (i.e., the lowest energy state), then the effect of the second order correction is such as to lower the energy of the ground state because $(E_0 - E_0') < 0$ for all $E_0' \neq E_0$ by definition.  This cannot be said for higher lying states.  There we see that the second order correction tends to produce a repulsion between neighboring energy levels. The sign of the overall energy shift, however, is not determined.

Let us examine a quantitative example of this method. We will reexamine the simple harmonic oscillator with dimensionless Hamiltonian

$$H_0 = \frac{1}{2} (p^2 + q^2), \tag{20}$$

and energies

$$E_0 = n + \frac{1}{2}, \quad n = 0,1,2,\ldots \tag{21}$$

We will take the perturbation as

$$H_1 = \gamma q^3, \quad (\gamma \text{ dimensionless}) \qquad (22)$$

making the system anharmonic.  The new dimensionless energies of the system are given approximately by

$$E_n \approx \left(n + \frac{1}{2}\right) + \langle \gamma q^3 \rangle_n + \sum_{n' \neq n} \frac{|\langle n | \gamma q^3 | n' \rangle|^2}{(n - n')}. \qquad (23)$$

Now, the first order energy correction vanishes in the unperturbed states because positive and negative position values occur symmetrically in $\psi_n^*(q')\psi_n(q')$.  Another way of arguing this is to say that the $q^3$ operator changes the parity of the state $|n\rangle$.  We can work out the necessary matrix elements of $q^3$ for the second order term as follows. Remember that

$$q = \frac{A + A^+}{\sqrt{2}}, \qquad (24)$$

where

$$A|n\rangle = \sqrt{n}\,|n - 1\rangle, \qquad (25)$$

$$A^+|n\rangle = \sqrt{n + 1}\,|n + 1\rangle. \qquad (26)$$

We now find successively:

$$q|n\rangle = \frac{1}{\sqrt{2}}\left[\sqrt{n}\,|n - 1\rangle + \sqrt{n + 1}\,|n + 1\rangle\right], \qquad (27)$$

$$q^2|n> = q[q|n>] = \frac{1}{2}\left[\sqrt{n|(n-1)}\ |n-2>\right.$$

$$\left. + (2n+1)|n> + \sqrt{(n+1)(n+2)}\ |n+2>\right], \quad (28)$$

$$q^3|n> = q[q^2|n>] = \frac{1}{2\sqrt{2}}\left[\sqrt{n(n-1)(n-2)}\quad |n-3>\right.$$

$$+ \sqrt{n}\ (3n)|n-1> + \sqrt{(n+1)}\ (3n+3)|n+1>$$

$$\left. + \sqrt{(n+1)(n+2)(n+3)}\ |n+3>\right]. \quad (29)$$

Therefore, we have from (23) that

$$E_n \approx \left(n + \frac{1}{2}\right) + \frac{\gamma^2}{8}\left[\frac{n(n-1)(n-2)}{3} + 9n^3 - 9(n+1)^3\right.$$

$$\left. + \frac{(n+1)(n+2)(n+3)}{-3}\right], \quad (30)$$

or

$$E_n = \left(n + \frac{1}{2}\right) - \frac{\gamma^2}{8}\ (30n(n+1) + 11). \quad (31)$$

What has happened to the energy levels?  Notice that the
correction term in (31) is always negative, lowering all of
the energies.  This lowering in energy increases in magnitude
as n increases.  In fact, for neighboring energy levels we
have

$$E_{n+1} - E_n = 1 - \frac{15}{2}\ \gamma^2\ (n+1). \quad (32)$$

Eqⁿ (32) implies that there is a value of n for which the
difference in energies is zero.  This is a backwards way of

finding out that our treatment of the perturbing Hamiltonian
can hardly be valid under these conditions. Eq$^{\underline{n}}$ (32) shows
that our perturbative treatment of $H_1$ must break down when

$$\gamma^2 n \sim 1. \tag{33}$$

Why has this happened? At higher energy levels, the system
is "sampling" larger q' (position) values. However, for any
fixed value of $\gamma$ in (22) there will be values of q' for which
$\gamma q'^3 > \frac{1}{2} q'^2$ for larger q'. Under these conditions the
"perturbation" will in fact be the dominant term in the
energy and a perturbative treatment is bound to be
inadequate.

We have left out a large class of problems in deriving
the result (19). We have specified that the energy levels of
our systems be non-degenerate. Many physical systems have
such degeneracies. (The hydrogen atom is one such system we
will study next semester.) Let us assume that we are trying
to solve for the energy levels of a Hamiltonian of the form
(17), but that there exists a k-fold degeneracy of the
unperturbed energy levels. In addition to the unshifted
energy label, $E_0$, there will now be another label which will
distinguish between these k states. Let's label such a state
as $|E_0 a>$ where a = 1,...,k. Of course the point is to find a
representation which diagonalizes the full Hamiltonian, H =
$H_0 + H_1$, the diagonal elements being the energy eigenvalues.
Now, it is reasonable (and justifiable) to assume that the
first order effect on the k members of the unperturbed

(degenerate) energy spectrum will just come from those states which are elements of the degenerate subspace; that is, we neglect the effect of any "distant" energy states. If this is so, then it is only necessary to diagonalize the perturbation, $H_1$, in the degenerate subspace. Then, the shifted energy levels will of course be given by

$$E_a = E_0 + <E_0a|H_1|E_0a> , \tag{34}$$

where $<E_0a|H_1|E_0a>$ is just the eigenvalue corresponding to the eigenvector labeled by "a". Thus, this is just a standard eigenvalue/eigenvector problem, but carried out entirely within the originally k-fold degenerate subspace. Then, if the diagonal elements of this matrix are all distinct, the degeneracy will have been lifted and we will have k distinct energy levels where before there was only one.  In this case one can then proceed to second order perturbation theory in a standard fashion, using the newly determined distinct eigenfunctions.  However, It may happen that not all the diagonal elements are distinct after $H_1$ is diagonalized, that some energy degeneracies remain. In order to proceed beyond this point in perturbation theory, it is necessary to use second-order degenerate perturbation theory (which takes into account the effects of "distant" states).  We will not pursue this subject further here as it occurs rather infrequently. (See Gottfried, Quantum Mechanics, problem 1, p.397 for a good problem along these lines.)

Let us now move on to talk about another useful approximation method:  the WKB semi-classical approximation. The Schrödinger equation for an arbitrary potential in one spatial dimension is

$$\left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] u(x) = Eu(x),$$ (35)

or

$$\left[ \frac{d^2}{dx^2} + \frac{2m}{\hbar^2} (E - V(x)) \right] u(x) = 0.$$ (36)

When we were solving (35) or (36) for flat potentials, as in the finite potential barrier problem, we defined the constant (for $E > V_0$, say)

$$k^2 = \frac{2m}{\hbar^2} (E - V_0),$$ (37)

for which the solutions to the Schrödinger equation were

$$u(x) \sim e^{\pm ikx}.$$ (38)

The wave number, k, is related to the deBroglie wavelength by

$$k = \frac{2\pi}{\lambda} = \frac{1}{\lambdabar} .$$ (39)

Following this lead, let us define a position dependent wavenumber by

$$k^2(x) \equiv \frac{2m}{\hbar^2} (E - V(x))$$ (40)

when $(E - V(x)) > 0$. If we think of the potential $V(x)$ in (39) as changing sufficiently slowly with x, then we might expect to get solutions of the form

$$u(x) \sim e^{\pm i \int^x dx' k(x')}, \tag{41}$$

where the lower limit on the integral is not yet specified. With this $u(x)$ we have*

$$\frac{1}{i} \frac{d}{dx} u(x) = \pm k(x)u(x), \tag{42}$$

and therefore

$$\frac{d^2}{dx^2} u(x) = -k^2(x)u(x) \pm i \frac{dk(x)}{dx} u(x), \tag{43}$$

which is just (36) if

$$\frac{1}{k^2(x)} \left| \frac{dk(x)}{dx} \right| \ll 1, \tag{44}$$

or, using an obvious definition of $\lambda(x)$:

$$\left| \frac{d\lambda(x)}{dx} \right| \ll 1. \tag{45}$$

In words, Eq$^{\underline{n}}$ (45) says that the change in the reduced wavelength because of the varying potential must be small compared to a change in x. This should be the case in many semi-classical applications where the particle energies are large compared to the potential.

---

* Here we are using the rule (Leibnitz)

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(x)dx = f[b(t)] \frac{db}{dt} - f[a(t)] \frac{da}{dt} .$$

To improve upon (41), let us define a new slowly varying wavefunction $\phi(x)$ such that

$$u(x) = e^{\pm i\int^x dx' k(x')}\phi(x). \tag{46}$$

We then have

$$\frac{d}{dx} u(x) = e^{\pm i\int^x dx' k(x')}\left[\frac{d}{dx} \pm ik(x)\right]\phi(x). \tag{47}$$

and similarly for the second derivative. The Schrödinger equation now takes the form

$$\left[\left(\frac{d}{dx} \pm 2ik(x)\right)\left(\frac{d}{dx}\right) \pm i\frac{dk(x)}{dx}\right]\phi(x) = 0, \tag{48}$$

which is still exact. Since $\phi(x)$ is supposedly a slowly varying function of x, let us neglect the second derivative term in (48). Then we have

$$\frac{\frac{d\phi}{dx}}{\phi} + \frac{1}{2}\frac{\frac{dk}{dx}}{k} = 0. \tag{49}$$

Integrating indefinitely gives

$$\phi(x) = \frac{C}{\sqrt{k(x)}}, \tag{50}$$

where C is an unspecified constant. Thus, we have approximately that

$$u(x) \simeq \frac{C}{\sqrt{k(x)}} e^{\pm i\int^x dx' k(x')}, \tag{51}$$

for $(E - V(x)) > 0$. The + sign in (51) represents a wave traveling to the right and the minus sign a wave traveling to the left.

Now let us think about the case $(E - V(x)) < 0$. We know that classical particles cannot penetrate into such regions because they have insufficient energy. Quantum mechanically, however, we know that nonzero wavefunctions are allowed and that they are given by real exponentials. We saw that these solutions are given by making the substitution $k \rightarrow \pm iK$ where

$K = \dfrac{\sqrt{2m(V_0 - E)}}{\hbar}$. Letting $K(x) = \dfrac{\sqrt{2m(V(x) - E)}}{\hbar}$ in (51), we find the WKB solutions

$$u(x) = \frac{C'}{\sqrt{K(x)}} \, e^{\pm \int^x dx' K(x')}, \tag{52}$$

valid for $(E - V(x)) < 0$.

We now have formulas for slowly varying potentials when $(E - V(x)) \gtrless 0$. However, the basis for the WKB description of wavefunctions breaks down near classical turning points, i.e., positions near where $(E - V(x)) = 0$. Consider a deBroglie wave of energy $E$ impinging on a potential barrier from the left, as shown below.
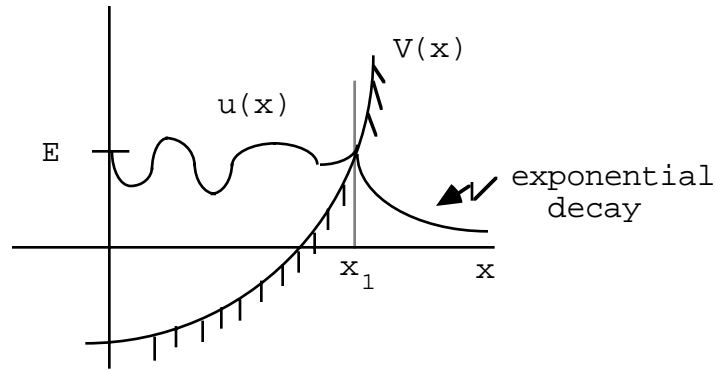
In this situation, we see that the local reduced wavelength, $\lambdabar(x)$, becomes infinite at $x = x_1$, and the condition (45) is violated.  However, we do have approximate solutions far to the left and right of the turning point in Eq$\underline{n}$ (51) and (52), respectively.  It is a question of trying to approximately solve the Schrödinger equation in the vicinity of the turning point and then matching this middle wavefunction to the WKB solutions on either side.  The approximate Schrödinger equation near the turning point is given by assuming an expansion of the potential

$$V(x) \approx E + (x - x_1) \left. \frac{\partial V}{\partial x} \right|_{x=x_1} \qquad (53)$$

near $x = x_1$.  Then the Schrödinger equation becomes

$$\left[ \frac{d^2}{dx^2} + \frac{2m}{\hbar^2} (x - x_1) \left. \frac{\partial V}{\partial x} \right|_{x=x_1} \right] u(x) = 0. \qquad (54)$$

The details of the mathematics that describe the solution of this equation and the matching of wavefunctions and their first derivatives at the turning points will not be recorded

here. (See Merzbacher, Quantum Mechanics, Ch. 7, for example). The results of such an analysis tells us that the connections between wavefunctions in the two regions are given by:

$$\underline{x < x_1} \qquad\qquad\qquad \underline{x > x_1}$$

$$\frac{1}{\sqrt{k(x)}} \; 2 \; \cos\left[ \int_{x_1}^{x} dx'k(x') + \frac{\pi}{4} \right] \; \leftrightarrow \quad \frac{1}{\sqrt{K(x)}} \; e^{-\int_{x_1}^{x} dx'K(x')}, \quad (55)$$

$$\frac{1}{\sqrt{k(x)}} \; \sin\left[ \int_{x_1}^{x} dx'k(x') + \frac{\pi}{4} \right] \; \leftrightarrow \; - \frac{1}{\sqrt{K(x)}} \; e^{\int_{x_1}^{x} dx'K(x')}. \quad (56)$$

Likewise, if we have a situation as in:
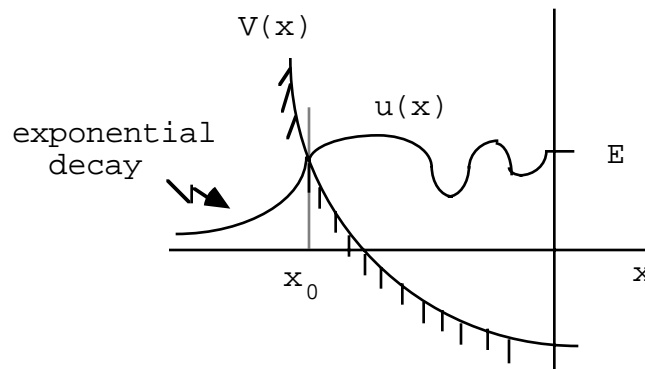


then we have the connections

$$\underline{x > x_0} \qquad\qquad\qquad \underline{x < x_0}$$

$$\frac{1}{\sqrt{k(x)}} \; 2 \; \cos\left[ \int_{x_0}^{x} dx'k(x') - \frac{\pi}{4} \right] \; \leftrightarrow \quad \frac{1}{\sqrt{K(x)}} \; e^{-\int_{x}^{x_0} dx'K(x')}, (57)$$

$$- \frac{1}{\sqrt{k(x)}} \; \sin\left[ \int_{x_0}^{x} dx'k(x') - \frac{\pi}{4} \right] \; \leftrightarrow \quad \frac{1}{\sqrt{K(x)}} \; e^{\int_{x}^{x_0} dx'K(x')}. \quad (58)$$

These sets of equations are called the <u>WKB connection</u> <u>formulas</u>. The double arrows mean a solution of one form in the given spatial region corresponds to the other form in the neighboring region.

There are several interesting applications of these formulas. One of these comes from considering the situation in the following figure:
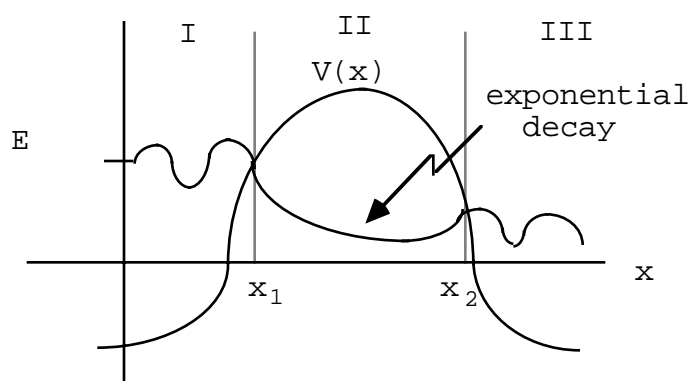


Here we are imagining an incident wave from the left impinging on the potential but having insufficient energy classically to overcome the barrier. However in quantum mechanics we know that there will be a finite probability that the particle reaches the region to the right of $x = x_2$ because of tunneling. Because of (51) and (52) the approximate solutions in regions I and III are:

$$u_I(x) = \frac{A}{\sqrt{k(x)}} \cos \left[ \int_{x_1}^{x} dx' k(x') + \frac{\pi}{4} \right]$$

$$+ \frac{B}{\sqrt{k(x)}} \sin \left[ \int_{x_1}^{x} dx' k(x') + \frac{\pi}{4} \right], \tag{59}$$

$$u_{III}(x) = \frac{E}{\sqrt{k(x)}} \cos\left[\int_{x_2}^{x} dx'k(x') - \frac{\pi}{4}\right]$$

$$+ \frac{F}{\sqrt{k(x)}} \sin\left[\int_{x_2}^{x} dx'k(x') - \frac{\pi}{4}\right]. \tag{60}$$

Our requirement of waves impinging from the left requires F = iE since $u_{III}(x)$ must be of the form of (51) with the upper, positive sign.  Now, instead of matching wavefunctions and their first derivatives at x = $x_1$ and $x_2$, it is only necessary to use the connection formulas.  Again, an overall normalization determines one of the constants, so we can, for example, divide everything through by A to isolate two ratios, $\frac{E}{A}$ and $\frac{B}{A}$ .  We must use the connection formulas from regions I and III to get two expressions for $u_{II}(x)$. Requiring consistency of these two expressions gives us the ratios

$$\frac{E}{A} = \frac{i}{2\theta} , \tag{61}$$

and

$$\frac{B}{A} = \frac{-i}{4\theta^2} , \tag{62}$$

where

$$\theta \equiv e^{\int_{x_1}^{x_2} dx'K(x')}. \tag{63}$$

Now the transmission coefficient is defined as the ratio of the absolute square of the coefficients of the right-traveling waves in regions III to I.  One can show that this means

$$T = \left| \frac{-2i \frac{E}{A}}{(1 - i \frac{B}{A})} \right|^2 . \tag{64}$$

Applying (64) to the barrier problem of Chapter 3, we find

$$T = e^{-4Ka}, \tag{65}$$

where we have neglected the $-\frac{1}{4\theta^2}$ term in the denominator of

(64) for $Ka \gg 1$. Eq$^{\underline{n}}$ (123) of Chapter 3 gives

$$T = \left[ \frac{1}{1 - \frac{1}{4} \left( \frac{K}{k_1} - \frac{k_1}{K} \right)^2} \right] e^{-4Ka}, \tag{65}$$

in the same limit. We see that the WKB method got the
exponential part of T correct, but has missed the overall
constant in front. However, a rectangular potential does not
satisfy well the WKB requirement for a slowly varying
potential.

Another application of the WKB method is to get
approximate bound state energies. Let us say that our
potential looks like:
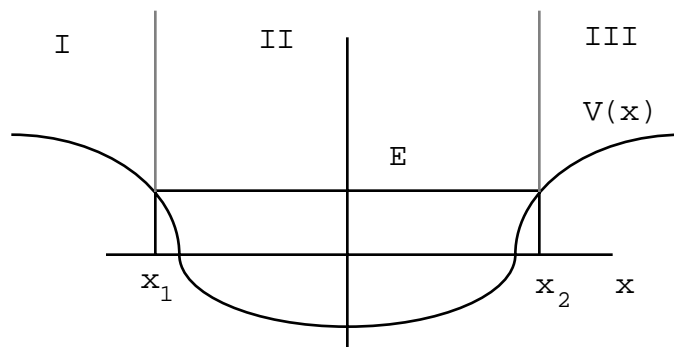
The energy specified corresponds to a bound state. In region I we must have an exponentially decreasing wavefunction:

$$u_I(x) = \frac{C}{\sqrt{k(x)}} \, e^{-\int_x^{x_1} dx' K(x')}. \tag{66}$$

By the connection formula (57), the solution in region II is

$$u_{II}(x) = \frac{C}{\sqrt{k(x)}} \, 2 \cos \left[ \int_{x_1}^{x} dx' k(x') - \frac{\pi}{4} \right]. \tag{67}$$

On the other hand, we must also have a decreasing exponential solution in region III:

$$u_{III}(x) = \frac{C'}{\sqrt{k(x)}} \, e^{-\int_{x_2}^{x} dx' K(x')}. \tag{68}$$

Using the connection formula (55) this then gives

$$u_{II}(x) = \frac{C'}{\sqrt{k(x)}} \, 2 \cos \left[ \int_{x_2}^{x} dx' k(x') + \frac{\pi}{4} \right]. \tag{69}$$

We see that (67) and (69) are only compatible when

$$\int_{x_1}^{x_2} dx' k(x') = \left( n + \frac{1}{2} \right) \pi. \tag{70}$$

and

$$\frac{C}{C'} = (-1)^n \, , \tag{71}$$

where $n = 0,1,2,\ldots$ Eq$^{\underline{n}}$ (70) determines approximately the bound state energies of the system.

Let us apply (70) to the infinite square well of Chapter 3, although rigorously we have no right to do this for such a

discontinuous potential. Using the notation of Chapter 3, we
find that

$$ka = \frac{\pi}{2} \left( n + \frac{1}{2} \right) , \quad n = 0,1,2,\ldots \tag{72}$$

$$\Rightarrow E = \frac{\hbar^2}{2m} \left( \frac{\pi \left( n + \frac{1}{2} \right)}{2a} \right)^2 . \tag{73}$$

Eq$^{\underline{n}}$s (22) and (25) of Chapter 3 can be combined to read

$$ka = \frac{\pi}{2} (n + 1), \quad n = 0,1,2,\ldots . \tag{74}$$

$$\Rightarrow E = \frac{\hbar^2}{2m} \left( \frac{\pi (n + 1)}{2a} \right)^2 . \tag{75}$$

for the exact energies.  We see that (73) has underestimated
the energies by an amount that decreases as n increases.
This makes sense since increasing n is in the direction of
classical energies where the WKB method is supposed to work
best.

   The constant $\frac{\pi}{2}$ on the right of (70) is special to bound
states in one dimension.  In general, the value of this
constant depends on the boundary conditions on the variable
being studied.  For the radial variable in a three
dimensional space this constant is zero.

**Problems**

1. Given the harmonic oscillator Hamiltonian,

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

and energies,

$$E_n = \hbar\omega(n+\frac{1}{2}), \quad n = 0,1,2,\ldots,$$

evaluate $<x^2>_n$ using Eq.(3) of Ch.5. [Hint: Carefully choose the parameter $\lambda$ to be one of the constants in H above.]

2. Prove that the constant $C(\lambda)$ in Eq.(4), Ch.5, is <u>real</u>.

3. Show Eqs.(61,62) of Ch.5.


**Other problems**

4. Consider a harmonic oscillator potential with a small bump at its bottom:



-a    a

$$H = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2 + V(x),$$

5.24
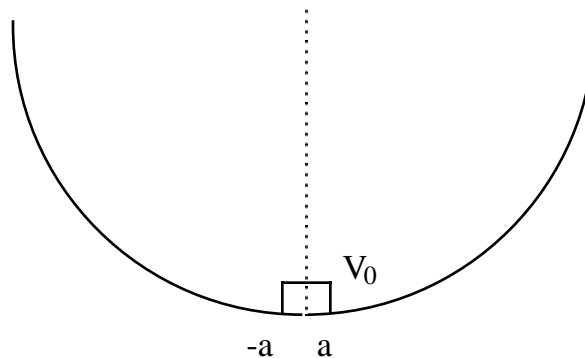
$$V(x) = \begin{cases} V_0, & |x| \le a \\ 0, & \text{elsewhere.} \end{cases}$$

Make the simplifying assumption that $a^2 \ll \dfrac{\hbar}{m\omega}$ in this

problem. The <u>unperturbed</u> energy eigenfunctions for this problem are (see (213) of Chapter 3)

$$u_n(x) = (\frac{m\omega}{\hbar})^{1/4} \frac{1}{\sqrt{\sqrt{\pi} \ 2^n n!}} e^{-m\omega x^2/2\hbar} \ H_n(\sqrt{\frac{m\omega}{\hbar}}x),$$

and the <u>unperturbed</u> energies are:

$$E_n = \hbar\omega(n+\frac{1}{2}), \quad n = 0,1,2,\ldots$$

(a) Find the first order effect of this perturbation on the lowest energy state:

$$E'_{n=0} = \frac{\hbar\omega}{2} + ?.$$

(b) Find the first order effect of this perturbation on the expectation value of $x^2$:

$$<x^2>_{n=0} = \frac{\hbar}{2m\omega} + ?.$$

[Notice, we do <u>not</u> know what the new <u>eigenstates</u> are.]

5.The eigenvalue equation for a system with a discrete, nondegenerate and complete set of states may be written formally as (Dirac notation; assume $<\psi_n|\psi_n> = 1$)

$$H_0 \ |\psi_n> = E_n \ |\psi_n>,$$

where $H_0$ is the unperturbed Hamiltonian and the old eigenenergies are given by $E_n$. A small potential, $\Delta V$, is added to $H_0$, $H = H_0 + \Delta V$. Show that the new ground state eigenvector, $|\psi>$, is given approximately by

$$|\psi> \; \approx \; |\psi_0> \; + \; \sum_{n=1} \; |\psi_n> \; \frac{<\psi_n|\Delta V|\psi_0>}{(E_0 - E_n)} \; .$$

6. Consider adding to the one-dimensional Hamiltonian,

$$H_0 \; = \; \frac{p_x^2}{2m} \; + \; V(x),$$

where the potential $V(x)$ is an even function of x ($V(x) = V(-x)$) the perturbation ($\gamma$ = constant)

$$H_1 \; = \; \gamma \, x.$$

(Physically, this could represent adding an electric field.)

(a)  On the basis of some physical property, argue that the new energies, to lowest order in $\gamma$, must be given by

$$E_n \; = \; E_n^{(0)} \; + \; \alpha_n \gamma^2,$$

where the $\alpha_n$ are constants and $E_n^{(0)}$ are the unperturbed energies. (Notice there is no <u>linear</u> term $\sim \gamma$ in $E_n$.)

(b)  Show that the constants $\alpha_n$ are given by the sum

$$\alpha_n \; = \; \sum_{n' \neq n} \frac{|<n|x|n'>|^2}{E_n^{(0)} - E_{n'}^{(0)}} \; .$$

(c)  Evaluate $\alpha_n$ for the <u>ground state</u> (n=0) of an harmonic oscillator. The energies are $E_n^{(0)} = \hbar\omega(n+\tfrac{1}{2})$. [Hint: Remember that (pages 3.34 and 3.35 of the notes)

$$x \; = \; \left(\frac{\hbar}{m\omega}\right)^{1/2} \frac{A + A^{\dagger}}{\sqrt{2}} \; ,$$

where

$$A|n> \; = \; \sqrt{n} \; |n-1>,$$
$$A^{\dagger}|n> \; = \; \sqrt{n+1} \; |n+1>.]$$

**CHAPTER SIX:   Generalization to Three Dimensions**

First step:  generalize some obvious results to 3-D. Starting point:  "grab bag" of results.

$$\vec{x}|\vec{x}'> = \vec{x}'|\vec{x}'> , \qquad \vec{p}|\vec{p}'> = \vec{p}'|\vec{p}'>. \qquad (1)$$

The <u>Cartesian</u> bases (we will also encounter a <u>spherical</u> basis) $|\vec{x}'>$ and $|\vec{p}'>$ are <u>direct</u> <u>products</u> of the basis states in the three orthogonal directions:  (One sometimes writes $|\vec{x}'> = |x_1'> \otimes |x_2'> \otimes |x_3'>$.)

$$\left. \begin{array}{ll} |\vec{x}'> = |x_1'>|x_2'>|x_3'>, & |\vec{p}'> = |p_1'>|p_2'>|p_3'> \\ \quad \equiv |x_1',x_2',x_3'> & \quad \equiv |p_1',p_2',p_3'> \end{array} \right\} \qquad (2)$$

(I will try to consistently label the three orthogonal space directions as 1, 2 and 3 rather than as x, y and z from now on.)  We also have (see (158) of Ch. 2)

$$<\vec{x}'|\vec{p}'> = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\vec{x}'\cdot\vec{p}'/\hbar} , \qquad (3)$$

$$<\vec{x}'|\vec{x}''> = \delta^3(\vec{x}' - \vec{x}''), \quad <\vec{p}'|\vec{p}''> = \delta^3(\vec{p}' - \vec{p}''), \qquad (4)$$

where

$$\left. \begin{array}{l} \delta^3(\vec{x}' - \vec{x}'') = \delta(x_1' - x_1'')\delta(x_2' - x_2'')\delta(x_3' - x_3''), \\ \delta^3(\vec{p}' - \vec{p}'') = \delta(p_1' - p_1'')\delta(p_2' - p_2'')\delta(p_3' - p_3''). \end{array} \right\} \qquad (5)$$

Also

$$1 = \int d^3x' |\vec{x}'><\vec{x}'|, \quad 1 = \int d^3p' |\vec{p}'><\vec{p}'|, \qquad (6)$$

6.2

with

$$d^3x' = dx'_1 dx'_2 dx'_3, \quad d^3p' = dp'_1 dp'_2 dp'_3. \tag{7}$$

The formal energy eignevalue problem is still stated as

$$H|a'> = E_{a'}|a'>, \tag{8}$$

where the $|a'>$ are a complete, orthogonal set of states:

$$\sum_{a'} |a'><a'| = 1, \quad <a'|a"> = \delta_{a'a"}. \tag{9}$$

(Eq$\underline{n}$ (9) assumes the energy eigenvalues are discrete and nondegenerate.  What would the analogous equations in the more general situation look like?)  Wavefunctions are given by the projections (see (174) of Ch. 2)

$$u_{a'}(\vec{x}') = <\vec{x}'|a'>, \tag{10}$$

which satisfy (using (6))

$$\int d^3x \; u^*_{a'}(\vec{x}) u_{a'}(\vec{x}) = 1. \tag{11}$$

Eq$\underline{n}$ (11) tells us the engineering dimensions of the $u_{a'}(\vec{x})$ are

$$[u_{a'}(\vec{x})] \sim [\text{length}]^{-3/2}. \tag{12}$$

For continuous spectra, we usually use a momentum rather than an energy basis to completely specify the state of the particle.  Then defining[*]

_____

[*] Many books use $u_{\vec{k}}(\vec{x}) \equiv \hbar^{3/2} u_{\vec{p}}(\vec{x})$ as the momentum eigenfunction, in which case it is dimensionless.

$$u_{\vec{p}}(\vec{x}) = <\vec{x}'|\vec{p}> \, , \tag{13}$$

we have (again from (6))

$$\int d^3x \; u_{\vec{p}}^{*}(\vec{x})u_{\vec{p}'}(\vec{x}) = \delta^3(\vec{p} - \vec{p}') \; . \tag{14}$$

We will limit ourselves to consideration of Hamiltonians of the form

$$H = \frac{\vec{p}^2}{2m} + V(\vec{x}) \; . \tag{15}$$

The form of the potential, $V(\vec{x})$, will determine the nature of the spatial basis to be used. For example, for $V(\vec{x}) = F(x_1) + F(x_2) + F(x_3)$ (as for the 3-D harmonic oscillator) we would use a Cartesian basis; for $V(\vec{x}) = V(r)$, where $r = |\vec{x}|$, one would use a spherical basis. We must use a basis in which the time independent Schrödinger equation separates; for example $u_{a'}(\vec{x}) = u_1(x_1)u_2(x_2)u_3(x_3)$ in Cartesian coordinates or $u_{a'}(\vec{x}) = u_1(r)u_2(\theta)u_3(\phi)$ in spherical coordinates. The time independent Schrödinger equation is, of course, just Eq$^{\underline{n}}$ (8) projected into an explicit basis.

We will continue to assume that

$$[x_i,p_i] = i\hbar \tag{16}$$

for each i = 1,2,3, where $x_i$ and $p_i$ are operators. However, what about $[x_i,x_j]$ for $i \neq j$? In the Cartesian basis

$$[x_1,x_2]|\vec{x}'> = (x_1x_2 - x_2x_1)|x_1',x_2',x_3'> = 0 \tag{17}$$

Since (17) is true for any $|\vec{x}'>$,

$$\Rightarrow [x_1,x_2] = 0 \ . \tag{18}$$

Similarly $[x_2,x_3] = [x_1,x_3] = 0$. Therefore

$$[x_i,x_j] = 0 \ . \tag{19}$$

for all $i,j$. We learned before in Ch.4 that $[A,B]$ is a measure of the "compatibility" of the operators A and B. Eq$^{\underline{n}}$ (19) above tells us that measurements of $x_i$ do not limit the precision of measurements of $x_j$ ($i \neq j$) for a particle. Thus the $x_i$ are simultaneously measurable for all states.

Let us generalize the unitary displacement operator we had before, also in Ch.4:

$$<x_1'|e^{ix_1''p_1/\hbar} \ = \ <x_1' + x_1''| \ . \tag{20}$$

Now we have

$$<x_1',x_2',x_3'|e^{ix_1''p_1/\hbar} \ = \ <x_1' + x_1'',x_2',x_3'|, \tag{21}$$

and similarly for displacements in the 2 and 3 directions. Now consider the following displacements:

$$x_1' + x_1'',x_2' + x_2'',x_3'$$



$$x_1',x_2',x_3' \qquad\qquad\qquad x_1' + x_1'',x_2',x_3'$$

We may get from $(x_1',x_2',x_3')$ to $(x_1' + x_1'',x_2' + x_2'',x_3')$ along

paths 1 or 2.  Along path 1

$$< x_1',x_2',x_3'|e^{ix_2''p_2/\hbar}\ e^{ix_1''p_1/\hbar}\ =\ < x_1' + x_1'',x_2' + x_2'',x_3'| \ . \qquad (22)$$

Along path 2:

$$< x_1',x_2',x_3'|e^{ix_1''p_1/\hbar}\ e^{ix_2''p_2/\hbar}\ =\ < x_1' + x_1'',x_2' + x_2'',x_3'| \ . \qquad (23)$$

The equivalence of these two operators tells us that

$$[p_1,p_2]\ =\ 0. \qquad (24)$$

This can obviously be done for the sets $(p_2,p_3)$ and $(p_1,p_3)$

also.  The conclusion is

$$[p_i,p_j]\ =\ 0. \qquad (25)$$

for all i,j.  Thus the $p_i$ are simultaneously measurable in

all states.  Notice that since (25) is true, we have

$$e^{ix_1'p_1/\hbar}\ e^{ix_2'p_2/\hbar}\ e^{ix_3'p_3/\hbar}\ =\ e^{i\vec{x}'\cdot\vec{p}/\hbar} \ , \qquad (26)$$

so that a general displacement can be written as

$$< \vec{x}'|e^{i\vec{x}''\cdot\vec{p}/\hbar}\ =\ < \vec{x}' + \vec{x}''|. \qquad (27)$$

What about mixed objects like $[x_1,p_2]$?  Consider the

infinitesimal change

$$< \vec{x}'|(1 + id\vec{x}''\cdot\vec{p}/\hbar)\ =\ < \vec{x}' + d\vec{x}''| \ . \qquad (28)$$

Multiplying both sides by the operator $\vec{x}$, we have

$$\langle \vec{x}' | (1 + i d\vec{x}'' \cdot \vec{p}/\hbar)\vec{x} = \langle \vec{x}' + d\vec{x}'' | (\vec{x}' + d\vec{x}''). \qquad (29)$$

Now do these two operations in the opposite order:

$$\langle \vec{x}' | \vec{x}(1 + i d\vec{x}'' \cdot \vec{p}/\hbar) = \langle \vec{x}' | (1 + i d\vec{x}'' \cdot \vec{p}/\hbar) \vec{x}'$$

$$= \langle \vec{x}' + d\vec{x}'' | \vec{x}'. \qquad (30)$$

Therefore, we have that

$$\langle \vec{x}' | [(1 + i d\vec{x}'' \cdot \vec{p}/\hbar), \vec{x}] = \langle \vec{x}' + d\vec{x}'' | d\vec{x}'', \qquad (31)$$

$$\Rightarrow \langle \vec{x}' | \frac{i}{\hbar} [d\vec{x}'' \cdot \vec{p}, \vec{x}] = \langle \vec{x}' | d\vec{x}'', \qquad (32)$$

$$\Rightarrow \frac{i}{\hbar} [d\vec{x}'' \cdot \vec{p}, \vec{x}] = d\vec{x}''. \qquad (33)$$

This statement becomes more transparent in component language:

$$\frac{i}{\hbar} \sum_k d x_k'' [p_k, x_j] = d x_j'', \qquad (34)$$

$$\Rightarrow [p_k, x_j] = \frac{\hbar}{i} \delta_{kj}. \qquad (35)$$

From the 1-D statement,

$$\langle x' | e^{i x'' p_x/\hbar} = \langle x' + x'' |, \qquad (36)$$

I derived the result (Ch. 4, Eq$^{\underline{n}}$ (153))

$$\langle x' | p_x = \frac{\hbar}{i} \frac{\partial}{\partial x'} \langle x' |. \qquad (37)$$

Likewise in our 3-D Cartesian basis given Eq$^{\underline{n}}$ (27) above, it is easy to show that

$$< \vec{x}' | \vec{p} = \frac{\hbar}{i} \vec{\nabla}' < \vec{x}' | , \qquad (38)$$

where $\vec{\nabla}'$ is the usual gradient (differential) operator:

$$\vec{\nabla}' = \sum_i \hat{e}_i \frac{\partial}{\partial x_i'} . \qquad (39)$$

Likewise one can show that

$$< \vec{p}' | \vec{x} = - \frac{\hbar}{i} \vec{\nabla}'_p < \vec{p}' | , \qquad (40)$$

where

$$\vec{\nabla}'_p = \sum_i \hat{e}_i \frac{\partial}{\partial p_i'} . \qquad (41)$$

Let us apply our Cartesian basis to a simple problem in 3-D.  Consider the infinite 3-D square well:



(Origin at center)

(a,a,-a)
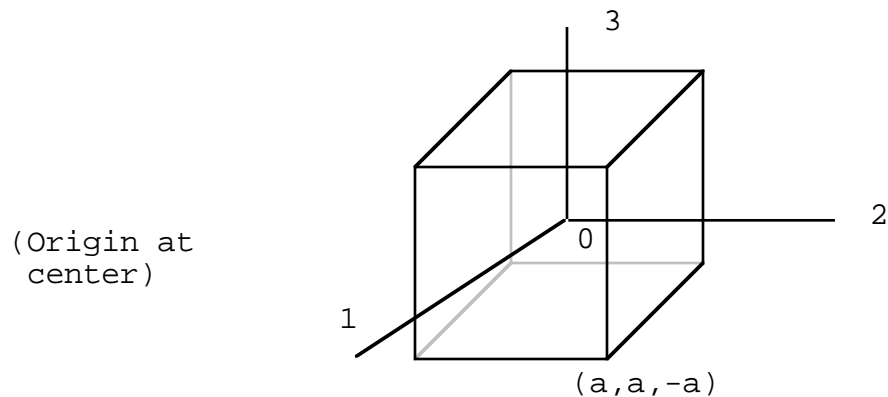
V = 0   inside          $-a \leq x_1 \leq a$
V = +∞ outside          $-a \leq x_2 \leq a$
                        $-a \leq x_3 \leq a$

The B.C. are

$$\left. \begin{array}{l} u(\pm a, x_2, x_3) = 0, \\ u(x_1, \pm a, x_3) = 0, \\ u(x_1, x_2, \pm a) = 0. \end{array} \right\} \qquad (42)$$

(In other words $u|_{\text{surface}} = 0$). Projected into the Cartesian basis, the energy eigenvalue condition $H|a'> = E_{a'}|a'>$ becomes

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right) u_{a'}(\vec{x}) = E_{a'}u_{a'}(\vec{x}). \qquad (43)$$

This is obviously separable in $x_1$, $x_2$ and $x_3$.   Let

$$u_{a'}(\vec{x}) = u_1(x_1)u_2(x_2)u_3(x_3). \qquad (44)$$

Then (43) may be put into the form $\left(u_1'' \equiv \dfrac{\partial^2 u}{\partial x_1^2}\ ,\ \text{etc.}\right)$

$$\frac{u_1''}{u_1} + \frac{u_2''}{u_2} + \frac{u_3''}{u_3} = -\frac{2mE_{a'}}{\hbar^2}\ , \qquad (45)$$

which means we may set

$$\left.\begin{aligned}
-\frac{\hbar^2}{2m}\,u_1'' &= E_1 u_1(x_1),\\
-\frac{\hbar^2}{2m}\,u_2'' &= E_2 u_2(x_2),\\
-\frac{\hbar^2}{2m}\,u_3'' &= E_3 u_3(x_3),
\end{aligned}\right\} \qquad (46)$$

where

$$E_1 + E_2 + E_3 = E_{a'}. \qquad (47)$$

Eqⁿs (46) and the B.C. (42) insure that the solution in each direction is identical to the one-dimensional case solved in Ch. 3.   Let me remind you of these solutions:

$$u_{n-}(x) = <x|n-> = \frac{1}{\sqrt{a}}\,\sin(k_{n-}x), \qquad (48)$$

$$u_{n+}(x) = <x|n+> = \frac{1}{\sqrt{a}}\,\cos(k_{n+}x), \qquad (49)$$

where

$$n = 1,2,3,\ldots \begin{cases} k_{n-} = \dfrac{n\pi}{a} & (50) \\[4mm] k_{n+} = \dfrac{(n - 1/2)\pi}{a} & (51) \end{cases}$$

and $E = \dfrac{\hbar^2 k^2}{2m}$ . Thus in 3-D, the solutions are

$$u_{a'}(\vec{x}) \equiv \langle \vec{x}' | a' \rangle = \langle x_1 | E_1 \rangle \langle x_2 | E_2 \rangle \langle x_3 | E_3 \rangle$$
$$= u_{n_1 P_1}(x_1) u_{n_2 P_2}(x_2) u_{n_3 P_3}(x_3), \qquad (52)$$

where each of the $u_{n_i P_i}$ are given in (48) and (49) with $P_{1,2,3} = \pm$ giving the parities of the state. Eq$^{\underline{n}}$ (47) says that the total energy is given by the sum of the $E_1$, $E_2$ and $E_3$ eigenenergies. This means that there are energy degeneracies. For example, consider energy levels for which $P_1 = P_2 = P_3 = -$,

$$E_- = \frac{\pi^2 \hbar^2}{2ma^2} (n_1^2 + n_2^2 + n_3^2). \qquad (53)$$

Although the lowest energy above is specified by $n_1 = n_2 = n_3 = 1$, the first excited state can be given as ($n_1 = 2$, $n_2 = 1$, $n_3 = 1$), ($n_1 = 1$, $n_2 = 2$, $n_3 = 1$) or ($n_1 = 1$, $n_2 = 1$, $n_3 = 2$).

Can you find a complete set of commuting operators that uniquely specify the independent states?

A much more useful coordinate basis in physics is a spherical basis in which the position of a particle is specified by the three numbers $r$, $\theta$, $\phi$:

$$\langle x_1', x_2', x_3' | \rightarrow \langle r, \theta, \phi |. \qquad (54)$$

Based on what we have seen before, we expect that

$$1 = \int d^3r \, |r,\theta,\phi\rangle\langle r,\theta,\phi|, \tag{55}$$

where

$$d^3r = r^2 \sin\theta \, drd\theta d\phi. \tag{56}$$

The range of these variables is as usual

$$\left.\begin{array}{l} r: \ 0 \rightarrow \infty, \\ \theta: \ 0 \rightarrow \pi, \\ \phi: \ 0 \rightarrow 2\pi, \end{array}\right\} \tag{57}$$

which picks out all points in coordinate space.

Just as the 3-D Cartesian basis,

$$\langle x_1', x_2', x_3'| = \langle x_1'|\langle x_2'|\langle x_3'|, \tag{58}$$

is a direct product of three Hilbert spaces, we expect that the spherical basis,

$$\langle r,\theta,\phi| = \langle r|\langle\theta|\langle\phi|, \tag{59}$$

is also a direct product of separate Hilbert spaces. And just as we have completeness in each Cartesian subspace,

$$\left.\begin{array}{l} 1_{x_1} = \int_{-\infty}^{\infty} dx_1' \, |x_1'\rangle\langle x_1'|, \\ 1_{x_2} = \int dx_2' \, |x_2'\rangle\langle x_2'|, \\ 1_{x_3} = \int dx_3' \, |x_3'\rangle\langle x_3'|, \end{array}\right\} \tag{60}$$

we demand that

$$1_r = \int_0^\infty dr r^2 |r><r| ,$$

$$1_\theta = \int_0^\pi d\theta \sin \theta \ |\theta><\theta| , \qquad\qquad (61)$$

$$1_\phi = \int_0^{2\pi} d\phi |\phi><\phi| ,$$

so that

$$1 = 1_1 \cdot 1_\theta \cdot 1_\phi = \int d^3r |r,\theta,\phi><r,\theta,\phi| \ . \qquad (62)$$

We will label the angular position as

$$|\hat{n}> \equiv |\theta,\phi> . \qquad\qquad (63)$$
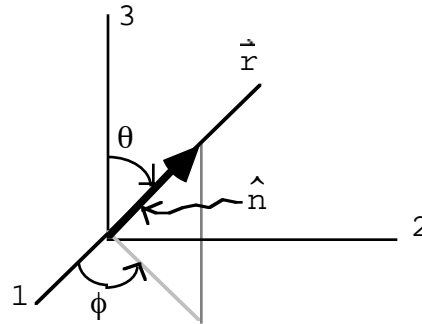


where $\hat{n}$ is a unit vector pointing in the $\vec{r}$ direction. Of course, there are many other bases possible, corresponding to cylindrical coordinates, elliptical coordinates, etc. It's clear that we have to be consistent in a given problem to stick with an initial choice, but other than this one is free to switch between various bases in order to simplify derivations and expressions. I will generally use $|\vec{x}>$ to denote a Cartesian basis and $|\vec{r}>$ to denote a spherical one.

6.12

Let us now introduce the quantum mechanical operator representing <u>orbital</u> angular momentum* :

$$\vec{L} \equiv \vec{x} \times \vec{p}. \tag{64}$$

Component-wise, we have

$$\left. \begin{array}{l} L_1 = x_2 p_3 - x_3 p_2, \\[1mm] L_2 = x_3 p_1 - x_1 p_3, \\[1mm] L_3 = x_1 p_2 - x_2 p_1. \end{array} \right\} \tag{65}$$

Notice that since $[x_i, p_j] = 0$ ($i \neq j$), the order of the operators in (65) does not matter.  Also notice that

$$\vec{L}^+ = \vec{L}, \tag{66}$$

i.e., it is Hermitian and therefore has real eigenvalues.

The various $L_i$ do not commute.  To see this, consider

$$\begin{aligned} [L_1, L_2] &= [x_2 p_3 - x_3 p_2, x_3 p_1 - x_1 p_3] \\[1mm] &= [x_2 p_3, x_3 p_1] + [x_3 p_2, x_1 p_3] \\[1mm] &= x_2 p_1 [p_3, x_3] + p_2 x_1 [x_3, p_3] \\[1mm] &= i\hbar (x_1 p_2 - x_2 p_1) = i\hbar L_3. \end{aligned} \tag{67}$$

Likewise

$$[L_1, L_3] = -i\hbar L_2, \tag{68}$$

$$[L_2, L_3] = i\hbar L_1. \tag{69}$$

---

* This is surely the classical form of the angular momentum, but an equivalent classical form is $\vec{L} = -\vec{p} \times \vec{x}$. Luckily, both of these produce the same <u>quantum</u> <u>mechanical</u> operator since only orthogonal components of $\vec{x}$, $\vec{p}$ are multiplied together. This is not always the case, however, and sometimes gives rise to "operator-ordering ambiguities".

It can be confirmed that

$$[L_i, L_j] = i\hbar \sum_k \varepsilon_{ijk} L_k . \tag{70}$$

is the general statement. ( $\varepsilon_{ijk}$ is the usual permutation symbol which is completely antisymmetric). This shows that the three components, $L_i$, are mutually incompatible observables. A quantum mechanical state can not, for example, be in an eigenstate of <u>both</u> $L_1$ <u>and</u> $L_2$. (There is one exception to this satement that we will discuss.) This is distinctly different from <u>linear</u> momentum for which we have seen

$$[p_i, p_j] = 0, \tag{71}$$

for all i,j.

We will now try to find the effect of the $L_i$ on a state $|\vec{x}'>$. Consider

$$\left(1 - i\left(\frac{\delta\phi}{\hbar}\right)L_3\right)|\vec{x}'> = \left(1 - i\left(\frac{\delta\phi}{\hbar}\right)(p_2 x_1' - p_1 x_2')\right)|\vec{x}'>, \tag{72}$$

where $\delta\phi$ is a positive, infinetismal quantity. Remember that

$$e^{i\vec{x}''\cdot\vec{p}/\hbar}|\vec{x}'> = |\vec{x}' + \vec{x}''>. \tag{73}$$

Let's choose $\vec{x}'' = (-\delta x'', 0, 0)$ where $\delta x''$ is also a positive, infinitesimal quantity. Then (73) implies that

$$\left(1 + i\frac{\delta x''}{\hbar}p_1\right)|\vec{x}'> = |x_1' - \delta x'', x_2', x_3'>. \tag{74}$$

Likewise for $\vec{x}'' = (0, \delta x'', 0)$ we get

$$\left(1 \; - \; i \; \frac{\delta x"}{\hbar} \; p_2\right) |\vec{x}' > \; = \; |x_1', x_2' \; + \; \delta x", x_3' > . \tag{75}$$

Since we may write

$$\left(1 \; - \; i \; \left(\frac{\delta\phi}{\hbar}\right) L_3\right) \; = \; \left(1 \; - \; i \; \left(\frac{\delta\phi}{\hbar}\right) x_1' p_2\right)\left(1 \; + \; i \; \left(\frac{\delta\phi}{\hbar}\right) x_2' p_1\right), \tag{76}$$

(because $\delta\phi$ is infinetismal) we get that

$$\left(1 \; - \; i \; \left(\frac{\delta\phi}{\hbar}\right) L_3\right) \; |\vec{x}' > \; = \; |x_1' \; - \; \delta\phi x_2', x_2' \; + \; \delta\phi x_1', x_3' > \tag{77}$$

The right hand side of (77) reveals that a rotation about the 3-axis has been performed. (See the following figure.) We are adopting the convention that this represents an <u>active</u> rotation of the physical system itself (rather than a passive rotation of the coordinate system in the opposite direction.) The rotation shown is defined to have $\delta\phi > 0$.
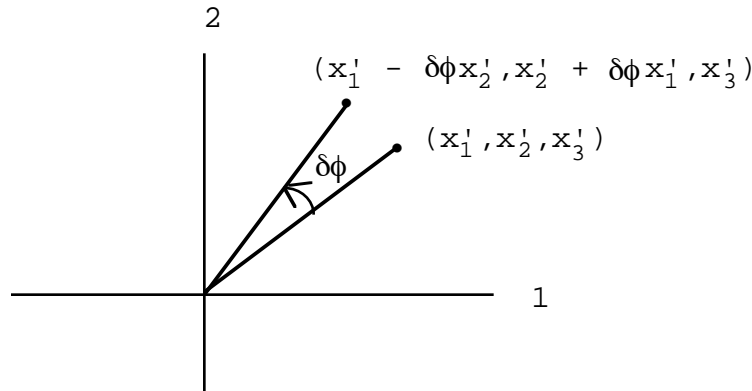


I have used the Cartesian basis to make these conclusions. In terms of a spherical basis, the effect of this operator is clearly

$$\left(1 \; - \; i \; \left(\frac{\delta\phi}{\hbar}\right) L_3\right) \; |r, \theta, \phi > \; = \; |r, \theta, \phi \; + \; \delta\phi > . \tag{78}$$

Since $\delta\phi$ is infinitesimal, we have

$$|r,\theta,\phi + \delta\phi> = |r,\theta,\phi> + \delta\phi\frac{\partial}{\partial\phi}|r,\theta,\phi>. \qquad (79)$$

Matching the coefficient of $\delta\phi$ on both sides of (78), we conclude that

$$L_3|r,\theta,\phi> = i\hbar\frac{\partial}{\partial\phi}|r,\theta,\phi>, \qquad (80)$$

or since r and $\theta$ play no role here, that

$$L_3|\phi> = i\hbar\frac{\partial}{\partial\phi}|\phi>. \qquad (81)$$

Equivalently,

$$<\phi|L_3 = -i\hbar\frac{\partial}{\partial\phi}<\phi|. \qquad (82)$$

Finite relations can also be produced using $L_3$. Any finite rotation, $\phi$, can always be imagined to consist of N identical partial rotations by an amount $\frac{\phi}{N}$. But in the limit $N \to \infty$ each of these partial rotations becomes infinitesimal. Thus, a finite rotation is accomplished by

$$\lim_{N\to\infty} \left(1 - i\left(\frac{\phi/N}{\hbar}\right)L_3\right)^N.$$

Applying the formula

$$\lim_{N\to\infty}\left(1 + \frac{x}{N}\right)^N = e^x, \qquad (83)$$

to the above gives

$$\lim_{N\to\infty} \left(1 - i \left(\frac{\phi/N}{\hbar}\right) L_3 \right)^N = e^{-iL_3\phi/\hbar}, \tag{84}$$

as the operator which performs finite rotations about the third axis. That is

$$e^{-iL_3\phi'/\hbar} |r,\theta,\phi> = |r,\theta,\phi + \phi'>. \tag{85}$$

Since $L_3$ is Hermitian, we recognize $e^{\pm iL_3\phi/\hbar}$ as a unitary operator. $L_3$ is called the <u>generator</u> of rotations about the third axis.

We will now find the effect of $L_1$ and $L_2$ on the $<\vec{r}|$ basis by a more cookbook-type approach. We have that

$$<\vec{r}| \vec{L} = <\vec{r}| \vec{x} \times \vec{p} = <\vec{r}| \vec{r} \times \vec{p}$$

$$= \vec{r} \times (<\vec{r}| \vec{p}) = \vec{r} \times \left( \frac{\hbar}{i} \vec{\nabla}_r <\vec{r}| \right). \tag{86}$$

Now since our basis is spherical, the gradient operator must be stated in spherical variables (which is symbolized by $\vec{\nabla}_r$):

$$\vec{\nabla}_r = \hat{e}_r \frac{\partial}{\partial r} + \hat{e}_\phi \frac{1}{r \sin\theta} \frac{\partial}{\partial \phi} + \hat{e}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} , \tag{87}$$

where $\hat{e}_r$, $\hat{e}_\phi$ and $\hat{e}_\theta$ are unit vectors pointing in the instantaneous r, $\phi$ and $\theta$ directions.
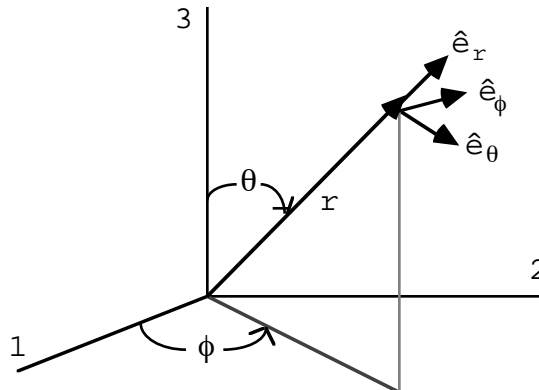
The picture informs us that

$$
\left.
\begin{array}{c}
\hat{e}_r \times \hat{e}_\theta = \hat{e}_\phi \ , \\[4pt]
\hat{e}_\theta \times \hat{e}_\phi = \hat{e}_r \ , \\[4pt]
\hat{e}_\phi \times \hat{e}_r = \hat{e}_\theta \ ,
\end{array}
\right\}
\tag{88}
$$

so that

$$
< \vec{r} \,| \, \vec{L} = \frac{\hbar}{i} \left[ \hat{e}_\phi \frac{\partial}{\partial \theta} - \hat{e}_\theta \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right] < \vec{r} \,| \ .
\tag{89}
$$

The $\hat{e}_r$, $\hat{e}_\theta$ and $\hat{e}_\phi$ can be related to unit vectors along $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$, in the above figure by

$$
\hat{e}_r = \sin \theta \cos \phi \ \hat{e}_1 + \sin \theta \sin \phi \ \hat{e}_2 + \cos \theta \ \hat{e}_3,
\tag{90}
$$

$$
\hat{e}_\phi = -\sin \phi \ \hat{e}_1 + \cos \phi \ \hat{e}_2 \ .
\tag{91}
$$

$$
\hat{e}_\theta = \cos \theta \cos \phi \ \hat{e}_1 + \cos \theta \sin \phi \ \hat{e}_2 - \sin \theta \ \hat{e}_3,
\tag{92}
$$

so when the basis in (89) is expressed in terms of the $\hat{e}_i$, we find

$$
< \vec{r} \,| \, L_1 = \frac{\hbar}{i} \left( -\sin \phi \frac{\partial}{\partial \theta} - \cos \phi \cot \theta \frac{\partial}{\partial \phi} \right) < \vec{r} \,| ,
\tag{93}
$$

$$
< \vec{r} \,| \, L_2 = \frac{\hbar}{i} \left( \cos \phi \frac{\partial}{\partial \theta} - \sin \phi \cot \theta \frac{\partial}{\partial \phi} \right) < \vec{r} \,| ,
\tag{94}
$$

$$
< \vec{r} \,| \, L_3 = \frac{\hbar}{i} \frac{\partial}{\partial \phi} < \vec{r} \,| \ .
\tag{95}
$$

Eqn (95) is, of course, consistent with (82) above.  We can also show from the above that

$$< \vec{r} \,| \, \vec{L}^2 \, = \, < \vec{r} \,| \, \left( L_1^2 \, + \, L_2^2 \, + \, L_3^2 \right)$$

$$= \, -\hbar^2 \left[ \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \, + \, \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \, \frac{\partial}{\partial \theta} \right) \right] < \vec{r} \,| , \quad (96)$$

$$\equiv \, L_{op}^2 \, < \vec{r} \,| . \qquad \qquad (97)$$

where we have defined the <u>differential</u> operator (as opposed to the Hilbert space operator, $\vec{L}^2$)

$$L_{op}^2 \, = \, -\hbar^2 \left[ \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \, + \, \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \, \frac{\partial}{\partial \theta} \right) \right] . \qquad (98)$$

Apart from an overall factor of $-\frac{1}{r^2 \hbar^2}$, this is just seen to be the angular part of the $\vec{\nabla}_r^2$ operator. Eq$^n$s (93) - (95) could also have been stated in terms of the angular basis $< \hat{n} |$ since the purely radial part of the basis plays no role in these considerations. These results will be useful in a moment.

I will now prove a useful identity for $\vec{L}^2$. We know that

$$L_i \, = \, \sum_{j,k} \varepsilon_{ijk} x_j p_k , \qquad \qquad (99)$$

where the order of the operators $x_j, p_k$ does not matter since $[x_j, p_k] = 0$ for $j \neq k$. Therefore,

$$\vec{L}^2 \, = \, \sum_i L_i^2 \, = \, \sum_i \left( \sum_{j,k} \varepsilon_{ijk} x_j p_k \sum_{\ell,m} \varepsilon_{i\ell m} x_\ell p_m \right)$$

$$= \, \sum_{j,k,\ell,m} \left( \sum_i \varepsilon_{ijk} \varepsilon_{i\ell m} \right) x_j p_k x_\ell p_m . \qquad (100)$$

One has

$$\sum_i \varepsilon_{ijk}\varepsilon_{i\ell m} = (\delta_{km}\delta_{j\ell} - \delta_{k\ell}\delta_{jm}), \tag{101}$$

so that we may write

$$\vec{L}^2 = \sum_{j,k,\ell,m} \left( \delta_{j\ell}\delta_{km}x_j \overbrace{(x_\ell p_k}^{p_k x_\ell} - i\hbar\delta_{\ell k})p_m - \delta_{k\ell}\delta_{jm}x_j p_k (\underbrace{p_m x_\ell}_{x_\ell p_m} + i\hbar\delta_{m\ell})) \right) \tag{102}$$

or

$$\vec{L}^2 = \vec{x}^2\vec{p}^2 - 2i\hbar\vec{x}\cdot\vec{p} - (\vec{x}\cdot\vec{p})(\vec{p}\cdot\vec{x}). \tag{103}$$

But (can you show it?)

$$\vec{p}\cdot\vec{x} = \vec{x}\cdot\vec{p} - 3i\hbar, \tag{104}$$

so

$$\vec{L}^2 = \vec{x}^2\vec{p}^2 - (\vec{x}\cdot\vec{p})^2 + i\hbar\vec{x}\cdot\vec{p}. \tag{105}$$

Notice that if $\vec{x}$ and $\vec{p}$ were regarded as numbers, the last term in (105) would not be present.

Now we can try to construct the differential equation implied by

$$H|a'> = E_{a'}|a'>, \tag{106}$$

in a problem with spherical symmetry, $H = \dfrac{\vec{p}^2}{2m} + V(r)$. Project both sides of (106) into the spherical basis $<\vec{r}|$:

$$<\vec{r}|H|a'> = E_{a'}<\vec{r}|a'>, \tag{107}$$

$$\Rightarrow \; <\vec{r}| \; \frac{\vec{p}^2}{2m} \; |a'> \; + \; V(r)<\vec{r}|a'> \; = \; E_{a'}<\vec{r}|a'>, \qquad (108)$$

where $u_{a'}(\vec{r}) = <\vec{r}|a'>$. We now have from (105) that

$$<\vec{r}|\vec{x}^2\vec{p}^2|a'> \; = \; <\vec{r}|\vec{L}^2|a'> \; + \; <\vec{r}|(\vec{x}\cdot\vec{p})^2|a'>$$
$$- \; i\hbar \; <\vec{r}|\vec{x}\cdot\vec{p}|a'>. \qquad (109)$$

We have that

$$<\vec{r}|\vec{x}\cdot\vec{p}|a'> \; = \; \vec{r} \cdot (<\vec{r}|\vec{p}|a'>) \; = \; \vec{r} \cdot \left( \frac{\hbar}{i} \vec{\nabla}_r \; <\vec{r}|a'> \right). \qquad (110)$$

with $\vec{\nabla}_r$ given by (87) above. Therefore

$$<\vec{r}|\vec{x}\cdot\vec{p}|a'> \; = \; \frac{\hbar}{i} \; r \; \frac{\partial}{\partial r} \; <\vec{r}|a'> \; = \; \frac{\hbar}{i} \; r \; \frac{\partial}{\partial r} \; u_{a'}(\vec{r}). \qquad (111)$$

Likewise

$$<\vec{r}|(\vec{x}\cdot\vec{p})^2|a'> \; = \; <\vec{r}|(\vec{x}\cdot\vec{p})(\vec{x}\cdot\vec{p})|a'>$$

$$= \; \left( \frac{\hbar}{i} \; r \; \frac{\partial}{\partial r} \right)\left( \frac{\hbar}{i} \; r \; \frac{\partial}{\partial r} \right) \; <\vec{r}|a'>$$

$$= \; -\hbar^2 \; \left( r^2 \; \frac{\partial^2}{\partial r^2} \; + \; r \; \frac{\partial}{\partial r} \right) \; u_{a'}(\vec{r}) \; , \qquad (112)$$

and

$$<\vec{r}|\vec{x}^2\vec{p}^2|a'> \; = \; r^2 \; <\vec{r}|\vec{p}^2|a'> \; . \qquad (113)$$

Using (111), (112) and (113) in (109), we find that (dividing by $r^2$ )

$$<\vec{r}|\vec{p}^2|a'> \; = \; -\hbar^2 \; \left( \frac{\partial^2}{\partial r^2} \; + \; \frac{2}{r} \; \frac{\partial}{\partial r} \right) \; u_{a'}(\vec{r}) \; + \; \frac{1}{r^2} \; <\vec{r}|\vec{L}^2|a'> . \qquad (114)$$

Now using (97), we get

$$< \vec{r} \,|\, \vec{p}^2 \,|\, a' > \; = \; \left[ - \frac{\hbar^2}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{L_{op}^2}{r^2} \right] u_{a'}(\vec{r}). \qquad (115)$$

Since

$$< \vec{r} \,|\, \vec{p}^2 \,|\, a' > \; = \; -\hbar^2 \, \vec{\nabla}_r^2 \, u_{a'}(\vec{r}), \qquad (116)$$

all we have really accomplished in (115) is to find an explicit expression for the $\vec{\nabla}_r^2$ operator in spherical coordinates (in an especially interesting way, however.) So, using (98):

$$\vec{\nabla}_r^2 \; = \; \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right). (117)$$

Therefore, returning to (108), we have the explicit radial Schrödinger equation:

$$\left[ - \frac{\hbar^2}{2mr^2} \left( \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) - \left( \frac{\vec{L}_{op}}{\hbar} \right)^2 \right) + V(r) \right] u_{a'}(\vec{r}) \; = \; E_{a'} u_{a'}(\vec{r}). \quad (118)$$

We could now proceed in a standard way to separate variables in (118) and find the eigenvalues and eigenvectors of the angular part of the problem using purely differential operator techniques (See, for example, Merzbacher, p.178 and on). Instead, let us proceed by considering what the set of quantum numbers {a'} above consists of. Notice

$$\left. \begin{array}{l} < \vec{r}\,' \,|\, V(r)\vec{L} \; = \; V(r') \, < \vec{r}\,' \,|\, \vec{L} \; = \; V(r')\vec{L}_{op} \, < \vec{r}\,' \,| \;, \\[2mm] < \vec{r}\,' \,|\, \vec{L}V(r) \; = \; \vec{L}_{op} V(r') \, < \vec{r}\,' \,| \; = \; V(r')\vec{L}_{op} \, < \vec{r}\,' \,|. \end{array} \right\} \qquad (119)$$

Therefore

$$[\vec{L}, V(r)] = 0. \tag{120}$$

Also we have that

$$\left[\vec{p}^2, L_j\right] = \left[\sum_i p_i^2, \sum_{k,\ell} \varepsilon_{jk\ell} x_k p_\ell\right] = \sum_{i,k,\ell} \left[p_i^2, x_k p_\ell\right] \varepsilon_{jk\ell}. \tag{121}$$

Now

$$[p_i^2, x_k p_\ell] = p_i[p_i, x_k p_\ell] + [p_i, x_k p_\ell]p_i$$

$$= p_i(\underbrace{x_k[p_i, p_\ell]}_{0} + \underbrace{[p_i, x_k]p_\ell}_{\frac{\hbar}{i}\delta_{ik}}) + (\underbrace{x_k[p_i, p_\ell]}_{0} + \underbrace{[p_i, x_k]p_\ell}_{\frac{\hbar}{i}\delta_{ik}})p_i \tag{122}$$

so that

$$[p_i^2, x_k p_\ell] = 2\frac{\hbar}{i} p_i p_\ell \delta_{ik}. \tag{123}$$

Using this in (121) gives

$$[\vec{p}^2, L_j] = \sum_{i,k,\ell} \varepsilon_{jk\ell} 2\frac{\hbar}{i} p_i p_\ell \delta_{ik} = 2\frac{\hbar}{i} \sum_{k,\ell} \varepsilon_{jk\ell} p_k p_\ell. \tag{124}$$

Notice in (124) we have an object which is symmetric in two summed indices $(k, \ell)$ multiplied into an object which is antisymmetric in the same two indices. The result is zero. We can see this in general as follows. Let's say we have the 2 index objects $A_{ij}$ and $B_{ij}$ and that

$$\left.\begin{array}{l} A_{ij} = -A_{ji}, \\ \text{and} \\ B_{ij} = B_{ji}. \end{array}\right\} \tag{125}$$

Then we have

$$\sum_{i,j} A_{ij}B_{ij} = -\sum_{i,j} A_{ji}B_{ji} \ . \tag{126}$$

But since we are free to rename our indices, this means

$$\sum_{i,j} A_{ij}B_{ij} = -\sum_{i,j} A_{ij}B_{ij} \ . \tag{127}$$

Anything which is equal to minus itself is zero, and so it is for the right hand side of (124):

$$[\vec{p}^2,\vec{L}] = 0. \tag{128}$$

so since $H = \dfrac{\vec{p}^2}{2m} + V(r)$, (120) and (128) imply

$$[H,\vec{L}] = 0. \tag{129}$$

Thus the $\vec{L}$ gives rise to good quantum numbers (See discussion in Ch.4). Eq$^{\underline{n}}$ (129) implies of course that

$$[H,\vec{L}^2] = 0. \tag{130}$$

What's more

$$[\vec{L}^2,L_i] = \left[\sum_j L_j^2,L_i\right] = \sum_j (L_j[L_j,L_i] + [L_j,L_i]L_j), \tag{131}$$

and using (70) above we get

$$[\vec{L}^2,L_i] = i\hbar \sum_{j,k} \varepsilon_{jik}(L_jL_k + L_kL_j). \tag{132}$$

This is again a situation in which a sum over two indices (j,k) is being performed on a symmetric object $[(L_jL_k + L_kL_j)]$ and an antisymmetric one ($\varepsilon_{jik}$). Therefore

$$[\vec{L}^2,\vec{L}] = 0. \tag{133}$$

Now one of the theorems I talked about, and partially proved in Ch.4, said essentially that:  $(A = A^+, B = B^+)$

A,B possess a common
complete set of             $\Leftrightarrow$  [A,B] = 0.
orthonormal eigenkets

Therefore from (129), (130) and (133) I may choose to characterize the set of quantum numbers {a'} in (118) (which is just an explicit version of (106) in a spherical basis) as eigenvalues of the set,

$$\{H,\vec{L}^2,L_3\} \ .$$

I could not, for example, add $L_1$ or $L_2$ to this list since $[L_{1,2},L_3] \neq 0$.  I did not have to choose $L_3$ in the above set; $L_1$ or $L_2$ would have done just as well. However, the choice of $L_3$ is simpler and conventional.  The above represent a complete set for a spinless particle subjected to a spherically symmetric potential.  (This is very nearly the case for an electron in the hydrogen atom.)

Thus, given the above choice of commuting observables, we take

$$|a'> = |n,a,b>, \qquad\qquad (134)$$

where n is a radial quantum number which depends on the nature of the potential (more about this later), and "a" and "b" are eigenvalues of $\vec{L}^2$ and $L_3$:

$$\vec{L}^2|a,b> = \hbar^2 a|a,b>, \qquad\qquad (135)$$

$$L_3|a,b> = \hbar b|a,b>, \tag{136}$$

where factors of $\hbar$ have been inserted for convenience. (a and b are real since $\vec{L}^2$ and $L_3$ are Hermitian.)

Let us find the eigenvalues and eigenvectors of $\vec{L}^2$ and $L_3$. Let us introduce

$$L_\pm = L_1 \pm iL_2, \tag{137}$$

$$L_+^+ = L_- \ . \ \text{(not Hermitian)} \tag{138}$$

We can now show that

$$[L_+, L_-] = 2\hbar L_3, \tag{139}$$

$$[L_3, L_\pm] = \pm\hbar L_\pm. \tag{140}$$

Also

$$[\vec{L}^2, L_\pm] = 0 \tag{141}$$

is obvious. Now consider

$$L_3(L_\pm|a,b>) = L_\pm(L_3 \pm \hbar)|a,b>$$
$$= \hbar(b \pm 1)(L_\pm |a,b>), \tag{142}$$

and

$$\vec{L}^2(L_\pm|a,b>) = L_\pm\hbar^2 a|a,b> = \hbar^2 a(L_\pm|a,b>). \tag{143}$$

Therefore, the $L_\pm$ are ladder operators in the "b" space; they raise or lower the value of this quantum number by one unit (similar to the operators A and $A^+$ in the harmonic oscillator problem.). The conclusion is that

$$L_\pm|a,b> = C_\pm|a,b \pm 1>, \tag{144}$$

or

$$<a,b|L_\mp = C_\pm^*<a,b \pm 1|, \tag{145}$$

where the $C_\pm$ are unknown constants.

Now we have that

$$L_-L_+ = (L_1 - iL_2)(L_1 + iL_2)$$

$$= L_1^2 + L_2^2 + i[L_1,L_2]$$

$$= \vec{L}^2 - L_3^2 - \hbar L_3 \tag{146}$$

so that

$$L_-L_+|a,b> = (\hbar^2 a - \hbar^2 b^2 - \hbar(\hbar b))|a,b>$$

$$= \hbar^2[a - b(b + 1)]|a,b> \tag{147}$$

But since

$$<a,b|L_-L_+|a,b> = |C_+|^2, \tag{148}$$

we have

$$|C_+|^2 = \hbar^2[a - b(b + 1)]. \tag{149}$$

We <u>choose</u> the arbitrary phase to be such that

$$C_+ = \hbar\sqrt{a - b(b + 1)} \quad . \tag{150}$$

Likewise

$$L_+L_- = \vec{L}^2 - L_3^2 + \hbar L_3 , \tag{151}$$

so that

$$L_+L_-|a,b> = \hbar^2[a - b(b - 1)]|a,b> , \qquad (152)$$

and since

$$|C_-|^2 = <a,b|L_+L_-|a,b>, \qquad (153)$$

we have

$$|C_-|^2 = \hbar^2[a - b(b - 1)], \qquad (154)$$

so

$$C_- = e^{i\phi}\hbar\sqrt{a - b(b - 1)}, \qquad (155)$$

where $\phi$ is an unknown phase. Actually $\phi$ is fixed from our previous choice for $C_+$ in (150). To see this consider

$$L_+[L_-|a,b> = e^{i\phi}\hbar\sqrt{a - b(b - 1)} \ |a,b - 1>] , \qquad (156)$$

$$\Rightarrow \quad L_+L_-|a,b> = e^{i\phi}\hbar^2\sqrt{a - b(b - 1)} \ \sqrt{a - (b - 1)b} \ |a,b>$$

$$= e^{i\phi}\hbar^2[a - b(b - 1)]|a,b> . \qquad (157)$$

Comparing (157) with (152) implies that $e^{i\phi} = 1$. Therefore, we have found that

$$L_\pm|a,b> = \hbar\sqrt{a - b(b \pm 1)} \ |a,b \pm 1> . \qquad (158)$$

Now, what are the allowed values of a and b? We can easily show that the expectation value of the square of a Hermitian operator is always nonnegative. Therefore since

$$L_1^2 + L_2^2 = \vec{L}^2 - L_3^2, \qquad (159)$$

we have

$$(L_1^2 + L_2^2)|a,b> = \hbar^2(a - b^2)|a,b>, \qquad (160)$$

and since, when multiplying on the left by $<a,b|$, the left hand side of (160) is guaranteed to be nonnegative, we have that

$$a - b^2 \geq 0. \qquad (161)$$

Now we know that the $L_\pm$ raise or lower the value of "b" in $|a,b>$ by ±1 unit while keeping the value of "a" unchanged. Hence, for a fixed "a" value, we have

$$-\sqrt{a} \leq b \leq \sqrt{a}. \qquad (162)$$

Therefore, for a given "a" value there is a largest value of "b"; let us call this $b_{max}$ ($b_{max}$ does not necessarily equal $\sqrt{a}$ since there may exist more restrictive conditions. (162) only shows b is bounded.) Thus by definition we must have

$$L_+|a,b_{max}> = 0. \qquad (163)$$

Likewise call the minimum value of b, for a given "a", $b_{min}$. Therefore

$$L_-|a,b_{min}> = 0. \qquad (164)$$

From (146) and (151) we have

$$L_\mp L_\pm = \vec{L}^2 - L_3(L_3 \pm \hbar). \qquad (165)$$

Thus, applying (165) to (163) and (164), we find

$$L_-L_+|a,b_{max}> = \hbar^2[a - b_{max}(b_{max} + 1)]|a,b_{max}> = 0, \left.\begin{array}{r}\\\\\end{array}\right\} \quad (166)$$

$$L_+L_-|a,b_{min}> = \hbar^2[a - b_{min}(b_{min} - 1)]|a,b_{min}> = 0.$$

The equations (166) together imply that

$$b_{max}(b_{max} + 1) = b_{min}(b_{min} - 1) , \quad (167)$$

or

$$(b_{max} + b_{min})(b_{max} - b_{min} + 1) = 0. \quad (168)$$

But since $b_{max} \geq b_{min}$, we get that

$$b_{max} = -b_{min}. \quad (169)$$

Let us say it takes $2\ell$ steps ($2\ell$ is a positive or zero integer) to go from $b = b_{min}$ to $b = b_{max}$ in steps of one:

$$b_{max} = b_{min} + 2\ell, \quad (170)$$

where the possible $\ell$ values are

$$\ell = 0 \quad \overset{X}{\frac{1}{2}},1,\overset{X}{\frac{3}{2}},\ldots \quad (171)$$

(The reason for the "X" above $\ell = \frac{1}{2},\frac{3}{2},\ldots$ will be discussed shortly).  Then because $b_{min} = -b_{max}$, we have

$$b_{max} = \ell, \quad (172)$$

and from either of Eq$^{\underline{n}}$s (166), we get

$$a = \ell(\ell + 1). \quad (173)$$

So, for a given $\ell$ value, we have the possible "b" values:

$$b = b_{min}, b_{min} + 1, \ldots, b_{max} - 1, b_{max}$$

or

$$b = -\ell, -\ell + 1, \ldots, \ell - 1, \ell$$

(174)

It is more conventional to relabel "b" as "m", called the "magnetic quantum number."  As an example, let us choose $\ell=3$. Then m can take on $2\ell + 1 = 7$ values:

$$\ell = 3 \begin{cases} \underline{\quad} & m = 3 & \downarrow L_- \\ \underline{\quad} & 2 \\ \underline{\quad} & 1 \\ \underline{\quad} & 0 \\ \underline{\quad} & -1 \\ \underline{\quad} & -2 \\ \underline{\quad} & -3 & \uparrow L_+ \end{cases}$$

There is a subtlety involved in the labeling of the eigenstates of $\vec{L}^2$ and $L_3$.  Since $\vec{L}^2$ and $L_3$ commute (and therefore are simultaneously measurable), we may regard the state $|\ell,m\rangle$ as a type of direct product, which would seem to imply that

$$|\ell,m\rangle \overset{?}{=} |\ell\rangle \otimes |m\rangle.$$

However, because $\vec{L}^2$ does not project entirely into $\theta$ space (see (93) and (94) above), the value of m enters the eigenvalue equation for $\vec{L}^2$ (see Eq$\underline{n}$ (188) below).  Therefore, we will define*

$$|\ell,m\rangle \equiv \underset{\substack{\text{projects into} \\ \theta \text{ space}}}{|\ell(m)\rangle} \otimes \underset{\substack{\text{projects into} \\ \phi \text{ space}}}{|m\rangle},$$

---

* Technically speaking, these states are separable but not true direct products.

$|\ell,m>$ is just a way of <u>labelling</u> the more proper object on
the right hand side. The $\ell(m)$ notation is supposed to
indicate that "$\ell$" is the quantum number associated with the $\theta$
eigenvalue equation, but that "m" enters this equation as a
parameter. The same sort of subtlety affects the labeling of
the Hilbert space description of the radial eigenstates; that
is, the value of $\ell$ enters the radial eigenvalue equation (see
Eq$^{\underline{n}}$ (253) below), and we shall define

$$|n,\ell,m> \equiv |n(\ell)> \otimes |\ell,m>.$$

These Hilbert spaces are such that

$$<r,\theta,\phi|n,\ell,m> = <r|n(\ell)> <\theta|\ell(m)> <\phi|m>$$
$$= u_{n\ell}(r)u_{\ell m}(\theta)u_m(\phi).$$

The eigenvalue equations for $u_m(\phi)$, $u_{\ell m}(\theta)$ and $u_{n\ell}(r)$ are given
by (180), (188) and (253) below, respectively.

Let us relabel our states as

$$|a,b> \rightarrow |\ell,m>. \tag{175}$$

We have therefore found the eigenvalues of $\vec{L}^2$ and $L_3$ as

$$\vec{L}^2|\ell,m> = \hbar^2\ell(\ell + 1)|\ell,m>, \tag{176}$$

$$L_3|\ell,m> = \hbar m|\ell,m>, \tag{177}$$

and

$$L_\pm|\ell,m> = \hbar\sqrt{(\ell \mp m)(\ell \pm m + 1)}\ |\ell,m \pm 1>. \tag{178}$$

Actually, "half-integer" values of $\ell$ ($\ell = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$) are not allowed. One way to see this is as follows. The eigenvalue equation for $L_3$ is

$$L_3 |m\rangle = m\hbar |m\rangle. \tag{179}$$

Projecting both sides of (179) into $\langle\phi|$ and calling $U_m(\phi) = \langle\phi|m\rangle$, we get

$$\frac{\hbar}{i} \frac{\partial}{\partial\phi} u_m(\phi) = m\hbar u_m(\phi). \tag{180}$$

The solution to (180) is

$$u_m(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi}, \tag{181}$$

which is normalized so that

$$\int_0^{2\pi} d\phi |u_m|^2 = 1 . \tag{182}$$

Now consider half-integer values of $\ell$. From (181) it would seem that

$$\langle\phi|m\rangle = \frac{1}{\sqrt{2\pi}} e^{im\phi}, \tag{183}$$

where m is also required, by (174), to take on half-integer values. Therefore given (183) we have

$$\langle\phi + 2\pi|m\rangle = -\frac{1}{\sqrt{2\pi}} e^{im\phi} . \tag{184}$$

But $\langle\phi|$ and $\langle\phi + 2\pi|$ pick out the <u>same</u> point in coordinate space. Therefore, the spatial wavefunctions of half-integer

$\ell$ are not single valued, a condition we must require for the transformation between the $|\phi>$ and $|m>$ bases.

There are other arguments as to why half-integer $\ell$ values are not allowed. The end result is to limit $\ell$ to the values

$$\ell = 0,1,2,3$$

and m to

$$m = -\ell, -\ell + 1, \ldots, \ell - 1, \ell$$

($2\ell + 1$ values) for each $\ell$.

Now we know the eigenvalues of $\vec{L}^2$ and $L_3$. We would like to find the explicit eigenvectors also. We write (the $Y_{\ell m}$ are called "spherical harmonics")

$$Y_{\ell m}(\theta,\phi) \equiv <\theta,\phi|\ell,m>$$
$$= <\theta|\ell(m)> <\phi|m>$$
$$\equiv u_{\ell m}(\theta)u_m(\phi). \qquad (185)$$

The eigenvalue equation for $u_m(\phi)$ is written down in (180), and it's normalized solution is (181). The eigenvalue equation for $u_{\ell m}(\theta)$ comes from

$$\vec{L}^2|\ell,m> = \hbar^2\ell(\ell + 1)|\ell,m>. \qquad (186)$$

We have to start out with $|\ell,m>$ in (186) (and <u>not</u> $|\ell(m)>$) since we don't know the effect of $\vec{L}^2$ on $<\theta|$ but only on $<\theta,\phi|$ from (97) and (98).

Projected into $<\theta,\phi|$ space (using (97) and (98) above gives

$$- \left[ \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left( \sin\theta \frac{\partial}{\partial\theta} \right) \right] u_{\ell m}(\theta) u_m(\phi)$$

$$= \ell(\ell + 1) u_{\ell m}(\theta) u_m(\phi), \qquad (187)$$

or, making the replacement $\frac{\partial^2}{\partial\phi^2} \rightarrow -m^2$ from (180), and then

dividing both sides by $u_m(\phi)$, we get

$$- \left[ \frac{-m^2}{\sin^2\theta} + \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left( \sin\theta \frac{\partial}{\partial\theta} \right) \right] u_{\ell m}(\theta) = \ell(\ell + 1) u_{\ell m}(\theta). \quad (188)$$

The solution to this equation determines the eignevectors $u_{\ell m}(\theta)$.

Rather than trying to solve (188) directly, we will generate the solutions by operator techniques, using the ladder operators $L_\pm$. We can construct all of the $|\ell, m>$ by considering $|\ell, \ell>$ and then applying $L_-$ $(\ell - m)$ times:

$$(L_-)^{\ell-m} |\ell, \ell> = \hbar\sqrt{2\ell} \; (L_-)^{(\ell-m)-1} |\ell, \ell - 1>$$

$$= \hbar^2\sqrt{2\ell} \; \sqrt{(2\ell - 1)2} \; (L_-)^{(\ell-m)-2} |\ell, \ell - 2>$$

$$= \hbar^3\sqrt{2\ell} \; \sqrt{(2\ell - 1)2} \; \sqrt{(2\ell - 2)3} (L_-)^{(\ell-m)-3} |\ell, \ell - 3>, \quad (189)$$

or, in general, after $(\ell - m)$ applications of $L_-$:

$$(L_-)^{(\ell-m)} |\ell, \ell>$$

$$= (\hbar)^{\ell-m} \underbrace{\sqrt{2\ell} \sqrt{(2\ell - 1)2} \ldots \sqrt{(2\ell - (\ell - m - 1))(\ell - m)}}_{(\ell - m \text{ factors})} \; |\ell, m>.$$

$$(190)$$

Let's look at some of the individual pieces that make up the overall factor in (190). We recognize the combination,

$$(2\ell)(2\ell-1)\ \ldots\ \underbrace{(2\ell-(\ell-m-1))}_{(\ell+m+1)}$$

in (190), which we can write as

$$\frac{(2\ell)(2\ell-1)\ \ldots\ (2)(1)}{(\ell+m)(\ell+m-1)\ \ldots\ (2)(1)} = \frac{(2\ell)!}{(\ell+m)!}\ . \tag{191}$$

We also have the combination

$$(1)(2)\ldots(\ell-m) = (\ell-m)!\ . \tag{192}$$

Therefore, we may write (190) as

$$(L_-)^{\ell-m}|\ell,\ell> = (\hbar)^{\ell-m}\sqrt{\frac{(2\ell)!(\ell-m)!}{(\ell+m)!}}\ |\ell,m>, \tag{193}$$

or, solving for $|\ell,m>$:

$$|\ell,m> = \frac{1}{(\hbar)^{\ell-m}}\sqrt{\frac{(\ell+m)!}{(2\ell)!(\ell-m)!}}\ (L_-)^{\ell-m}|\ell,\ell>. \tag{194}$$

I remind you of the effect of $L_1$ and $L_2$ (compare with (93) and (94); here I am using the $<\hat{n}| = <\theta,\phi|$ notation):

$$<\hat{n}|L_1 = \frac{\hbar}{i}\left(-\sin\phi\ \frac{\partial}{\partial\theta} - \cos\phi\ \cot\theta\ \frac{\partial}{\partial\phi}\right)<\hat{n}|, \tag{195}$$

$$<\hat{n}|L_2 = \frac{\hbar}{i}\left(\cos\phi\ \frac{\partial}{\partial\theta} - \sin\phi\ \cot\theta\ \frac{\partial}{\partial\phi}\right)<\hat{n}|. \tag{196}$$

So therefore

$\langle \hat{n}|L_+ =$

$$\frac{\hbar}{i}\left((-\sin\phi + i\cos\phi)\frac{\partial}{\partial\theta}\quad -(\cos\phi + i\sin\phi)\cot\theta\,\frac{\partial}{\partial\phi}\right)\langle \hat{n}|, \quad (197)$$

where the first group is $ie^{i\phi}$ and the second is $e^{i\phi}$.

or

$$\langle \hat{n}|L_+ = \hbar e^{i\phi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\,\frac{\partial}{\partial\phi}\right)\langle \hat{n}|. \quad (198)$$

Also

$\langle \hat{n}|L_- =$

$$\frac{\hbar}{i}\left((-\sin\phi - i\cos\phi)\,\frac{\partial}{\partial\theta} - (\cos\phi - i\sin\phi)\cot\theta\,\frac{\partial}{\partial\phi}\right)\langle \hat{n}|. \quad (199)$$

or

$$\langle \hat{n}|L_- = -\hbar e^{-i\phi}\left(\frac{\partial}{\partial\theta} - i\cot\theta\,\frac{\partial}{\partial\phi}\right)\langle \hat{n}|. \quad (200)$$

Now consider

$$\langle \hat{n}|L_-|\ell,m\rangle = -\hbar e^{-i\phi}\left(\frac{\partial}{\partial\theta} - i\cot\theta\,\frac{\partial}{\partial\phi}\right)\langle \hat{n}|\ell,m\rangle. \quad (201)$$

where $\langle \hat{n}|\ell,m\rangle = Y_{\ell m}(\hat{n})$.  We know that

$$Y_{\ell m}(\theta,\phi) = u_{\ell m}(\theta)\,\frac{1}{\sqrt{2\pi}}\,e^{im\phi}, \quad (202)$$

where $m = 0,\pm 1,\pm 2,\ldots,\pm\ell$ $(2\ell + 1$ values$)$.  Therefore

$$<\hat{n}|L_-|\ell,m> = -\frac{\hbar}{\sqrt{2\pi}}\ e^{i(m-1)\phi}\left(\frac{d}{d\theta} + m\cot\theta\right)u_{\ell m}(\theta).\qquad(203)$$

Consider the identity:

$$(\sin\theta)^{1-m}\ \frac{d}{d\cos\theta}\ [(\sin\theta)^m F(\theta)]$$

$$= (\sin\theta)^{1-m}\ \underbrace{\frac{d\theta}{d\cos\theta}}\ \underbrace{\frac{d}{d\theta}[(\sin\theta)^m F(\theta)]}$$

$$\underbrace{-\frac{1}{\sin\theta}}\qquad \underbrace{(\sin\theta)^m\ \frac{dF(\theta)}{d\theta} + m(\sin\theta)^{m-1}\cos\theta\ F(\theta)}$$

$$= -\left(\frac{d}{d\theta} + m\cot\theta\right)F(\theta).\qquad\qquad(204)$$

But the right hand side of (204) is the same as the structure in (203), so we may make the replacement

$$<\hat{n}|L_-|\ell,m> =$$

$$\frac{\hbar}{\sqrt{2\pi}}\ e^{i(m-1)\phi}\ \sin(\theta)^{1-m}\ \frac{d}{d\cos\theta}\ [(\sin\theta)^m\ u_{\ell m}(\theta)].\qquad(205)$$

Now consider

$$<\hat{n}|L_-L_-|\ell,m> = -\hbar e^{-i\phi}\left(\frac{\partial}{\partial\theta} - i\cot\theta\ \frac{\partial}{\partial\phi}\right)<\hat{n}|L_-|\ell,m>$$

$$= \frac{-(\hbar)^2}{\sqrt{2\pi}}\ e^{i(m-2)\phi}\left(\frac{d}{d\theta} + (m-1)\cot\theta\right)$$

$$\cdot\left[(\sin\theta)^{1-m}\ \frac{d}{d\cos\theta}\left((\sin\theta)^m u_{\ell m}(\theta)\right)\right].\qquad(206)$$

Employing (204) again in (206) (with $m \to m-1$) gives

6.38

$$< \hat{n} | (L_-)^2 | \ell, m >$$

$$= \frac{(\hbar)^2}{\sqrt{2\pi}} \; e^{i(m-2)\phi} (\sin \theta)^{2-m} \frac{d^2}{d \cos \theta^2} \left[ (\sin \theta)^m u_{\ell m}(\theta) \right]. \quad (207)$$

Seeing the pattern that seems to have developed, we may now prove by induction that

$$< \hat{n} | (L_-)^k | \ell, m >$$

$$= \frac{(\hbar)^k}{\sqrt{2\pi}} \; e^{i(m-k)\phi} (\sin \theta)^{k-m} \left( \frac{d}{d \cos \theta} \right)^k \left[ (\sin \theta)^m u_{\ell m}(\theta) \right]. \quad (208)$$

Let's set $m = \ell$ and $k = \ell - m$ in (208):

$$< \hat{n} | (L_-)^{\ell - m} | \ell, \ell >$$

$$= \frac{(\hbar)^{\ell - m}}{\sqrt{2\pi}} \; e^{im\phi} (\sin \theta)^{-m} \left( \frac{d}{d \cos \theta} \right)^{\ell - m} \left[ (\sin \theta)^\ell u_{\ell\ell}(\theta) \right]. \quad (209)$$

But from (194)

$$< \hat{n} | \ell, m > = (\hbar)^{m-\ell} \sqrt{\frac{(\ell + m)!}{(2\ell)!(\ell - m)!}} \; < \hat{n} | (L_-)^{\ell - m} | \ell, \ell >, \quad (210)$$

and so we find $(Y_{\ell m}(\hat{n}) \equiv < \hat{n} | \ell, m >)$

$$Y_{\ell m}(\theta, \phi) = \frac{e^{im\phi}}{\sqrt{2\pi}} \sqrt{\frac{(\ell + m)!}{(2\ell)!(\ell - m)!}} \; (\sin \theta)^{-m} \left( \frac{d}{d \cos \theta} \right)^{\ell - m}$$

$$\left[ (\sin \theta)^\ell u_{\ell\ell}(\theta) \right]. \quad (211)$$

Therefore, we will have a general expression for <u>all</u> the $Y_{\ell m}(\hat{n})$ if we can find the explicit expression for $u_{\ell \ell}(\theta)$.  Now remember that

$$L_+|\ell,\ell> = 0, \tag{212}$$

so that

$$<\hat{n}|L_+|\ell,\ell> = \hbar e^{i\phi}\left(\frac{\partial}{\partial\theta} + i \cot \theta \frac{\partial}{\partial\phi}\right)\underbrace{<\hat{n}|\ell,\ell>}, \tag{213}$$

$$\frac{e^{i\ell\phi}}{\sqrt{2\pi}} u_{\ell \ell}(\theta)$$

which gives us the first order differential equation:

$$\left[\frac{d}{d\theta} - \ell \cot \theta\right] u_{\ell \ell}(\theta) = 0. \tag{214}$$

It's easy to check that the solution to (214) is

$$u_{\ell \ell}(\theta) = C_\ell (\sin \theta)^\ell . \tag{215}$$

(Do it.)  $C_\ell$ is an unknown constant which is determined by the normalization condition

$$\int d\Omega_{\hat{n}}\left|u_{\ell \ell}(\theta) \frac{e^{i\ell\phi}}{\sqrt{2\pi}}\right|^2 = 1. \tag{216}$$

(Eq$^{\underline{n}}$ (216) can be viewed as saying the probability of seeing the particle somewhere in angular space is unity.)
Explicitly, this gives

$$|C_\ell|^2 \int_0^\pi d\theta \, \sin\theta \, \sin^{2\ell}\theta = 1, \qquad (217)$$

or

$$|C_\ell|^2 \int_{-1}^{1} d(\cos\theta)\sin^{2\ell}\theta = 1, \; . \qquad (218)$$

Setting

$$x = \cos\theta \qquad (219)$$

we then get

$$|C_\ell|^2 \int_{-1}^{1} (dx)(1 - x^2)^\ell = 1. \qquad (220)$$

The integral in (220) can be done by parts (this will be a homework problem) to yield

$$|C_\ell|^2 \left( 2 \, \frac{(2^\ell \ell!)^2}{(2\ell + 1)!} \right) = 1. \qquad (221)$$

It is conventional to choose the phase such that

$$C_\ell = \frac{(-1)^\ell}{2^\ell \ell!} \sqrt{\frac{(2\ell + 1)!}{2}} \; , \qquad (222)$$

and so

$$u_{\ell\ell}(\theta) = \frac{(-1)^\ell}{2^\ell \ell!} \sqrt{\frac{(2\ell + 1)!}{2}} \, \sin^\ell\theta \; . \qquad (223)$$

Using (223) in (211) now gives the general result

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

$$\times \left(\frac{d}{d \cos \theta}\right)^{\ell-m} [(\sin \theta)^{2\ell}]. \qquad (224)$$

Because of the normalization of the spherical basis, Eq$^{\underline{n}}$(62), we have

$$< \ell',m' | \left[ \int d\Omega_{\hat{n}} |\hat{n}> < \hat{n}| = 1 \right] |\ell,m>, \qquad (225)$$

which means that

$$\int d\Omega_{\hat{n}} Y^*_{\ell'm'}(\hat{n}) Y_{\ell m}(\hat{n}) = \delta_{\ell\ell'} \delta_{mm'} . \qquad (226)$$

Eq$^{\underline{n}}$ (226) is completeness for spherical harmonics in angular space.  We also have

$$< \hat{n} | \left[ \sum_{\ell,m} |\ell,m> < \ell,m| = 1 \right] | \hat{n}'> , \qquad (227)$$

or

$$\sum_{\ell,m} Y_{\ell m}(\hat{n}) Y^*_{\ell m}(\hat{n}') = < \hat{n} | \hat{n}'>, \qquad (228)$$

which expresses completeness of the $|\ell,m>$ basis states.  The $< \hat{n} | \hat{n}'>$ is a <u>spherical</u> Dirac delta function.  We can get an explicit form for it by requiring that

$$1_{\hat{n}} = \int d\Omega_{\hat{n}} |\hat{n}> < \hat{n} | \int d\Omega_{\hat{n}'} |\hat{n}'> < \hat{n}' | ,$$

$$= \int d\Omega_{\hat{n}} d\Omega_{\hat{n}'} |\hat{n}> < \hat{n} | \hat{n}'> < \hat{n}' | ,$$

$$= \int d(\cos \theta) d(\cos \theta') d\phi d\phi' |\theta,\phi> < \theta,\phi | \theta',\phi'> < \theta',\phi' | . \qquad (229)$$

Comparing this with the original expression

$$1_{\hat{n}} = \int d(\cos \theta) d\phi |\theta,\phi> <\theta,\phi| \ , \tag{230}$$

means that we may take

$$<\hat{n}|\hat{n}'> = \delta(\cos \theta - \cos \theta')\delta(\phi - \phi'). \tag{231}$$

[Using the delta function rule (see Jackson, p.30)

$$\delta(f(x)) = \sum_i \frac{1}{\left|\frac{df}{dx}(x_i)\right|} \delta(x - x_i),$$

where the sum is over the simple zeros of $f(x)$, located at $x = x_i$, we may write

$$\delta(\cos \theta - \cos \theta') = \frac{1}{\sin \theta'} \delta(\theta - \theta'),$$

where both $\theta$ and $\theta'$ are assumed to be in the range from 0 to $\pi$.]   Therefore

$$\sum_{\ell=0}^{\infty} \sum_{m=\ell}^{\ell} Y_{\ell m}(\theta,\phi) Y_{\ell m}^*(\theta',\phi') = \delta(\cos \theta - \cos \theta')\delta(\phi - \phi'). \tag{232}$$

[Notice we haven't really proven either (226) or (232); the proofs require more sophisticated analysis.]   A useful connection is

$$Y_{\ell 0}(\theta,\phi) = \frac{(-1)^{\ell}}{2^{\ell}\ell!} \sqrt{\frac{(2\ell + 1)}{4\pi}} \left(\frac{d}{d \cos \theta}\right)^{\ell} (\sin \theta)^{2\ell}$$

$$= \sqrt{\frac{(2\ell + 1)}{4\pi}} \; P_\ell(\cos \theta), \qquad (233)$$

where $P_\ell(x)$ is called a "Legendre polynomial."  (Notice that when $m = 0$, there is no $\phi$ dependence in $Y_{\ell 0}(\theta,\phi)$.)  By definition then

$$P_\ell(x) = \frac{1}{2^\ell \ell!} \left(\frac{d}{dx}\right)^\ell (x^2 - 1)^\ell . \qquad (234)$$

I am not going to go through all of the explicit steps, but in the same way that I showed (194) to be true by operating $(\ell - m)$ times with $L_-$ on $|\ell,\ell>$ to give $|\ell,m>$, we can also start at the other end and operate $(\ell + m)$ times with $L_+$ on $|\ell,-\ell>$ to give $|\ell,m>$.  The result is

$$|\ell,m> = (\hbar)^{-(\ell+m)} \sqrt{\frac{(\ell - m)!}{(2\ell)!(\ell + m)!}} \; (L_+)^{\ell+m} \, |\ell,-\ell>. \qquad (235)$$

We can also show that

$$<\hat{n}|(L_+)^{\ell+m} |\ell,-\ell> = \frac{(\hbar)^{(\ell+m)}}{\sqrt{2\pi}} \, (-1)^{\ell+m} \, e^{im\phi}$$

$$\cdot \, (\sin \theta)^m \left(\frac{d}{d \cos \theta}\right)^{\ell+m} \left[ (\sin \theta)^\ell \, u_{\ell-\ell}(\theta) \right]. \qquad (236)$$

From our earlier $Y_{\ell m}(\theta,\phi)$ expression, (224), we can show that

$$Y_{\ell-\ell}(\theta,\phi) = \frac{e^{-i\ell\phi}}{2^\ell \ell!} \sqrt{\frac{(2\ell + 1)!}{4\pi}} \, (\sin \theta)^\ell , \qquad (237)$$

from which we can identify

$$u_{\ell-\ell}(\theta) = \frac{1}{2^\ell \ell!} \sqrt{\frac{(2\ell + 1)!}{2}} \quad (\sin \theta)^\ell \ . \tag{238}$$

(Notice that unlike (223) there is no factor of $(-1)^\ell$ here.) The above steps then lead in the same manner as before to the alternate expression

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

$$\cdot \left(\frac{d}{d \cos \theta}\right)^{\ell+m} \left[ (\sin \theta)^{2\ell} \right] \ . \tag{239}$$

You should now go back to Eq$^{\underline{n}}$ (224) and carefully compare it to (239) above.) By using the expression (224) when $m \geq 0$ and the expression (239) when $m \leq 0$, one may also write, for example

$$Y_{\ell m}(\theta,\phi) = \frac{(-1)^{(m-|m|)/2}(-1)^\ell}{2^\ell \ell!} e^{im\phi} \sqrt{\frac{(2\ell + 1)(\ell + |m|)!}{4\pi(\ell - |m|)!}}$$

$$\times (\sin \theta)^{-|m|} \left(\frac{d}{d \cos \theta}\right)^{\ell-|m|} \left[ (\sin \theta)^{2\ell} \right] \ . \tag{240}$$

We can read off from (240) the symmetry property

$$(-1)^m Y_{\ell-m}(\theta,-\phi) = Y_{\ell m}(\theta,\phi) \ , \tag{241}$$

but since

$$Y_{\ell m}^*(\theta,\phi) = Y_{\ell m}(\theta,-\phi) \ , \tag{242}$$

we may write (241) as

$$(-1)^m Y^*_{\ell,-m}(\theta,\phi) = Y_{\ell m}(\theta,\phi) \ . \tag{243}$$

We can tie this discussion into the parity operator (introduced in Chapter 3) for which, by definition

$$\langle \hat{n}| \mathbb{P} = \langle -\hat{n}| \ , \tag{244}$$

where we may take

$$\langle \hat{n}| = \begin{cases} \langle \pi - \theta, \phi + \pi| , & \text{if } \phi < \pi \\ \langle \pi - \theta, \phi - \pi| , & \text{if } \phi \geq \pi \ . \end{cases} \tag{245}$$

Now we have that

$$\begin{aligned} \cos(\pi - \theta) &= -\cos\theta, \\ \sin(\pi - \theta) &= \sin\theta, \end{aligned} \tag{246}$$

which helps us to see from (240) that

$$Y_{\ell m}(\pi - \theta, \phi \pm \pi) = Y_{\ell m}(\theta,\phi)e^{\pm im\pi}\underbrace{(-1)^{\ell-|m|}}_{(-1)^{\ell}}. \tag{247}$$

Therefore

$$\langle \hat{n}| \mathbb{P} |\ell,m\rangle = \langle -\hat{n}|\ell,m\rangle = (-1)^{\ell}\langle \hat{n}|\ell,m\rangle. \tag{248}$$

Since (248) is true for all $\langle \hat{n}|$, we have that

$$\mathbb{P} |\ell,m\rangle = (-1)^{\ell}|\ell,m\rangle \ . \tag{249}$$

The words that go with Eq$^{\underline{n}}$ (249) say "the parity of the state $|\ell,m\rangle$ is $(-1)^{\ell}$."

We now return to the Schrödinger equation in spherical coordinates, Eq$\underline{n}$ (118).  We now know that

$$u_{a'}(\vec{r}) = u_{n\ell}(r)Y_{\ell m}(\theta,\phi), \tag{250}$$

and also that

$$L^2_{op} <\hat{n}|\ell,m> = <\hat{n}|\vec{L}^2|\ell,m> = \hbar^2\ell(\ell + 1)<\hat{n}|\ell,m>, \tag{251}$$

or

$$L^2_{op}Y_{\ell m}(\theta,\phi) = \hbar^2\ell(\ell + 1)Y_{\ell m}(\theta,\phi), \tag{252}$$

so that (118) is equivalent to

$$\left[-\frac{\hbar^2}{2mr^2}\left(\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \ell(\ell + 1)\right) + V(r)\right]u_{n\ell}(r) = E_{n\ell}u_{n\ell}(r). \tag{253}$$

Notice that the magnetic quantum number, m, does not enter in (253).  This equation determines the energy levels of the system; therefore, the energies are independent of m for a problem with spherical symmetry and we have a 2$\ell$ + 1 fold degeneracy (at least) of each energy level labeled by (n$\ell$). Also note that, as stated earlier, $\ell$ enters this equation simply as a *parameter*; the *quantum number* determined by this equation is "n".  In the next chapter we will examine solutions to (253) for various forms for the potential V(r).

## Problems

1. Answer the question in the notes, bottom of page 6.9.

2. Using expressions (93)-(95) of the notes, show that (96) is true.

3.(a) Given an operator U with the properties

$$[\vec{L}^2,U] = 0$$
$$[L_3,U] = -\ell\hbar U,$$

   show that

$$U|\ell,\ell\rangle = \text{const.}|\ell,0\rangle.$$

(b) Given an operator V with the properties

$$[L_+,V] = 0$$
$$[L_3,V] = \hbar V,$$

   show that

$$V|\ell,\ell\rangle = \text{const.}|\ell+1,\ell+1\rangle.$$

(c) Given an operator W with the properties,

$$[L_-,W] = 0,$$
$$[L_3,W] = -\hbar W,$$

   find:

$$W|\ell,-\ell\rangle = ?$$

4. Show Eq.(139) and Eq.(140) of the text (Ch.6).

5. Do the integral in Eq.(220) of Ch.6 of the notes.

6. Using Eq.(224) of Ch.6, write out the explicit forms for the spherical harmonics:

$$Y_{00},\ Y_{11},\ Y_{10},\ Y_{1-1}.$$

7. <u>Prove</u>:  The expectation value of the square of a
Hermitian operator is nonnegative.  (I used this on p.6.27
of the notes.  This is essentially a one-line proof.)

8. Prove Eq.(208) of Ch.6 by induction. (That is, assume it
is true for k, and use this to show it then holds for the
k+1 case.)

**Other  Problems**

9. The wavefunction of a bound particle is given by

$$\Psi(\vec{r},0) = xz\Psi(r),$$

where $r = |\vec{r}|$.

    (a)  If $\vec{L}^2$ is measured at t = 0, what value is found?

    (b)  What possible values of $L_z$ will measurement find at
t = 0, and with what probability will they occur? [Hint:
See Table 9.1, p.369, of Liboff.]

10.(a)  Evaluate:
$$[L_3,\phi] = ?$$

($\phi$ is an operator whose eigenvalue is the spherical
azimuthal angle:
$$\phi|\phi'> = \phi'|\phi'>.)$$

    (b)  Apply (a) to evaluate the quantity:

$$e^{-iL_3\phi'/\hbar} \ \phi \ e^{iL_3\phi'/\hbar} = ?$$

($\phi'$ is a number.) [If you can't figure out part (a), I will give you the answer, but you will then get no credit for (a).]

**Other problems**

11. Assume a particle has an orbital angular momentum with $L_3' = \hbar m$ and $(\vec{L}^2)' = \hbar^2 \ell(\ell + 1)$. Show that in this state:

    (a) $\qquad\qquad\qquad\qquad$ $\langle L_x \rangle = \langle L_y \rangle = 0,$

    (b) $\qquad\qquad$ $\langle L_x^2 \rangle = \langle L_y^2 \rangle = \frac{1}{2}\hbar^2(\ell(\ell + 1) - m^2).$

12. Can one measure a particle's momentum, $\vec{p}$, and angular momentum, $\vec{L}$, along the <u>same</u> coordinate axis simultaneously? What quantity must I compute in order to answer this question? Compute it!

**CHAPTER 7:    The Three Dimensional Radial Equation**


Let us recap the situation.  We started out as usual by projecting

$$H|a'> = E_{a'}|a'>, \tag{1}$$

where

$$H = \frac{\vec{p}^2}{2m} + V(r), \tag{2}$$

into a spherical basis:

$$<\vec{r}| \left( \frac{\vec{p}^2}{2m} + V(r) \right) |a'> = E_{a'}<\vec{r}|a'>. \tag{3}$$

Using the definition

$$u_{a'}(\vec{r}) \equiv <\vec{r}|a'>, \tag{4}$$

and the result

$$<\vec{r}|\vec{p}^2|a'> = \left[ -\frac{\hbar^2}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{L_{op}^2}{r^2} \right] u_{a'}(\vec{r}), \tag{5}$$

we got

$$\left[ -\frac{\hbar^2}{2mr^2} \left( \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) - \left( \frac{L_{op}^2}{\hbar^2} \right) \right) + V(r) \right] u_{a'}(\vec{r}) = E_{a'} u_{a'}(\vec{r}). \tag{6}$$

We know the solution to (6) can be written

$$u_{a'}(\vec{r}) = u_{n\ell}(r) Y_{\ell m}(\theta,\phi). \tag{7}$$

and the radial eigenvalue equation we are to solve is

$$\left[ -\frac{\hbar^2}{2mr^2} \left( \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) - \ell(\ell + 1) \right) + V(r) \right] u_{n\ell}(r) = E_{n\ell} u_{n\ell}(r). \quad (8)$$

We can cast this equation into a more convenient form by introducing

$$u_{n\ell}(r) = \frac{R_{n\ell}(r)}{r}, \quad (9)$$

for which we get

$$\left[ -\frac{\hbar^2}{2m} \left( \frac{d^2}{dr^2} - \frac{\ell(\ell + 1)}{r^2} \right) + V(r) \right] R_{n\ell}(r) = E_{n\ell} R_{n\ell}(r). \quad (10)$$

In this form, the radial 3-D eigenequation looks very much like a 1-D problem (see Eq$\underline{n}$ (14) of Chapter 3) with an effective potential

$$V_{eff}(r) = \begin{cases} V(r) + \dfrac{\hbar^2}{2m} \dfrac{\ell(\ell + 1)}{r^2} & , \ r > 0 \\ \\ \infty & , \ r \le 0 . \end{cases} \quad (11)$$

The "centrifugal energy" term, $\dfrac{\hbar^2}{2m} \dfrac{\ell(\ell + 1)}{r^2}$, should be familiar to you from clasical mechanics, where it has the form $\dfrac{\vec{\ell}^2}{2mr^2}$ where $\vec{\ell}$ is the classical relative angular momentum. If this _were_ a one dimensional problem, the fact that the potential becomes infinite at r = 0 would imply that the particle would be unable to penetrate to the r < 0 region and we would expect that

$$R(0) = 0, \qquad\qquad (12)$$

would be the correct boundary condition.  Actually, the

correct boundary condition (based upon the requirement that H

be Hermitian) is somewhat more complicated.  Eqⁿ (12) will be

sufficient for our purposes here.  (For more discussion see

Merzbacher, 2nd ed., p.201.)

We are going to discuss the solution to Eqⁿ (8) in the

following 5 situations:

(a)  The free particle, $V(r) = 0$.

(b)  The infinite spherical well,

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

(c)  The "deuteron" (really a finite potential well),

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

(d)  The Coulomb problem

$$V(r) = - \frac{Ze^2}{r} \ .$$

(e)  A combination of (b) and (d) I call the "confined

Coulombic model" where

$$V(r) = \begin{cases} \infty & , \ r \geq a \\ - \dfrac{\xi}{r} & , \ r < a \ . \end{cases}$$

7.4

## (a)   The Free Particle

The description of a free particle in spherical coordinates is not trivial.  We talked extensively about the description of the 1-D free particle in Chapter 2.  The solution to the 1-D time-independent Schrödinger equation

$$- \frac{\hbar^2}{2m} \frac{d^2 u_{p_x'}(x)}{dx^2} = E u_{p_x'}(x) \tag{13}$$

with $u_{p_x'}(x) = \langle x | p_x' \rangle$ subject to the normalization

$$\int_{-\infty}^{\infty} dp_x' |p_x'\rangle \langle p_x'| = 1,$$

$$\Rightarrow \int_{-\infty}^{\infty} dp_x' u_{p_x'}(x) u_{p_x'}^*(x') = \delta(x - x'), \tag{14}$$

is just

$$u_{p_x'}(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ixp_x'/\hbar}, \tag{15}$$

where $p_x' > 0$ solutions represent particles moving in the $+x$ direction, and similarly for $p_x' < 0$. [We could also, if we wish, adopt the energy normalization condition (recall Eq$\underline{\text{n}}$ (175) of Chapter 2)

$$\sum_{\substack{p_x' \geq 0 \\ p_x' < 0}} \int_0^{\infty} dE |E\rangle \langle E| = 1,$$

$$\Rightarrow \sum_{\substack{p_x' \geq 0 \\ p_x' < 0}} \int_0^{\infty} dE \, u_E(x) u_E^*(x') = \delta(x - x'),$$

which gives

$$u_E(x) = \frac{m}{\hbar\sqrt{2\pi k}} \cdot \begin{cases} e^{ikx} & , \ p_x > 0 \\ e^{-ikx} & , \ p_x < 0 \ . \end{cases}]$$

In spherical coordinates, we want to find solutions to (we relabel $n \to k$ in this problem)

$$-\frac{\hbar^2}{2m}\left[\frac{d^2}{dr^2} - \frac{\ell(\ell+1)}{r^2}\right] R_{k\ell}(r) = \pm\frac{\hbar^2 k^2}{2m} R_{k\ell}(r), \tag{16}$$

subject to the boundary condition (12) and where $\ell = 0,1,2,\ldots$ . We have defined

$$k \equiv \frac{\sqrt{2mE}}{\hbar} \ . \tag{17}$$

The $\pm$ sign on the right hand side of (16) corresponds to $E = \pm\frac{\hbar^2 k^2}{2m}$ ; that is, to whether E is positive or negative. Defining a new dimensionless variable

$$p \equiv kr, \tag{18}$$

we may write (16) as

$$\left[\frac{d^2}{dp^2} - \frac{\ell(\ell+1)}{p^2} \pm 1\right] R_{k\ell}(p) = 0 \ . \tag{19}$$

We can eliminate the possibility of negative energies here. Consider the case $\ell = 0$, E < 0:

$$\left[\frac{d^2}{dp^2} - 1\right] R_{k\ell}(p) = 0 \tag{20}$$

The solutions to (20) are real exponentials:

$$R_{k\ell}(p) \sim e^{\pm p} . \tag{21}$$

But these solutions do not satisfy the boundary condition (12). Similarly for the other $\ell \neq 0$ values. Thus, the physically relevant solutions in this problem have $E \geq 0$ and the equation we must solve is (except for $E = 0$ exactly)

$$\left[\frac{d^2}{dp^2} - \frac{\ell(\ell + 1)}{p^2} + 1\right] R_{k\ell}(p) = 0 . \tag{22}$$

Let us look at the $\ell = 0$ case of Eq$^{\underline{n}}$ (22):

$$\left[\frac{d^2}{dp^2} + 1\right] R_{k0}(p) = 0. \tag{23}$$

The solutions to (23) are of course

$$R_{k0}(p) \sim \sin p, \cos p. \tag{24}$$

However, only the first possibility on the right satisfies the boundary condition (12).

In order to find the solutions for $\ell \neq 0$, let us define

$$R_{k\ell}(p) \equiv (-1)^{\ell} p^{\ell+1} X_{k\ell}(p). \tag{25}$$

Let us find the equation that $X_{k\ell}(p)$ satisfies. We have that

$$\frac{dR_{k\ell}}{dp} = (-1)^{\ell} \left\{(\ell + 1)p^{\ell}X_{k\ell} + p^{\ell+1}\frac{dX_{k\ell}}{dp}\right\} , \tag{26}$$

and

$$\frac{d^2 R_{k\ell}}{dp^2} = (-1)^{\ell} \left\{ (\ell + 1)p^{\ell-1} X_{k\ell} + 2(\ell + 1)p^{\ell} \frac{dX_{\ell}}{dp} + p^{\ell+1} \frac{d^2 X_{k\ell}}{dp^2} \right\}. \quad (27)$$

Therefore, we have

$$\left[ \frac{d^2}{dp^2} - \frac{\ell(\ell + 1)}{p^2} + 1 \right] R_{k\ell}(p)$$

$$= (-1)^{\ell} \left\{ \ell(\ell + 1)p^{\ell-1} X_{k\ell} + 2(\ell + 1)p^{\ell} \frac{dX k_{\ell}}{dp} \right.$$

$$\left. + p^{\ell+1} \frac{d^2 X_{k\ell}}{dp^2} - \frac{\ell(\ell + 1)}{p^2} p^{\ell+1} X_{k\ell} + p^{\ell+1} X_{k\ell} \right\}$$

$$= (-1)^{\ell} p^{\ell+1} \left[ \frac{d^2 X_{k\ell}}{dp^2} + \frac{2(\ell + 1)}{p} \frac{dX_{k\ell}}{dp} + X_{k\ell} \right]. \quad (28)$$

Thus, the equation satisfied by $X_{k\ell}$ is

$$\left[ \frac{d^2}{dp^2} + \frac{2(\ell + 1)}{p} \frac{d}{dp} + 1 \right] X_{k\ell}(p) = 0 . \quad (29)$$

Differentiating Eqⁿ (29) yields

$$X_{k\ell}''' + \frac{2(\ell + 1)}{p} X_{k\ell}'' - \frac{2(\ell + 1)}{p^2} X_{k\ell}' + X_{k\ell}' = 0, \quad (30)$$

where the primes denote differentiation with respect to p.
Now let's substitute

$$X_{k\ell}' = p\psi_{k\ell} \ ,$$

$$X_{k\ell}'' = \psi_{k\ell} + p \ \psi_{k\ell}' \ ,$$

$$X_{k\ell}''' = 2\psi_{k\ell}' + p\psi_{k\ell}'' \ ,$$

$$(31)$$

into (30). We get

$$\left( 2\psi_{k\ell}' + p\psi_{k\ell}'' \right) + \frac{2(\ell + 1)}{p} \left( \psi_{k\ell} + p \ \psi_{k\ell}' \right)$$

$$- \frac{2(\ell + 1)}{p^2} (p\psi_{k\ell}) + p\psi_{k\ell} = 0, \qquad (32)$$

or

$$\psi_{k\ell}'' + \frac{2(\ell + 2)}{p} \psi_{k\ell}' + \psi_{\ell} = 0. \qquad (33)$$

Comparing with Eq$^{\underline{n}}$ (29) above, we thus conclude that we may choose

$$\psi_{k\ell}(p) = X_{k\ell+1}(p). \qquad (34)$$

From the first of Eq$^{\underline{n}}$s (31) we then have

$$X_{k\ell+1}(p) = \frac{1}{p} \frac{d}{dp} X_{k\ell}(p). \qquad (35)$$

Therefore by induction

$$X_{k\ell}(p) = \left( \frac{1}{p} \frac{d}{dp} \right)^{\ell} X_{k0}(p). \qquad (36)$$

This impies that (from Eq$^{\underline{n}}$ (25))

$$R_{k\ell}(p) = (-1)^{\ell} p^{\ell+1} \left( \frac{1}{p} \frac{d}{dp} \right)^{\ell} X_{k0}(p). \qquad (37)$$

But from (24) and (25) (for $\ell = 0$) we have

$$X_{k0}(p) \sim \frac{\sin p}{p} . \tag{38}$$

Eq$^{\text{n}}$ (38) and (37) determine, outside of normalization, the full set of solutions to (22) with the boundary condition (12). To connect to standard definitions, let us introduce the "spherical Bessel functions", $j_\ell(p)$ (which are solutions of (8)):

$$j_\ell(p) \equiv (-p)^\ell \left(\frac{1}{p} \frac{d}{dp}\right)^\ell \left[\frac{\sin p}{p}\right] . \tag{39}$$

The first few $j_\ell$ are illustrated in the top part of Figure 10.3, p.410 of the book, Introductory Quantum Mechanics, by Liboff. (Notice that $j_0(x)$ goes to unity as x goes to zero.)

What we have accomplished is to solve Eq$^{\text{n}}$ (8)

$$- \frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) - \frac{\ell(\ell + 1)}{r^2}\right] u_{k\ell}(r) = \frac{\hbar^2 k^2}{2m} u_{k\ell}(r). \tag{40}$$

subject to the boundary condition (12). More abstractly, since (strictly speaking, the rhs is $<r|k(\ell)>$)

$$u_{k\ell}(r) = <r|k\ell>, \tag{41}$$

Eq$^{\text{n}}$ (40) is equivalent to

$$<r| \left(\frac{p_r^2}{2m} + \frac{\hbar^2 \ell(\ell + 1)}{2mr^2}\right) |k\ell> = \frac{\hbar^2 k^2}{2m} <r|k\ell>, \tag{42}$$

where we have defined the Hilbert space operator $p_r$ (the same as in the homework)

$$<r|p_r \equiv \frac{\hbar}{i} \frac{1}{r} \frac{\partial}{\partial r} r <r| \ . \tag{43}$$

Since (42) is true for all $<r|$ (except possibly at $r = 0$) we then get

$$\left(\frac{p_r^2}{2m} + \frac{\hbar^2 \ell(\ell + 1)}{2mr^2}\right) |k\ell> = \frac{\hbar^2 k^2}{2m} |k\ell>. \tag{44}$$

as the abstract equation that $|k\ell>$ satisfies. Notice that

for a given value of k (and therefore the energy, E) there is an <u>infinite</u> degeneracy in $\ell$ and m. Physically, this

corresponds to the fact that a particle with a given energy

in 3-D has an infinite number of directions in which to

travel, as opposed to the 2-fold degeneracy in 1-D space.

　　　The $|k\ell>$ are assumed to give a complete description and

so we assume

$$\int_0^\infty dk k^2 |k\ell> <k\ell| = 1_r \tag{45}$$

for each $\ell$ value. Other normalizations of the states $|k\ell>$

are possible. Eq$^{\underline{n}}$ (45) is written down in anology to the

first of Eq$^{\underline{n}}$s (61) of the last Chapter, just like

$\int dp_1 |p_1> <p_1| = 1$ is the momentum space version of

$\int dx_1 |x_1> <x_1| = 1$. The consequence of (45) is that

$$1_r \cdot 1_r = \int dk' k'^2 |k'\ell><k'\ell| \int dk k^2 |k\ell><k\ell|$$

$$= \int dk' dk \ k'^2 k^2 |k'\ell><k'\ell|k\ell><k\ell| . \qquad (46)$$

Comparison of (46) and (45) reveals that

$$<k'\ell|k\ell> = \frac{1}{k^2} \delta(k - k') . \qquad (47)$$

In the same way from

$$1_r = \int dr r^2 |r><r| \qquad (48)$$

we get

$$<r'|r> = \frac{1}{r^2} \delta(r - r') . \qquad (49)$$

Thus, the complete set of radial and angular eigenkets for the free particle is given by:

$$|k,\ell,m> \equiv \underset{\text{r-space}}{|k\ell>} \otimes \underset{\theta\text{-space}}{|\ell(m)>} \otimes \underset{\phi\text{-space}}{|m>} . \begin{pmatrix} 0 < k < \infty \\ \ell = 0,1,2,\ldots \\ m = -\ell,\ldots,\ell \end{pmatrix}$$

Of course, as pointed out at the beginning of this Chapter there is the alternate complete set,

$$|p_1',p_2',p_3'> \equiv |p_1'> \otimes |p_2'> \otimes |p_3'>, \quad (-\infty < p_1',p_2',p_3' < \infty)$$

based upon a Cartesian description.  Because both sets are complete, they should be expandable in terms of each other. That is, a Cartesian plane wave should be expressable in spherical harmonies, and vice versa.  We will not attempt to show this here, but this subject is covered in more advanced

treatments.  (See for example Gottfried, 1st ed., p.91

onward.)

If we adopt (45) as our normalization (similar to Eq$^{\underline{n}}$

(14) above in the 1-D case) then we require

$$\left.\begin{array}{c} < r \,\Big|\, \left[\int_0^\infty dk k^2 \,|\,k\ell > < k\ell\,| \;\; = \;\; 1\right] \Big|\, r' > \\[3mm] \Rightarrow \int_0^\infty dk k^2 u_{k\ell}(r) u_{k\ell}^*(r') \;\; = \;\; \frac{1}{r^2}\,\delta(r - r') \end{array}\right\}. \qquad (50)$$

We have found that

$$u_{k\ell}(r) \;\; = \;\; C_\ell\, j_\ell(kr) \qquad\qquad (51)$$

where $C_\ell$ is an unknown normalization factor.  Putting (51)

into (50) gives

$$\int_0^\infty dk k^2 |C_\ell|^2 j_\ell(kr) j_\ell(kr') \;\; = \;\; \frac{1}{r^2}\,\delta(r - r'). \qquad (52)$$

Eq$^{\underline{n}}$ (52) determines the normalization factor $C_\ell$.  The $C_\ell$

cannot be a function of k since by the normalization
condition (50) the $u_{k\ell}(r)$ are dimensionless.  In order to get

an explicit expression, it is necessary to perform the

integration of the left of (52).  The mathematics involved in

doing this is beyond the level of this class.  I will simply

quote the necessary result:

$$\int_0^\infty dk k^2 j_\ell(kr) j_\ell(kr') \;\; = \;\; \frac{\pi}{2r^2}\,\delta(r - r'). \qquad (53)$$

(You will, however, be asked to confirm the above result for the $\ell$ = 0 case. Eq$^{\underline{n}}$ (53) is, mathematically, only true for the case r,r' > 0.)

Comparing (53) and (52), we find that

$$|C_\ell|^2 = \frac{2}{\pi} , \tag{54}$$

for which we choose

$$C_\ell = \sqrt{\frac{2}{\pi}} . \tag{55}$$

This implies finally that

$$u_{k\ell}(r) = \sqrt{\frac{2}{\pi}} \ j_\ell(kr). \tag{56}$$

The complete free particle wavefunction is then

$$u_{k\ell m}(\vec{r}) = \sqrt{\frac{2}{\pi}} \ j_\ell(kr)Y_{\ell m}(\theta,\phi). \tag{57}$$

The entire content of their completeness can be stated as

$$< r | \left[ \int dk \ k^2 \sum_{\ell,m} |k,\ell,m> <k,\ell,m| \ = \ 1 \right] |r'> ,$$

$$\Rightarrow \int_0^\infty dk \ k^2 \sum_{\ell,m} u_{k\ell m}(\vec{r})u_{k\ell m}^*(\vec{r}\,') = \delta(\vec{r} - \vec{r}), \tag{58}$$

and as

$$< k\ell',\ell',m' | \left[ \int d^3r |\vec{r} > < \vec{r} | \ = \ 1 \right] |k\ell,\ell,m> ,$$

$$\Rightarrow \int d^3r \ u_{k'\ell'm'}^*(\vec{r})u_{k\ell m}(\vec{r}) = \delta_{\ell\ell'}\delta_{mm'} \frac{1}{k^2} \delta(k - k'). \tag{59}$$

## (b)  The  Infinite  Spherical  Well

The solutions for the free particle resulted in a continuous spectrum, as it should. We now imagine putting a

particle into a spherical well that has a potential which rises to ∞ at the surface.  This means the particle is trapped in the region r < a and has zero probability of escaping from this spherical region.  In the 1-D case we subjected the free particle solutions,

$$u(x) \sim \sin(kx), \cos(kx),$$

to the boundary conditions

$$u(x)\big|_{x = \pm a} = 0, \tag{60}$$

to find the allowed energies.  We do a similar thing here in our radial 3-D description.  The free particle solutions are given by (57), and our boundary condition

$$u(\vec{r})\big|_{r = a} = 0, \tag{61}$$

determines the energy levels as being solutions of

$$j_\ell(ka) = 0. \tag{62}$$

Eq$^{\underline{n}}$ (62) allows only certain discrete values of ka as solutions.  These are called <u>zeros</u> of the Bessel junction $j_\ell(ka)$.  Let's look at the $\ell = 0$ case:

$$j_0(ka) = 0, \tag{63}$$

$$\Rightarrow \quad \frac{\sin(ka)}{ka} = 0,$$

$$\Rightarrow \quad ka = n\pi, \quad n = 1,2,3,\ldots \tag{64}$$

In the $\ell = 1$ case, however, we have

$$j_1(ka) = 0, \tag{65}$$

$$\Rightarrow \quad \sin(ka) = (ka)\cos(ka).$$

The transcendental Eq$^{\underline{n}}$ (65) can be solved numerically to yield approximately

$$ka = 4.493, 7.725, 10.90, \ldots \qquad (66)$$

for the first three zeros.  Tables of the zeros of Bessel functions are, for example, published in the National Bureau of Standards, Handbook of Mathematical Functions.  The energies of these states are then given by

$$E = \frac{\hbar^2 (ka)^2}{2ma^2} . \qquad (67)$$

A convenient label that catagorizes the solutions is the number of nodes (or zeros) in the radial wavefunction (for $r \neq 0$).  The lowest energy solution for each $\ell$ value has a single node at $r = a$ (the surface).  The next highest energy solution has two nodes, the next has three nodes, and so on. We will use $n = 1, 2, 3, \ldots$ to label these solutions. (Notice that only in the case $\ell = 0$ is ka proportional to n.) Defining

$$u_{n\ell}(r) \equiv N_\ell \, j_\ell(k_n r), \qquad (68)$$

where $k_n$ represents a solution with n nodes and N is a normalization factor determined by

$$\int_0^a dr\, r^2 |u_{n\ell}(r)|^2 = 1, \qquad (69)$$

we may give a schematic representation of these wavefunctions as below:
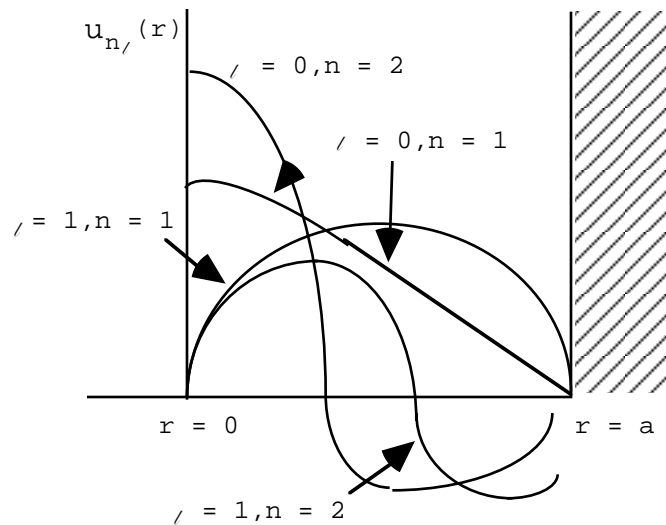
(Notice that only for $\ell = 0$ states is $u_{n\ell}(0) \neq 0$.)  The

normalization factor in (68) can be shown to be given by

$$N_\ell^2 = \frac{2}{a^3 j_{\ell+1}^2(k_n a)} . \tag{70}$$

Some conventional terminology to denote these states is given below.

$\ell = 0$ <u>s state</u>  "sharp"

$\ell = 1$ <u>p state</u>  "principle"

$\ell = 2$ <u>d state</u>  "diffuse"

$\ell = 3$ <u>f state</u>

from the days of exp'$\ell$ spectral analysis.

For $\ell = 4, 5, \ldots$, the letter designation becomes alphabetical.

In addition, one also sees the notation:

$\ell = 0$, $n = 1$:  <u>1s shell</u>

$\ell = 0$, $n = 2$:  <u>2s shell</u>

$\ell = 1$, $n = 1$:  <u>1p shell</u>

$\ell = 1$, $n = 2$:  <u>2p shell</u> .

This is called spectroscopic notation.  The number in front
of the letter in this case is just n, the number of radial
(r ≠ 0) nodes.  It also indicates the energy level <u>order</u>.
That is, the 1s state is the lowest energy $\ell$ = 0 state, the
2s state has the next highest energy for $\ell$ = 0, and so on.

   Since the energies of the states are proportional to the
dimensionless quantity $(ka)^2$, a listing of the first few
levels gives an idea of the separations involved.

| State | 1s | 1p | 1d | 2s | 1f |
|---|---|---|---|---|---|
| $(ka)^2$ | 9.87 | 20.14 | 33.21 | 39.48 | 48.83 |
| | 2p | 1g | 2d | 1h | 3s |
| | 59.68 | 66.96 | 82.72 | 87.53 | 88.83 |

Each energy level, of course, is in general degenerate:
remember the $(2\ell + 1)$ degeneracy in m of the $\ell^{th}$ level.

   One of the reasons the spherical well is so interesting
is because it turns out to represent a crude, and yet fairly
accurate, model of the nucleus.  Such models assume that the
interaction between a single nucleon (either a proton or
neutron) and all the other nucleons can be represented by an
external potential, V(r).  This is called the <u>single-particle</u>
or <u>shell</u> model of the nucleus and is described in all
introductory nuclear physics books.  This picture is, of
course, a gross phenomenological oversimpification, but such
simplifications are in fact necessary because of the extreme
complexity of the problem.  The ultimate justification for
the assumption that the nucleons move practically independent

of one another is due to a property of matter we will study

more in Chapter 9.

Using the assumption that we can only place <u>two neutrons

and two protons in each energy level</u>, we can calculate the

number of nucleons held by each shell.  This assumption is

completely mysterious at this point but will be explained when

we come to Chapter 9.  The assumption of two <u>electrons</u> per

energy level in <u>atomic</u> spectroscopy was known by early workers

as "zweideutigkut," which is German for "two-valuedness."

| shell | # protons(Z) $= 2(2\ell + 1)$ | # neutrons(N) $= 2(2\ell + 1)$ | additive total |
|---|---|---|---|
| <u>1s</u> ($\ell = 0$) | 2 | 2 | 2 |
| <u>1p</u> ($\ell = 1$) | 6 | 6 | 2 + 6 = 8 |
| <u>1d</u> ($\ell = 2$) | 10 | 10 | 8 + 10 = 18 |
| <u>2s</u> ($\ell = 0$) | 2 | 2 | 18 + 2 = 20 |
| <u>1f</u> ($\ell = 3$) | 14 | 14 | 20 + 14 = 34 |
| <u>2p</u> ($\ell = 1$) | 6 | 6 | 34 + 6 = 40 |
| <u>1g</u> ($\ell = 4$) | 18 | 18 | 40 + 18 = 58 |

Because of the fact that jumps in energy occur when a given

shell is filled, we might expect that nuclei with Z or N equal

to the numbers in the right hand column above should be

particularly stable.  This is similar to the closed-shell effect

for the Nobel gases in atomic physics.  On the basis of the

above simple model then, we might exect to have particularly

stable nuclei whenever Z or N equals:

2,8,18,20,34,40,58,...

Instead, experiment shows that the favored stable nuclei have Z

or N equal to:

$$2,8,20,28,50,82, \text{ and } 126.$$

These Z values correspond to helium, oxygen, calcium, nickel, tin and lead nuclei. (A Z = 126 nucleus is experimentally unknown.) These are called the nuclear "magic numbers." Nuclei which have <u>both</u> Z and N equal to a magic number are especially stable. An example is a type of calcium nucleus:

$$Ca_{20(=Z)}^{40(=Z+N=A)} .$$
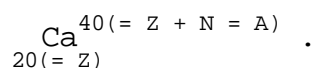
The Z value determines which element, but each element can have a number of N values, called <u>isotopes</u>.

Of course, the assumption of an infinite spherical potential is partly to blame for the fact that the above set of predicted and observed magic numbers do not agree. A more realistic model is to assume some sort of <u>finite</u> potential well. (We will study this possibility briefly when the deuteron is discussed.) This is important to allow for the known exponential tail in the nucleon distribution rather than a sharp cutoff at some spherical radius, a. Another important ingredient in a successful theory of the atomic nucleus is an assumption known as <u>LS coupling</u> or <u>spin-orbit coupling</u>. This phenomenon will be explained when we come to the hydrogen atom. The end result of such calculations is an energy level diagram, which is actually different for each nuclear isotype. A schematic representation of such an energy level diagram will be passed out in class. The spectroscopic notation on the right hand side of this figure (in the case of $\ell$s coupling) has not

been explained yet.

A real, fundamental theory of strong interactions has existed since about 1972 and is called <u>Quantum Chromodynamics</u> (QCD). The reason it took such a long time to uncover this theory is because the observed particles in a nucleus, the proton and neutron, are actually composite objects made up of 3 <u>quarks</u>. Just as the relativistic field theory <u>Quantum Electrodynamics</u> (QED) describes the interactions of electrons and photons (particles of light), QCD describes the interactions of quarks and other force-mediating particles called <u>gluons</u>. It is mentally comforting to know that a fundamental theory of nuclear forces is known. However, QCD is actually of little practical value in most nuclear physics problems because of the extreme complexity of the systems involved. However, it is extremely gratifying to know that the true, underlying physics is known.

**(c)   The "Deuteron"**

Up to the present, we have only been investigating the behavior of a single, isolated particle subject to some simple external potential. More realistically, the simpliest sytems in nature are two particle systems. Continuing the above discussion of the atomic nucleus, we will concentrate here on a phenomenological description of the simpliest nuclear subsystem, the deuteron, which consists of a single proton and neutron. Of course, from the more fundamental point of view of quark dynamics, this is already an extremely complicated system consisting of six quarks. In order to get

a practical description, we will ignore this reality and treat the neutron and proton as fundamental objects. Physics is a multi-layered discipline. It turns out for many applications that we can ignore the deeper layers of reality (usually at higher energies) if we wish to describe a system at some reasonable level of accuracy.

Since we are considering a system consisting of two independent particles, it is reasonable to suppose we can introduce independent position and momentum operators for each particle. That is, we introduce momentum operators $\vec{p}_1, \vec{p}_2$ and position operators $\vec{x}_1, \vec{x}_2$ such that

$$[\vec{p}_\alpha, \vec{p}_\beta] = 0, \quad [\vec{x}_\alpha, \vec{x}_\beta] = 0 \tag{71}$$

where $\alpha$ and $\beta$ are now <u>particle</u> labels. Our direct product states can be taken as

$$|\vec{p}_1', \vec{p}_2'> = |\vec{p}_1'> \otimes |\vec{p}_2'> , \tag{72}$$

$$|\vec{x}_1', \vec{x}_2'> = |\vec{x}_1'> \otimes |\vec{x}_2'> . \tag{73}$$

We will also assume that position and momentum operators from different particles commute,

$$[p_{i\alpha}, x_{j\beta}] = \frac{\hbar}{i} \delta_{ij} \delta_{\alpha\beta} . \tag{74}$$

Just as in classical dynamics, we can imagine introducing a different set of coordinates, based on the center of mass of the system, that is useful to locate the positions of our two particles. Given particles at positions

$\vec{x}_1'$ and $\vec{x}_2'$ relative to a fixed coordinate system, we can also locate their positions given the center of mass vector $\vec{X}'$ and relative position $\vec{x}'$.  The relationship between these coordinates is illustrated below:
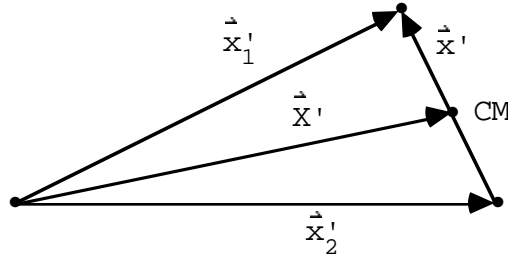


Based upon these relationships, we define operators for these quantities as follows

$$\vec{x} = \vec{x}_1 - \vec{x}_2 , \qquad (75)$$

$$\vec{X} = \frac{m_1\vec{x}_1 + m_2\vec{x}_2}{(m_1 + m_2)} . \qquad (76)$$

Just as we introduced a new basis to represent a particle's position in spherical coordinates, we imagine an alternate description based upon knowledge of $\vec{X}'$ and $\vec{x}'$:

$$< \vec{x}_1',\vec{x}_2'| \; \rightarrow \; < \vec{x}',\vec{X}'| \; . \qquad (77)$$

Such a description is only allowed if $\vec{x}'$ and $\vec{X}'$ may be simultaneously measured, that is, if $\vec{x}$ and $\vec{X}$ commute.  This is easy to show using (71):

$$[X_i,x_j] = \left[\frac{(m_1x_{i1} + m_2x_{i2})}{(m_1 + m_2)} , \; x_{j1} - x_{j2}\right] = 0 \qquad (78)$$

Thus we may take

$$\langle \vec{x}\,', \vec{X}\,' | \ = \ \langle \vec{x}\,' | \langle \vec{X}\,' | \tag{79}$$

where we define (we imagine Cartesian bases for now)

$$\langle \vec{x}\,' | \, \vec{x} \ = \ \langle \vec{x}\,' | \, \vec{x}\,' \ , \tag{80}$$

$$\langle \vec{X}\,' | \vec{X} \ = \ \langle \vec{X}\,' | \vec{X}\,' \ . \tag{81}$$

Along with this coordinate description, we introduce center of mass momentum operators

$$\vec{P} \ = \ \vec{p}_1 \ + \ \vec{p}_2 \ , \tag{82}$$

$$\vec{p} \ = \ \frac{m_2 \vec{p}_1 \ - \ m_1 \vec{p}_2}{(m_1 \ + \ m_2)} \ . \tag{83}$$

We have

$$[p_i , P_j] \ = \ \left[ \frac{m_2 p_{i1} \ - \ m_2 p_{i2}}{(m_1 \ + \ m_2)} \ , \ p_{j1} \ - \ p_{j2} \right] = 0, \tag{84}$$

from Eq$^{\underline{n}}$ (71). Therefore, a new momentum description

$$\langle \vec{p}\,'_1 , \vec{p}\,'_2 | \ \rightarrow \ \langle \vec{p}\,' , \vec{P}\,' | \ . \tag{85}$$

where

$$\langle \vec{p}\,' , \vec{P}\,' | \ = \ \langle \vec{p}\,' | \ \langle \vec{P}\,' | \ , \tag{86}$$

and

$$\langle \vec{p}\,' | \vec{p} \ = \ \langle \vec{p}\,' | \vec{p}\,' \ , \tag{87}$$

$$\langle \vec{P}\,' | \vec{P} \ = \ \langle \vec{P}\,' | \vec{P}\,' \ , \tag{88}$$

is allowed.  We assume both the CM coordinate description,
Eq$^n$ (79), and the CM momentum description, Eq$^n$ (86), are
complete.

Based upon (74), we also have that

$$[P_i, x_j] = [X_i, p_j] = 0 .\tag{89}$$

However, we will also get

$$[P_i, X_j] = \left[ p_{i1} + p_{i2}, \frac{m_1 x_{j1} + m_2 x_{j2}}{(m_1 + m_2)} \right]$$

$$= \frac{m_1}{m_1 + m_2} [p_{i1}, x_{j1}] + \frac{m_2}{m_1 + m_2} [p_{i2}, x_{j2}] = \frac{\hbar}{i} \delta_{ij} ,\tag{90}$$

$$[p_i, x_j] = \left[ \frac{m_2 p_{i1} - m_1 p_{i2}}{(m_1 + m_2)} , x_{j1} - x_{j2} \right]$$

$$= \frac{m_2}{m_1 + m_2} [p_{i1}, x_{j1}] + \frac{m_1}{m_1 + m_2} [p_{i2}, x_{j2}] = \frac{\hbar}{i} \delta_{ij} ,\tag{91}$$

Based upon (90) and (91) we may show that

$$< \vec{X}' | \vec{P} = \frac{\hbar}{i} \vec{\nabla}_{X'} < \vec{X}' | ,\tag{92}$$

$$< \vec{x}' | \vec{p} = \frac{\hbar}{i} \vec{\nabla}_{x'} < \vec{x}' | .\tag{93}$$

by reasoning similar to that in Chapter 4 leading up to Eq$^n$
(153).

Solving for $\vec{p}_1$ and $\vec{p}_2$ from (82) and (83), we find

$$\vec{p}_1 = \frac{m_1}{m_1 + m_2} \vec{P} + \vec{p} ,\tag{94}$$

$$\vec{p}_2 = \frac{m_2}{m_1 + m_2} \vec{P} - \vec{p} ,\tag{95}$$

so that for the kinetic energy we get

$$\frac{\vec{P}_1^{\,2}}{2m_1} + \frac{\vec{P}_2^{\,2}}{2m_2} = \frac{\vec{P}^{\,2}}{2M} + \frac{\vec{p}^{\,2}}{2\mu} \tag{96}$$

where $M = m_1 + m_2$ and $\mu = \dfrac{m_1 m_2}{M}$ (the usual formula for reduced

mass).  Assuming the interaction of the two particles depends

only upon the distance between the two particles, the

Hamiltonian becomes

$$H = \frac{\vec{P}^{\,2}}{2M} + \frac{\vec{p}^{\,2}}{2\mu} + V(r) \tag{97}$$

where, in a spherical coordinate description, "r" represents

an operator giving the magnitude of the distance between the

particles.  Let's call

$$H_{ext} = \frac{\vec{P}^{\,2}}{2M} \quad , \tag{98}$$

and

$$H_{int} = \frac{\vec{p}^{\,2}}{2\mu} + V(r) \quad . \tag{99}$$

Notice that

$$[H_{ext}, H_{int}] = \left[ \frac{\vec{P}^{\,2}}{2M} \, , \, \frac{\vec{p}^{\,2}}{2\mu} + V(r) \right] = 0 \quad , \tag{100}$$

so that our system can be taken to have simultaneous

eigenvalues of $H_{ext}$ and $H_{int}$.  The eigenvalue equation for $H_{ext}$

is just

$$H_{ext} \, |\vec{P}\,'> = \frac{\vec{P}\,'^{\,2}}{2M} \, |\vec{P}\,'>, \tag{101}$$

so that

$$< \vec{X}' | H_{ext} | \vec{P}' > = \frac{\vec{P}'^2}{2M} < \vec{X}' | \vec{P}' >, \qquad (102)$$

or

$$-\frac{\hbar^2}{2M} \vec{\nabla}_{X'}^2 u_{\vec{P}'}(\vec{X}') = \frac{\vec{P}'^2}{2M} u_{\vec{P}'}(\vec{X}'), \qquad (103)$$

where $u_{\vec{P}'}(\vec{X}') = < \vec{X}' | \vec{P}' >$ and we have used Eq$\underline{n}$ (92) above. The solution to (103) is just

$$u_{\vec{P}'}(\vec{X}') = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\vec{P}' \cdot \vec{X}'/\hbar}, \qquad (104)$$

that is, just free particle plane waves normalized such that

$$\int d^3X \; u_{\vec{P}'}^*(\vec{X}) u_{\vec{P}''}(\vec{X}) = \delta(\vec{P}' - \vec{P}''), \qquad (105)$$

as usual. This shows that the center of mass of the system behaves like a free particle, taking on continuous energy values.

The interesting part of the two body problem is the eigenvalue equation for $H_{int}$. We switch to a spherical polar description for this sector

$$< \vec{x}' | \rightarrow < \vec{r} | , \qquad (106)$$

and write the eigenvalue equation as

$$< \vec{r} | H_{int} | a' > = E_{a'} < \vec{r} | a' > , \qquad (107)$$

or using the form, Eq$\underline{n}$ (99), for $H_{int}$ , we have

$$\left[ -\frac{\hbar^2}{2\mu} \vec{\nabla}_r^2 + V(r) \right] u_{a'}(\vec{r}) = E_{a'} u_{a'}(\vec{r}), \tag{108}$$

where $u_{a'}(\vec{r}) = \langle \vec{r} | a' \rangle$.  This is of the same form as the one-body Schrödinger equation we have already considered, except that m is replaced by $\mu$.

The above has been a short introduction to two-body formalism and has not been directed toward the deuteron problem as such. Let me now recite some of the known experimental facts concerning the deuteron.  It has been determined that the binding energy of the deuteron is

$$E = -2.226 \text{ Mev.} \tag{109}$$

[Remember that

> 1ev = energy given to an electronic
>       charge accelerated through 1 volt

 or

> 1ev = $1.602 \times 10^{-12}$ erg.

We have 1Mev = $10^6$ ev, of course.  Since energies can be measured in Mev, mass can be measured in units of Mev/$c^2$.]
Eq$\underline{n}$ (109) implies that it takes 2.226 Mev of energy to separate the proton and neutron in deuterium to infinity.  In addition, it is known that there are no excited energy states of the deuteron, that is, there is only a single bound state of this system.  One (non-fundamental) theory says the force between nucleons (or at least the long range part of the potential) is due to pion exchange between them.  The energy-time uncertainty principle (Ch. 3) reads

$$\Delta E t_c \approx \hbar \tag{110}$$

The energy of a system of two nucleons fluctuates when we consider events that take place over short periods of time. Because of this, one nucleon can spontaneously give rise to another particle, called a pion, which has a mass

$$m_{\pi^+} \simeq 140 \text{ Mev/c}^2 . \tag{111}$$

Setting $\Delta E = m_{\pi^+}c^2$, then we learn from (110) that this pion is allowed to "live" during a time

$$t_c \approx \frac{\hbar}{m_{\pi^+}c^2} . \tag{112}$$

If we assume this emitted pion travels close to the speed of light, we see that it has a range of

$$ct_c \approx \frac{\hbar}{m_{\pi^+}c} . \tag{113}$$

The right hand side of (113) is just the pion's reduced Compton wavelength (see Ch. 2) which has a numerical value of

$$\frac{\hbar}{m_{\pi^+}c} \approx 1.43 \times 10^{-15} \text{ m} = 1.43 \text{ fm} \tag{114}$$

$$(1 \text{ fermi} = 10^{-15} \text{ m} = 10^{-15} \text{ cm}).$$

A useful numerical relationship easily memorized is

$$1\text{fm}^{-1} = (197.3 \text{ Mev})/\hbar c.$$

Therefore, if the second nucleon is within this appropriate range, the pion can be reabsorbed by it and a force between the two nucleons can be transmitted. This picture of the origin of the strong nuclear force is due to Hidiki Yukawa.

It seems to explain the short range of the strong nuclear
force.  In its details it cannot be completely correct since
we now know that all of these particles, protons, neutrons
and pions, are made up of more fundamental objects.  (In the
simpliest possible interpretation, a pion is just a bound
state of a quark and an anti-quark.)  The emitted and
absorbed pion in the discussion is called a "virtual"
particle.  The extent to which this picture can be relied
upon is currently a matter of debate in nuclear physics.

Assuming a spherically symmetric force the radial equation
we need to solve is (just Eq$^n$ (10) above with
$m \rightarrow \mu$).  (I will sometimes leave the quantum number labels off
of the radial eigenfunctions until it is clear what they are.)

$$\left[ -\frac{\hbar^2}{2\mu}\left(\frac{d^2}{dr^2} - \frac{\ell(\ell+1)}{r^2}\right) + V(r)\right] R(r) = ER(r). \qquad (115)$$

What about the potential, $V(r)$?  The remarks about the range of
the nuclear force suggests the simpliest possible picture below:
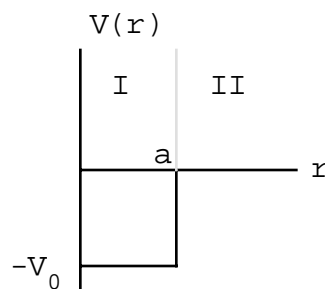


That is

$$V(r) = \begin{cases} 0, & r \geq a \\ -V_0, & r < a \ . \end{cases} \qquad (116)$$

We expect that the well radius, a, is on the order of 1.43 fm. In order to produce a binding energy of about 2.2 Mev, and to have only a single bound state, what should the well depth, $V_0$, be?

We have labeled the region of space with $r \leq a$ as "I" and $r \geq a$ as "II" in the above figure. In I we need to solve

$$- \frac{\hbar^2}{2\mu} \left( \frac{d^2}{dr^2} - \frac{\ell(\ell + 1)}{r^2} \right) R(r) = (E + V_0)R(r) \ . \qquad (117)$$

This is the same form as the free particle problem (Eq$\underline{n}$ (16) above) except that $E \rightarrow E + V_0$ here. Following the same reasoning as before, we learn that $E + V_0$ must be a positive quantity and that the solutions are of the form $j_\ell(Ka)$ where

$$K \equiv \frac{\sqrt{2\mu(E + V_0)}}{\hbar} \ . \qquad (118)$$

In the region labeled as II, we must instead solve

$$- \frac{\hbar^2}{2\mu} \left( \frac{d^2}{dr^2} - \frac{\ell(\ell + 1)}{r^2} \right) R(r) = ER(r) \ , \qquad (119)$$

but now because we are specifically interested in a bound state we require $E < 0$. (There are also $E > 0$ solutions. These describe scattering.) In order to find the appropriate solutions to (119) in this case, let us for the moment specialize to $\ell = 0$. In dimensionless form with

$$\left. \begin{array}{l} \rho \equiv K'r, \\[1em] K' \equiv \dfrac{\sqrt{-2\mu E}}{\hbar} = \dfrac{\sqrt{2\mu|E|}}{\hbar}, \end{array} \right\} \qquad (120)$$

the $\ell = 0$ version of (119) reads

$$\left[ \frac{d^2}{d\rho^2} - 1 \right] R(\rho) = 0. \tag{121}$$

We encountered this equation before ((20) above) and found that its solutions were of the form

$$R(\rho) \sim e^{\pm\rho} . \tag{122}$$

These were discarded because at the origin they do not obey (12). However, we are now solving (121) in region II, which does not include the origin. Region II does, however, include spatial infinity where the wavefunction must vanish,

$$R(\infty) = 0 . \tag{123}$$

This is actually a necessary but not sufficient condition that the wavefunction must satisfy. The real requirement on the wavefunction is that it be square integrable, which here means that $R(\rho)$ must fall off faster than $\rho^{-1/2}$ as $\rho \to \infty$. (See the similar 1-D discussion on p.2.33.) The condition (123) applied to (122) is enough to pick out the unique solution

$$R(\rho) \sim e^{-\rho} . \tag{124}$$

This brings the present discussion up to the same point as the free particle discussion leading to the unique solution $R_{k0}(\rho) \sim \sin \rho$. We can repeat the rest of that discussion, which leads to the solutions of the form (see Eq$^{\underline{n}}$ (37))

$$\rho^{\ell+1} \left( \frac{1}{\rho} \frac{d}{d\rho} \right)^\ell X_{\ell=0}(\rho) . \tag{125}$$

Before, for the free particle we had $X_{\ell=0}(\rho) \sim \dfrac{\sin \rho}{\rho}$, which

lead to the use of the spherical Bessel functions

$$j_\ell(\rho) = (-\rho)^\ell \left( \frac{1}{\rho} \frac{d}{d\rho} \right)^\ell \left( \frac{\sin \rho}{\rho} \right) , \qquad (126)$$

as solutions of the radial equation.  However, in this case

we have $X_{\ell=0}(\rho) \sim \dfrac{e^{-\rho}}{\rho}$, and we can take our solutions to be

proportional to the <u>spherical Hankel functions of the first</u>

<u>kind</u> (with imaginary arguments), $h_\ell^{(1)}(i\rho)$, defined as:

$$h_\ell^{(1)}(i\rho) \equiv -(i\rho)^\ell \left( \frac{1}{\rho} \frac{d}{d\rho} \right)^\ell \left( \frac{e^{-\rho}}{\rho} \right) . \qquad (127)$$

Of course, the solutions in Regions I and II have to be

joined smoothly, which means in this case that $R(\rho)$ and its

first derivative must connect continuously at $r = a$.  These

conditions arise here in 3 dimensions completely analogously

to the continuity conditions discussed in the notes on pgs.

3.25, 3.26 for 1-D.  Therefore, our explicit solutions are (A

and B are normalization constants)

$$u(r) = \begin{cases} Aj_\ell(Kr), & r \leq a \\[2mm] Bh_\ell^{(1)}(iK'r), & r \geq a \end{cases} \qquad (128)$$

and our requirements of continuity in $u(r)$ and $u'(r)$ at $r = a$

read:

$$Aj_\ell(Ka) = Bh_\ell^{(1)}(iK'a), \qquad (129)$$

$$Aj_\ell'(Ka) = Bh_\ell^{(1)'}(iK'a). \tag{130}$$

Dividing (130) by (129) gives us the statement

$$\frac{j_\ell'(Ka)}{j_\ell(Ka)} = \frac{h_\ell^{(1)'}(iK'a)}{h_\ell^{(1)}(iK'a)} \ . \tag{131}$$

Eq$^n$ (131) is the eigenvalue equation for the energies, and for finite well depth, $V_0$, has only a finite number of solutions. Let us explicitly investigate (131) in the case $\ell = 0$. Since

$$j_0(Ka) = \frac{\sin Ka}{Ka} \ , \tag{132}$$

$$h_0^{(1)}(iK'a) = -\frac{e^{-K'a}}{K'a} \ , \tag{133}$$

eq$^n$ (131) reads

$$\frac{\left(\dfrac{\sin Kr}{r}\right)'}{\left(\dfrac{\sin Kr}{r}\right)}\Bigg|_{r=a} = \frac{\left(\dfrac{e^{-K'r}}{r}\right)'}{\left(\dfrac{e^{-K'r}}{r}\right)}\Bigg|_{r=a} \tag{134}$$

or

$$(Ka)\cot Ka = -(K'a) \tag{135}$$

Notice also that from (118) and (120)

$$(Ka)^2 + (K'a)^2 = \frac{2\mu V_0 a^2}{\hbar^2} \ . \tag{136}$$

Therefore, we may solve the eigenvalue problem, given a value of $V_0$, by plotting (135) and (136) simultaneously on axes labeled by (Ka) and (K'a). This is done below.
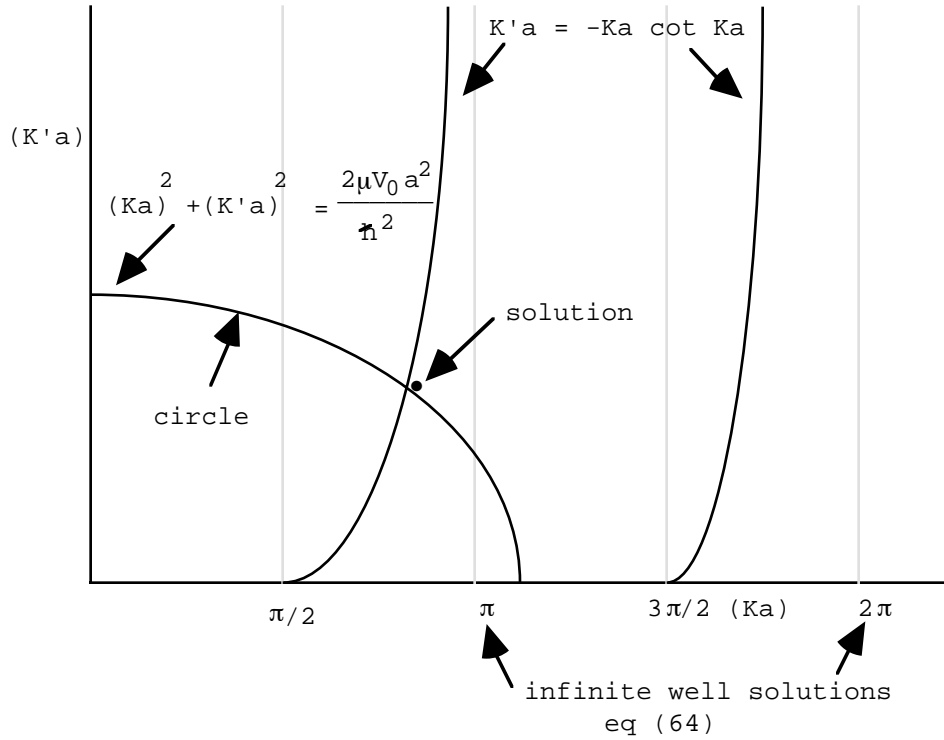
Notice that, based upon the above figure, we can determine the number of bound s-states for a given well depth quite easily. When we have, for example

$$\frac{2\mu V_0 a^2}{\hbar^2} < \left( \frac{\pi}{2} \right)^2, \tag{137}$$

then there will be no bound s-states because the two sets of curves in the figure never intersect. If

$$\left( \frac{\pi}{2} \right)^2 < \frac{2\mu V_0 a^2}{\hbar^2} < \left( \frac{3\pi}{2} \right)^2, \tag{138}$$

there will be a single bound s-state. In general it is easy to see that

$$\left( n - \frac{1}{2} \right)^2 \pi^2 < \frac{2\mu V_0 a^2}{\hbar^2} < \left( n + \frac{1}{2} \right)^2 \pi^2, \tag{139}$$

is the condition on the well depth such that n bound s-states

exist.  Of course for the deuteron we would like to choose $V_0$

such that a single bound state exists.  This would seem to

mean that we should require (138) to hold for the well depth,

$V_0$.  However, this is only a condition to have a single bound

s-state; there can also be bound states with $\ell$ = 1,2,3...

One can show that the first $\ell$ = 1 state becomes bound when

$$\frac{2\mu V_0 a^2}{\hbar^2} > \pi^2 \; . \tag{140}$$

One can also show that, if $(V_0)_\ell$ represents the minimum well

depth to bind a state of angular momentum $\ell$, then

$$(V_0)_0 < (V_0)_1 < (V_0)_2 \; ... \tag{141}$$

This is actually the converse of a theorem for energies that

appears in a problem in Quantum Mechanics, V.I, 1st ed.,

p.367 by Messiah.  (The theorem in Messiah is easily proved

by considering that $\vec{\ell}^2/2\mu$ is a positive definite operator.)

Thus, the condition to produce a single bound state is

$$\left( \frac{\pi}{2} \right)^2 < \frac{2\mu V_0 a^2}{\hbar^2} < \pi^2. \tag{142}$$

In terms of our estimate that a $\simeq$ 1.43 fm, eq$^{\underline{n}}$ (81) says that

($\mu \simeq 1/2 \; m_{proton}$, $m_{proton}$ = 940 Mev/c$^2$).

$$52 \; \text{Mev} < V_0 < 206 \; \text{Mev}. \tag{143}$$

According to our simple theory, this single bound state

must have $\ell$ = 0.  In reality, because the neutron-proton

potential is not exactly spherically symmetric (because of the so-called <u>tensor force</u>), the deuteron is actually a mixture of $\ell$ = 0 and a small amount of $\ell$ = 2. (The angular momentum $\vec{L}^2$ does not give rise to a good quantum number.)

The wave function of the lowest bound s-state looks like:
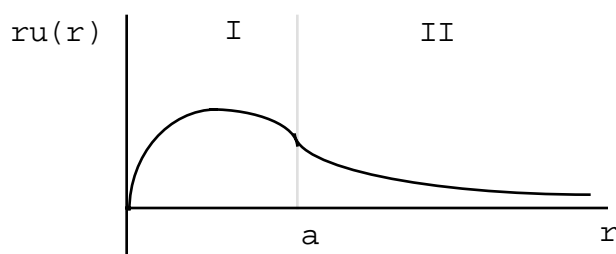


Notice the exponential tail in the classically excluded Region II.

We have not found explicit formulas for the normalization factors A and B in (128), but they are, as usual, determined from

$$\int_0^\infty dr\, r^2 |u(r)|^2 = 1 \ . \tag{144}$$

Counting the point at infinity as a node (since u(r) = 0 there), the lowest bound s-state has one radial node, the next highest (if bound) has two nodes, and so on. Similarly for the higher $\ell$ values. Therefore, the bound state quantum numbers can be taken to be n$\ell$, where n represents the number of radial nodes, as in the infinite spherical well also. This is the first explicit case we have studied where the eigenenergies an have <u>both</u> a discrete (E < 0) and a continuous (E > 0) spectrum. We will not study the

scattering solutions.  Suffice it to say that if we desired

to write completeness for these sets of radial wavefunctions,

we would have to have both a discrete sum over the bound

states as well as an integral over the scattering states as

in

$$\sum_{bound} |n\ell> <n\ell| \; + \; \int_0^a dk k^2 |k\ell> <k\ell| \; = \; 1 \; . \tag{145}$$

(k has the same meaning above as in the free particle

problem.)

Of course the model considered here for the deuteron is

very crude and cannot fit all the experimental facts.  In

particular, the characterization of the potential as a finite

spherical well with the estimates (114) and (143) is quite

crude.  Better models are based upon the use of a potential

which turns repulsive at small enough distances (the "hard

core"), have an exponential tail at large distances (to

simulate the finite range pion exchange potential), and

include the previously mentioned tensor force.

**(d)  The  Coulomb  Problem**

The static potential between two electric charges $e_1$ and

$e_2$ is

$$V(r) \; = \; \frac{e_1 e_2}{r} \; , \tag{146}$$

where r is the distance between them.  If we take $e_1$ = -e and

$e_2$ = Ze (e = $|e|$ is the magnitude of charge on an electron)

we have

$$V(r) = -\frac{Ze^2}{r} \quad . \tag{147}$$

"Z" is the number of protons in the atomic nucleus.  Such a potential describes "Hydrogen-like" systems.  The radial eigenvalue equation is given by (115) above:

$$\left[ \frac{\hbar^2}{2\mu} \left( -\frac{d^2}{dr^2} + \frac{\ell(\ell+1)}{r^2} \right) - \frac{Ze^2}{r} \right] R(r) = ER(r). \tag{148}$$

Let us consider solutions to (148) in the neighborhood of r = 0.  For r sufficiently small, the equation R(r) satisfies is

$$\left[ -\frac{d^2}{dr^2} + \frac{\ell(\ell+1)}{r^2} \right] R(r) \approx 0 \quad . \tag{149}$$

Actually, any potential such that $\lim\limits_{r \to 0} r^2 V(r) = 0$ will satisfy (149) near r = 0.  (This includes all the cases that will be studied in this Chapter.)  Let us try a power law solution to (149):

$$u(r) = Cr^a \quad . \tag{150}$$

This implies that

$$a(a - 1) = \ell(\ell + 1) \quad . \tag{151}$$

We therefore have two possible solutions:

$$a = \ell + 1, \ -\ell \tag{152}$$

The solutions $\sim \dfrac{1}{r^\ell}$ do not satisfy the boundary condition (12) above.  Therefore near the origin, we expect the R(r)

function to behave like $r^{\ell+1}$ .   Actually, for the Coulomb

potential the $\ell$ = 0 case has to be handled separately to

confirm that R(r) ~ r near the origin.   For the $\ell$ = 0 case we

have

$$\left[ \frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + \frac{Ze^2}{r} \right] R(r) \approx 0 \tag{153}$$

near the origin.   Assuming R(r) = Cr in this vicinity, the

first order correction is then determined by

$$\frac{d^2 R_{corr.}(r)}{dr^2} \simeq - \frac{2Z}{a_0 r} (Cr) \tag{154}$$

where $a_0 = \frac{\hbar^2}{\mu e^2}$ .   (The "Bohr radius" of eq$^{n}$ (30) of chapter 2,

but with m → $\mu$.)   The correction induced by (154) is clearly

of order $r^2$, giving us

$$R(r) \simeq Cr \left( 1 - \frac{Zr}{a_0} \right) \tag{155}$$

Clearly, this process of correcting the lowest order solution

can be continued to higher orders, showing that the

assumption R(r) ≈ Cr is self-consistent.

Now let us consider the opposite extreme, r → ∞ in

(148).   In this case we expect that

$$- \frac{\hbar^2}{2\mu} \frac{d^2 R(r)}{dr^2} \approx ER(r) . \tag{156}$$

We will choose to look at the bound state solutions for which

E = $-|E|$.   (Scattering solutions also exist for E > 0 as for

the deuteron.)   The solutions to (156) are of the form

7.40

$$R(r) \sim e^{-kr}, e^{kr} \ . \tag{157}$$

Only the first possibility statisfies the boundary condition,

(123). $\left(k = \dfrac{\sqrt{2\mu|E|}}{\hbar}\right.$ , as usual.) Summing up what we have

learned about solutions to (148), we can state that:

$$\left. \begin{array}{ll} r \to 0: & R(r) \sim r^{\ell+1} \\[2ex] r \to \infty: & R(r) \sim e^{-kr} \end{array} \right\} \tag{158}$$

Based upon (158), we guess for the $\ell = 0$ wavefunction that

$$R(r) = Cre^{-kr} \tag{159}$$

Substituting (159) into (148), we learn that we indeed have a
solution if

$$k = \frac{Z}{a_0} \ . \tag{160}$$

Expressing k in terms of E, we can show that this means

$$E = -\frac{\hbar^2 Z^2}{2\mu a_0^2} = -\frac{\mu e^4 Z^2}{2\hbar^2} \ . \tag{161}$$

We estimated this result before in Chapter 2, p.2.13 when
Z = 1. The only difference between the above and our
previous estimate is the use of the reduced mass, $\mu$, in place
of the electron mass, m. A conventional way of stating the
energy of atomic states is to give the E/hc value, which has
units of inverse distance. The result (161) gives for
Hydrogen

$$\left(\frac{E}{hc}\right)_{\text{theory}} = 109677.6 \pm .9 \text{ cm}^{-1} \ , \tag{162}$$

whereas experiment tells us (1971, Masui)

$$\left(\frac{E}{nc}\right)_{expt} = 109677.587 \pm .005 \text{ cm}^{-1} . \qquad (163)$$

The agreement is quite good.  The error bars in (162) come
from uncertainites in the values of the fundamental constants
that come into eq$\underline{^n}$ (161).  (I used some older numbers from
E.R. Cohen and B.N. Taylor, <u>The 1973 Least-Squares Adjustment</u>
<u>of the Fundamental Constants</u>, to get (162).  There are higher
order corrections to (162) from relativistic effects I have
not included.  The result (163) is also an older experimental
result I got from the same reference.)  To normalize the
result (159) we demand that

$$1 = \int_0^\infty dr R^2(r) = C^2 \int_0^\infty dr r^2 e^{-2Zr/a_0}, \qquad (164)$$

but

$$\int_0^\infty dx x^2 e^{-ax} = 2a^{-3}, \qquad (165)$$

so that

$$C = 2 \left(\frac{Z}{a_0}\right)^{3/2} . \qquad (166)$$

Therefore

$$u(r) = R(r)r^{-1} = 2 \left(\frac{Z}{a_0}\right)^{3/2} e^{-Zr/a_0}. \qquad (167)$$

The full $\ell = 0$ wavefunction $\left(Y_{00} = \dfrac{1}{\sqrt{4\pi}}\right)$ is

$$u(r)Y_{00} = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{3/2} e^{-Zr/a_0} . \qquad (168)$$

We will now solve for the wavefunctions in general. We will do this using a method I learned about from Julian Schwinger. We begin in a seemingly strange place. Let us remind ourselves of the basic facts concerning the one dimensional harmonic oscillator (covered in Chapter 3, p. 3.33 onward). The Hamiltonian is

$$H = \frac{p_x^2}{2m} + \frac{1}{2} m\omega^2 x^2 . \tag{169}$$

By introducing dimensionless variables

$$\left. \begin{array}{l} H = \dfrac{H}{\hbar\omega} , \\[2em] p = \dfrac{p_x}{\sqrt{m\omega\hbar}} , \\[2em] q = \sqrt{\dfrac{m\omega}{\hbar}}\, x , \end{array} \right\} \tag{170}$$

the Hamiltonian becomes

$$H = \frac{1}{2} (p^2 + q^2). \tag{171}$$

The eigenvalue equation is $\left( <q|p = \frac{1}{i} \frac{d}{dq} <q| \right)$

$$\left[ -\frac{1}{2} \frac{d^2}{dq^2} + \frac{1}{2} q^2 \right] u_n(q) = E_n u_n(q), \tag{172}$$

where

$$E_n = \left( n + \frac{1}{2} \right) , \quad n = 0,1,2,\ldots \tag{173}$$

The normalization we chose was (see the comments in Ch.3 about switching variables to x instead of q)

$$\int_{-\infty}^{\infty} dq' u_n(q') u_n(q') = 1, \tag{174}$$

and we found explicitly that

$$u_n(q) = \frac{1}{\sqrt{\sqrt{\pi}\, 2^n n!}}\; e^{q^2/2} \left(-\frac{d}{dq}\right)^n e^{-q^2}. \tag{175}$$

A Taylor series for a general function $f(q)$ may be written

$$f(q + q') = \sum_{n=0}^{\infty} \frac{1}{n!} \left(q'\frac{\partial}{\partial q}\right)^n f(q). \tag{176}$$

Now multiply both sides of (175) by $\dfrac{\lambda^n}{\sqrt{n!}}$ (where $\lambda$ is a real number) and sum over n:

$$\sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} u_n(q) = \frac{1}{\pi^{1/4}} e^{q^2/2} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{\lambda}{\sqrt{2}}\frac{d}{dq}\right)^n e^{-q^2}. \tag{177}$$

Comparing the rhs of (177) to the lhs side of (176) we can identify $q' = -\dfrac{\lambda}{\sqrt{2}}$ , and we find that

$$\sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} u_n(q) = \frac{1}{\pi^{1/4}} e^{q^2/2} e^{-(q - \lambda/\sqrt{2})^2}, \tag{178}$$

$$= \frac{1}{\pi^{1/4}} e^{-q^2/2 + \sqrt{2}\lambda q - \lambda^2/2}. \tag{179}$$

The right hand side of (178) or (179) is called a generating function for the $u_n(q)$ because by doing a specified mathematical operation, we may generate all the

$u_n(q)$'s.  Reversing the reasoning that lead to these results,
we find that that mathematical operation is

$$u_n(q) = \frac{1}{\pi^{1/4}} \frac{1}{\sqrt{n!}} \left(\frac{d}{d\lambda}\right)^n \left[ e^{-q^2/2 + \sqrt{2}\,\lambda q - \lambda^2/2} \right] \Big|_{\lambda=0} . \qquad (180)$$

Now that we have that under our belts, let's consider
the two-dimensional harmonic oscillator.  Using the same
dimensionless form as before, our Hamiltonian is

$$\mathsf{H} = \frac{1}{2} (p_1^2 + p_2^2) + (q_1^2 + q_2^2) . \qquad (181)$$

Defining

$$\mathsf{H}_1 = \frac{1}{2} (p_1^2 + q_1^2), \ \mathsf{H}_2 = \frac{1}{2} (p_2^2 + q_2^2) \qquad (182)$$

we easily see that

$$[\mathsf{H}_1, \mathsf{H}_2] = 0. \qquad (183)$$

The simultaneous eigenkets of $\mathsf{H}_1$ and $\mathsf{H}_2$ are

$$|n_1, n_2 > \ = \ |n_1> \otimes \ |n_2> \qquad (184)$$

where

$$\left. \begin{array}{c} \mathsf{H}_1|n_1> \ = \ \left(n_1 + \frac{1}{2}\right)|n_1> \\[3mm] \mathsf{H}_2|n_2> \ = \ \left(n_2 + \frac{1}{2}\right)|n_2> \end{array} \right\} \qquad (185)$$

so that ($n_1, n_2 = 0, 1, 2, \ldots$ independently)

$$\mathsf{H}|n_1, n_2> \ = \ (n_1 + n_2 + 1)|n_1, n_2> \qquad (186)$$

Calling $n = n_1 + n_2$, the energy degeneracy of the energy level labelled by n is clearly $n + 1$.  The eigenvalue equation (186) projected into a coordinate basis is:

$$\left[ -\frac{1}{2} \left( \frac{\partial^2}{\partial q_1^2} + \frac{\partial^2}{\partial q_2^2} \right) + \frac{1}{2} \left( q_1^2 + q_2^2 \right) \right] u_{n_1 n_2}(q_1, q_2) = (n + 1) u_{n_1 n_2}(q_1, q_2), \quad (187)$$

where of course

$$< q_1, q_2 | n_1, n_2 > \; = \; < q_1 | n_1 > < q_2 | n_2 >, \quad (188)$$

or

$$u_{n_1 n_2}(q_1, q_2) \; = \; u_{n_1}(q_1) u_{n_2}(q_2) \; . \quad (189)$$

Let us now do the same thing for the two dimensional harmonic oscillator as we just got through doing for the one dimensional case, that is, find a generating function for the wavefunctions.  It is now easy to see that the statement analogous to (179) is

$$\sum_{n_1, n_2 = 0}^{\infty} \frac{\lambda_1^{n_1} \lambda_2^{n_2}}{\sqrt{n_1! n_2!}} \; u_{n_1 n_2}(q_1, q_2)$$

$$= \frac{1}{\sqrt{\pi}} \exp \left[ -\frac{1}{2} \left( q_1^2 + q_2^2 \right) + \sqrt{2} \left( q_1 \lambda_1 + q_2 \lambda_2 \right) - \frac{1}{2} \left( \lambda_1^2 + \lambda_2^2 \right) \right]. \quad (190)$$

Although these wavefunctions are a complete solution to the problem, we wish to find the eigenfunctions in terms of the plane polar coordinates $\rho$ and $\phi$, defined in the diagram, in preparation for solving the Coulomb problem.
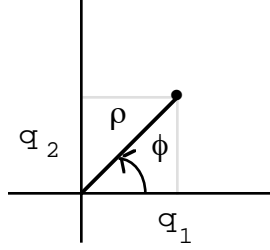
The transformation equations are

$$\left.\begin{array}{l} q_1 = \rho \cos \phi \\[2mm] q_2 = \rho \sin \phi \end{array}\right\} . \tag{191}$$

(In terms of the physical coordinates $x_1$ and $x_2$,
$\rho^2 = \frac{m\omega}{\hbar} \left( x_1^2 + x_2^2 \right).$)  Let us introduce the complex numbers,

$$\left.\begin{array}{l} \lambda_+ = \frac{1}{\sqrt{2}} (\lambda_1 - i\lambda_2), \\[3mm] \lambda_- = \frac{1}{\sqrt{2}} (\lambda_1 + i\lambda_2). \end{array}\right\} \tag{192}$$

Notice that $\lambda_+^* = \lambda_-$ and that

$$\lambda_+ \lambda_- = \frac{1}{2} \left( \lambda_1^2 + \lambda_2^2 \right). \tag{193}$$

The generating function (190) can be written in terms of $\rho$ and $\phi$ as

$$\frac{1}{\sqrt{\pi}} \exp\left[ -\frac{1}{2} \left( q_1^2 + q_2^2 \right) + \sqrt{2} (\lambda_1 q_1 + \lambda_2 q_2) - \frac{1}{2} \left( \lambda_1^2 + \lambda_2^2 \right) \right]$$

$$= \frac{1}{\sqrt{\pi}} \exp\left[ -\frac{1}{2} \rho^2 + \rho \left( \lambda_+ e^{i\phi} + \lambda_- e^{-i\phi} \right) - \lambda_+ \lambda_- \right] \tag{194}$$

By expanding the left hand side of (194) in a Taylor series in $\lambda_1$ and $\lambda_2$, we have learned that the coefficients, outside of a numerical factor, <u>are</u> the wavefunctions in the $q_1, q_2$

basis.  We now do the same thing in $\lambda_+$ and $\lambda_-$.  That is, we expand the right hand side of (194) in powers of $\lambda_+$ and $\lambda_-$, thereby defining a complete set of solutions in the $\rho,\phi$ basis:

$$\frac{1}{\sqrt{\pi}} \exp\left[- \frac{1}{2} \rho^2 + \rho(\lambda_+ e^{i\phi} + \lambda_- e^{-i\phi}) - \lambda_+ \lambda_-\right]$$

$$\equiv \sum_{n_+,n_-=0}^{\infty} \frac{\lambda_+^{n_+} \lambda_-^{n_-}}{\sqrt{n_+! n_-!}} u_{n_+ n_-}(\rho,\phi). \qquad (195)$$

Since the left hand side of (195) is equal to the right hand side of (190), we have that

$$\sum_{n_1,n_2=0}^{\infty} \frac{\lambda_1^{n_1} \lambda_2^{n_2}}{\sqrt{n_1! n_2!}} u_{n_1 n_2}(q_1,q_2) = \sum_{n_+,n_-=0}^{\infty} \frac{\lambda_+^{n_+} \lambda_-^{n_-}}{\sqrt{n_+! n_-!}} u_{n_+ n_-}(\rho,\phi). \qquad (196)$$

But, actually, we can say more than this.  Let us imagine picking out all the terms of the left of (196) such that $n_1 + n_2 = n$, where n is fixed.  (There will be n + 1 such terms.) These terms must be equal to the terms on the right of (196) such that $n_+ + n_- = n$ also, since $\lambda_+$ and $\lambda_-$ are just linear combinations of $\lambda_1$ and $\lambda_2$.  Therefore, we have the more specific statement that

$$\sum_{n_1+n_2=n} \frac{\lambda_1^{n_1} \lambda_2^{n_2}}{\sqrt{n_1! n_2!}} u_{n_1 n_2}(q_1,q_2) = \sum_{n_+ + n_- = n} \frac{\lambda_+^{n_+} \lambda_-^{n_-}}{\sqrt{n_+! n_-!}} u_{n_+ n_-}(\rho,\phi). \qquad (197)$$

(The notation, $\sum_{n_1+n_2=n}$, means to sum over all values of $n_1, n_2 = 0,1,2,\ldots$ such that $n_1 + n_2 = n$, where n is fixed.)

Eq$^{\underline{n}}$ (197) says that the $u_{n_1 n_2}(q_1,q_2)$ and the $u_{n_+ n_-}(\rho,\phi)$ are just

linear combinations of each other.  Since we know that the $u_{n_1 n_2}(q_1, q_2)$ represent a complete set of basis functions, it follows that the $u_{n_+ n_-}(\rho, \phi)$ are also a complete set, but in a different coordinate basis.  Although we know that the $u_{n_+ n}(\rho, \phi)$ are complete, they may not be orthonormal.  I prove in the Appendix to this Chapter that they actually <u>are</u> orthonormal.

We now will find explicit expressions for the $u_{n_+ n_-}$.  By operating on both sides of (195) with $\dfrac{1}{\sqrt{n_+!}}\left(\dfrac{d}{d\lambda_+}\right)^{n_+}$, evaluated at $\lambda_+ = 0$, we get

$$\sum_{n_-=0}^{\infty} \frac{\lambda_-^{n_-}}{\sqrt{n_-!}}\, u_{n_+ n_-} = \frac{1}{\sqrt{\pi n_+!}}\, \exp\!\left(-\frac{1}{2}\,\rho^2 + \rho e^{-i\phi}\lambda_-\right)(\rho e^{i\phi} - \lambda_-)^{n_+} . \qquad (198)$$

Now we operate with $\dfrac{1}{\sqrt{n_-!}}\left(\dfrac{d}{d\lambda_-}\right)^{n_-}$, evaluated at $\lambda_- = 0$, on both sides of (198) to get

$$u_{n_+ n_-} = \frac{1}{\sqrt{\pi n_+! n_-!}}\, e^{-1/2\,\rho^2}\left(\frac{d}{d\lambda_-}\right)^{n_-}\left[e^{\rho e^{-i\phi}\lambda_-}\,(\rho e^{i\phi} - \lambda_-)^{n_+}\right]\Bigg|_{\lambda_-=0}. \qquad (199)$$

Let us start to simplify this expression by defining a new parameter z such that

$$\lambda_- = \rho e^{i\phi}(1 - z). \qquad (200)$$

Replacing $\lambda_-$ by z in (199), we get

$$u_{n_+n_-} = \frac{1}{\sqrt{\pi n_+! n_-!}} \, e^{-1/2\rho^2} \left(- \frac{1}{(\rho e^{i\phi})} \frac{d}{dz}\right)^{n_-} \left[ e^{\rho^2(1-z)} (\rho e^{i\phi} z)^{n_+} \right] \Bigg|_{z=1}$$

$$= \frac{1}{\sqrt{\pi n_+! n_-!}} \, e^{1/2\rho^2} (\rho e^{i\phi})^{n_+-n_-} \left(- \frac{d}{dz}\right)^{n_-} \left[ e^{-z\rho^2} z^{n_+} \right] \Bigg|_{z=1} . \qquad (201)$$

Notice that we may write

$$\left(- \frac{d}{dz}\right)^{n_-} \left[ z^{n_+} e^{-\rho^2 z} \right] \Bigg|_{z=1} = \rho^{-2n_+} \left(- \frac{d}{dz}\right)^{n_-} \left[ (\rho^2 z)^{n_+} e^{-\rho^2 z} \right] \Bigg|_{z=1}$$

$$= \rho^{2(n_- - n_+)} \left(- \frac{d}{d(\rho^2 z)}\right)^{n_-} \left[ (\rho^2 z)^{n_+} e^{-\rho^2 z} \right] \Bigg|_{z=1} \qquad (202)$$

We can now set z = 1 in the last expression to get

$$\left(- \frac{d}{dz}\right)^{n_-} \left[ z^{n_+} e^{-\rho^2 z} \right] \Bigg|_{z=1} = \rho^{2(n_- - n_+)} \left(- \frac{d}{d\rho^2}\right)^{n_-} \left[ \rho^{2n_+} e^{-\rho^2} \right] . \qquad (203)$$

Our simplified expression for $u_{n_+n_-}$ is

$$u_{n_+n_-}(\rho, \phi) = \frac{1}{\sqrt{\pi n_+! n_-!}} \, e^{i(n_+-n_-)\phi} \rho^{n_--n_+} \, e^{1/2\rho^2} \left(- \frac{d}{d\rho^2}\right)^{n_-} \left[ \rho^{2n_+} e^{-\rho^2} \right] . (204)$$

Let us introduce the quantity

$$L_k^\alpha(x) \equiv x^{-\alpha} e^x \frac{1}{k!} \left(\frac{d}{dx}\right)^k \left[ x^{k+\alpha} e^{-x} \right] , \qquad (205)$$

where k is an integer (0,1,2,..) and $\alpha$ is an arbitrary real quantity.  These are called <u>Laguerre polynomials</u> when $\alpha = 0$ or <u>associated Laguerre polynomials</u> for $\alpha \neq 0$. (They are

polynomials of order k.) We can write our $u_{n_+ n_-}$ functions in terms of them as follows:

$$u_{n_+ n_-}(\rho, \phi) = \frac{1}{\sqrt{2\pi}} e^{i(n_+ - n_-)\phi} P_{n_+ n_-}(\rho) , \qquad (206)$$

where

$$P_{n_+ n_-}(\rho) = \sqrt{2 \frac{n_-!}{n_+!}} e^{-\rho^2/2} (-1)^{n_-} \rho^{n_+ - n_-} L_{n_-}^{n_+ - n_-}(\rho^2) . \qquad (207)$$

The quantity $\frac{1}{\sqrt{2\pi}} e^{i(n_+ - n_-)\phi}$ in (206) is the two dimensional analog of the spherical harmonics. Notice also that the argument of the Laguerre function is $\rho^2$, not $\rho$. The associated Laguerre polynomials have the symmetry property that

$$L_k^\alpha(x) = \frac{(\alpha + k)!}{k!} (-1)^\alpha x^{-\alpha} L_{\alpha+k}^{-\alpha}(x) , \qquad (208)$$

as long as k = 0,1,2,... and $\alpha$ = -k,-k + 1,-k + 2,... . Using (208) in (207), we get an alternative form for $P_{n_+ n_-}(\rho)$:

$$P_{n_+ n_-}(\rho) = \sqrt{2 \frac{n_+!}{n_-!}} e^{-\rho^2/2} (-1)^{n_+} \rho^{n_- - n_+} L_{n_+}^{n_- - n_+}(\rho^2) . \qquad (209)$$

Just like when we had two alternative expressions for the $Y_{\ell m}(\theta, \phi)$ and combined them into a third, more convenient, expression, we now do the same thing here with the alternative expressions (207) and (209). We choose to use (207) when $n_+ - n_- \geq 0$ and (209) when $n_+ - n_- \leq 0$. We can now write

$$P_{n_+ n_-}(\rho) = \sqrt{2 \dfrac{\left(\dfrac{n_+ + n_-}{2} - \dfrac{|n_+ - n_-|}{2}\right)!}{\left(\dfrac{n_+ + n_-}{2} + \dfrac{|n_+ - n_-|}{2}\right)!}} \ e^{-\rho^2/2} (-1)^{\frac{n_+ + n_-}{2} - \frac{|n_+ - n_-|}{2}}$$

$$\times \ \rho^{|n_+ - n_-|} \ L^{|n_+ - n_-|}_{\frac{n_+ + n_-}{2} - \frac{|n_+ - n_-|}{2}} (\rho^2) \qquad\qquad (210)$$

The expression (210) is somewhat awkward.  Let us define

$$m = n_+ - n_-,$$

$$n_r = \frac{n_+ + n_-}{2} - \frac{|m|}{2}, \qquad\qquad\qquad \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \qquad (211)$$

and relabel $P_{n_+ n_-}$ as $P_{n_r m}$.  We find

$$P_{n_r m}(\rho) = \sqrt{2 \frac{n_r!}{(n_r + |m|)!}} \ e^{-\rho^2/2} (-1)^{n_r} \rho^{|m|} L^{|m|}_{n_r}(\rho^2) \ . \quad (212)$$

The quantum number $n_r$ gives the number of nodes or zeros in $L^{|m|}_{n_r}(\rho^2)$ (or $P_{n_r m}(\rho)$) (excluding the point at infinity).  (We labeled the radial wavefunctions of the finite or infinite spherical wells in the same way.)  Our relabeled, complete radial wavefunctions are thus

$$u_{n_r m}(\rho, \phi) = \frac{1}{\sqrt{2\pi}} \ e^{im\phi} P_{n_r m}(\rho), \qquad\qquad (213)$$

with $P_{n_r m}(\rho)$ given above.

   We now have two different basis sets in which to uniquely characterize the states of our two dimensional system.  The explicit forms for the Cartesian wavefunctions are given in (175) and (189) above, and their quantum numbers are $n_1$ and $n_2$ where each can take on the values 0,1,2,...

independently.  Likewise, the explicit radial wavefunctions
are given in (212) and (213), and their quantum numbers are
specified by $n_r$ and m, where $n_r$ = 0,1,2,... and m is any
positive or negative integer (or zero).  Thus, our Hilbert
space is spanned either by the basis

$$|n_1, n_2 >,$$

which has energies

$$H|n_1, n_2 > = \hbar\omega(n_1 + n_2 + 1)|n_1, n_2 >, \qquad (214)$$

or by the basis

$$|n_r, m >,$$

which has energies

$$H|n_r, m > = \hbar\omega(2n_r + |m| + 1)|n_r, m > . \qquad (215)$$

(We have from (197) above that $n_+ + n_- = 2n_r + |m|$
characterizes the energy of the states.)  Although both
characterizations are complete, in general there is not a
one-to-one correspondence between thse different basis states
because of the degeneracies in energy.  In either
characterization, the possible energies of the system are
E = $\hbar\omega(n + 1)$ where n = 0,1,2,..., and the degeneracy of the
$n^{\underline{th}}$  level is given by n + 1.  An explicit energy diagram of
the different basis states is given below.

| Cartesian basis | | Radial basis |
|---|---|---|
| $n_1$ = 0, $n_2$ = 2 | | $n_r$ = 0, m = 2 |
| $n_1$ = 1, $n_2$ = 1 | | $n_r$ = 0, m = -2 |
| $\underline{n_1 = 2, n_2 = 0}$ | E = 3$\hbar\omega$ | $\underline{n_r = 1, m = 0}$ |

$n_1 = 0,\ n_2 = 1$                                                   $n_r = 0,\ m = 1$

$\underline{n_1 = 1,\ n_2 = 0}$                    $E = 2\hbar\omega$        $\underline{n_r = 0,\ m = -1}$

$\underline{n_1 = n_2 = 0}$                    $E = \hbar\omega$        $\underline{n_r = m = 0}$

In general, the (n + 1) Cartesian or radial states for which $E = \hbar\omega(n + 1)$ are linear combinations of each other. These are given by Eq$\underline{^n}$ (197) above. (There is, of course, a one-to-one correspondence between the two nondegenerate ground states.)

Just to get a feeling for the correspondence of the two sets of states, let's work some of the relationships out explicitly. The Cartesian basis ground state is given by

$$u_{00}(q_1,q_2) = u_0(q_1)u_0(q_2), \tag{216}$$

where (from (175))

$$u_0(q) = \frac{1}{\pi^{1/4}}\ e^{-q^2/2}. \tag{217}$$

Therefore

$$u_{00}(q_1,q_2) = \frac{1}{\sqrt{\pi}}\ e^{-(q_1^2 + q_2^2)/2}$$

$$\left(= \frac{1}{\sqrt{\pi}}\ e^{-\rho^2/2}\right), \tag{218}$$

On the other had we have

$$u_{00}(\rho,\phi) = \frac{1}{\sqrt{\pi}}\ e^{-\rho^2/2}\ L_0^0(\rho^2), \tag{219}$$

but it's easy to show that

$$L_0^\alpha(\rho^2) = 1, \tag{220}$$

for any $\alpha$, so that

7.54

$$u_{00}(\rho,\phi) = \frac{1}{\sqrt{\pi}} e^{-\rho^2/2}. \qquad (221)$$

Eq$^{\underline{n}}$s (218) and (221) are identical.

For the E = 2$\hbar\omega$ energy level we have the linearly independent Cartesian set

$$\left. \begin{array}{l} u_{10}(q_1,q_2) = u_1(q_1)u_0(q_2) \ , \\[2mm] u_{01}(q_1,q_2) = u_0(q_1)u_1(q_2) \ . \end{array} \right\} \qquad (222)$$

From (175) we have that

$$u_1(q) = \frac{\sqrt{2}}{\pi^{1/4}} q e^{-q^2/2}, \qquad (223)$$

so that

$$\left. \begin{array}{l} u_{10}(q_1,q_2) = \sqrt{\frac{2}{\pi}} \ q_1 e^{-(q_1^2 + q_2^2)/2} \\[4mm] \qquad\qquad \left( = \sqrt{\frac{2}{\pi}} \ \rho \cos\phi \ e^{-\rho^2/2} \right), \\[6mm] u_{01}(q_1,q_2) = \sqrt{\frac{2}{\pi}} \ q_2 e^{-(q_1^2 + q_2^2)/2} \\[4mm] \qquad\qquad \left( = \sqrt{\frac{2}{\pi}} \ \rho \sin\phi \ e^{-\rho^2/2} \right). \end{array} \right\} \qquad (224)$$

The radial basis set is

$$u_{01}(\rho,\phi) = \frac{1}{\sqrt{2\pi}} e^{i\phi} \sqrt{2} e^{-\rho^2/2} \rho L_0^1(\rho^2),$$

$$= \frac{1}{\sqrt{\pi}} \rho e^{i\phi} e^{-\rho^2/2},$$

$$u_{0-1}(\rho,\phi) = \frac{1}{\sqrt{2\pi}} e^{-i\phi} \sqrt{2} e^{-\rho^2/2} \rho L_0^1(\rho^2),$$

$$= \frac{1}{\sqrt{\pi}} \rho e^{-i\phi} e^{-\rho^2/2}.$$

$$(225)$$

The relationship between these two sets is given by (197) with n = 1:

$$\sum_{n_1+n_2=1} \frac{\lambda_1^{n_1} \lambda_2^{n_2}}{\sqrt{n_1! n_2!}} u_{n_1 n_2} = \sum_{n_++n_-=1} \frac{\lambda_+^{n_+} \lambda_-^{n_-}}{\sqrt{n_+! n_-!}} u_{n_r m} , \qquad (226)$$

which gives

$$\lambda_1 u_{10}(q_1,q_2) + \lambda_2 u_0(q_1,q_2) = \lambda_+ u_{01}(\rho,\phi) + \lambda_- u_{0-1}(\rho,\phi). \qquad (227)$$

However, expressing $\lambda_+$ and $\lambda_-$ in terms of $\lambda_1$ and $\lambda_2$ using (192) and matching coefficients of $\lambda_1$ and $\lambda_2$ on both sides of (227), we find that

$$u_{10}(q_1,q_2) = \frac{1}{\sqrt{2}} (u_{01}(\rho,\phi) + u_{0-1}(\rho,\phi)),$$

$$u_{01}(q_1,q_2) = \frac{i}{\sqrt{2}} (u_{0-1}(\rho,\phi) - u_{01}(\rho,\phi)).$$

$$(228)$$

Plugging in our explicit expressions above, we see that equations (228) are identically satisfied.

What explicit equations do the $u_{n_r m}(\rho, \phi)$ solve? Our two dimensional Schrödinger equation reads

$$\left[ -\frac{1}{2} \left( \frac{\partial^2}{\partial q_1^2} + \frac{\partial^2}{\partial q_2^2} \right) + \frac{1}{2} \left( q_1^2 + q_2^2 \right) \right] u_{n_1 n_2} = (n_1 + n_2 + 1) u_{n_1 n_2}, \quad (229)$$

in Cartesian coordinates. We change variables to $\rho$ and $\phi$:

$$\left. \begin{aligned} \frac{\partial^2}{\partial q_1^2} + \frac{\partial^2}{\partial q_2^2} &= \frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} , \\ q_1^2 + q_2^2 &= \rho^2 . \end{aligned} \right\} \quad (230)$$

The $u_{n_r m}(\rho, \phi)$ satisfy

$$\left[ -\frac{1}{2} \left( \frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} \right) + \frac{1}{2} \rho^2 \right] u_{n_r m} = (2n_r + |m| + 1) u_{n_r m}. \quad (231)$$

The variables $\phi$ and $\rho$ separate in (231) and according to (213) we may replace

$$\frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} \rightarrow -\frac{m^2}{\rho^2} . \quad (232)$$

In (231) we also replace $u_{n_r m}(\rho, \phi)$ by $P_{n_r m}(\rho)$. Thus $P_{n_r m}(\rho)$ satisfies

$$\left[ -\frac{1}{2} \left( \frac{d^2}{d\rho^2} + \frac{1}{\rho} \frac{d}{d\rho} - \frac{m^2}{\rho^2} \right) + \frac{1}{2} \rho^2 \right] P_{n_r m}(\rho) = (2n_r + |m| + 1) P_{n_r m}(\rho), \quad (233)$$

or

$$\left[\frac{d^2}{d\rho^2} + \frac{1}{\rho}\frac{d}{d\rho} - \frac{m^2}{\rho^2} + 2(2n_r + |m| + 1) - \rho^2\right] P_{n_r m}(\rho) = 0. \quad (234)$$

Let's now come to the point. The Coulomb problem, eq$^{\underline{n}}$ (148), can be cast into the form:

$$\left[\frac{d^2}{dr^2} - \frac{\ell(\ell + 1)}{r^2} + \frac{2Z}{a_0 r} - \frac{2\mu}{\hbar^2}|E|\right] R(r) = 0, \quad (235)$$

where we have set $E = -|E|$ for bound states. In order to establish a connection between eq$^{\underline{n}}$s (234) and (235), let us let

$$\rho^2 = \lambda r \quad (236)$$

in the two dimensional oscillator equation. It is easy to establish that

$$\left.\begin{aligned}\frac{d^2}{d\rho^2} &= \frac{4}{\lambda}\left[\frac{1}{2}\frac{d}{dr} + r\frac{d^2}{dr^2}\right],\\[2mm]\frac{1}{\rho}\frac{d}{d\rho} &= \frac{2}{\lambda}\frac{d}{dr}.\end{aligned}\right\} \quad (237)$$

Then, multiplying by $\frac{\lambda}{4r}$ , we find that (234) may be cast into the form

$$\left[\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{m^2}{4r^2} + \frac{\lambda(2n_r + |m| + 1)}{2r} - \frac{\lambda^2}{4}\right] P_{n_r m}\left(\sqrt{\lambda r}\right) = 0. \quad (238)$$

Eq$^{\underline{n}}$ (238) is of the same form of (235) except for the $\frac{1}{r}\frac{d}{dr}$ term. We can get rid of it by writing the equation satisfied by $\sqrt{r}\, P_{n_r m}$:

$$\left[\frac{d^2}{dr^2} - \frac{m^2 - 1}{4r^2} + \frac{\lambda(2n_r + |m| + 1)}{2r} - \frac{\lambda^2}{4}\right] \left[\sqrt{r}\ P_{n_r m}\left(\sqrt{\lambda}r\right)\right] = 0.$$

(239)

Casting our eyes back upon (235), we see that since we know the explicit solutions to (239), we can also work out the solutions to (235) given the following correspondence:

| 2-D oscillator | Coulomb |
|---|---|
| $(m^2 - 1)/4$ | $\ell(\ell + 1)$ |
| $\lambda(2n_r + |m| + 1)/2$ | $2Z/a_0$ |
| $\lambda^2/4$ | $2\mu|E|/\hbar^2$ |

The first line above tells us what $|m|$ corresponds to, the last two lines tell us what $\lambda$ and $n_r$ correspond to. Thus, from the first line we get

$$m^2 \rightarrow 4\ell(\ell + 1) + 1 = (2\ell + 1)^2 \quad,$$

so that

$$|m| \rightarrow |2\ell + 1| = 2\ell + 1.$$

(240)

The arrow symbol "$\rightarrow$" means "corresponds to". We should view (240) simply as a correspondence between parameters; that is, as far as eq$^{\underline{n}}$s (235) and (239) are concerned, $|m|$ and $\ell$ are just parameters which can take on any values. (Separate eigenvalue equations determine the values of m and $\ell$.) From the third line of the correspondence, we get

$$\lambda \rightarrow \frac{\sqrt{8\mu|E|}}{\hbar} \quad.$$

(241)

($\lambda$ must, from (236), be a positive quantity.)  Then, from the second line above, we get

$$\lambda n_r \rightarrow \frac{2Z}{a_0} - (\ell + 1)\lambda$$

from (240).  Then, using (241), we find the correspondence

$$n_r \rightarrow \frac{2Z\hbar}{a_0\sqrt{8\mu|E|}} - (\ell + 1). \qquad (242)$$

Now, according to the $\vec{L}^2$ eigenvalue equation, Eq$^{\underline{n}}$ (188) of Chapter 6, $\ell$ takes on values $0,1,2,\ldots$ .  Also, $n_r$ is <u>required</u>

by its eigenvalue equation (either (234), (238) or (239)) to take on values $0,1,2,\ldots$ also.  Therefore, the quantity

$\dfrac{2Z\hbar}{a_0\sqrt{2\mu|E|}}$ is also restricted to integer values such (242) is

satisfied.  In other words we have ($n = n_r + \ell + 1$)

$$\frac{2Z\hbar}{a_0\sqrt{8\mu|E|}} = n \qquad (243)$$

where, for a fixed $\ell$, we must have $n \geq \ell + 1$ for the integer $n$ (i.e., $n_r \geq 0$).  Eq$^{\underline{n}}$ (243) determines the energy levels of the Coulomb problem!  Solving for E, we get

$$E = -|E| = -\frac{Z^2\hbar^2}{2\mu a_0^2 n^2} , \qquad (244)$$

$$= -\frac{Z^2 e^2}{2a_0 n^2} . \qquad (245)$$

The integer n is called the <u>principal quantum number</u> since it completely determines the energy value. Reinserting n into (241) and (242), we find that the complete correspondence between the two dimensional harmonic oscillator and the (three dimensional) Coulomb problem is specified by

$$
\left.
\begin{aligned}
|m| &\to 2\ell + 1, \\[2mm]
\lambda &\to \frac{2Z}{a_0 n}\,, \\[2mm]
n_r &\to n - (\ell + 1).
\end{aligned}
\right\}
\qquad (246)
$$

The result (246) is very useful because it is only necessary to make the above substitutions in the quantity $\sqrt{r}\, P_{n_r m}\left(\sqrt{\lambda}\, r\right)$ (see eq$\underline{n}$ (239)) to read off the Coulomb radial eigenfunctions. (The angular eigenfunctions are, of course, just the $Y_{\ell m}(\theta,\phi)$ as in any central force problem.) Making the substitutions (246) into the explicit form for $P_{n_r m}$, Eq$\underline{n}$ (212), we find for these eigenfunctions

$$
R_{n\ell}(r) = N \sqrt{\frac{a_0 n}{Z}\,\frac{(n - \ell - 1)!}{(n + \ell)!}}\;(-1)^{n-\ell-1}\; e^{-Zr/a_0 n}
$$

$$
\times \left(\frac{2Zr}{a_0 n}\right)^{\ell+1} L_{n-\ell-1}^{2\ell+1}\left(\frac{2Zr}{a_0 n}\right).
\qquad (247)
$$

N is an unknown normalization constant which will be evaluated shortly. The phase factor $(-1)^{n-\ell-1}$ is unimportant and may be discarded if desired. (Because of the symmetry

property (208), there is an alternate form of these
wavefunctions that can be written down.)

We can evaluate the constant N as follows. We require
that (remember that u(r) = R(r)/r from (9))

$$\int_0^\infty dr \, |R_{n\ell}(r)|^2 = 1 \ . \tag{248}$$

We may also effectively evaluate the integral on the left
hand side of (248) by appealing to the two dimensional
harmonic oscillator.

Consider the integral

$$I \equiv \int_0^\infty d\rho \rho^3 (P_{n_r m}(\rho))^2 \ . \tag{249}$$

This integral may be evaluated using the result (3) from
Chapter 5:

$$\left\langle \frac{\partial H}{\partial \lambda} \right\rangle_E = \frac{\partial E}{\partial \lambda} \ . \tag{250}$$

The two dimensional harmonic oscillator Hamiltonian is

$$H = \frac{p_1^2 + p_2^2}{2m} + \frac{m\omega^2}{2} \left( x_1^2 + x_2^2 \right) , \tag{251}$$

and our orthonormality condition was (see Eq$\underline{^n}$s (A11) of the
Appendix and (206) above)

$$\int_0^\infty d\rho \rho \, P_{n_r m}(\rho) P_{n_r' m'}(\rho) = \delta_{n_r n_r'} \delta_{mm'} \tag{252}$$

when expressed in polar coordintes. (Remember,

$\rho^2 = \dfrac{m\omega}{\hbar} \left( x_1^2 + x_2^2 \right)$. Picking $\lambda = \omega$ for use in (250), we find that

$$\frac{1}{\hbar} \left\langle \frac{\partial H}{\partial \omega} \right\rangle_{n_r m} = \left\langle \rho^2 \right\rangle_{n_r m} \qquad (253)$$

but

$$\left\langle \rho^2 \right\rangle_{n_r m} = I, \qquad (254)$$

and

$$\frac{1}{\hbar} \frac{\partial E}{\partial \omega} = (2n_r + |m| + 1). \qquad (255)$$

Putting (253), (254) and (255) together, we find explicitly that

$$\int_0^\infty d\rho \rho^3 (P_{n_r m}(\rho))^2 = 2n_r + |m| + 1 . \qquad (256)$$

Now making the substitutions $\rho = (\lambda r)^{1/2}$ and (246) into (256), we get the statement that

$$\frac{1}{2} \left( \frac{2Z}{a_0 n} \right)^2 \int_0^\infty dr\, r \left( P_{\substack{n_r = n - (\ell+1) \\ m = \ell + 1}} \left( \sqrt{\frac{2Zr}{a_0 n}} \right) \right)^2 = 2n . \qquad (257)$$

However, in (247) we have defined

$$R_{n\ell}(r) \equiv N r^{1/2} \, P_{\substack{n_r = n - (\ell+1) \\ m = \ell + 1}} \left( \sqrt{\frac{2Zr}{a_0 n}} \right), \qquad (258)$$

so that (257) implies that

$$\int_0^\infty dr R_{n\ell}^2(r) = \frac{nN^2}{\left(\frac{Z}{a_0 n}\right)^2} \; . \qquad (259)$$

In order to reconcile (259) with (248) we can choose

$$N = \frac{Z}{a_0 n^{3/2}} \; . \qquad (260)$$

Thus, the fully normalized radial Coulomb wavefunctions are determined as:

$$R_{n\ell}(r) = \sqrt{\frac{Z}{a_0 n^2} \frac{(n - \ell - 1)!}{(n + \ell)!}} \; (-1)^{n-\ell-1} \; e^{-Zr/a_0 n}$$

$$\times \left(\frac{2Zr}{a_0 n}\right)^{\ell+1} L_{n-\ell-1}^{2\ell+1}\left(\frac{2Zr}{a_0 n}\right) , \qquad (261)$$

and the three dimensional wavefunctions are

$$u_{n\ell m}(\vec{r}) = \frac{R_{n\ell}(r)}{r} \; Y_{\ell m}(\theta,\phi) \; . \qquad (262)$$

They satisfy orthonormality:

$$\int d^3 r \; u_{n\ell m}^*(\vec{r}) u_{n'\ell'm'}(\vec{r}) = \delta_{nn'}\delta_{\ell\ell'}\delta_{mm'} \; . \qquad (263)$$

However, because there are scattering solutions (E > 0) that we have not included, the set in (262) is not complete.

We can connect with our earlier special solution for the ground state (eq$\underline{^n}$ (168)) now by putting n = 1, $\ell$ = 0 in (262) and (261). We get

$$u_{100}(\vec{r}) = 2 \left(\frac{Z}{a_0}\right)^{3/2} e^{-Zr/a_0} L_0^1 \frac{1}{\sqrt{4\pi}}$$

$$= \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{3/2} e^{-Zr/a_0} \quad , \tag{264}$$

exactly as we had before.

Notice that since the energy levels, Eqⁿ (245), are given by n <u>alone</u>, we now have a degeneracy in both m <u>and</u> $\ell$ for the Coulomb potential. This is called an <u>accidental degeneracy</u>. (This term seems to be reserved for situations where the <u>dynamics</u> and not the <u>symmetry</u> determines the degeneracy. Remember, it was the fact that we are working with a spherically symmetric central force that caused the m degeneracy.) Because $\ell + 1 \leq n$, we have the following classifications of the first few energy levels:

| n | $\ell$ | notation | #states | total |
|---|--------|----------|---------|-------|
| 1 | 0 | 1s | 1 | 1 |
| 2 | 0 | 2s | 1 | 4 |
|   | 1 | 2p | 3 |   |
| 3 | 0 | 3s | 1 |   |
|   | 1 | 3p | 3 | 9 |
|   | 2 | 3d | 5 |   |
| . | . | . |   | . |
| . | . | . |   | . |
| . | . | . |   | . |

One can easily show that the degeneracy of the nᵗʰ energy level is $n^2$. Schematically, the Coulomb energy levels look as follows:    $\left(\omega_0 = \frac{Z^2 e^4}{2 \hbar^4}\right)$

continuous { XXXXXXXXXXXXXXXXXXX E = 0

$$E_3 = -\omega_0/9$$

$$E_3 = -\omega_0/4$$

discrete

$$E_1 = -\omega_0$$
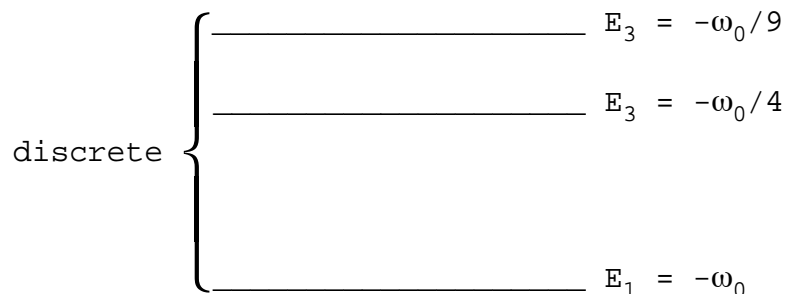
Using the same mysterious "zweideutigheit" rule as before, applied to electrons now, ("at most two electrons to each energy level") we can, on the basis of the Coulomb solution (neglecting the repulsive interactions of the electrons among themselves), begin to get a crude idea of the structure of atomic energy levels.  We use the spectroscopic notation introduced before with the addition of a superscript to tell us how many electrons are in a given atomic shell. (The principle quantum number, n, does <u>not</u> represent the number of nodes in the radial wavefunction here but simply indicates the ordering of the energy levels for an ideal hydrogen atom.  The actual number of radial nodes is given by n - $\ell$ - 1 since the wavefunction (262) is proportional to $e^{-Zr/a_0 n} L_{n-\ell-1}^{2\ell+1} \left(\frac{2Zr}{a_0 n}\right)$.  (I am not counting the point at r = $\infty$ as a node.)  The <u>atomic configurations</u> of the first 19 elements are given below.

| Element | Z | total config | last electron |
|---|---|---|---|
| H | 1 | 1s | 1s |
| He (inert) | 2 | $(1s)^2$ | 1s |
| Li | 3 | He(2s) | 2s |
| Be | 4 | $He(2s)^2$ | 2s |

7.66

| | | | |
|---|---|---|---|
| B | 5 | Be(2p) | 2p |
| C | 6 | Be(2p)$^2$ | 2p |
| N | 7 | Be(2p)$^3$ | 2p |
| 0 | 8 | Be(2p)$^4$ | 2p |
| F | 9 | Be(2p)$^5$ | 2p |
| Ne (inert) | 10 | Be(2p)$^6$ | 2p |
| Na | 11 | Ne(3s) | 3s |
| Mg | 12 | Ne(3s)$^2$ | 3s |
| Al | 13 | Mg(3p) | 3p |
| Si | 14 | Mg(3p)$^2$ | 3p |
| P | 15 | Mg(3p)$^3$ | 3p |
| S | 16 | Mg(3p)$^4$ | 3p |
| Cl | 17 | Mg(3p)$^5$ | 3p |
| Ar (inert) | 18 | Mg(3p)$^6$ | 3p |
| K | 19 | Ar(4s) | 4s |

There are obviously regularities in the order in which these energy levels are filled.  In order to try to explain these, I will refer to a schematic energy level diagram.  (Also to be passed out in class.)

Notice from the list on the previous page that the 2s shell is filled before the 2p shell and the 3s shell is filled before the 3p shell.  These facts are also brought out in the above mentioned energy level diagram.  In our ideal Coulomb solution, the 2s,2p and the 3s,3p energy levels are degenerate.  In real atoms, with their complicated interactions between the many particles involved, these energy degeneracies are split.  We can get a qualitative idea of the ordering of the split levels from the following considerations.  If we imagine adding electrons to an atomic nucleus with a fixed charge, the first electrons added will be the most strongly bound and therefore will tend to have wavefunctions that are largest near the nucleus.  As more

loosely bound electrons are added, the inner electrons tend
to <u>shield</u> some of the nucleus' electric charge from the outer
electrons.  Loosely speaking, this means that for electrons
in outer shells, a wavefunction which has a larger magnitude
near the nucleus will, in effect, sample a larger attractive
central charge.  Thus, the 2s energy levels will be filled
before the 2p shell since on the average the 2s electron's
wavefunction has a smaller separation from the nucleus than
the 2p electron does.  Likewise for the 3s and 3p energy
levels.  These splittings are between levels which have the
same value of n.  This shielding effect eventually means that
some energy levels with higher values of n will fill up
before levels with a lower n value.  The first example of
this is Potassium (K) in which the outer electron is in a 4s
state rather than 3d.

It is the order in filling up atomic shells that puts
the "Period" in the Periodic Table.  Obviously, there are
many other things to observe about the atomic structure of
the various elements, but we will not dwell on them here.  We
will, however, have more to say about the energy levels of
the real hydrogen atom in the next chapter.

## (e)   The "Confined Coulombic Model"

Let us contrast the schematic energy level diagrams
found in the Coulomb and infinite spherical well cases.
Plotting the energies and the assumed potentials

simultaneously, we get the following diagrams.  (We

generalize the Coulomb solution to the form $V(r) = -\dfrac{\xi}{r}$ ; let

us keep an open mind as to the meaning of the constant $\xi$.)



Coulomb                                   infinite spherical well
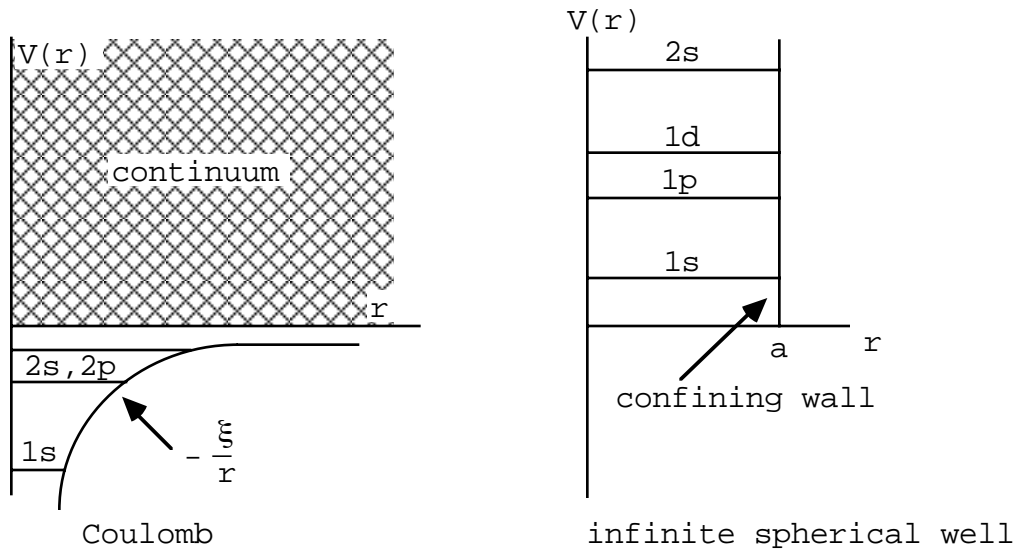
   In the Coulomb case, we must supply the boundary

condition that $R(r) \xrightarrow[r\to\infty]{} 0$, and in the infinite spherical well

case we must require that $R(a) = 0$.  For both we have that

$R(0) = 0$.  In the Coulomb problem, the discrete energies,

which are negative, are given by (Eq$\underline{n}$ (245) with $Ze^2 \to \xi$)

$$E_n = -\frac{\xi^2 \mu}{2\hbar^2 n^2} \cdot \tag{265}$$

In the infinite well case the energies are positive and are

given by (eq$\underline{n}$ (67)).

$$E = \frac{\hbar^2 (ka)^2}{2ma^2} , \tag{266}$$

where the dimensionless quantity ka is given by the condition

(62). We used the infinite spherical well as a starting point

for describing nuclear dynamics and we developed the Coulomb

solution as a way to begin to understand some physics of

atomic systems.   I would like to point out now that there is

a way of connecting these two seemingly different situations

as special cases of another model, which I am calling the

confined Coulombic model.   The assumed potential of this

composite model looks as follows.

$$-\frac{\xi}{r}$$
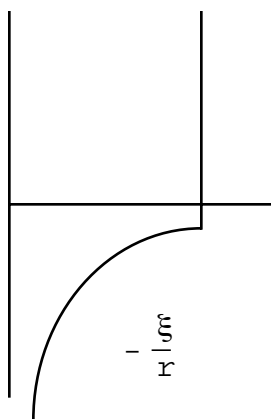
    Just like the spherical well problem, we are imagining an

infinite confining wall to exist at r = a, where we will

assume that the radial wavefunction vanishes, R(a) = 0.

Inside the wall, instead of there being a flat potential, we

are postulating an attractive Coulombic potential,

V(r) = $-\frac{\xi}{r}$.   Because of the confining boundary condition, we

expect that the eigenenergies will always be discrete like the

spherical well.   Then, in the limit a → ∞, we would also

expect that the allowed energies become more Coulomb-like.

That is, even though for "a" large but finite, although the

E > 0 energies remain discrete, they should become denser.

Likewise, the E < 0 discrete states will be better and better

reproduced as "a" becomes larger.   On the other hand, as "a"

becomes smaller we would expect that the energies of the states to be essentially due to the kinetic energy of confinement. That is, as we squeeze a quantum mechanical particle into a smaller volume we would, as a consequence of having decreased the particle's position uncertainty, also expect that the particle's momentum (and energy) to increase. This means, for small enough "a", that the value of the energies will be essentially independent of the Coulomb potential and will approach the infinite spherical well energies arbitrarily closely. We may well ask what use this model is in the real physical world; I will touch upon this later.

The equation we need to solve is eq$^{\underline{n}}$ (10),

$$\left[ - \frac{\hbar^2}{2m} \left( \frac{d^2}{dr^2} - \frac{\ell(\ell + 1)}{r^2} \right) + V(r) \right] R(r) = ER(r), \qquad (267)$$

where the potential is

$$V(r) = \begin{cases} \infty, & r \geq a \\ \\ -\dfrac{\xi}{r}, & r < a, \end{cases} \qquad (268)$$

and the boundary conditions are, essentially, that R(0) = 0 and R(a) = 0. In general, there are both positive and negative energy solutions to eq$^{\underline{n}}$ (267). (However, if the confinement radius becomes small enough, there will only be positive energy solutions.) We will only examine the E > 0

solutions here. Defining $\rho = kr$ with $k = \sqrt{\frac{2mE}{\hbar}}$ as usual, we may cast (267) into the form

$$\left[\frac{d^2}{d\rho^2} - \frac{\ell(\ell + 1)}{\rho^2} + \frac{\xi k}{E\rho} + 1\right] R(\rho) = 0 \qquad (269)$$

We can cast (269) into a more standard form by defining a function $\omega(\rho)$ such that

$$R(\rho) = \rho^{\ell+1} e^{-i\rho} \omega(\rho). \qquad (270)$$

We can now work out the first and second derivative as follows:

$$\frac{dR}{d\rho} = \rho^{\ell+1} e^{-i\rho}\left(\frac{(\ell + 1)}{\rho} \omega - i\omega + \frac{d\omega}{d\rho}\right), \qquad (271)$$

$$\frac{d^2R}{d\rho^2} =$$

$$\rho^{\ell+1} e^{-i\rho}\left(\frac{\ell(\ell + 1)}{\rho^2} \omega - \frac{2i(\ell + 1)}{\rho} \omega - \omega + \frac{2(\ell + 1)}{\rho} \frac{d\omega}{d\rho} - 2i \frac{d\omega}{d\rho} + \frac{d^2\omega}{d\rho^2}\right). \qquad (272)$$

The differential equation satisfied by $\omega(\rho)$ can therefore be written

$$\frac{d^2\omega}{d\rho^2} + 2\left(\frac{(\ell + 1)}{\rho} - i\right) \frac{d\omega}{d\rho} + \left(\frac{\xi k}{E\rho} - \frac{2i(\ell + 1)}{\rho}\right) \omega = 0. \quad (273)$$

We now define a new $x$ such that $x = i\rho$. Eq$^{\underline{n}}$ (273) can now be written

$$-\frac{d^2\omega}{dx^2} + 2\left(\frac{-(\ell + 1)}{x} + 1\right)\frac{d\omega}{dx} + \left(\frac{i\xi k}{Ex} + \frac{2(\ell + 1)}{x}\right)\omega = 0. \qquad (274)$$

or as

$$\frac{x}{2}\frac{d^2\omega}{dx^2} + (\ell + 1 - x)\frac{d\omega}{dx} + \left(-(\ell + 1) - \frac{i\xi k}{2E}\right)\omega = 0. \qquad (275)$$

We compare (275) with the differential equation satisfied by the <u>confluent hypergeometric functions</u>:

$$z\frac{d^2F}{dz^2} + (b - z)\frac{dF}{dz} - aF = 0. \qquad (276)$$

One often writes the solution to (276) as F(a,b,z).  The solution to (276) which is finite at z = 0 is given by the series solution

$$F(a,b,z) = 1 + \frac{az}{b} + \frac{a(a + 1)z^2}{b(b + 1)2!} + \frac{a(a + 1)(a + 2)}{b(b + 1)(b + 2)}\frac{z^3}{3!} + \dots \qquad (277)$$

It is understood in (277) that b is not zero or a negative integer; F(a,b,z) is undefined for these values.  Because (276) is a <u>second order</u> differential equation, there is another, linearly independent solution which, however, we will not be interested in because it will not be able to satisfy the boundary condition R(0) = 0.  There is a more compact notation for writing (277).  Let us define

$$\left.\begin{array}{l}(a)_n \equiv a(a + 1)(a + 2) \dots (a + n - 1), \\[2mm] (a)_0 \equiv 1.\end{array}\right\} \qquad (278)$$

Then, (277) can be written as

$$F(a,b,z) = \sum_{n=0}^{\infty} \frac{(a)_n z^n}{(b)_n n!} . \qquad (279)$$

The equations (275) and (276) become identical upon making the identifications

$$\left.\begin{array}{l} z = 2x, \\[2mm] a = \ell + 1 + \dfrac{i\xi k}{2E} , \\[2mm] b = 2\ell + 2. \end{array}\right\} \qquad (280)$$

The solution to (273) is therefore of the form

$$\omega(\rho) = CF\left(\ell + 1 + \frac{i\xi k}{2E} , 2\ell + 2, 2i\rho\right), \qquad (281)$$

and the radial wavefunction, $R(r)$ for this problem looks like

$$R(r) = N(kr)^{\ell+1} e^{-ikr} F\left(\ell + 1 + \frac{i\xi k}{2E} , 2\ell + 2, 2ikr\right). \qquad (282)$$

(Amazingly, the function $e^{-ikr} F\left(\ell + 1 + \dfrac{i\xi k}{2E} , 2\ell + 2, 2ikr\right)$ is completely real!) The quantity $N$ is the normalization factor, which we will not determine here. The condition that determines the positive energy levels, similar to, but more complicated than (62), is thus

$$e^{-ika} F\left(\ell + 1 + \frac{i\xi k}{2E} , 2\ell + 2, 2ika\right) = 0. \qquad (283)$$

Unfortunately, the energy eigenvalue condition cannot be simplified more than this. Now there are two dimensionless quantities in the arguments of F in (283). Taking $ka$ as one of these quantities, since we may write

$$\frac{\xi k}{2E} = \frac{(\xi ma/\hbar^2)}{ka} , \qquad (284)$$

we learn that $\xi ma/\hbar^2$ is the other dimensionless quantity.

Choosing a value of $\xi ma/\hbar^2$, Eq$\underline{n}$ (283) then determines the

possible values of ka, and therefore the energies.

Qualitatively, a plot of ka as a function of $\xi ma/\hbar^2$, looks

like the following for the 1s state:
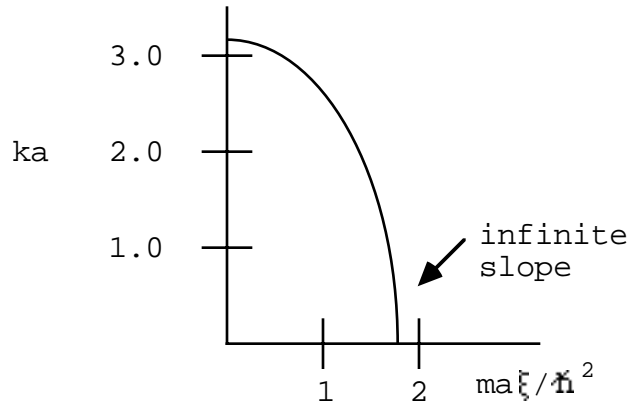


The place where ka = 0 in this graph is where the condition

(283) specializes to $k \to 0$.  In fact, one can show that

(just use the form (279) above)

$$\lim_{k\to 0} F\left(\ell + 1 + i\ \frac{\xi m/\hbar^2}{k}\ ,\ 2\ell + 2,\ 2ika\right) \to$$

$$(2\ell + 1)!\, J_{2\ell+1}\left(2\ \sqrt{\frac{2ma\xi}{\hbar^2}}\right)\left(\frac{2ma\xi}{\hbar^2}\right)^{-\ell-1/2}, \qquad (285)$$

where $J_n(x)$  (n = $2\ell$ + 1 and x = $2\ \sqrt{\dfrac{2ma\xi}{\hbar^2}}$ above) is a <u>Bessel</u>

<u>function</u> of order n.  Do not get them confused with the

earlier spherical Bessel functions we discussed in the free

particle and spherical well cases.  The relationship between

these two types of Bessel functions is given by

$$j_n(x) = \sqrt{\frac{\pi}{2x}} \; J_{n+1/2}(x). \tag{286}$$

A formula for the Bessel function $J_n(x)$ is

$$J_n(x) = \left(\frac{1}{2}\, x\right)^n \sum_{k=0}^{\infty} \frac{\left(-\frac{1}{4}\, x^2\right)^k}{k!\,(n+k)!}. \tag{287}$$

A more general formula for the order n not restricted to zero

or positive integers is also available (eq. 9.1.10 of the

National Bureau of Standards), but will not be dealt with

here.   Thus, the condition (283), restricted to k = 0 (that

is, solutions with exactly zero energy) becomes

$$J_{2\ell+1}\left(2\sqrt{\frac{2ma\xi}{\hbar^2}}\right) = 0. \tag{288}$$

The first zero of $J_1(x)$ for $x \neq 0$ occurs when

$$x = 3.83171... \tag{289}$$

which means the curve in the above figure crosses the $\dfrac{\xi ma}{\hbar^2}$

axis when

$$\frac{ma\xi}{\hbar^2} = 1.83525... \tag{290}$$

One can also show that the slope on this graph at this point

becomes infinite.   (This fact is true in general for the

other energy states.)

As stated at the beginning of this section, the confined

Coulombic model provides a connection between the Coulomb and

spherical well solutions in the limits $a \to \infty$ and $a \to 0$, respectively.  Since we are working with $E > 0$ here, let us consider the $a \to 0$ limit of the eigenvalue condition (283). Let us imagine fixing the values of $m$ and $\xi$ in the $a \to 0$ limit.  Then we have that

$$\lim_{a \to 0} \frac{ma\xi}{\hbar^2} = 0, \tag{291}$$

but from the qualitative behavior seen in the above graph, we would expect that

$$\lim_{a \to 0} ka = \text{fixed number.} \tag{292}$$

(Can you understand why Eq$^{\underline{n}}$ (292) is really just a statement of the uncertainty relation in the small "a" limit for this system?)  This equation may seem somewhat mysterious to you until you realize that $k$ is actually an implicit function of "a"; the eigenvalue condition (283) determines this dependence.  Thus, when a is made small enough to make

$$\frac{\xi ma}{\hbar^2} \ll ka, \tag{293}$$

(the system does not have to become relativistic to do this), then we get

$$\lim_{a \to 0} F\left(\ell + 1 + \frac{i\xi ma/\hbar^2}{ka} , 2\ell + 2, 2ika\right) \to F(\ell + 1, 2\ell + 2, 2ika). \tag{294}$$

However, from an identity, Eq$^{\underline{n}}$  13.6.4 of the NBS handbook, we have

$$F(\ell + 1, 2\ell + 2, 2ika) = \Gamma\left(\frac{3}{2} + \ell\right)e^{ika}\left(\frac{1}{2} \ ka\right)^{-\ell-1/2} \left(\frac{2ka}{\pi}\right)^{1/2} j_\ell(ka). \quad (295)$$

$\Gamma(x)$ (where $x = \frac{3}{2} + \ell$ above) is a famous classical function

called the <u>gamma</u> <u>function</u>.  Its form and value are immaterial

for the argument here except to say that $\Gamma(x) \neq 0$ for $x > 0$.

Thus, we learn that in the small "a" limit that the

eigenvalue condition just becomes

$$j_\ell(ka) = 0. \qquad\qquad (296)$$

This constitutes the recovery of the result (62) for the

infinite spherical well.  One can also show, although it

won't be done here, that the $a \to \infty$ limit of the eigenvalue

condition for $E < 0$ (which is slightly different from (283)

above) gives back the Coulomb energies.

   Since one is able to recover the bound state Coulomb

energies or the infinite spherical well energies as special

cases of this model, this means there must be a one-to-one

correspondence between these levels.  It is easy to establish

what this correspondence is, based upon the values of $\ell$, the

angular momentum quantum number, and the number of nodes in

the radial wavefunctions, neither of which can change as the

confinement radius is adjusted.  Remembering that the number

of nodes in the radial Coulomb wavefunctions is given by

$n - \ell$ (where "n" is the principle quantum number), we get the

following correspondence between the $E < 0$ Coulomb bound

states and the $E > 0$ infinite spherical well solutions.

(Remember, the number in front of the spectroscopic letter in

the Coulomb case gives the value of the principle quantum
number, while the number preceeding the letter in the
infinite well case just tells us the number of radial nodes.)



```
           Coulomb              infinite well

              4f _____                  3s
               .        \      /------------
               .         \    /      .
               .          \  /       .
                           \/
              3d _____  /\ _____ 2p
degenerate {  3p _____ \/  _____ 1f
              3s _____ /\   _____ 2s
              2p _____ /  \   _____ 1d
degenerate {  2s _____/    \   \_____ 1p
              1s _____ 1s
```
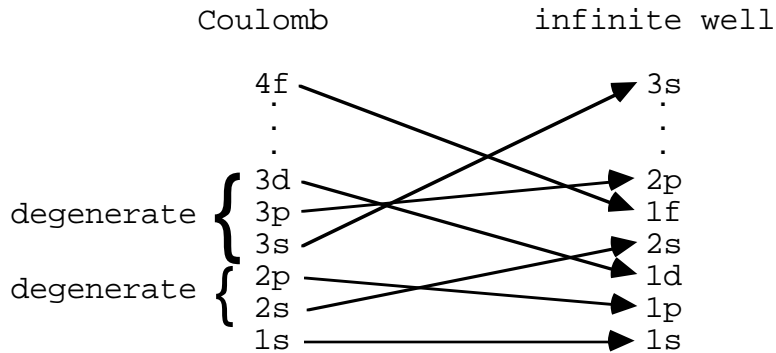
There are many other aspects to this problem that I
haven't addressed here.  I will stop with my short survey of
this model here, except to point out the type of system this
model is supposed to portray.  Earlier, I mentioned quarks as
being the building blocks of protons and neutrons.
Experimentally, 5 types of quarks are known (really 15 types
of quarks if we count the fact that each quark has a three-
fold symmetry called "color").  They are called up (u), down
(d), strange (s), charmed (c), and bottom (b).  (A sixth type
of quark, top (t), is not known experimentally, but is
necessary theoretically).  These are listed in the order of
their masses; the u quark has the lightest mass and the b
quark is the most massive.  The concept of the "mass" of a
quark is a somewhat fuzzy concept because isolated quarks are
never seen in nature.  This fact of nature is called
confinement.  The model presented here is a crude mock-up of
some quark systems which consist of a heavy quark (say, c or

b type) and a light quark (u or d).  (Actually, these systems
consist of a quark-<u>anti</u>quark pair, but the <u>antiquark</u> has the
identical mass of the same type of <u>quark</u>.)  The infinite
potential barrier in the model presented is a way of
mimmicking the confinement of quarks, while the Coulomb
potential in the interior is a way of modeling the short
range gluon interactions between the quarks.  In this model,
the heavy quark, relatively unaffected by the dynamics of its
light partner, resides in the center of the spherical well,
giving rise to the Coulomb potential.  The model given above
is, of course, nonrelativistic.  A relativistic version,
based upon solution of the <u>Dirac equation</u>, provides a more
realistic description of such systems.

I close this epic chapter with a copy of a paper on
energy level displacements in this model.  This is purely for
your amusement.

### APPENDIX A

I will derive here the orthonormalization condition on
the $u_{n_+n_-}(\rho,\phi)$, alluded to on p.48 of the present Chapter.  We
start out with the generating function result, eq$^{\underline{n}}$ (195):

$$\sum_{n_+,n_-=0}^{\infty} \frac{\lambda_+^{n_+}\lambda_-^{n_-}}{\sqrt{n_+!n_-!}}\, u_{n_+n_-} = \frac{1}{\sqrt{\pi}}\, \exp\left[-\frac{1}{2}\,\rho^2 + \rho\left(\lambda_+e^{i\phi} + \lambda_-e^{-i\phi}\right) - \lambda_+\lambda_-\right].$$

$$(A1)$$

I now multiply (A1) by its complex conjugate, except that $\lambda_+$
and $\lambda_-$ are replaced by the independent parameters $\lambda_+'$ and $\lambda_-'$ .
I also integrate:

$$\sum_{\substack{n_+,n_-=0 \\ n'_+,n'_-=0}}^{\infty} \frac{\lambda_+^{n_+}\lambda_-^{n_-}\lambda_-'^{n'_+}\lambda_+'^{n'_-}}{\sqrt{n_+!n_-!n'_+!n'_-!}} \int d^2q \; u_{n_+n_-} u^*_{n'_+n'_-}$$

$$= \frac{1}{\pi}\int d^2q \exp\left[-\rho^2 + \rho\left((\lambda_+ + \lambda'_+)e^{i\phi} + (\lambda_- + \lambda'_-)e^{-i\phi}\right) - \lambda_+\lambda_- - \lambda'_+\lambda'_-\right].$$

(A2)

I now change variables in the integral on the right in (A2) from $\rho,\phi$ to $q_1,q_2$ in order to do the integral. I get:

$$(\text{r.h.s. of (A2)}) = \frac{1}{\pi}\int dq_1 dq_2$$

$$\exp\left[-(q_1^2 + q_2^2)+(\lambda_+ + \lambda'_+)(q_1 + iq_2)+(\lambda_- + \lambda'_-)(q_1 - iq_2)- \lambda_+\lambda_- - \lambda'_+\lambda'_-\right].$$

(A3)

I now complete the square in $q_1$ in order to do the integral:

$$-q_1^2 + q_1(\lambda_+ + \lambda_- + \lambda'_+ + \lambda'_-)$$
$$= -(q_1 - \frac{1}{2}(\lambda_+ + \lambda_- + \lambda'_+ + \lambda'_-))^2 + \frac{1}{4}(\lambda_+ + \lambda_- + \lambda'_+ + \lambda'_-)^2$$

(A4)

The relevant $q_1$ part of the integral in (A3) is

$$\int_{-\infty}^{\infty} dq_1 \exp\left[-q_1^2 + q_1(\lambda_+ + \lambda_- + \lambda'_+ + \lambda'_-)\right]$$

$$= \underbrace{\int_{-\infty}^{\infty} dx e^{-x^2}}_{\sqrt{\pi}} \exp\left[\frac{1}{4}(\lambda_+ + \lambda_- + \lambda'_+ + \lambda'_-)^2\right].$$
(A5)

Similarly for $q_2$:

$$\int_{-\infty}^{\infty} dq_2 \, \exp\left[ -q_2^2 + iq_2(\lambda_+ + \lambda_+' - \lambda_- - \lambda_-') \right]$$

$$= \int_{-\infty}^{\infty} dx e^{-x^2} \exp\left[ -\frac{1}{4}(\lambda_+ + \lambda_+' - \lambda_- - \lambda_-')^2 \right]. \qquad (A6)$$

$$\underbrace{\phantom{xxxxxx}}$$
$$\sqrt{\pi}$$

Thus, we have

(r.h.s of (A2))

$$= \exp\left[ \frac{1}{4}(\lambda_+ + \lambda_+' + \lambda_- + \lambda_-')^2 \right.$$

$$\left. - \frac{1}{4}(\lambda_+ + \lambda_+' - \lambda_- - \lambda_-')^2 - \lambda_+\lambda_- - \lambda_+'\lambda_-' \right], \qquad (A7)$$

$$= \exp\left[ \lambda_+\lambda_-' + \lambda_+'\lambda_- \right]. \qquad (A8)$$

We may expand (A8) as

$$\exp\left[ \lambda_+\lambda_-' + \lambda_+'\lambda_- \right] = \sum_{\substack{n_+=0 \\ n_-=0}}^{\infty} \frac{(\lambda_+\lambda_-')^{n_+}}{n_+!} \frac{(\lambda_+'\lambda_-)^{n_-}}{n_-!} . \qquad (A9)$$

Our results up to this point can be summarized as

$$\sum_{\substack{n_+,n_-=0 \\ n_+',n_-'=0}}^{\infty} \frac{\lambda_+^{n_+}\lambda_-'^{n_+'} \lambda_-^{n_-} \lambda_+'^{n_-'}}{\sqrt{n_+!n_-!n_+'!n_-'!}} \int d^2q \, u_{n_+n_-} u_{n_+'n_-'}^* = \sum_{n_+,n_-=0}^{\infty} \frac{\lambda_+^{n_+} \lambda_-'^{n_+} \lambda_-^{n_-} \lambda_+'^{n_-}}{n_+!n_-!} \qquad (A10)$$

Comparing equal powers of $\lambda_+$, $\lambda_-'$, $\lambda_-$, $\lambda_+'$ on either side of (A10), we get that

$$\int d^2q \, u_{n_+n_-} u_{n_+'n_-'}^* = \delta_{n_+n_+'}\delta_{n_-n_-'}. \qquad (A11)$$

7.82

This is the desired result.  This result is useful when we normalize the Coulomb wave function in (252) of the present Chapter.

## Problems

1. The Hamiltonian of a three-dimensional harmonic oscillator is

$$H = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + \frac{1}{2} m\omega^2(x^2 + y^2 + z^2).$$

(a) Given the wavefunction and energy levels of the one-dimensional harmonic oscillator as given in Ch. 3 of the notes, find the wavefunctions and allowed energies for the three dimensional case. [Hint: The Schrodinger equation is separable in the three dimensions. Just use, don't re-do, the Ch. 3 calculation.]

(b) What is the degree of degeneracy of the first three energy levels?

(c) Find a formula which expresses the degree of degeneracy of the $n^{th}$ energy level. [Hint: The summation formula,

$$\sum_{i=1}^{n} i = \frac{n}{2}(n+1),$$

is useful here.]

2. Get explicit forms for $j_0(p)$, $j_1(p)$, and $j_2(p)$ from Eq. 7.39 of the script.

3. Using the form Eq. 7.39, derive the recurrence relation

$$\frac{dj_\ell(x)}{dx} = -j_{\ell+1}(x) + \frac{\ell}{x} j_\ell(x).$$

4. Verify Eq. 7.53 in the case $\ell=0$. [Hint: Consider the quantity $[\delta(r-r') - \delta(r+r')]$ where

$$\delta(r-r') \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \; e^{ik(r-r')}.]$$

5. This will be a rather long problem that will, hopefully, lead you to result Eq.(70) of Ch.7 in the notes. The radial eigenvalue equation in spherical coordinates is (R = r u(r))

$$-\frac{\hbar^2}{2m} [R'' - \frac{\ell(\ell+1)}{r^2} R] + V(r)R = E R. \tag{1}$$

   (a)  First, take the derivative of (1) with respect to E and multiply by R; call this (2).  Then multiply (1) by $\partial R/\partial E$ and call this (3). Subtract (2) − (3) to get

$$-\frac{\hbar^2}{2m} [R \frac{\partial R''}{\partial E} - R'' \frac{\partial R}{\partial E}] = R^2. \tag{4}$$

   (b)  Considering that the radial normalization condition for the finite spherical well is (assume R is real)

$$\int_0^a dr \; R^2 = 1, \tag{5}$$

show that (4) implies

$$-\frac{\hbar^2}{2m} [R' \frac{\partial R}{\partial E} - R \frac{\partial R'}{\partial E}] \Big|_{r=a} = 1, \tag{6}$$

given that R(0) = 0 and that R'(0) is finite.

   (c)  Now assuming (N(E) real)

$$R(r) = N(E)j_\ell(k_n r)r, \tag{7}$$

where $k_n$ (= $\sqrt{2mE}/\hbar$) is determined by

$$j_\ell(k_n a) = 0, \tag{8}$$

show that (6) specializes to

$$\frac{N^2a^3}{2} \left(\frac{dj_\ell(x)}{dx}\right)^2 \Big|_{x=k_na} = 1. \tag{9}$$

(d)  Now, using the recursion relation from prob.2 above, show that (9) gives

$$N^2 = \frac{2}{a^3 \, j_{\ell+1}^2 \, (k_na)} \, , \tag{10}$$

when (8) is also used.

6. Check out the result (10) of the last problem by finding $N^2$ explicitly for $\ell = 0$ by doing the normalization integral explicitly.  Compare with (10) specialized to $\ell = 0$.

7. Show for the two body problem that

$$\vec{L} = \vec{L}_{cm} + \vec{\ell},$$

where

$$\vec{L} = \vec{x}_1 \times \vec{p}_1 + \vec{x}_2 \times \vec{p}_2,$$

$$(1,2 \text{ are particle labels})$$

$$\vec{L}_{cm} = \vec{X} \times \vec{P},$$

$$\vec{\ell} = \vec{x} \times \vec{p}.$$

All these quantities are operators.

8. Using the definitions in prob.1 above, show that

$$[(L_{cm})_i, \ell_j] = 0,$$

so that these two quantities may be specified simultaneously.

9. Show that the $\ell = 1$ eigenvalue equation (131) of Ch.7 can be written as

$$\frac{1}{(\kappa'a)^2} + \frac{1}{(\kappa'a)} = -\frac{1}{(\kappa a)^2} + \frac{\cot(\kappa a)}{(\kappa a)}.$$

Argue on the basis of this equation that the well depth that binds the first $\ell = 1$ state is (see Eq.(140) of the notes)

$$\frac{2\mu V_0 a^2}{\hbar^2} = \pi^2.$$

10. Find the normalization factors A and B (in Eq.(128)) for the $\ell = 0$ deuteron bound state. Show that

$$|A|^2 = \frac{2\kappa'}{1+\kappa'a},$$

$$|B|^2 = \frac{2\kappa'(\sin \kappa a)\ e^{2\kappa'a}}{1+\kappa'a}.$$

Be sure to require continuity in u(r) at r=a.

11. Show that:

$$\sum_{n=0}^{\infty} \frac{\lambda^n n}{\sqrt{n!}}\ u_n(q) = \lambda(\sqrt{2}q - \lambda)\frac{1}{\pi^{1/4}}\ e^{-q^2/2 + \sqrt{2}\lambda q - \lambda^2/2}$$

12. (a) Show by induction that

$$(\frac{d}{dx})^n\ e^{-x}\ f(x) = e^{-x}\ (\frac{d}{dx} - 1)^n\ f(x)$$

for an arbitrary function f(x).
    (b) Using (a) and given $(L_n(x) \equiv L_n^{(0)}(x))$

$$L_n(x) = \frac{e^x}{n!}\ (\frac{d}{dx})^n\ e^{-x}\ x^n,$$

show that

$$L_n(x) = \frac{1}{n!} \left(\frac{d}{dx} - 1\right)^n x^n.$$

   (c) Using the binomial theorem, show that the result in (b) then implies

$$L_n(x) = \sum_{k=0}^{n} \frac{(-1)^k \, n! \, x^k}{(k!)^2 \, (n-k)!} \, .$$

13. (a) Show by induction that

$$\left(\frac{d}{dx} - 1\right)^{n+1} x \, f(x) = x\left(\frac{d}{dx} - 1\right)^{n+1} f(x) + (n+1)\left(\frac{d}{dx} - 1\right)^n f(x).$$

   (b) Using (a), and given that

$$L_{n+1}(x) = \frac{1}{(n+1)!} \left(\frac{d}{dx} - 1\right)^{n+1} x \, x^n,$$

show that

$$(n+1)(L_{n+1} - L_n) = x \left(\frac{d}{dx} - 1\right)L_n.$$

14. (a) Show that

$$\frac{d}{dx} (L_{n+1} - L_n) + L_n = 0.$$

[Hint: We can write

$$L_{n+1} = \frac{1}{(n+1)!} \left(\frac{d}{dx} - 1\right)^n \left(\frac{d}{dx} - 1\right)x^{n+1}.]$$

   (b) Using 14(a) and 13(b), now show that $L_n(x)$ satisfies the differential equation

$$\left(x \frac{d^2}{dx^2} + (1-x) \frac{d}{dx} + n\right)L_n(x) = 0.$$

15. (a) Write down the <u>radial</u> equation for the 3-dimensional harmonic oscillator.  By introducing the dimensionless variable $\rho = (\frac{m\omega}{\hbar})^{1/2}r$, show that this equation may be put into the form $(V(r) = \frac{1}{2}m\omega^2 r^2 = \frac{1}{2}\hbar\omega\rho^2)$

$$[\frac{d^2}{d\rho^2} - \frac{\ell(\ell+1)}{\rho^2} - \rho^2 + 2(n+\frac{3}{2})] \ R_{n\ell}(\rho) = 0.$$

   (b)  Starting with (7.234), show that the 2-dimensional harmonic oscillator radial equation may be written as $(\rho^2 = \frac{m\omega}{\hbar}(x^2+y^2))$

$$[\frac{d^2}{d\rho^2} - \frac{(m^2-\frac{1}{4})}{\rho^2} + 2(|m|+2n_r+1) - \rho^2] \ [\vec{\nabla}\rho \ P_{n_r m}(\rho)] = 0.$$

Establish the correspondences between the various quantum numbers in the two cases.  Then, based upon the explicit form $P_{n_r m}(\rho)$, given in Eq.(7.212) of the notes, write down the implied form for the 3-D wavefunction $R_{n\ell}(\rho)$. (Do not worry about the normalization of the $R_{n\ell}(\rho)$.)

16. In the text I used the relation

$$< \frac{\partial H}{\partial\lambda} >_E = \frac{\partial E}{\partial\lambda} \ ,$$

to evaluate an integral necessary to normalize the Coulomb solutions.  Apply the same technique to the Coulomb Hamiltonian,

$$H = \frac{\vec{p}^2}{2\mu} - \frac{Ze^2}{r} \ ,$$

to show that

$$< \frac{1}{r} >_E = \frac{Z}{a_0 n^2} \ .$$

17. Find the E < 0 eigenvalue condition for the confined Coulombic model. (This replaces Eq.(283) for E > 0.)

18. The first zero of the Bessel function $J_1(x)$ occurs at x = 3.83171.... Look up the next 3 zeros of this function. Give a _physical_ _interpretation_ of these zeros in the context of the confined Coulombic model.

**Other  Problems**

19. A free particle wavefunction is given at t=0 by the ket (I am using the $|k,\ell,m>$ notation in my notes),

$$|\psi(k),0> \equiv C_1|k,0,0> + \frac{C_2}{\sqrt 5} (|2k,1,0> + 2i|2k,1,1>).$$

    (a)  What is $<\vec{r}|\psi(k),0>$?  (Write it out as _explicitly_ as possible as a function of r, $\phi$, $\theta$.)
    (b)  What is $|\psi(k),t>$?
    (c)  What is the expectation value of the energy for this particle?
    (d)  What possible values of $L_z$ will measurement find at t=0, and with what probabilities?
[Please show some intermediate steps or explain your reasoning in reaching your answers in parts (c) and (d).]

20. Use the differential relation (assume true),

$$\frac{d}{dx} L_{p+1}^{q-1}(x) = - L_p^q(x),$$

to show that the Laguerre polynomials (my definition)

$$L_p^q(x) \equiv \frac{x^{-q}}{p!} e^x \left(\frac{d}{dx}\right)^p x^{p+q} e^{-x},$$

are related to the alternate definition (Liboff's)

$$\mathcal{L}_p^q(x) \equiv (-1)^q \left(\frac{d}{dx}\right)^q \mathcal{L}_{q+p}^0(x),$$

$$\mathcal{L}_n^0(x) \equiv e^x \left(\frac{d}{dx}\right)^n x^n e^{-x},$$

by the equation:

$$L_p^q(x) = \frac{\mathcal{L}_p^q(x)}{(p+q)!} .$$

21. (a) Derive the energy eigenvalue condition (E < 0) for the $\ell=1$ state of a <u>finite</u> spherical well of radius "a".
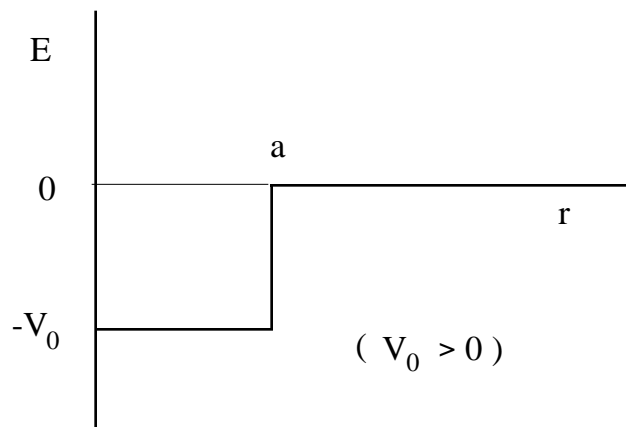


Use the definitions $\kappa' = \sqrt{-2mE}/\hbar$, $\kappa = \sqrt{2m(E+V_0)}/\hbar$, and simplify your expression as much as possible.

(b) Write down the $\ell=1$ <u>infinite</u> square well (V(r) = 0, r < a; V(r) = ∞, r ≥ a) energy eigenvalue condition (E > 0) as a function of $k = \sqrt{2mE}/\hbar$ and "a".

(c) By taking some appropriate limit, show how the expression in (a) reduces to the (b) eigenvalue condition. (Make the finite square well become infinite.)

22. Within the confined Coulombic model, use the <u>Heisenberg</u> <u>uncertainty principle</u> to <u>estimate</u> the confinement radius, $r_c$, which raises the energy of the 1s state of hydrogen, originally $E_{1s} = -\dfrac{\xi^2 m}{2\hbar^2}$ for $a \to \infty$, to $E_{1s} = 0$. (Consider $\xi$ and m fixed parameters.) Compare your estimate with the <u>exact</u> answer, given somewhere in the text.( Where is it, where is it??)

23. Find the value of the expectation value of r in the state where $\ell=n-1$ (it's maximum value). That is, show that

$$<n,\ell=n-1|r|n,\ell=n-1> = a_0 n(n + \tfrac{1}{2}),$$

where $a_0$ is the Bohr radius. [Hints: The explicit normalization condition for the radial part of the $\ell=n-1$ hydrogenic type wavefunctions is given on p.438 of Liboff as

$$\int_0^\infty dr\, r^2 |R_{n,\ell=n-1}(r)|^2 = \int_0^\infty dr\, r^{2n} \left(\frac{2Z}{a_0 n}\right)^{2n+1} \exp\left(-\frac{2Zr}{a_0 n}\right) \frac{1}{2n[(2n-1)!]}$$

$$= 1.$$

The integral we want is

$$<n,\ell=n-1|r|n,\ell=n-1> = \int_0^\infty dr\, r^3 |R_{n,\ell=n-1}(r)|^2.$$

The hint is to generate the integral we want by taking a derivative.]

24. Find an alternate form of the hydrogenic wavefunctions given in equation (261) (p. 7.62) of the notes. Do this by going back to p. 7.50 of the notes and using the following procedure:

**1)** Get an expression equivalent to (210) by using (209) when $n_+ - n_- \geq 0$, and (207) when $n_+ - n_- \leq 0$.

7.92

**2)** In your expression change to the new quantum numbers $n_r$ and m given in (211).

**3)** Use the relations (258) and (260) in the notes to find the new form of $R_{n\ell}(r)$.

25. The three dimensional harmonic oscillator differential equation was given in the problem set as $(V(r) = \frac{1}{2}m\omega^2 r^2 = \frac{1}{2}\hbar\omega\rho^2$ where $\rho = (\frac{m\omega}{\hbar})^{1/2}r)$

$$[\frac{d^2}{d\rho^2} - \frac{\ell(\ell+1)}{\rho^2} - \rho^2 + 2(n+\frac{3}{2})]R_{n\ell}(\rho) = 0.$$

Find the approximate $\rho$-dependence of $R_{n\ell}(\rho)$:

(a) near the origin, $\rho \ll 1$,

and

(b) far from the origin, $\rho \gg 1$.

## CHAPTER 8:   Addition of Angular Momenta

Let me remind you of the commutation properties associated with spin $\frac{1}{2}$ .   There were four independent operators in this case, which we chose to be the unit symbol 1, and the three $\sigma_i$ given by

$$\left. \begin{array}{l} \sigma_1 = |->< +| + |+><-|, \\[2mm] \sigma_2 = i(|->< +| - |+><-|), \\[2mm] \sigma_3 = |+><+| - |-><-|. \end{array} \right\} \qquad (1)$$

They have the properties that (Eq$^n$s (131)-(133) of Chapter 1)

$$\left. \begin{array}{ll} \sigma_k^2 = 1 & (k = 1,2,3), \\[1mm] \sigma_k\sigma_\ell = -\sigma_\ell\sigma_k & (k \neq \ell), \\[1mm] \sigma_k\sigma_\ell = i\sigma_m & (k,\ell,m \ \underline{\text{cyclic}}). \end{array} \right\} \qquad (2)$$

I later asserted that the entire content of their algebra, given in (2), is combined in the statement that (Ch.1, Eq$^n$ (149))

$$\sigma_i\sigma_j = 1\delta_{ij} + i \sum_k \varepsilon_{ijk}\sigma_k. \qquad (3)$$

An immediate consequence of (3) is

$$[\sigma_i,\sigma_j] = 2i \sum_k \varepsilon_{ijk}\sigma_k. \qquad (4)$$

The "crucial connection" that allowed us to tie our Process Diagram formalism to the real world property of electron spin was (Eq$^n$ (134) of Ch.1)

$$S_i = \frac{\hbar}{2} \sigma_i . \tag{5}$$

In terms of the $S_i$, the commutation relation (4) reads

$$[S_i, S_j] = i\hbar \sum_k \varepsilon_{ijk} S_k. \tag{6}$$

These relations are exactly of the same form as the commutation relations, given by Eq$^{\underline{n}}$ (70) of Chapter 6, for orbital angular momentum:

$$[L_i, L_j] = i\hbar \sum_k \varepsilon_{ijk} L_k. \tag{7}$$

Remember, these commutation relations lead directly to the statements

$$\left. \begin{aligned} \vec{L}^2 |\ell, m> &= \hbar^2 \ell(\ell + 1) |\ell, m>, \\ L_3 |\ell, m> &= \hbar m |\ell, m>. \end{aligned} \right\} \tag{8}$$

The physical picture associated with the statements in (8) is that of an orbiting particle (or an orbiting system of two particles). The allowed values of $\ell$ are $\ell = 0, 1, 2, \ldots$ . The cases $\ell = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$ were not consistent with the requirement

$$<\phi| = <\phi + 2\pi| . \tag{9}$$

Because the spin operators, $S_i$, satisfy the same commutation properties as the $L_i$, we must have analogous statements to (8) for $\vec{S}^2$ and $S_3$. Of course we know that

$$\vec{S}^2 = S_1^2 + S_2^2 + S_3^2 = \frac{\hbar^2}{4} \left( \sigma_1^2 + \sigma_2^2 + \sigma_3^2 \right) = \frac{3\hbar^2}{4} \mathbb{1}, \tag{10}$$

from (3), and that

$$S_3|\sigma_3'> = \frac{\hbar}{2}\sigma_3|\sigma_3'> = \frac{\hbar}{2}\sigma_3'|\sigma_3'>, \qquad (11)$$

where $\sigma_3' = \pm 1$. Relabeling our states by $m_S = \frac{1}{2}\sigma_3' = \pm\frac{1}{2}$,

Eq$^n$s (10) and (11) lead to

$$\left. \begin{array}{l} \vec{S}^2|m_S> = \hbar^2\left(\frac{3}{4}\right)|m_S>, \\[2mm] S_3|m_S> = \hbar m_S|m_S>. \end{array} \right\} \qquad (12)$$

The structure of these equations is the same as (8) above for

$\ell = \frac{1}{2}$. Now, however, since the origin of spin is an

intrinsic internal property of the particle, the requirement

(9) is not relevant. Thus, although there are no systems in

nature for which $\ell = \frac{1}{2}$ , $\frac{3}{2}$ , $\frac{5}{2}$ ,..., there are no such

restrictions on the value of the spin angular momentum.

Generalizing (12), we expect realizations in nature of

particle spins such that

$$\left. \begin{array}{l} \vec{S}^2|S,m_S> = \hbar^2 S(S+1]|S,m_S>, \\[2mm] S_3|S,m_S> = \hbar m_S|S,m_S>, \end{array} \right\} \qquad (13)$$

where $S = 0, \frac{1}{2}$ , $1$ , $\frac{3}{2}$ , $2$ ,... . In terms of Process

Diagrams, $S = \frac{1}{2}$ represents the two physical outcome case,

$S = 1$ represents the three outcome case, and so on. Only the

lower intrinsic spin values seem to be realized by

fundamental particles in nature.

Electrons and quarks have spin $\frac{1}{2}$ $\left(S = \frac{1}{2}\right.$ in $\left.(13)\right)$. It is

not presently known whether they are truly fundamental

objects without composite structure or whether they are also

non-elementary objects. Many particles which were originally

thought to be fundamental are now known to be composite. For

example, the proton is a composite structure containing three

quarks. The _total_ angular momentum of this particle, which

involves both the intrinsic spin of the quarks as well as the

relative angular momentum between them, is measured to be

$\frac{1}{2}$. In order to begin to understand such composite objects,

it is necessary to learn how to combine or add the angular

momenta of the subsystems.

Consider two systems which possess angular momentum

states given by

System 1: $|j_1, m_1>$,

System 2: $|j_2, m_2>$,

where we consider all possible values $j_1, j_2 = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ .

It is immaterial for this discussion whether this angular

momentum is due to spin or orbital motion. An obvious basis

for the description of this composite system is given by a

direct product,

$$|j_1, j_2; m_1, m_2> = |j_1, m_1>|j_2, m_2>. \qquad (14)$$

A complete set of operators whose eigenvalues completely characterize this state are just given by $\vec{J}_1^2$, $J_{1z}$, $\vec{J}_2^2$, $J_{2z}$, for which, of course,

$$
\left.
\begin{aligned}
\vec{J}_1^2 |j_1,j_2;m_1,m_2> &= \hbar^2 j_1(j_1 + 1)|j_1,j_2;m_1,m_2>, \\
J_{1z}|j_1,j_2;m_1,m_2> &= \hbar m_1|j_1,j_2;m_1,m_2>, \\
\vec{J}_2^2|j_1,j_2;m_1,m_2> &= \hbar^2 j_2(j_2 + 1)|j_1,j_2;m_1,m_2>, \\
J_{2z}|j_1,j_2;m_1,m_2> &= \hbar m_2|j_1,j_2;m_1,m_2>.
\end{aligned}
\right\} \tag{15}
$$

(I will often label Cartesian components "x,y,z" in this Chapter for clarity.)  Since our composite state is just a direct product of the individual states, it follows that ("1" and "2" are the particle labels)

$$[J_{1i},J_{2j}] = 0. \tag{16}$$

We can imagine a situation where a measurement on systems 1 and 2 is less convenient than a measurement on the total system.  There should be a more appropriate description for such a characterization.  Let us define the operator

$$\vec{J} = \vec{J}_1 + \vec{J}_2, \tag{17}$$

(which really means $\vec{J} = \vec{J}_1 \otimes 1 + 1 \otimes \vec{J}_2$) and call it the "total angular momentum."  Then, if the components $J_{1i}$ and $J_{2i}$ separately satisfy the commutation relations (like $\vec{L}$ or $\vec{S}$)

$$
\left.
\begin{aligned}
[J_{1i},J_{1j}] &= i\hbar \sum_k \varepsilon_{ijk}J_{1k}, \\
[J_{2i},J_{2j}] &= i\hbar \sum_k \varepsilon_{ijk}J_{2k},
\end{aligned}
\right\} \tag{18}
$$

we also have that the total angular momentum components satisfy the same relations (if Eq$^{\underline{n}}$ (16) holds):

$$[J_i, J_j] = i\hbar \sum_k \varepsilon_{ijk} J_k. \tag{19}$$

A consequence of (19) is that we may characterize the composite states of the total system by the eigenvalues of $\vec{J}^2$ and $J_z$, just as the commutation relations (7) lead to Eq$^{\underline{n}}$ (8). Therefore, we postulate the existence of an alternate set of states $|j_1, j_2; j, m>$ such that

$$\left.\begin{aligned}
\vec{J}^2 |j_1, j_2; j, m> &= \hbar^2 j(j+1) |j_1, j_2; j, m>, \\
J_z |j_1, j_2; j, m> &= \hbar m |j_1, j_2; j, m>.
\end{aligned}\right\} \tag{20}$$

The notation used anticipates the fact that the operators $\vec{J}_1^2$ and $\vec{J}_2^2$ commute with $\vec{J}^2$ and $J_z$; i.e., that

$$\left.\begin{aligned}
[\vec{J}_{1,2}, \vec{J}^2] &= 0, \\
[\vec{J}_{1,2}^2, J_z] &= 0.
\end{aligned}\right\} \tag{21}$$

Thus, the eigenvalues of $\vec{J}$, $J_z$, $\vec{J}_1^2$ and $\vec{J}_2^2$ can all be specified simultaneously. These operators represent an alternate complete set of operators that can be used to characterize the system.

Given the values of $j_1$ and $j_2$, the statement of completeness for the $|j_1, j_2; m_1, m_2>$ basis is

$$\sum_{m_1, m_2} |j_1, j_2; m_1, m_2> <j_1, j_2; m_1, m_2| = 1. \tag{22}$$

In more detail, this means that

$$\sum_{m_1} |j_1,m_1> <j_1,m_1| \ \sum_{m_2} |j_2,m_2> <j_2,m_2| \ = \ 1_1 \cdot 1_2. \qquad (23)$$

Using (22) then, we can formally express a basis state $|j_1,j_2;jm>$ in terms of the $|j_1,j_2;m_1,m_2>$ states:

$$|j_1,j_2;j,m> \ = \ \sum_{m_1,m_2} |j_1,j_2;m_1,m_2> \ <j_1,j_2;m_1,m_2|j_1,j_2;j,m>. \quad (24)$$

The quantities $<j_1,j_2;m_1,m_2|j_1,j_2;j,m>$ are called the "Clebsch-Gordon" coefficients. We will work out the values of some of these quantities in some special cases a little later.

Now we ask: What are the relationships between the quantum numbers $j_1$, $j_2$, $m_1$ and $m_2$ (which specify the eigenvalues of the $\vec{J}_1^{\,2}$, $\vec{J}_2^{\,2}$, $J_{1z}$, $J_{2z}$ set of operators) and the quantum numbers $j$ and $m$? We know that

$$J_z \ = \ J_{1z} \ + \ J_{2z}, \qquad (25)$$

from taking the z-component of (17). We therefore can write

$$<j_1,j_2;j,m|J_z - J_{1z} - J_{2z}|j_1,j_2;m_1,m_2> \ = \ 0. \qquad (26)$$

Allowing $J_{1z}$ and $J_{2z}$ to act to the right and $J_z$ to act to the left, we then get

$$(m - m_1 - m_2)<j_1,j_2;jm|j_1,j_2;m_1,m_2> \ = \ 0. \qquad (27)$$

That is, for all states that have a nonzero overlap, $<j_1,j_2;j,m|j_1,j_2;m_1,m_2> \ne 0$, we must have that

$$m = m_1 + m_2. \tag{28}$$

Thus, the quantized z-components of the individual angular momenta add to give the total value, m.  This is not surprising if we imagine the operator equation (17) to be a vector statement for numbers rather than a relation between operators.  Now remember that $m_1$ and $m_2$ can only take on values separated by integer intervals between $j_1$ and $-j_1$ or $j_2$ and $-j_2$, respectively.  That is:

$$m_1 = j_1, j_1 - 1, \ldots, -j_1 + 1, -j_1 \qquad (2j_1 + 1 \text{ values}),$$
$$m_2 = j_2, j_2 - 1, \ldots, -j_2 + 1, -j_2 \qquad (2j_2 + 1 \text{ values}).$$

Thus the maximum positive values, for fixed $j_1$ and $j_2$, are

$$(m_1)_{max} = j_1, \quad (m_2)_{max} = j_2. \tag{29}$$

Eq$^{\underline{n}}$ (29) then implies

$$(m)_{max} = j_1 + j_2. \tag{30}$$

One can show that the state associated with $(m)_{max} = j_1 + j_2$ has j $= j_1 + j_2$ also.  This corresponds to the maximum value of j.  The next lowest value of j that is conceivable would be $j = j_1 + j_2 - \frac{1}{2}$.  However, such a state would have

$m = j_1 + j_2 - \frac{1}{2}, j_1 + j_2 - \frac{3}{2}, \ldots, -j_1 - j_2 + \frac{1}{2}$, none of which are consistent with $m = m_1 + m_2$.  The only allowed values of j appear to be

$$j = j_1 + j_2, j_1 + j_2 - 1, j_1 + j_2 - 2, \ldots \ . \tag{31}$$

What is the lower limit of this process?  We require that the number of states in each basis (for given $j_1, j_2$ values) be the same.  In the $|j_1, j_2; m_1, m_2>$ basis, the number of distinct quantum states is given by $(2j_1 + 1)(2j_2 + 1)$ by counting the allowed values of $m_1$ and $m_2$.  This must also be the number of states in the $|j_1, j_2; j, m>$ basis.  Fixing the value of $j$ in $|j_1, j_2; j, m>$, we have $(2j + 1)$ m values.  Therefore, we require

$$\sum_{j=j_{min}}^{j_{max}} (2j + 1) = (2j_1 + 1)(2j_2 + 1) \qquad (32)$$

where $j_{max} = j_1 + j_2$, but $j_{min}$ is unknown.  We will try to solve (32) for $j_{min}$.  We can always write

$$\sum_{j=j_{min}}^{j_{max}} j = \sum_{j=1}^{j_{max}} j - \sum_{j=1}^{j_{min}-1} j . \qquad (33)$$

Using the fact that

$$\sum_{i=1}^{n} i = \frac{1}{2} n(n + 1), \qquad (34)$$

we then find that (we are covering the $j_{max,min}$ = integer case here; try to construct the analagous statements when $j_{max,min}$ are half-integers)

$$\sum_{j=j_{min}}^{j_{max}} j = \frac{1}{2} j_{max}(j_{max} + 1) - \frac{1}{2}(j_{min} - 1)j_{min}$$

$$= \frac{1}{2} \left[ j_{max}^2 + j_{max} - j_{min}^2 + j_{min} \right] . \qquad (35)$$

We therefore get

$$\sum_{j=j_{min}}^{j_{max}} (2j + 1) = j_{max}^2 + 2j_{max} - j_{min}^2 + 1. \tag{36}$$

Setting (36) equal to $(2j_1 + 1)(2j_2 + 1)$ and solving for $j_{min}^2$, we now find that

$$j_{min}^2 = 1 - (2j_1 + 1)(2j_2 + 1) + 2(j_1 + j_2) + (j_1 + j_2)^2$$

$$= (j_1 - j_2)^2. \tag{37}$$

Since $j_{min}$ must be a positive number, we can then write

$$j_{min} = |j_1 - j_2|. \tag{38}$$

Thus, the sum in (32) is such that

$$|j_1 - j_2| \le j \le j_1 + j_2, \tag{39}$$

which is called the "triangle inequality."  This result is also not surprising if we think of $j_1$ and $j_2$ as being the "magnitudes" of the "vectors" $\vec{J}_1$ and $\vec{J}_2$.  Then the case $j = j_1 + j_2$ would correspond to $\vec{J}_1$ and $\vec{J}_2$ being "parallel" while $j = |j_1 - j_2|$ would be associated with $\vec{J}_1$ and $\vec{J}_2$ being "antiparallel".

Now that we understand how the quantum numbers $j$ and $m$ arise (given fixed $j_1$ and $j_2$ values), we can write the statement of completeness in the $|j_1,j_2;j,m>$ basis which is is analogous to (22).  It is

$$\sum_{j=|j_1-j_2|}^{j_1+j_2} \sum_{m=-j}^{j} |j_1,j_2;j,m> < j_1,j_2;j,m| = 1 . \tag{40}$$

As a concrete example of the relationships between these two bases, let us consider the simple case of adding the angular momenta of two spin $\frac{1}{2}$ objects. That is, we set $j_1 = \frac{1}{2}$ and $j_2 = \frac{1}{2}$. The four states that arise in the $|j_1,j_2;m_1,m_2>$ basis are (I denote the state $|\frac{1}{2},\frac{1}{2};m_1,m_2>$ as $|m_1,m_2>$ for convenience in this context):

$$|\tfrac{1}{2},\tfrac{1}{2}> \ , \ |-\tfrac{1}{2},-\tfrac{1}{2}>,$$

$$|-\tfrac{1}{2},\tfrac{1}{2}> \ , \ |\tfrac{1}{2},-\tfrac{1}{2}>.$$

The four states in the $|j_1,j_2;j,m>$ basis are (the state $|\frac{1}{2},\frac{1}{2};j,m>$ is denoted here by $|j,m>$):

$$j = 1 \begin{cases} |1,1> \\ |1,0> \\ |1,-1> \end{cases} \quad j = 0 \ |0,0> \ .$$

These states must be linear combinations of each other. In order to find the explicit connections between them, let us recall some of the results of Chapter 6. We introduced there the operators $L_\pm = L_x \pm iL_y$ and found that these were raising ($L_+$) or lowering ($L_-$) operators on the quantum number m in the state $|\ell,m>$. These results depended <u>only</u> on the commutation properties of the $L_i$. Now since the total angular momentum operators, $J_i$, satisfy the same algebra, we have identical results for the $|j,m>$. That is, introducing

$$J_\pm = J_x \pm iJ_y \quad (J_x = J_{1x} + J_{2x},\ J_y = J_{1y} + J_{2y}), \tag{41}$$

we then have

$$J_\pm |j,m> = \hbar\sqrt{(j \mp m)(j \pm m + 1)} \; |j,m \pm 1>, \qquad (42)$$

and similarly for $J_{1\pm} = J_{1x} \pm iJ_{1y}$ on $|j_1,m_1>$ and $J_{2\pm} = J_{2x} \pm iJ_{2y}$ on $|j_2,m_2>$.

Going back to (24) and choosing $m = j_1 + j_2$ (it's maximum value), we see that the sum on the right is only a single term, and we get

$$|1,1> = |\tfrac{1}{2},\tfrac{1}{2}> <\tfrac{1}{2},\tfrac{1}{2}|1,1>. \qquad (43)$$

The Clebsch-Gordon coefficient $<\tfrac{1}{2},\tfrac{1}{2}|1,1>$ can be set equal to one, in order to maintain the usual normalization conditions on the states. Now let us apply the operator $J_- = J_{1-} + J_{2-}$ to both sides of (43):

$$\Rightarrow \sqrt{(1 + 1)(1 - 1 + 1)} \; |1,0>$$

$$= \sqrt{\left(\tfrac{1}{2} + \tfrac{1}{2}\right)\left(\tfrac{1}{2} - \tfrac{1}{2} + 1\right)} \; |-\tfrac{1}{2},\tfrac{1}{2}> + \sqrt{\left(\tfrac{1}{2} + \tfrac{1}{2}\right)\left(\tfrac{1}{2} - \tfrac{1}{2} + 1\right)} \; |\tfrac{1}{2},-\tfrac{1}{2}>$$

$$\Rightarrow |1,0> = \frac{1}{\sqrt{2}}\left[|-\tfrac{1}{2},\tfrac{1}{2}> + |\tfrac{1}{2},-\tfrac{1}{2}>\right]. \qquad (44)$$

Now apply it again to both sides of (44):

$$\Rightarrow \sqrt{(1 + 0)(1 - 0 + 1)} \; |1,-1>$$

$$= \frac{1}{\sqrt{2}}\left[\sqrt{\left(\tfrac{1}{2} + \tfrac{1}{2}\right)\left(\tfrac{1}{2} - \tfrac{1}{2} + 1\right)} \; |-\tfrac{1}{2},-\tfrac{1}{2}> + \sqrt{\left(\tfrac{1}{2} + \tfrac{1}{2}\right)\left(\tfrac{1}{2} - \tfrac{1}{2} + 1\right)} \; |-\tfrac{1}{2},-\tfrac{1}{2}>\right]$$

$$\Rightarrow \quad |1,-1> = |-\tfrac{1}{2},-\tfrac{1}{2}> \quad . \tag{45}$$

We could also have <u>started</u> with (45) and applied
$J_+ = J_{1+} + J_{2+}$ to deduce (44) and (43).

   The state $|0,0>$ is not yet determined.  Given that
$m = m_1 + m_2$, its most general form in this case is

$$|0,0> = C_1 |-\tfrac{1}{2},\tfrac{1}{2}> + C_2 |\tfrac{1}{2},-\tfrac{1}{2}> \quad . \tag{46}$$

Now the states $|0,0>$ and $|1,0>$ must be orthogonal.  This
means that

$$< 1,0|0,0> = \frac{1}{\sqrt{2}} \left[ C_1 < -\tfrac{1}{2},\tfrac{1}{2}|-\tfrac{1}{2},\tfrac{1}{2}> + C_2 < \tfrac{1}{2},-\tfrac{1}{2}|\tfrac{1}{2},-\tfrac{1}{2}> \right]$$

$$\Rightarrow \quad 0 = \frac{1}{\sqrt{2}} \left[ C_1 + C_2 \right]$$

$$\Rightarrow \quad C_1 = -C_2 \tag{47}$$

If we then normalize such that $< 0,0|0,0> = 1$ we can then
choose $C_1 = \dfrac{1}{\sqrt{2}}$ , and we get finally

$$|0,0> = \frac{1}{\sqrt{2}} \left[ |\tfrac{1}{2},-\tfrac{1}{2}> - |-\tfrac{1}{2},\tfrac{1}{2}> \right] \quad . \tag{48}$$

Collecting our results together, we have

$$\left.\begin{aligned}
|1,1 > &= |\tfrac{1}{2},\tfrac{1}{2}>, \\[2mm]
|1,0 > &= \frac{1}{\sqrt{2}} \left[ |-\tfrac{1}{2},\tfrac{1}{2}> + |\tfrac{1}{2},-\tfrac{1}{2}> \right], \\[2mm]
|1,-1 > &= |-\tfrac{1}{2},-\tfrac{1}{2}>,
\end{aligned}\right\} \tag{49}$$

for the j = 1 states (which are called the "triplet" states because m takes on 3 values) and

$$|0,0> = \frac{1}{\sqrt{2}} \left[ \, |\tfrac{1}{2},-\tfrac{1}{2} > \; - \; |-\tfrac{1}{2},\tfrac{1}{2} > \right],$$ (50)

for the j = 0 state (called a "singlet" for obvious reasons).

It is now easy to show, for this same system of two spin $\frac{1}{2}$ objects, that the statements of completeness in (22) and (40) are equivalent.

Since either basis set is complete, there is no reason at this point to prefer one description of the composite system to the other. However, when a Hamiltonian is specified for the system, there is in general no reason why the <u>individual</u> third components should be conserved. This would be expressed by $[H,J_{1z}] \neq 0$ and $[H,J_{2z}] \neq 0$. We would only expect in general that the <u>total</u> third component would be a constant of the motion, i.e., that $[H,J_z] = 0$. Thus, the $|j_1,j_2;j,m>$ states are usually the more relevant ones for interacting composite systems.

In order to investigate this point some more, let's now examine the addition of an <u>orbital</u> and a <u>spin $\frac{1}{2}$</u> angular momentum. This is the situation that occurs in the hydrogen atom. We set $\vec{J}_1 = \vec{L}$ and $\vec{J}_2 = \vec{S}$, so that

$$\vec{J} = \vec{L} + \vec{S}.$$ (51)

To set the notation, I list the mathematical properties of these operators:

$$[L_i, S_j] = 0, \tag{52}$$

$$\vec{L}^2 |\ell, m_\ell> = \hbar^2 \ell(\ell + 1)|\ell, m_\ell>, \tag{53}$$

$$L_3 |\ell, m_\ell> = \hbar m_\ell |\ell, m_\ell>, \tag{54}$$

$$\vec{S}^2 |m_s> = \frac{3\hbar^2}{4} |m_s>, \tag{55}$$

$$S_3 |m_s> = \hbar m_s |m_s>. \tag{56}$$

(I have gone back to a numeric labelling of the Cartesian components.)  The composite state, an eigenfunction of $\vec{L}^2$, $\vec{S}^2$, $L_3$, and $S_3$ will be denoted as follows:

$$|m_\ell, m_s> \equiv |\ell, m_\ell> |m_s>. \tag{57}$$

The number of such states, for fixed $\ell$ value, is $2(2\ell + 1)$. The properties of the states of total angular momentum are, of course

$$\vec{J}^2 |j, m> = \hbar^2 j(j + 1)|j, m>, \tag{58}$$

$$J_3 |j, m> = \hbar m |j, m>, \tag{59}$$

where, by the triangle inequality, we have for the allowed values of j,

$$\left. \begin{array}{l} j = \ell - \frac{1}{2} \text{ or } \ell + \frac{1}{2} , \ \ell > 0 \\[2em] j = \frac{1}{2} \qquad\qquad\qquad , \ \ell = 0. \end{array} \right\} \tag{60}$$

The number of states is also $2(2\ell + 1)$ for these two cases:

$$\# \text{ states} = \left(2\left(\ell - \frac{1}{2}\right) + 1\right) + \left(2\left(\ell + \frac{1}{2}\right) + 1\right), \quad \ell > 0 \\ \# \text{ states} = 2, \qquad\qquad\qquad\qquad\qquad\qquad \ell = 0 \;. \quad \Biggr\} \quad (61)$$

The $|j,m>$ are eigenfunctions of $\vec{L}^2$ and $\vec{S}^2$ also in addition to $\vec{J}^2$ and $J_3$.

We start out our investigation of the relationship between two sets of states by writing down the most general possible connection given that $m = m_\ell + m_S$:

$$|\ell + \tfrac{1}{2}, m> = C_1 |m - \tfrac{1}{2}, \tfrac{1}{2}> + C_2 |m + \tfrac{1}{2}, -\tfrac{1}{2}> , \qquad (62)$$

$$|\ell - \tfrac{1}{2}, m> = C_3 |m - \tfrac{1}{2}, \tfrac{1}{2}> + C_4 |m + \tfrac{1}{2}, -\tfrac{1}{2}> . \qquad (63)$$

The normalization conditions

$$<\ell + \tfrac{1}{2}, m | \ell + \tfrac{1}{2}, m> \; = \; <\ell - \tfrac{1}{2}, m | \ell - \tfrac{1}{2}, m> \; = 1,$$

give (we can choose all these constants real)

$$C_1^2 + C_2^2 = 1, \qquad\qquad (64)$$

$$C_3^2 + C_4^2 = 1. \qquad\qquad (65)$$

These mean that we may set

$$C_1 = \cos \alpha, \; C_2 = \sin \alpha, \qquad\qquad (66)$$

$$C_3 = \cos \beta, \; C_4 = \sin \beta. \qquad\qquad (67)$$

We must also have $<\ell + \tfrac{1}{2}, m | \ell - \tfrac{1}{2}, m> = 0$ (the states are orthogonal), which means that

$$\cos \alpha \cos \beta + \sin \alpha \sin \beta = 0 \tag{68}$$

or

$$\cos(\alpha - \beta) = 0. \tag{69}$$

Choosing

$$\beta = \alpha + \frac{\pi}{2} \tag{70}$$

(this makes $C_3$ = -sin $\alpha$ and $C_4$ = cos $\alpha$) will satisfy (69). Thus, we now only have one undetermined constant, say cos $\alpha$, to determine.

In order to determine this remaining constant, let's use the raising and lowering operators already introduced when we added two spin $\frac{1}{2}$ angular momenta. We have, of course, that

$$\vec{J}^2 = \vec{L}^2 + \vec{S}^2 + 2\vec{L} \cdot \vec{S}. \tag{71}$$

(We can write $\vec{L} \cdot \vec{S} = \vec{S} \cdot \vec{L}$ here since all components of $\vec{L}$ and $\vec{S}$ commute.) We can then derive the result that

$$2\vec{L} \cdot \vec{S} = 2L_3 S_3 + L_+ S_- + L_- S_+ , \tag{72}$$

where, as usual

$$\left. \begin{array}{l} L_\pm = L_1 \pm iL_2, \\[2mm] S_\pm = S_1 \pm iS_2. \end{array} \right\} \tag{73}$$

Now operating on both sides of (62) with $\vec{J}^2$, we find that

$$\left(\ell + \frac{3}{2}\right)\left(\ell + \frac{1}{2}\right) \, |\ell + \frac{1}{2}, m >$$

$$= \cos \alpha \left[ \ell(\ell + 1) + \frac{3}{4} + 2\left(m - \frac{1}{2}\right)\frac{1}{2}\right] \; |m - \frac{1}{2} , \frac{1}{2} >$$

$$+ \cos \alpha \left[ \sqrt{\left(\ell - m + \frac{1}{2}\right)\left(\ell + m + \frac{1}{2}\right)} \; \sqrt{\left(\frac{1}{2} + \frac{1}{2}\right)\left(\frac{1}{2} - \frac{1}{2} + 1\right)} \; \right] \; |m + \frac{1}{2}, -\frac{1}{2} >$$

$$+ \sin \alpha \left[ \ell(\ell + 1) + \frac{3}{4} + 2\left(m + \frac{1}{2}\right)\left(-\frac{1}{2}\right) \right] \; |m + \frac{1}{2}, -\frac{1}{2} >$$

$$+ \sin \alpha \left[ \sqrt{\left(\ell + m + \frac{1}{2}\right)\left(\ell - m + \frac{1}{2}\right)} \; \sqrt{\left(\frac{1}{2} + \frac{1}{2}\right)\left(\frac{1}{2} - \frac{1}{2} + 1\right)} \; \right] \; |m - \frac{1}{2}, \frac{1}{2} > .$$

$$(74)$$

More simply, this is the same as (dividing both sides by $\left(\ell + \frac{3}{2}\right)\left(\ell + \frac{1}{2}\right)$)

$$|\ell + \frac{1}{2}, m >$$

$$= \frac{|m - \frac{1}{2}, \frac{1}{2} >}{\left(\ell + \frac{3}{2}\right)\left(\ell + \frac{1}{2}\right)} \left\{ \cos \alpha \left[ \ell(\ell + 1) + \frac{1}{4} + m\right] + \sin \alpha \left[ \sqrt{\left(\ell + \frac{1}{2}\right)^2 - m^2} \; \right] \right\}$$

$$+ \frac{|m + \frac{1}{2}, -\frac{1}{2} >}{\left(\ell + \frac{3}{2}\right)\left(\ell + \frac{1}{2}\right)} \left\{ \sin \alpha \left[ \ell(\ell + 1) + \frac{1}{4} - m\right] + \cos \alpha \left[ \sqrt{\left(\ell + \frac{1}{2}\right)^2 - m^2} \; \right] \right\} .$$

$$(75)$$

Comparing (75) with the starting point, Eq$^{\underline{n}}$ (62), since the left hand sides of these equations are identical, we have that (remember $C_1 = \cos \alpha$, $C_2 = \sin \alpha$)

$$\cos \alpha = \frac{\cos \alpha \left[ \ell(\ell + 1) + \frac{1}{4} + m\right] + \sin \alpha \sqrt{\left(\ell + \frac{1}{2}\right)^2 - m^2}}{\left(\ell + \frac{3}{2}\right)\left(\ell + \frac{1}{2}\right)} , \qquad (76)$$

$$\sin \alpha = \frac{\sin \alpha \left[ \ell(\ell + 1) + \frac{1}{4} - m \right] + \cos \alpha \sqrt{\left(\ell + \frac{1}{2}\right)^2 - m^2}}{\left(\ell + \frac{3}{2}\right)\left(\ell + \frac{1}{2}\right)} . \qquad (77)$$

We will only need one of these equations to solve for $\cos \alpha$ or $\sin \alpha$. Dividing both sides of (76) by $\cos \alpha$ and then solving for $\tan \alpha$ gives

$$\tan \alpha = \frac{\left(\ell + \frac{3}{2}\right)\left(\ell + \frac{1}{2}\right) - \left[ \ell(\ell + 1) + \frac{1}{4} + m \right]}{\sqrt{\left(\ell + \frac{1}{2}\right)^2 - m^2}} . \qquad (78)$$

We can simplify this to

$$\tan \alpha = \frac{\ell - m + \frac{1}{2}}{\sqrt{\left(\ell + \frac{1}{2}\right)^2 - m^2}} = \sqrt{\frac{\ell + \frac{1}{2} - m}{\ell + \frac{1}{2} + m}} . \qquad (79)$$

Solving for $\cos \alpha$ then, we get

$$\cos^2 \alpha = \frac{1}{1 + \tan^2 \alpha} = \frac{\ell + \frac{1}{2} + m}{2\ell + 1} , \qquad (80)$$

$$\Rightarrow \cos \alpha = \sqrt{\frac{\ell + \frac{1}{2} + m}{2\ell + 1}} , \qquad (81)$$

where, by convention, we choose the positive root. (Eq$^{\text{n}}$ (70) is another conventional choice.) Given (81), we then get

$$\sin \alpha = \sqrt{\frac{\ell + \frac{1}{2} - m}{2\ell + 1}} , \qquad (82)$$

where the positive root is now <u>determined</u> because tan $\alpha \geq 0$ from (79) and we have chosen cos $\alpha \geq 0$ in (81). Our explicit connections between the states of the total angular momentum and composite orbital/spin states are therefore

$$|\ell + \tfrac{1}{2},m> = \sqrt{\frac{\ell + \frac{1}{2} + m}{2\ell + 1}} \; |m - \tfrac{1}{2},\tfrac{1}{2}>$$

$$+ \sqrt{\frac{\ell + \frac{1}{2} - m}{2\ell + 1}} \; |m + \tfrac{1}{2},-\tfrac{1}{2}> \, , \qquad (83)$$

$$|\ell - \tfrac{1}{2},m> = -\sqrt{\frac{\ell + \frac{1}{2} - m}{2\ell + 1}} \; |m - \tfrac{1}{2},\tfrac{1}{2}>$$

$$+ \sqrt{\frac{\ell + \frac{1}{2} + m}{2\ell + 1}} \; |m + \tfrac{1}{2},-\tfrac{1}{2}> \, . \qquad (84)$$

By projecting these kets into the spin/angular space $<\theta,\phi,m_S|$ and using the explicit matrix representation of spin $\tfrac{1}{2}$ talked about on pgs. 1.85-1.87 of the notes, Eq$^{\underline{n}}$s (83) and (84) are seen to be equivalent to

$$\mathcal{Y}_\ell^{j=\ell+1/2,m}(\theta,\phi) = \sqrt{\frac{\ell + \frac{1}{2} + m}{2\ell + 1}} \; Y_{\ell,m-1/2}(\theta,\phi)\psi_+$$

$$+ \sqrt{\frac{\ell + \frac{1}{2} - m}{2\ell + 1}} \; Y_{\ell,m+1/2}(\theta,\phi)\psi_- \, , \qquad (85)$$

$$\mathcal{Y}_\ell^{j=\ell-1/2,m}(\theta,\phi) = -\sqrt{\frac{\ell + \frac{1}{2} - m}{2\ell + 1}} \; Y_{\ell,m-1/2}(\theta,\phi)\psi_+$$

$$+ \sqrt{\frac{\ell + \frac{1}{2} + m}{2\ell + 1}} \; Y_{\ell,m+1/2}(\theta,\phi)\psi_- \; , \qquad (86)$$

where we have defined

$$\left[ \mathcal{Y}_{\ell}^{j=\ell\pm1/2,m}(\theta,\phi) \right]_{m_S} = \; < \theta,\phi,m_S | \ell \pm \frac{1}{2},m > \qquad (87)$$

where the two values of $m_S = \pm \frac{1}{2}$ are being used as matrix row labels.  (We associate $m_S = \frac{1}{2}$ with the top row and $m_S = -\frac{1}{2}$ with the bottom row.)  The $\psi_{\pm}$ are the column matrices (as in Eq$\underline{\text{n}}$ (211), Chapter 1)

$$\psi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; \psi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \qquad (88)$$

More compactly, both (85) and (86) may be written as

$$\mathcal{Y}_{\ell}^{j=\ell\pm1/2,m}(\theta,\phi) = \begin{array}{c} m_S = \frac{1}{2} \\[3em] m_S = -\frac{1}{2} \end{array} \left( \begin{array}{c} \pm \sqrt{\dfrac{\ell + \frac{1}{2} \pm m}{2\ell + 1}} \; Y_{\ell,m-1/2}(\theta,\phi) \\[2em] \sqrt{\dfrac{\ell + \frac{1}{2} \mp m}{2\ell + 1}} \; Y_{\ell,m+1/2}(\theta,\phi) \end{array} \right). \quad (89)$$

The top signs go with the case $j = \ell + \frac{1}{2}$ and the bottom signs with $j = \ell - \frac{1}{2}$ .  The $\mathcal{Y}_{\ell}^{j,m}(\theta,\phi)$ are called <u>spin-angle</u> <u>functions</u>.

We now have two complete sets of eigenfunctions in which to describe a situation where spin $\frac{1}{2}$ and orbital angular momentum are being added.  The set $|\ell,s;m_{\ell},m_s >$ ($S = \frac{1}{2}$ here) is an eigenvector of $\vec{L}^2$, $\vec{S}^2$, $L_3$ and $S_3$.  The other set, denoted

as $|\ell,s;j,m>$ and giving rise to the spin-angle functions

above, are eigenvectors of $\vec{J}^2$, $\vec{L}^2$, $\vec{S}^2$, and $J_3$. Which set

should we use in a given problem? Mathematically, it doesn't

matter, but computationally it makes a lot of difference.

Let us go back to the problem of the hydrogen atom in order

to get some experience in these matters. The following

considerations will be illustrative of both the use of spin-

angle basis states as well as the perturbation theory

development in Chapter 5.

First, let us get a sense of the speed involved in the

motion of an electron in the hydrogen atom ground state. The

Bohr radius is

$$a_0 = \frac{\hbar^2}{me^2} . \tag{90}$$

I argued back in Chapter 2, Eq$^n$ (34), that from the

uncertainty principle

$$p_r a_0 \sim \hbar , \tag{91}$$

for the ground state. Setting $p_r = mv$ and solving for v from

(90) and (91) then gives

$$\frac{v}{c} \sim \frac{e^2}{\hbar c} . \tag{92}$$

We came upon the constant $\alpha = \frac{e^2}{\hbar c}$ back on page 1.29 where it

was called the "fine structure constant." It is a pure

number with the approximate value

$$\alpha = \frac{1}{137.036} \ . \tag{93}$$

The result (92) implies a small relativistic correction to the energy levels calculated in Chapter 7, the so-called "fine structure."

The relativistic kinetic energy is

$$E^2 = \vec{p}^2 c^2 + m^2 c^4 \ , \tag{94}$$

from which we find that

$$E - mc^2 \approx \frac{p^2}{2m} - \frac{1}{2mc^2} \left(\frac{\vec{p}^2}{2m}\right)^2 . \tag{95}$$

The first term is the usual kinetic energy term and the second represents the lowest order relativistic correction. From (92) we would expect this correction to be of the order

$$\frac{\dfrac{1}{2mc^2}\left(\dfrac{\vec{p}^2}{2m}\right)^2}{\dfrac{\vec{p}^2}{2m}} \sim \left(\frac{v}{c}\right)^2 \sim \alpha^2, \tag{96}$$

relative to the unperturbed ground state energy.

There is another correction to the hydrogen atom energy levels of the same order of magnitude. It comes about because the electron, moving in the electric field of the nucleus, experiences an _effective_ magnetic field given by (this comes from Maxwell's equations)

$$\vec{H}_{eff} = - \frac{\vec{v}}{c} \times \vec{E} \ , \tag{97}$$

where the electric field, $\vec{E}$, is

$$\vec{E} = -\vec{\nabla}\phi(r). \tag{98}$$

$\phi(r)$ is the (central) electrostatic potential, which in the case of the Coulomb law, is given by

$$\phi(r) = \frac{Ze}{r}. \tag{99}$$

(We imagine Z protons in the nucleus; "e" is the <u>magnitude</u> of the electron's charge.) The electron's magnetic moment is given by Eq$^{n}$s (42) and (44) of Chapter 1:

$$\vec{\mu} = -\frac{e}{mc}\vec{S}. \tag{100}$$

(The relation (100) is in fact not exactly true but has some small corrections, the most important of which was calculated by Julian Schwinger.) Given (97) and (100), it is reasonable to expect that there will be an interation of the form $-\vec{\mu}\cdot\vec{H}_{eff}$. (See Eq$^{n}$ (29) of Chapter 1). Now, using (97), (98) and (100), one can show that this interaction can be written as

$$-\vec{\mu}\cdot\vec{H}_{eff} = -\frac{e}{m^2c^2}\frac{1}{r}\frac{d\phi}{dr}(\vec{L}\cdot\vec{S}), \tag{101}$$

where we have used the classical form $\vec{L} = \vec{r}\times(m\vec{v})$. The result (101) is actually too large by a factor of two. The reason is that we have not been careful in using the correct relativistic kinematics in evaluating the interaction. The

extra necessary factor of $\frac{1}{2}$ is called the "Thomas precession

factor" (after the English physicist L. H. Thomas).  It is

difficult to justify kinematically but extremely easy to

recover from the Dirac equation, which is the <u>relativistic</u>

equation satisfied by electrons.  An interaction of the form

of (101) says that the electron's spin will interact with its

own angular momentum.  Such an effect is called L-S coupling.

(This is the same L-S coupling mentioned in the deuteron

discussion in Chapter 7.)

In the above discussion, the quantities $\vec{p}$, $\vec{L}$, $\vec{S}$, etc.

are classical quantities.  As usual we replace these

quantities by their quantum mechanical operators.  (Notice

that $\vec{L} \cdot \vec{S} = \vec{S} \cdot \vec{L}$ for $\vec{S}$, $\vec{L}$ as operators so there is no

ambiguity in the replacement.)  Therefore, we write our

slightly corrected hydrogen atom Hamiltonian as

$$H = H_0 + H_{rel} + H_{LS} \tag{102}$$

where $H_0 = \dfrac{\vec{p}^2}{2\mu} - \dfrac{Ze^2}{r}$ was the original Hamiltonian, and $H_{rel}$ and

$H_{LS}$ are given by (letting $m \rightarrow \mu$ where appropriate makes

little difference here)

$$H_{rel} = - \frac{1}{2mc^2} \left( \frac{\vec{p}^2}{2m} \right)^2, \tag{103}$$

$$H_{LS} = - \frac{e}{2m^2c^2} \frac{1}{r} \frac{d\phi}{dr} (\vec{L} \cdot \vec{S}). \tag{104}$$

The only thing to do with $H_{rel}$ and $H_{LS}$ is to treat them as

perturbations as in the development in Chapter 5.  We must

use the degenerate perturbation theory outlined there since,

in fact, each energy level specified by n has a $2n^2$-fold

degeneracy.  (The factor of two comes from considering the

two values of electron spin, $m_s = \pm \frac{1}{2}$ .)  From Eq$\underline{n}$ (34) of

Chapter 5, the perturbed energy levels are given by

$$E_a = E_0 + <E_0 a | H_1 | E_0 a> \qquad (105)$$

to first order.  Remember, the label "a" in (105) was used to

distinguish between states with the same energy.  Remember

also that the basis to be used in (105) was one in which the

perturbing Hamiltonian, $H_1$, is diagonal in the degenerate

subspace specified by a.

In our case, this degenerate subspace is the one

specified by the quantum numbers $\ell$, $m_\ell$ and $m_s$ from the

complete set $\{\vec{L}^2, \vec{S}^2, L_3, S_3\}$ or the quantum numbers $\ell$, j and m

from the alternate set $\{\vec{J}^2, \vec{L}^2, \vec{S}^2, J_3\}$.  (The radial states

differing by the principle quantum number n are never

degenerate.)  Now $H_{rel}$ commutes with all of the operators in

both sets.  The other perturbation, $H_{LS}$, fails to commute with

all of the operators in $\{\vec{L}^2, \vec{S}^2, L_3, S_3\}$ and its perturbation

matrix is not diagonal in this set.  However, the

perturbation $H_{rel} + H_{LS}$ is diagonal in the set $\{\vec{J}^2, \vec{L}^2, \vec{S}^2, J_3\}$

and we may use (105) directly to find the new energies.  If,

for example, one were unaware of the $|\ell, s; j, m>$ set of states,

we could still find the first order perturbed energies, but
we couldn't start with Eq$^{\underline{n}}$ (105) above.  We would be forced
to diagonalize the perturbation in the

$|\ell,s;m_\ell,m_s>$ basis.  This would essentially repeat the work we
already did in finding the explicit relationship between the
two basis sets.  This would eventually give us the same
answers but with a lot of redundant work.  As Sakurai says in
Modern Quantum Mechanics:  "You have to be either a fool or a
masochist to use the $L_z,S_z$ eigenkets as base kets for this
problem."

    We thus have for the perturbative energy shifts

$$\Delta E_{n\ell sjm} = <n\ell sjm|(H_{rel} + H_{LS})|n\ell sjm>, \qquad (106)$$

where we are defining the separable state

$$|n\ell sjm> \equiv |n\ell> \otimes |\ell,s;j,m> . \qquad (107)$$

The $|n\ell>$ provide the hydrogen atom radial basis found in
Chapter 7:

$$\frac{R_{n\ell}(r)}{r} = u_{n\ell}(r) = <r|n\ell> . \qquad (108)$$

Now we know that (I continue to ignore in this Chapter the
distinction between m and $\mu$)

$$\left(\frac{\vec{p}^2}{2m} - \frac{Ze^2}{r}\right) |n\ell sjm> = E_n|n\ell sjm>, \qquad (109)$$

gives the unperturbed energy eigenvalues.  Therefore

$$\frac{\vec{p}^2}{2m} \; |n\ell sjm> \; = \; \left(E_n \; + \; \frac{Ze^2}{r}\right) \; |n\ell sjm>, \tag{110}$$

which implies that

$$<n\ell sjm|H_{rel}|n\ell sjm> \; = \; <n\ell sjm|- \; \frac{1}{2mc^2} \left(E_n \; + \; \frac{Ze^2}{r}\right)^2 \; |n\ell sjm>$$

$$= \; - \; \frac{1}{2mc^2} \; <n\ell |\left(E_n \; + \; \frac{Ze^2}{r}\right)^2 \; |n\ell>. \tag{111}$$

In order to evaluate (111), we will need to know the

expectation values $<\frac{1}{r}>_{n\ell}$ and $<\frac{1}{r^2}>_{n\ell}$ . Actually, from a

homework problem in the last Chapter, we know that

$$<\frac{1}{r}>_{n\ell} \; = \; \frac{Z}{a_0 n^2} \; . \tag{112}$$

We can use the same technique as displayed in this problem

(and the discussion of normalization of the hydrogen atom

eigenfunctions in Chapter 7) to also find $<\frac{1}{r^2}>_{n\ell}$ . Our

unperturbed Hamiltonian in the radial eigenspace can be

written as

$$H \; = \; \frac{p_r^2}{2m} \; + \; \frac{\hbar^2 \ell(\ell + 1)}{2mr^2} \; - \; \frac{Ze^2}{r} \tag{113}$$

with the $p_r$ operator defined as in Eq$^{\underline{n}}$ (43) of Chapter 7.

Then, choosing $\lambda = \ell$ in Eq$^{\underline{n}}$ (3) of Chapter 5, we get that

$$<\frac{\partial H}{\partial \ell}>_{n\ell} \; = \; \frac{\partial E_n}{\partial \ell} \; = \; \frac{\partial}{\partial \ell} \left(- \; \frac{Z^2\hbar^2}{2ma_0^2(n_r + \ell + 1)^2}\right) \; , \tag{114}$$

$$\Rightarrow \quad \left\langle \frac{\hbar^2 (2\ell + 1)}{2mr^2} \right\rangle_{n\ell} = \frac{Z^2 \hbar^2}{a_0^2 n^3} \, , \tag{115}$$

$$\Rightarrow \quad \left\langle \frac{1}{r^2} \right\rangle_{n\ell} = \frac{Z^2}{a_0^2} \frac{1}{\left(\ell + \frac{1}{2}\right) n^3} \, . \tag{116}$$

Putting the pieces together, we finally evaluate

$$< n\ell \,|\, H_{rel} \,|\, n\ell > = -\frac{1}{2mc^2} \left(\frac{Z^2 e^2}{a_0 n^2}\right)^2 \left(\frac{n}{\ell + \frac{1}{2}} - \frac{3}{4}\right)$$

$$= \frac{Z^2 E_n}{n^2} \, \alpha^2 \left(\frac{n}{\ell + \frac{1}{2}} - \frac{3}{4}\right). \tag{117}$$

Thus, we have from (117) (for n = 1, Z = 1) that

$$\frac{<H_{rel}>_{00}}{E_0} \sim \alpha^2 , \tag{118}$$

confirming (96) above.

We also have that

$$\vec{J}^2 = (\vec{L} + \vec{S})^2 = \vec{L}^2 + \vec{S}^2 + 2\vec{L} \cdot \vec{S} , \tag{119}$$

and therefore

$$\vec{L} \cdot \vec{S} = \frac{1}{2} (\vec{J}^2 - \vec{L}^2 - \vec{S}^2), \tag{120}$$

$$\Rightarrow \vec{L} \cdot \vec{S} \,|\ell, s; j, m> = \frac{\hbar^2}{2} \left(j(j + 1) - \ell(\ell + 1) - \frac{3}{4}\right)|\ell, s; j, m>. \tag{121}$$

This allows us to write

$$< n\ell s j m \,|\, H_{LS} \,|\, n\ell s j m > =$$

$$- \frac{e\hbar^2}{4m^2c^2} \left[ j(j + 1) - \ell(\ell + 1) - \frac{3}{4} \right] < n\ell | \frac{1}{r} \frac{d\phi}{dr} | n\ell >. \qquad (122)$$

Since $\phi(r) = \frac{Ze}{r}$ (Eq$^{\underline{n}}$ (99)) for the Coulomb potential, we need

to find $\left< \frac{1}{r^3} \right>_{n\ell}$ in (122).  There is an easy way of doing

this.  Again, from Eq$^{\underline{n}}$ (43) of Chapter 7

$$< r | p_r = \frac{\hbar}{i} \frac{1}{r} \frac{\partial}{\partial r} r < r |. \qquad (123)$$

Therefore, for any function $f(r)$ we have (I am not being

careful to distinguish the use of "r" as an operator or

eigenvalue here)

$$< r | p_r f(r) = \frac{\hbar}{i} \frac{1}{r} \frac{\partial}{\partial r} (r < r | f(r))$$

$$= \frac{\hbar}{i} \left( \frac{f(r)}{r} < r | + f(r) \frac{\partial}{\partial r} < r | + f'(r) < r | \right), \qquad (124)$$

and

$$< r | f(r) p_r = \frac{\hbar}{i} f(r) \left( \frac{1}{r} < r | + \frac{\partial}{\partial r} < r | \right), \qquad (125)$$

which implies that

$$< r | [p_r, f(r)] = \frac{\hbar}{i} \frac{\partial f(r)}{\partial r} < r |, \qquad (126)$$

$$\Rightarrow [p_r, f(r)] = \frac{\hbar}{i} \frac{\partial f(r)}{\partial r}. \qquad (127)$$

If we now take $f(r) = H \left( = \dfrac{p_r^2}{2m} + \dfrac{\hbar^2 \ell(\ell + 1)}{2mr^2} - \dfrac{Ze^2}{r} \right)$, and

evaluate the expectation value of the left hand side of

(127), we get

$$< n\ell \,|\, [p_r, H] \,|\, n\ell > \; = \; < n\ell \,|\, p_r H - H p_r \,|\, n\ell >$$

$$= \; (E_n - E_n) < n\ell \,|\, p_r \,|\, n\ell > \; = \; 0. \qquad (128)$$

Since

$$\left< \dfrac{\partial H}{\partial r} \right>_{n\ell} \; = \; - \dfrac{\hbar^2 \ell(\ell + 1)}{m} \left< \dfrac{1}{r^3} \right>_{n\ell} + Ze^2 \left< \dfrac{1}{r^2} \right>_{n\ell}, \qquad (129)$$

one has that

$$\left< \dfrac{1}{r^3} \right>_{n\ell} \; = \; \dfrac{Ze^2 n}{\hbar^2 \ell(\ell + 1)} \left< \dfrac{1}{r^2} \right>_{n\ell}, \qquad (130)$$

or using (116) above, that

$$\left< \dfrac{1}{r^3} \right>_{n\ell} \; = \; \left( \dfrac{Z}{na_0} \right)^3 \dfrac{1}{\ell \left( \ell + \dfrac{1}{2} \right)(\ell + 1)}. \qquad (131)$$

Notice that the expectation value in (131) diverges if we set

$\ell = 0$. It is easy to understand how this comes about. We

found in Ch.7 that near the origin, the radial function $R(r)$

behaved as $R(r) \sim r^{\ell+1}$. Therefore for the integral in (131)

we have when $\ell = 0$

$$\lim_{r \to 0} \dfrac{R^2(r)}{r^3} \; \sim \; \dfrac{1}{r}. \qquad (132)$$

The integral of $r^{-1}$ diverges logarithmically when the lower limit is r = 0. Therefore, the perturbative treatment given here breaks down for s-states.

We now have that

$$< n\ell sjm|H_{LS}|n\ell sjm > \ = \ - \ \frac{Z^2 E_n}{2n} \ \alpha^2 \ \frac{\left[ j(j \ + \ 1) \ - \ \ell(\ell \ + \ 1) \ - \ \frac{3}{4} \right]}{\ell\left(\ell \ + \ \frac{1}{2}\right)(\ell \ + \ 1)} \ , \quad (133)$$

confirming that the order of magnitude of the energy shift from $H_{LS}$ is the same as from $H_{rel}$. Since

$$\frac{\left[ j(j \ + \ 1) \ - \ \ell(\ell \ + \ 1) \ - \ \frac{3}{4} \right]}{\ell\left(\ell \ + \ \frac{1}{2}\right)(\ell \ + \ 1)} \ = \ \begin{cases} \dfrac{1}{\left(\ell \ + \ \frac{1}{2}\right)\left(j \ + \ \frac{1}{2}\right)} \ , \quad j \ = \ \ell \ + \ \frac{1}{2} \\[20pt] - \dfrac{1}{\left(\ell \ + \ \frac{1}{2}\right)\left(j \ + \ \frac{1}{2}\right)} \ , \quad j \ = \ \ell \ - \ \frac{1}{2} \end{cases}$$

$$(134)$$

we can combine (117) and (133) together as (top sign is for $j \ = \ \ell \ + \ \frac{1}{2}$ , bottom sign is for $j \ = \ \ell \ - \ \frac{1}{2}$ )

$$\Delta E_{nj} \ = \ \frac{Z^2 E_n}{n^2} \ \alpha^2 \ \left[ \frac{n}{\ell \ + \ \frac{1}{2}} \ \mp \ \frac{n}{2\left(\ell \ + \ \frac{1}{2}\right)\left(j \ + \ \frac{1}{2}\right)} \ - \ \frac{3}{4} \right]$$

$$= \ \frac{Z^2 E_n}{n^2} \ \alpha^2 \ \left[ \frac{n}{j \ + \ \frac{1}{2}} \ - \ \frac{3}{4} \right]. \quad\quad (135)$$

Eq$\underline{n}$ (135) is our final result for the fine structure splitting in a Coulomb field and holds for both cases,

$j = \ell \pm \frac{1}{2}$ .   These levels, which before were degenerate, are

now split by a small amount.  For the n = 2, $\ell$ = 0,1 levels

(the 2s,2p states), we have the following energy diagram:

**before**

**perturbation**

$2p(j=\frac{3}{2})$

$2s,2p(j=\frac{3}{2})$

_____

(8 states)

**after  perturbation**

$2p(j=\frac{3}{2})$ (4 states)

_____

$2s,2p$ $(j=\frac{1}{2})$ (4 states)

_____
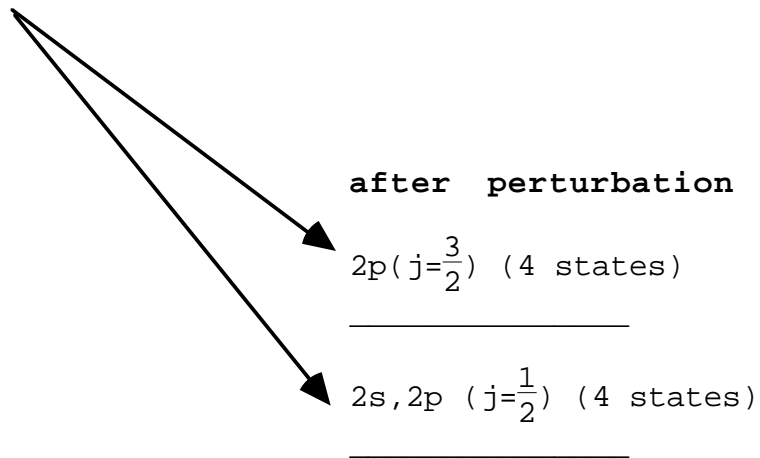
Actually, one can show that (135), even in the case of $\ell$ = 0,

is also a result of the Dirac theory, at least to order $\alpha^2$.

It was shown experimentally in the hydrogen atom that the

$2s\left(j = \frac{1}{2}\right)$ and $2p\left(j = \frac{1}{2}\right)$ levels are actually split by a small

amount.  The same person who discovered this effect, Willis

Lamb, was also the first person to publish the correct

quantum mechanical derivation of the effect, which is called

the Lamb shift. (Ask me to tell you about the famous footnote

number 13 in Feynman's Lamb shift paper.)

**Problems**

1. Show for $\vec{J} = \vec{J}_1 + \vec{J}_2$ ($[J_{1i},J_{2j}] = 0$) that (see Eqs.(21) of Ch. 8)

    (a) $[\vec{J}_1^{\,2}, \vec{J}^{\,2}] = 0$,

    (b) $[\vec{J}_1^{\,2}, J_z] = 0$.

2. Show that

$$\vec{J}_1 \cdot \vec{J}_2 = J_{1z}J_{2z} + \frac{1}{2}(J_{1+}J_{2-} + J_{1-}J_{2+})$$

where

$$J_{1\pm} \equiv J_{1x} \pm i\, J_{1y},$$

$$J_{2\pm} \equiv J_{2x} \pm i\, J_{2y}.$$

(Eq. (72) of Ch. 8 is a special case of this where $\vec{J}_1 = \vec{L}$ and $\vec{J}_2 = \vec{S}$.)

3. Using Eqs. (97), (98), and (100), show (101) holds. (All equations in Ch. 8.)

4. Given the Hamiltonian (see Eq. (104))

$$H = V_1(r) + V_2(r)\, \vec{L} \cdot \vec{S},$$

which of the operators $\vec{L}^{\,2}$, $\vec{S}^{\,2}$, $L_3$, $S_3$ fail to commute with H? What is the physical meaning of this? [Remember that $[\vec{L},V(r)] = 0$, as derived in Eq. (120) of Ch. 6.]

**Other Problems**

5. Suppose two $\ell = 1$ electrons (p electrons) in an atom are found in a $|1,1;2,1>$ state ($\ell_1 = \ell_2 = 1$; $\ell = 2$, $m = 1$).

   (a)  Assuming

$$|1,1;2,2> = |1,1>_1 \; |1,1>_2,$$

show that

$$|1,1;2,1> = \sqrt{\frac{1}{2}} \; |1,1>_1 \; |1,0>_2 + \sqrt{\frac{1}{2}} \; |1,0>_1 \; |1,1>_2.$$

   (b)  What is the probability that $\ell'_{1z} = \hbar$, $\ell'_{2z} = 0$?

   (c)  Evaluate explicitly the angular-dependent amplitude:
$$<\theta_1,\phi_1;\theta_2,\phi_2|1,1;2,1> \; = \; ?$$

6. Add $j_1 = \dfrac{1}{2}$ to $j_2 = \dfrac{3}{2}$.

   (a)  List the quantum numbers of the states in the uncoupled representation.

   (b)  List the quantum numbers of the states in the coupled representation.

   (c)  Write the coupled state $|\dfrac{3}{2},\dfrac{1}{2};2,-2>$ in terms of uncoupled states.

7. Imagine the electron has a small <u>electric dipole</u> moment $(\vec{p})$ in addition to its usual <u>magnetic dipole</u> moment $(\vec{\mu})$. If this electric dipole is proportional to spin, $\vec{p} = \gamma_p \; \vec{S}$, then we would expect the electron in a hydrogen atom to feel a small, additional perturbation given by

$$H' \; = \; -\vec{p} \cdot \vec{E} \; = \; \frac{-Ze}{r^3} \; \vec{p} \cdot \vec{r} \; = \; \frac{-Ze\gamma_p}{r^3} \; (\vec{S} \cdot \vec{r})$$

$$(\vec{r} \; = \; x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}})$$

8.36

Calculate:

(a)   $[S_z,H'] = ?$

(b)   $[L_z,H'] = ?$

(c)   $[J_z,H'] = ?$

**On the basis of your calculation, which of these operators gives rise to a conserved quantum number?**
[Note:   Actually, there are many reasons to believe $\vec{p} = 0$ exactly for the electron.]

8. Suppose two $\ell = 1$ electrons (p electrons) in an atom are found in a $|1,1;2,1\rangle$ state ($\ell_1 = \ell_2 = 1$; $\ell = 2$, $m = 1$).

(a)   Assuming (coupled state notation: $|\ell_1,\ell_2;\ell,m\rangle$)

$$|1,1;2,2\rangle = |1,1\rangle_1 |1,1\rangle_2,$$

show that

$$|1,1;2,1\rangle = \sqrt{\tfrac{1}{2}} |1,1\rangle_1 |1,0\rangle_2 + \sqrt{\tfrac{1}{2}} |1,0\rangle_1 |1,1\rangle_2.$$

(b)   What is the probability that $\ell'_{1z} = \hbar$, $\ell'_{2z} = 0$?

(c)   Evaluate explicitly the angular-dependent amplitude:

$$\langle\theta_1,\phi_1;\theta_2,\phi_2|1,1;2,1\rangle = ?$$

9. In terms of the principal quantum number "n", there is a $2n^2$ degeneracy of the $n^{\underline{th}}$ (unperturbed) energy level in the hydrogen atom. For n=2, this means there are 8 linearly independent states with the same energy. Specify the quantum numbers of these 8 states using,

(a) the "uncoupled basis": $|\ell,s;m_\ell,m_s\rangle$,

(b) the "coupled basis": $|\ell,s;j,m\rangle$.

10. Specify the quantum numbers of all the unperturbed states of the hydrogen atom which have n = 3 and ℓ = 1, using:

(a) the "uncoupled basis": $|\ell,s; m_\ell,m_s>$,

(b) the "coupled basis": $|\ell,s; j,m>$.

(c) An additional piece of the Hamiltonian is added to $H_0$ ($= \frac{\vec{p}^2}{2m} - \frac{Ze^2}{r}$) which has the form,

$$H' = \frac{eB_z}{2mc} (L_z + 2S_z),$$

which represents the interaction of the atomic electron with an external magnetic field, $B_z$, pointing along z. Find the effect of H' on the degenerate n = 3, ℓ = 1 energy levels. [You will have to make an appropriate choice of basis to do this calculation.]

11. (Like Prob. 9.32 of Liboff.) Using the relations

$$\vec{L}^2 = \vec{L}_1^2 + \vec{L}_2^2 + 2\vec{L}_1\cdot\vec{L}_2,$$
$$\vec{L}_1\cdot\vec{L}_2 = \frac{1}{2}(L_{1+}L_{2-} + L_{1-}L_{2+}) + L_{1z}L_{2z},$$

verify the ℓ,m values of the following coupled angular momentum eigenstate for two p electrons (the coupled state notation is the same as in problem 9 above, which is different from Liboff):

$$|2,0;1,1> = \sqrt{\tfrac{1}{6}}|1,1>_1|1,-1>_2 + \sqrt{\tfrac{2}{3}}|1,0>_1|1,0>_2 + \sqrt{\tfrac{1}{6}}|1,-1>_1|1,1>_2,$$

12. Prove that the coupled state $|j_1,j_2;j,m>$ corresponding to $m = j_1+j_2$ (it's maximum value) has $j=j_1+j_2$. Do this as follows.

8.38

   (a) First, argue that

$$\langle j_1, j_2; j, m | \vec{J}^2 - (\vec{J}_1 + \vec{J}_2)^2 | j_1, j_2; m_1, m_2 \rangle = 0.$$

   (b) By setting $m_1 = j_1$, $m_2 = j_2$, $m = j_1 + j_2$ above, now
show that this expression is equivalent to

$$[j(j+1) - (j_1+j_2)(j_1+j_2+1)]\langle j_1, j_2; j, j_1+j_2 | j_1, j_2; j_1, j_2 \rangle = 0,$$

and therefore conclude that $j = j_1 + j_2$.

## Chapter 9:   Spin and Statistics

There are two subjects mentioned previously involving apparently unrelated phenomena that I would like to bring back to our attention.  One is the "zweideutigkeit" of Chapter 7.  We encountered the rule "at most two neutrons and two protons in each energy level" in connection with the simple model of the nucleus presented there.  We also encountered the rule "at most two electrons to each energy level" in the atomic model presented.  The other subject was brought up at the very beginning of this course in regard to experimental indications of a need for a new type of mechanics to replace Newtonian mechanics for microscopic systems.  We had defined the molar specific heat at constant volume by

$$C_V = \left( \frac{\partial \overline{E}}{\partial T} \right)_V \tag{1}$$

where $\overline{E}$ was the average internal energy per mole and T is temperature.  We saw that the universal prediction of Dulong-Petit,

$$C_V = 3R \tag{2}$$

(R = kNa) did not hold for all materials, especially diamond. The law did seem to hold true for Copper and Silver, at least near room temperature.  However, I pointed out that there is still a paradox associated with these materials.  If each atom of Copper or Silver gave up one or more valence

electrons, we would expect there to be an electronic component to the specific heat.  This is not observed at room temperatures.

The origin of these two mysteries can be explained, as we will see, by a deep connection found in nature between a particle's spin and the type of statistics obeyed by a collection of identical particles.  Specifically, it has been established that:

**\*Systems of identical particles with zero or positive integer spin have symmetrical wave functions and are said to obey Bose-Einstein statistics.  Such particles are called <u>bosons</u>.**

**\*Systems of identical particles with half-integer spin have anti-symmetrical wavefunctions and are said to obey Fermi-Dirac statistics. Such particles are called <u>fermions</u>.**

This connection between the spin of a particle (which might be composite) and the wavefunction of the system is a cornerstone of relativistic quantum mechanics.  A consistent relativistic description in fact <u>requires</u> such a connection. The connection between statistics and spin was first formulated by Wolfgang Pauli.

What is meant by the above statements regarding symmetrical and anti-symmetrical wavefunctions?  Let us consider several simple examples.  First, consider a system

of two identical particles which are, however, in distinct

quantum states.   These are individually described by the

single-particle states:

$$\text{Particle 1:} \quad |a'>_1,$$
$$\text{Particle 2:} \quad |a''>_2.$$

The labels a' and a" (a' ≠ a") stand for some quantum

number or set of numbers.   These could include, for example,

$\vec{J}^2$, $J_3$, energy, etc.   We have temporarily labeled the kets

associated with the first or second particle with a

subscript.    This composite state will be denoted by

$$|a'>_1 \ |a''>_2 \ \equiv \ |a',a''>. \tag{3}$$

However, if the particles are indistinguishable, it is not

possible to know which particle is in a given state.

Therefore, another possible state of the system is specified

by:

$$\text{Particle 1:} \quad |a''>_1,$$
$$\text{Particle 2:} \quad |a'>_2.$$

The composite state is

$$|a''>_1 \ |a'>_2 \ \equiv \ |a'',a'>. \tag{4}$$

The true physical composite state of the system, which can be

neither (3) or (4) since they distinguish between the

particles, must somehow be a mixture of these two

possibilities.   Quantum mechanics says that the physically

realizable states of such a system depends upon the spin of

the particles involved.  If we are dealing with two identical

bosons, the true composite state would be symmetric under the

interchange of particle labels:

$$\frac{1}{\sqrt{2}} \left( |a',a''> + |a'',a'> \right) .$$

If we are dealing with two identical fermions, the state

would be anti-symmetric under label interchange:

$$\frac{1}{\sqrt{2}} \left( |a',a''> - |a'',a'> \right) .$$

The factors of $\frac{1}{\sqrt{2}}$ above are included for normalization.

There is also, of course, an arbitrary overall phase

involved.

     Now let us consider the case of three identical

particles, two of which are in a quantum state a', the other

being in a state described by a".  The possible states of the

system are:

$$|a',a',a''>, \ |a',a'',a'>, \ |a'',a',a'> .$$

The bosonic composite state associated with this example is

just

$$|2,1>_b^{a'a''} \equiv \frac{1}{\sqrt{3}} \left( |a',a',a''> + |a',a'',a'> + |a'',a',a'> \right) .$$

This state is symmetric under the interchange of any two

particle labels. When we try to construct an antisymmetrical

combination for this example, we discover that it can not be

done. This is in fact what will happen any time more than one

particle is in a given state. One can not build a completely
antisymmetric state (under the interchange of any two
particle labels) from composite states which themselves are
already partly symmetric. Physically, for a system of
identical fermions, this means at most one fermion can occupy
a given state of the system. This simple fact has enormous
consequences in nature, and is intimately related to the
"zweideutigkeit" and suppressed electronic component of
specific heat phenomena discussed above.

   As a last example, consider a state of three particles
in three different quantum states given by a', a" and a"'.
The possible distinct combinations are six in number:

$$|a',a",a"'>, \quad |a",a',a"'>, \quad |a',a"',a">,$$
$$|a",a"',a'>, \quad |a"',a',a">, \quad |a"',a",a'>.$$

The completely symmetric and antisymmetric combinations
appropriate to a system of bosons or fermions, respectively,
are given by: (overall phases are not important)

$$|1,1,1 >_b^{a'a"a"'} \equiv \frac{1}{\sqrt{6}} (|a',a",a"'> + |a",a',a"'> + |a',a"',a">)$$

$$+ (|a",a"',a'> + |a"',a',a"> + |a"',a",a'>),$$

and

$$|1,1,1 >_f^{a'a"a"'} \equiv \frac{1}{\sqrt{6}} (|a',a",a"'> - |a",a',a"'> - |a',a"',a">)$$

$$+ (|a",a"',a'> + |a"',a',a"> - |a"',a",a'>) .$$

One can recover the results of the previous example of a symmetric combination of three identical particles in two quantum states a' and a" by setting a"' = a' in the first of these.  However, when we try to set two quantum states equal in the anti-symmetric combination, we get a complete cancellation of terms, telling us that in fact no such state exists.  The number of possible states of "n" particles in "g" separate single-particle quantum states is given by the expressions

$$\frac{(g + n - 1)!}{n! \ (g-1)!} \text{ for bosons}$$

and

$$\frac{(g)!}{n! \ (g-n)!} \text{ for fermions.}$$

Can you find a way of motivating these expressions? (Notice that the fermion expression makes no sense for n > g.)

Notice in the above examples we are normalizing these states consistently.  The normalization factors are just given by the inverse of the square root of the number of distinct terms in the sums.  If we let $n_i$ = number of particles in quantum state i, and we let $n = \sum_i n_i$ be the total number of particles in the system, then clearly we have that

$$\text{\# terms in F-D state} = n!, \tag{5}$$

and the normalization factor for a system of n fermions will be $\frac{1}{\sqrt{n!}}$ .  For bosons we can show

$$\text{\# terms in B-E state} = \frac{n!}{\prod_i n_i!} \ . \qquad (6)$$

Notice that

$$\frac{n!}{\prod_i n_i!} \leq n! \ . \qquad (7)$$

The only time the equality holds is when $n_i = 1$ for all i.
Eq$^{\underline{n}}$ (6) is a consequence of the grouping that occurs when $n_i$
originally distinct particles are forced to be in a symmetric
state.  For example, the symmetric state of three particles
in distinct states involves 3! = 6 terms, as above.  When two
of the particles occupy the same state, however, we get an
expression with $\frac{3!}{2!}$ = 3 terms, which is just specified by (6)
with n = 3, $n_1$ = 2, $n_2$ = 1.  (n = $n_1$ + $n_2$).  Generalizing
these results, one can show that the symmetric state of n
identical bosons, with quantum state occupation numbers $n_i$,
is given by the somewhat sketchy expression,

$$\left(\frac{n!}{\prod_i n_i!}\right)^{-1/2} \sum_{\substack{\text{distinct} \\ \text{permutations}}} |\overbrace{a_1, a_1}^{n_1}, \dots, \overbrace{a_2, a_2}^{n_2}, \dots, \dots >,$$

whereas for fermions it is given by the completely
antisymmetric expression,

$$\frac{1}{\sqrt{n!}} \ \det \begin{pmatrix} |a_1>_1 & |a_1>_2 & \cdots & a_1>_n \\ |a_2>_1 & |a_2>_2 & \cdots & a_2>_n \\ \vdots & \vdots & & \vdots \end{pmatrix} \ .$$

(Technically speaking, these expressions assume that the number of allowed states of the composite system is finite.)

We may view these general expressions for boson or fermion wavefunctions as defining a new set of basis states. Instead of talking about the physical attributes of particles labeled as 1,2,3,...,n, one now talks about having a system of $n_1$ particles of type 1, $n_2$ particles of type 2, etc., with $\sum_i n_i = n_1$. That is, instead of specifying the properties of numbered particles, we count the number of particles with a specified property. We will call this new basis a _particle occupation basis_. It will be denoted as:

$$|n_1,n_2,n_3,\ldots >.$$

It will be assumed to be complete in the usual sense,

$$\sum_{n_1,n_2,\ldots} |n_1,n_2,\ldots > <n_1,n_2,\ldots| = 1,$$

and is built up out of the single particle basis states as described before. As this form suggests, the eigenvalues of this basis are not particle _properties_ but particle _numbers_.

We define an operator, $a_i^+$, such that for any state $|\ldots n_i \ldots >$ (the symbol "~" means "proportional to"),

$$a_i^+ |\ldots n_i \ldots > \ \sim\ |\ldots n_i + 1 \ldots >. \qquad (8)$$

Taking the Hermitian conjugate of (8) and multiplying on the right by a general state $|\ldots n_i' \ldots >$, we see that the only non-zero result happens when $n_i' = n_i + 1$. This means that

$$a_i |\ldots n_i + 1 \ldots > \ \sim\ |\ldots n_i \ldots > \ . \qquad (9)$$

We supplement this with the additional definition

$$a_i \ |\ldots 0_i \ldots > = 0, \qquad (10)$$

where $0_i$ labels the _zero-particle_ or _vacuum state_ for the particle of type i.

Since the effect of the operator $a_i^+$ is to increase the particle occupation number of state i by one, it is called the _creation_ operator. Similarly, since from (9) we see that $a_i$ reduces the occupation number of state i by one, it is called the _annihilation_ operator. These statements are made in the sense of $a_i^+$ and $a_i$ acting on kets. When acting on bras, we have that

$$< \ldots n_i \ldots | a_i \ \sim\ < \ldots n_i + 1 \ldots | \ , \qquad (11)$$

$$< \ldots n_i + 1 \ldots | a_i^+ \ \sim\ < \ldots n_i \ldots | \ , \qquad (12)$$

and

$$< \ldots 0_i \ldots | a_i^+ \ = 0, \qquad (13)$$

which follow by taking the Hermitian conjugate of (8), (9) and (10).

Now from (8) and (9) we know that

$$a_i^+ a_i |\ldots n_i \ldots > \; \sim \; |\ldots n_i \ldots >. \tag{14}$$

For the state $n_i = 0$ we have

$$a_i^+ a_i |\ldots 0_i \ldots > \; = 0, \tag{15}$$

which follows from (10). Based upon (14) and (15), we define the effect of the <u>number operator</u>,

$$N_i = a_i^+ a_i, \tag{16}$$

on the occupation basis to be

$$N_i |\ldots n_i \ldots > \; = n_i |\ldots n_i \ldots >. \tag{17}$$

There are as many such operators as there are states in the system. The total number operators is given by

$$N = \sum_i N_i . \tag{18}$$

Its eigenvalue is n, the total number of particles in the system.

Let us derive some commutation properties of these quantities. Since $a_i$ is an annihilation operator, we have that

$$N \, a_i |\ldots n_i \ldots > \; = (n - 1)a_i |\ldots n_i \ldots >$$
$$= a_i (N - 1) |\ldots n_i \ldots >. \tag{19}$$

The statement (19) being true for any state $|\ldots n_i \ldots >$ then implies

$$[a_i, N] = a_i. \tag{20}$$

Since we have that

$$N^+ = \left( \sum_i a_i^+ a_i \right)^+ = N,\qquad(21)$$

the adjoint of (20) is

$$[a_i^+, N] = -a_i^+ .\qquad(22)$$

Now consider

$$N_j a_i |\ldots n_i \ldots n_j \ldots > = \left\{ \begin{array}{c} n_j \\ (n_j - 1) \end{array} \right\} a_i |\ldots n_i \ldots n_j \ldots >,\qquad(23)$$

the top result holding for $i \neq j$, the bottom result for $i = j$.  One can write both results at once by

$$N_j a_i |\ldots n_i \ldots n_j \ldots > = (n_j - \delta_{ij}) a_i |\ldots n_i \ldots n_j \ldots >$$
$$= a_i (N_j - \delta_{ij}) |\ldots n_i \ldots n_j \ldots >.\qquad(24)$$

Eq$^{\underline{n}}$ (24) holding true for all occupation kets implies

$$[a_i, N_j] = a_i \delta_{ij}.\qquad(25)$$

This is consistent with (20) because

$$[a_i, N] = \sum_j [a_i, N_j] = \sum_j \delta_{ij} a_i = a_i.\qquad(26)$$

Taking the adjoint of (25) gives

$$[a_i^+, N_j] = -a_i^+ \delta_{ij} .\qquad(27)$$

Another test of consistency of these relations is to check and see if $N_i$ and $N_j$, which are assumed to have simultaneous eigenvalues, commute:

$$[N_i, N_j] = [a_i^+ a_i, N_j] = a_i^+ [a_i, N_j] + [a_i^+, N_j] a_i$$

$$= a_i^+ a_i \delta_{ij} - a_i^+ a_i \delta_{ij} = 0. \tag{28}$$

We now require that

$$a_i^+ a_j^+ | \ldots n_i \ldots n_j \ldots > = \lambda_{ij} a_j^+ a_i^+ | \ldots n_i \ldots n_j \ldots >, \tag{29}$$

that is, that the states produced by $a_i^+ a_j^+$, or alternatively by $a_j^+ a_i^+$, be the same except for an overall normalization factor, $\lambda_{ij}$. We assume this constant is independent of the occupation numbers $n_i, n_j$ and is symmetric in i and j. (Such an assumption need not be made. See the treatment in Merzbacher, <u>Quantum</u> <u>Mechanics</u>, 2<u>nd</u> edition, Ch. 20.) Then (29) implies

$$a_i^+ a_j^+ = \lambda_{ij} a_j^+ a_i^+ . \tag{30}$$

Since (30) is true for all i,j, it leads to

$$(\lambda_{ij})^2 = 1. \tag{31}$$

This means we have

$$[a_i, a_j] = [a_i^+, a_j^+] = 0, \tag{32}$$

when $\lambda_{ij} = 1$ and

$$\{a_i, a_j\} = \{a_i^+, a_j^+\} = 0, \tag{33}$$

when $\lambda_{ij} = -1$. In (33) we are encountering the anticommutator:

$$\{A, B\} = AB + BA. \tag{34}$$

Now from (25) we may write

$$a_i(a_j^+ a_j) - (a_j^+ a_j)a_i = \delta_{ij}a_i, \qquad (35)$$

which becomes, with the use of (32) and (33)

$$[a_i, a_j^+]a_j = \delta_{ij}a_i , \qquad (36)$$

when $\lambda_{ij} = 1$ and

$$\{a_i, a_j^+\}a_j = \delta_{ij}a_i , \qquad (37)$$

when $\lambda_{ij} = -1$. Assuming that the action of the operators $[a_i, a_j^+]$ and $\{a_i, a_j^+\}$, when acting on states, doesn't depend on the occupation numbers of the states in the cases $\lambda_{ij} = 1$ and $\lambda_{ij} = -1$, respectively (this assumption may also be avoided; see Merzbacher, op. cit.), we conclude that

$$[a_i, a_j^+] = \delta_{ij} , \qquad (38)$$

when $\lambda_{ij} = 1$ and

$$\{a_i, a_j^+\} = \delta_{ij} , \qquad (39)$$

when $\lambda_{ij} = -1$.

Notice that the commutation properties of $a_i$ and $a_i^+$ in (38) and the eigenvalue-eigenvector statement for $N_i = a_i^+ a_i$ in (17) are mathematically equivalent to the ladder operators $A$ and $A^+$ that we defined for the harmonic oscillator in Chapter 3. For those operators we had

$$[A, A^+] = 1, \qquad (40)$$

$$A^+ A|n> = n|n>. \qquad (41)$$

The difference is in the physical interpretation of these statements. Eq$^{\underline{n}}$s (40) and (41) are results for the <u>energy levels</u> of a single particle. Eq$^{\underline{n}}$s (17) and (38) refer to the <u>occupation numbers</u> of a single physical state. However, because these mathematical systems are identical, we may take over the results previously derived. In particular, we showed in Chapter 3 that

$$A|n> = \sqrt{n} \ |n - 1>, \tag{42}$$

$$A^+|n> = \sqrt{n + 1} \ |n + 1>, \tag{43}$$

and

$$|n> = \frac{(A^+)^n}{\sqrt{n!}} \ |0>, \tag{44}$$

where the state $|0>$ represents the ground state. Similarly, we have that

$$a_i| \ldots n_i \ldots > = \sqrt{n_i} \ | \ldots n_i - 1 \ldots >, \tag{45}$$

$$a_i^+| \ldots n_i \ldots > = \sqrt{n_i + 1} \ | \ldots n_i + 1 \ldots >, \tag{46}$$

and

$$| \ldots n_i \ldots > = \frac{(a_i^+)^{n_i}}{\sqrt{n_i!}} \ | \ldots 0_i \ldots >, \tag{47}$$

for each physical property labeled by i. We identify this situation ($\lambda_{ij} = 1$) as the case of Bose-Einstein statistics. The most familiar particle to which these considerations apply is the photon, but the creation and annihilation

operators of any zero or integer spin particle must obey the same relations.

We have not encountered a system before for which anticommutation relations like (33) and (39) hold.  Notice that if i = j, (33) implies that

$$(a_i^+)^2 = 0. \tag{48}$$

Thus, any attempt to put two or more particles in the same quantum state fails in the case $\lambda_{ij} = -1$.  We immediately recognize this situation as describing Fermi-Dirac statistics. We expect for this case that all the occupation numbers, $n_i$, of the state $|...n_i...>$ can only take on values 0 or 1.  This is confirmed from the algebra since from (39) for i = j

$$a_i a_i^+ + a_i^+ a_i = 1, \tag{49}$$

$$\Rightarrow \quad a_i a_i^+ = 1 - N_i. \tag{50}$$

Now, by multiplying on the left by $N_i = a_i^+ a_i$, we learn that

$$N_i(1 - N_i) = 0. \tag{51}$$

Applying this null operator on any state $|...n_i...>$ then gives us that

$$n_i = \{0,1\} \tag{52}$$

for all i.

In order to keep things simple, we will restrict our attention to one and two physical property systems in the

Fermi-Dirac case.  We hypothesize that for the simplest case
of a single physical property that

$$a^+|0> = C_1|1>, \qquad (53)$$

$$\Rightarrow \quad <0|a = C_1^*<1|, \qquad (54)$$

and that

$$a|1> = C_2|0>,$$

$$\Rightarrow \quad <1|a^+ = C_2^*<0|. \qquad (55)$$

From the number operator relation

$$a^+a|1> = |1> \qquad (56)$$

we then have that

$$C_2C_1 = 1. \qquad (57)$$

Multiplying (56) on the left by $<1|$ also gives us

$$|C_2|^2 = 1. \qquad (58)$$

Both of these relations are satisfied if we choose

$$C_1 = C_2 = 1, \qquad (59)$$

resulting in

$$a^+|0> = |1>, \qquad (60)$$

$$a|1> = |0>. \qquad (61)$$

These statements imply that

$$a = |0><1| \qquad (62)$$

$$\Rightarrow \quad a^+ = |1><0|. \qquad (63)$$

In our old language of measurement symbols, we would say that $a = |01|$ and $a^+ = |10|$. However, unlike our previous applications, we are not specifying physical properties but occupation numbers in the states. Thus, this simplest fermion system is just another manifestation of the two-physical-outcome formalism of Chapter 1. Using (62) and (63) in

$$aa^+ + a^+a = 1, \tag{64}$$

then reveals this as just the statement of completeness for this two level system:

$$aa^+ + a^+a = \sum_{n=\{0,1\}} |n><n|. \tag{65}$$

Raising the complexity a notch, we now consider a two physical property identical fermion system. Following the above, we may choose

$$a_1^+|0,0> = |1,0>, \tag{66}$$

$$a_2^+|0,0> = |0,1>, \tag{67}$$

and

$$a_1|1,0> = |0,0>, \tag{68}$$

$$a_2|0,1> = |0,0>. \tag{69}$$

Notice that

$$a_1^+a_2^+ |0,0> = -a_2^+a_1^+ |0,0>, \tag{70}$$

because of (33). There are now two possible definitions of the state $|1,1>$:

$$|1,1> = \pm a_2^+ a_1^+ |0,0>. \qquad (71)$$

We see that the <u>order</u> in which the states are raised or lowered is of crucial importance in this case. In general, one may choose for identical fermions that

$$a_i |\ldots n_i \ldots > = \begin{cases} 0, & n_i = 0 \\ \pm |\ldots 0_i \ldots >, & n_i = 1 \end{cases} \qquad (72)$$

and therefore

$$a_i^+ |\ldots n_i \ldots > = \begin{cases} \pm |\ldots 1_i \ldots >, & n_i = 0 \\ 0, & n_i = 1. \end{cases} \qquad (73)$$

(Eq$^{\underline{n}}$s (72) and (73) preserve the statement $N_i |\ldots n_i \ldots > = n_i |\ldots n_i \ldots >$ for $n_i$ = 0 or 1.) Rather than specifying an arbitrary (but eventually necessary) convention to fix the signs in (72) and (73), we will find it possible to live peaceably with this ambiguity in the present study. (See the homework problem, however.)

Fermi-Dirac statistics explain the "zweideutigkeit" of atomic and nuclear physics. In the early days of atomic spectroscopy, physicists were unaware of electron spin, and it appeared as if an unexplained rule "two electrons per energy level" was operating. Actually, since electrons as spin $\frac{1}{2}$ particles obey Fermi-Dirac statistics, atomic systems are really built up with one electron per <u>state</u>, but the two orientations of electron spin are very nearly degenerate in energy. The fact that electrons, neutrons and protons are

all spin $\frac{1}{2}$ particles determines to a large extent the nature

of our universe.  If these particles had integer spin, they

would obey Bose-Einstein statistics and there would be no

restriction on the number of particles in a given state.  The

ground states of all such systems would closely resemble one

another, removing the incredible diversity seen in atomic and

nuclear systems.

　　　Other useful quantities can be built out of the

occupation basis creation and annihilation operators.  The

total particle number operator

$$N = \sum_{i = \text{all states}} a_i^+ a_i, \tag{74}$$

can be generalized to the form

$$F^{(1)} = \sum_{i,j} a_i^+ a_j f_{ij}, \tag{75}$$

to represent various additive single-particle properties.  We

interpret the $f_{ij}$ as being matrix elements of some single-

particle property operator, f:

$$f_{ij} = <i|f|j>. \tag{76}$$

For example, if we set f = H where H is the single-particle

Hamiltonian, then

$$f_{ij} = <i|H|j> = \varepsilon_i \delta_{ij}, \tag{77}$$

and $F^{(1)}$ will represent the total energy operator of the

system:

$$F^{(1)} | \ldots n_k \ldots > \; = \; \left( \sum_i \varepsilon_i n_i \right) | \ldots n_k \ldots >. \tag{78}$$

Another useful operator in this basis that can be used to represent an additive two-particle property (like potential energy) is of the form

$$F^{(2)} \; = \; \frac{1}{2} \sum_{i,j,k,l} a_i^+ a_j^+ a_k a_l \; V_{ijlk}, \tag{79}$$

where, again, the $V_{ijlk}$ are a set of matrix elements. Notice the somewhat unusual ordering of the indices on $V_{ijlk}$ relative to the order of the creation and annihilation operators, as well as the conventional factor of $\frac{1}{2}$ which is included to avoid double counting because of a symmetry of the $V_{ijlk}$ (upcoming in Eq$^{\underline{n}}$ (84)). We choose the $V_{ijlk}$ as matrix elements of some two-particle operator, $V$,

$$V_{ijlk} \equiv < i,j | V | l,k >, \tag{80}$$

where as usual

$$\left. \begin{array}{l} | i,j > \; = \; | i >_1 \otimes | j >_2, \\ < i,j | \; = \; < i |_1 \otimes < j |_2 \,. \end{array} \right\} \tag{81}$$

We have used a two-particle basis previously in describing the deuteron in Chapter 7. A simple example of a two-body operator would be $V = V(r)$, where $r$ is the relative distance operator between the two particles. (A more realistic but also more complicated potential in that case would also include the previously mentioned tensor force, which depends on the relative spin orientations of the two particles.) A

diagrammatic way of visualizing the two-particle matrix element $V_{ijlk}$ is contained in the following picture:

$< i,j|V|l,k >:$
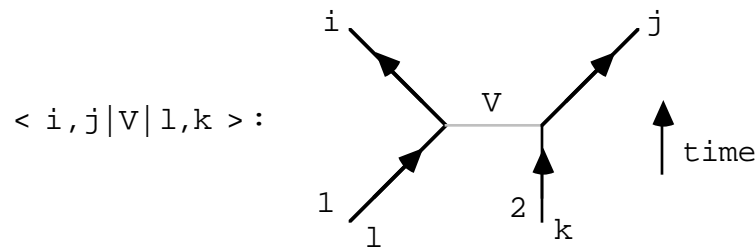


The particle to the left has been labeled as particle 1 and the dotted line in the diagram drawn perpendicularly to the time axis, represents the interaction of the two particles through the two body potential, V. The time order of events associated with the matrix element $< i,j|V|l,k >$ are being read off from right to left in this interpretation. (The same implicit time ordering of events occurs in our earlier right to left interpretation of measurement symbols in Chapter 1. However, the time direction in the associated Process Diagrams were also right to left, as opposed to the bottom to top time direction choice in the above.) This picture of the interaction of two particles is called a Feynman diagram and can be used to represent a weak scattering event between two particles. The above is actually a nonrelativistic interpretation of such a diagram. In the usual relativistic diagram, the interaction, V, would not necessarily take place at a single instant in time, but would be summed over all time intervals.

As mentioned previously, these matrix elements satisfy certain identities. First of all, from the meaning of Hermitian conjugation, we have

$$< i,j|V|l,k >^* \ = \ < l,k|V^+|i,j >. \tag{82}$$

If $V = V^+$ then

$$< i,j|V|l,k >^* \ = \ <l,k|V|i,j >. \tag{83}$$

Another identity follows from the fact that a relabeling of identical particles (which we assume these are) does not affect the value of the matrix element $V_{ijlk}$. Thus, under the substitution $1 \leftrightarrow 2$ we have

$$< i,j|V|l,k > \ = \ < j,i|V|k,l >. \tag{84}$$

A diagrammatic interpretation of (83) is (only the <u>directions</u> of the lines have any significance here)
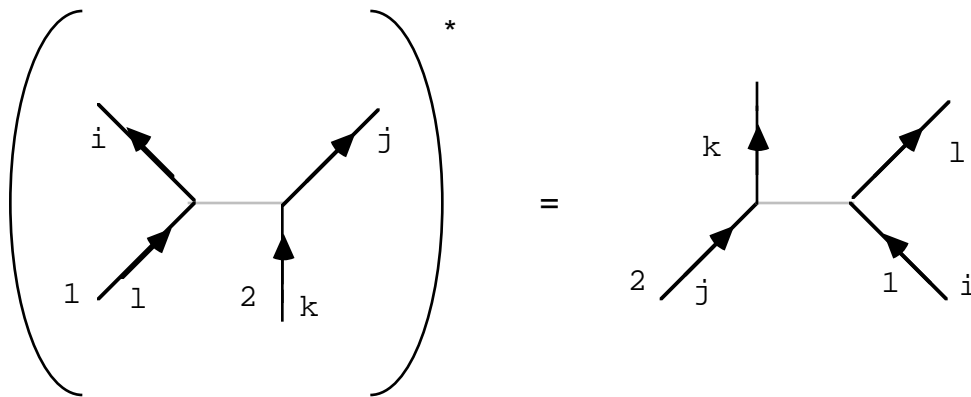


and the statement (84) can be visualized as
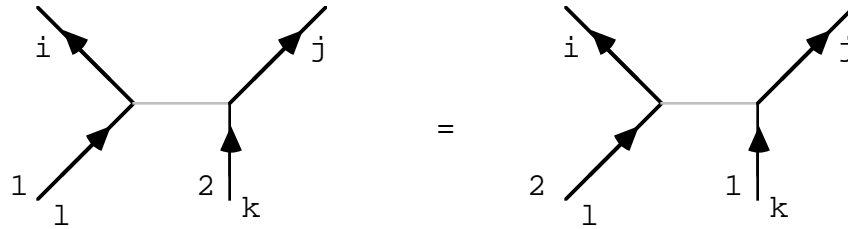
We are imagining that the i,j,k,l include the momentum state of the particle which, if we think of these particles as being contained in a box, will be discrete.

We are getting closer to an application of these ideas. Consider a gas consisting of many particles. The characteristics of such systems depends upon the statistics of the particles involved. We will study such systems under the assumption of <u>detailed</u> <u>balance</u>. To explain what this means, consider an interaction between two distinguishable particles: (1, 1', 2 and 2' are particular distinct values of the state variables i, j, k and $\ell$)
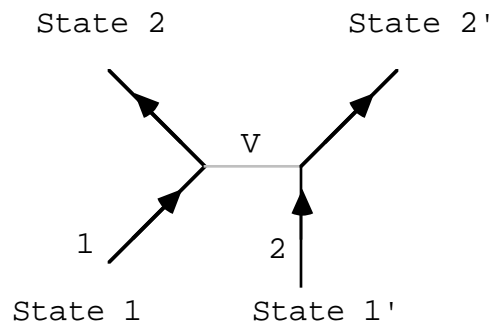


In the present (classical) context, the rate for this reaction to be taking place somewhere in our gas will be proportional to the product of the number of particles of types 1 and 1' present,

$$\text{rate}(1 + 1' \rightarrow 2 + 2') \sim n_1 n_{1'} . \tag{85}$$

The rate for the reversed collision,
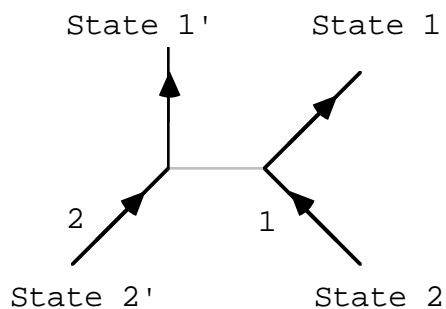


is also given by

$$\text{rate}(2 + 2' \rightarrow 1 + 1') \sim n_2, n_{2'} . \qquad (86)$$

Detailed balance implies that

$$\text{rate}(1 + 1' \rightarrow 2 + 2') = \text{rate}(2 + 2' \rightarrow 1 + 1'). \qquad (87)$$

That is, in equilibrium the rates of a reaction and it's inverse must be equal. Assuming the same proportionality constants in (85) and (86), this means that

$$n_1 n_{1'} = n_2 n_{2'} , \qquad (88)$$

which also implies

$$\ln \frac{1}{n_1} + \ln \frac{1}{n_{1'}} = \ln \frac{1}{n_2} + \ln \frac{1}{n_{2'}} . \qquad (89)$$

Eq$^{\text{n}}$ (89) has the appearance of a conservation law for some scalar quantity. It is plausible that (89) is an expression of energy conservation for this situation. Then, comparing (89) with such a statement,

$$\varepsilon_1 + \varepsilon_{1'} = \varepsilon_2 + \varepsilon_{2'} \qquad (90)$$

we tentatively conclude that for any species i

$$\ell n \, \frac{1}{n_i} = \alpha + \beta \varepsilon_i \, , \tag{91}$$

where $\alpha$ and $\beta$ are unknown constants, or that

$$n_i = e^{-\alpha - \beta \varepsilon_i} \, . \tag{92}$$

Eq$^n$ (92) becomes an expression of the classical statistics of distinguishable particles, Maxwell-Boltzmann statistics (see Chapter 1) if we take

$$\beta = \frac{1}{kT} \tag{93}$$

where k is the Boltzmann constant and T is absolute temperature. The quantity $\alpha$ in (92), called the <u>chemical potential</u>, can be thought of as a scale factor fixed by the number of particles, n, in the system:

$$\sum_i n_i = \sum_i e^{-\alpha - \beta \varepsilon_i} = n. \tag{94}$$

Although the above argument was very sketchy, it has led to correct conclusions about the type of statistics obeyed by distinguishable particles. Let us see if we can repeat it for the two quantum mechanical cases of indistinguishable particles.

Using V in (79) as an expression of the two-body potential between identical particles, we relabel

$$F^{(2)} \rightarrow H_{int} = \frac{1}{2} \sum_{i,j,k,\ell} a_i^+ a_j^+ a_k a_\ell \, <i,j|V|\ell,k>. \tag{95}$$

The key to this discussion is the relation

$$\text{rate}(1 \rightarrow 2) \sim |<2|H_{int}|1>|^2. \tag{96}$$

which is a partial statement of what is called "Fermi's golden rule."  (This rule comes about from time-dependent perturbation theory, which we have not studied.) The states $<2|$ and $|1>$ above are the appropriate particle occupation basis states for the reaction in question. In order to find the effect of $H_{int}$ on the particle occupation states, let us go back to our previous expressions for the effects of creation and annihilation operators.  From (45) and (46)

$$a_2^+ a_1 |n_1,n_2> = \sqrt{n_2 + 1} \sqrt{n_1} \ |n_1 - 1, n_2 + 1> \tag{97}$$

for the Bose-Einstein case.  We may make a similar statement for Fermi-Dirac statistics.  We may write both statements in (72) as

$$a_i |\ldots n_i \ldots > = \pm \sqrt{n_i} \ |\ldots n_i - 1 \ldots > \tag{98}$$

where $n_i = 0,1$ only, and we see the usual Fermi-Dirac sign ambiguity. In addition, (73) may be written

$$a_i^+ |\ldots n_i \ldots > = \pm \sqrt{1 - n_i} \ |\ldots n_i + 1 \ldots >. \tag{99}$$

Putting (98) and (99) together we then get

$$a_2^+ a_1 |n_1,n_2> = \pm \sqrt{1 - n_2} \sqrt{n_1} \ |n_1 - 1, n_2 + 1>, \tag{100}$$

similar to (97) for Bose-Einstein particles.  We write (97) and (100) together as

$$a_2^+ a_1 |n_1,n_2> = (S.F.) \sqrt{1 \pm n_2} \sqrt{n_1} \; |n_1 - 1, n_2 + 1>, \qquad (101)$$

where the top sign is taken for Bosons and the bottom sign for fermions. The quantity "(S.F.)" (meaning "sign factor") is only to be considered for fermions. Eq$^{\underline{n}}$ (101) implies that

$$a_2^+ a_{2'}^+ a_1 a_{1'} |n_1, n_{1'}, n_2, n_{2'}> = (S.F) \sqrt{1 \pm n_2} \sqrt{1 \pm n_{2'}}$$
$$\times \sqrt{n_1} \sqrt{n_{1'}} \; |n_1 - 1, n_{1'} - 1, n_2 + 1, n_{2'} + 1>. \qquad (102)$$

where the $\pm$ signs are to be interpreted as above.

We are now in a position to partially determine the rate for the reaction $1 + 1' \rightarrow 2 + 2'$:

$$\text{rate}(1 + 1' \rightarrow 2 + 2') \sim$$
$$|<n_1-1, n_{1'}-1, n_2+1, n_{2'}+1 | H_{int} | n_1, n_{1'}, n_2, n_{2'} >|^2. \qquad (103)$$

We are assuming in general that there are $n_1, n_{1'}, n_2, n_{2'}$ particles in the distinct modes $1, 1', 2, 2'$, respectively, in the initial state. Letting the labels i, j, k and l in (95) take on all possible values, we have that

$$<n_1-1, n_{1'}-1, n_2+1, n_{2'}+1 | H_{int} | n_1, n_{1'}, n_2, n_{2'} >$$

$$= \frac{1}{2} <n_1-1, n_{1'}-1, n_2+1, n_{2'}+1 | \Big\{ a_2^+ a_{2'}^+ a_1 a_{1'} \; <2,2'|V|1',1>$$
$$+ a_2^+ a_{2'}^+ a_1 a_1 \; <2,2'|V|1,1'> + a_2^+ a_{2'}^+ a_1 a_{1'} \; <2',2|V|1',1>$$
$$+ a_2^+ a_{2'}^+ a_1 a_1 \; <2',2|V|1,1'> \Big\} | n_1, n_{1'}, n_2, n_{2'} >. \qquad (104)$$

However, from (84) we have that

$$<2,2'|V|1,1'> = <2',2|V|1',1> \qquad (105)$$

and

$$< 2,2'|V|1',1 > = < 2',2|V|1,1' >. \qquad (106)$$

Also, because we are assuming the modes 1, 1', 2, 2' are all distinct, we easily see that

$$a_2^+ a_{2'}^+ a_{1'} a_1 = \pm\, a_2^+ a_{2'}^+ a_1 a_{1'} \qquad (107)$$

$$a_{2'}^+ a_2^+ a_1 a_{1'} = \pm\, a_2^+ a_{2'}^+ a_1 a_{1'} \qquad (108)$$

where the top signs are for bosons, the bottom signs for fermions. For either case

$$a_{2'}^+ a_2^+ a_{1'} a_1 = a_2^+ a_{2'}^+ a_1 a_{1'}. \qquad (109)$$

The expression in (104) now simplifies to

$$< n_1-1, n_{1'}-1, n_2+1, n_{2'}+1 | H_{int} | n_1, n_{1'}, n_2, n_{2'} >$$

$$= < n_1-1, n_{1'}-1, n_2+1, n_{2'}+1 | \left\{ a_2^+ a_{2'}^+ a_1 a_{1'} | n_1, n_{1'}, n_2, n_{2'} > \right.$$

$$\left. \cdot\, [< 2,2'|V|1',1 > \pm < 2,2'|V|1,1' >] \right\}. \qquad (110)$$

Then, using the result (102), we get that

$$< n_1-1, n_{1'}-1, n_2+1, n_{2'}+1 | H_{int} | n_1, n_{1'}, n_2, n_{2'} >$$

$$= (S.F.)\ [< 2,2'|V|1',1 > \pm < 2,2'|V|1,1' >]$$

$$\cdot\, \sqrt{(1 \pm n_2)(1 \pm n_{2'}) n_1 n_{1'}}. \qquad (111)$$

Using our previous diagrammatic conventions, the quantity in square brackets on the right of (111) may be represented by
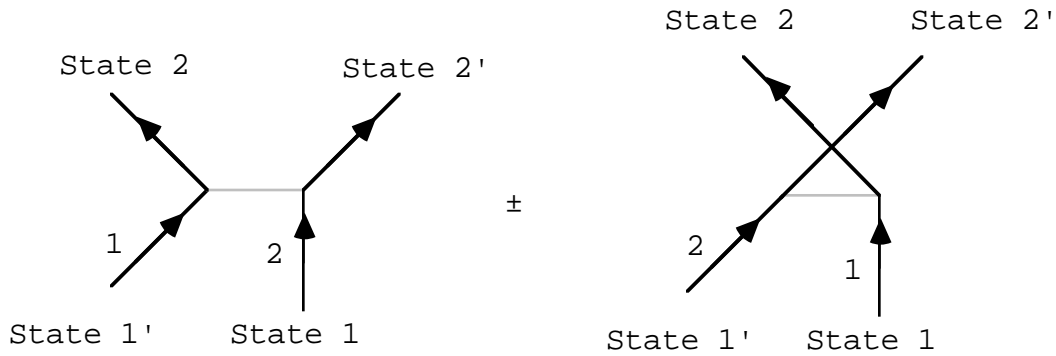
The figure to the right above is called a "crossed diagram" for obvious reasons.

Now Fermi's golden rule, Eq$^{\underline{n}}$ (103), reads:

$$\text{rate}(1 + 1' \rightarrow 2 + 2') \sim |<2,2'|V|1',1> \pm <2,2'|V|1,1'>|^2$$

$$\times (1 \pm n_2)(1 \pm n_{2'})n_1 n_{1'} . \tag{112}$$

The rate for the reverse action is simply given by a relabeling of the above:

$$\text{rate}(2 + 2' \rightarrow 1 + 1') \sim |<1,1'|V|2',2> \pm <1,1'|V|2,2'>|^2$$

$$\times (1 \pm n_1)(1 \pm n_{1'})n_2 n_{2'} . \tag{113}$$

From the identities (83) and (84) we now have that

$$<2,2'|V|1',1>^* = <1,1'|V|2',2>, \tag{114}$$

and

$$<2,2'|V|1,1'>^* = <1,1'|V|2,2'>, \tag{115}$$

which reveals that

$$|<2,2'|V|1',1> \pm <2,2'|V|1,1'>|^2$$

$$= |<1,1'|V|2',2> \pm <1,1'|V|2,2'>|^2 . \tag{116}$$

The statement of detailed balance for this reaction, Eq$^n$ (87), now gives us

$$(1 \pm n_2)(1 \pm n_{2'})n_1 n_{1'} = (1 \pm n_1)(1 \pm n_{1'})n_2 n_{2'} \qquad (117)$$

in contrast to the classical result, Eq$^n$ (88).

Assuming the same overall proportionality constant in the rate as in the classical result, we see that B-E statistics "encourages" the reaction relative to the classical rate, while the minus signs associated with fermions shows that F-D statistics "discourages" the reaction by not permitting it if certain states are already occupied.

Performing the same mathematical steps as in the classical case, we first write (117) as

$$\ln\left(\frac{1}{n_1} \pm 1\right) + \ln\left(\frac{1}{n_{1'}} \pm 1\right) = \ln\left(\frac{1}{n_2} \pm 1\right) + \ln\left(\frac{1}{n_{2'}} \pm 1\right), \qquad (118)$$

which we assume is an expression of energy conservation when the system is in equilibrium.  Thus, for each species i we now have

$$\ln\left(\frac{1}{n_i} \pm 1\right) = \alpha + \beta\varepsilon_i \qquad (119)$$

which gives

$$n_i = \frac{1}{e^{\alpha+\beta\varepsilon_i} + 1} \qquad (120)$$

in the F-D case and

$$n_i = \frac{1}{e^{\alpha+\beta\varepsilon_i} - 1} \qquad (121)$$

in the B-E case.  As before, $\alpha$ is fixed by the total number of particles present and $\beta = (kT)^{-1}$.  Notice that both (120) and (121) go over to the classical result, (92), when $\alpha \gg 1$.  We assume that the energies of the nonrelativistic gas particles are given by

$$\varepsilon_i = \frac{\vec{p}_i^2}{2m} .$$ (122)

Given the results (92), (120) and (121) in the classical, Fermi-Dirac and Bose-Einstein cases respectively, we then have for the total number and energy of these systems that

$$n = \sum_i n_i ,$$ (123)

$$\varepsilon = \sum_i n_i \varepsilon_i .$$ (124)

In order to evaluate these expressions, it will be necessary to replace the sums in (123) and (124) by approximate integrals.  First, let us enumerate exactly the possible momentum states of an ideal gas of noninteracting particles contained in some finite volume, V. We want to take the boundary conditions on the wavefunctions of the particles in the gas to be as general, and yet as simple, as possible. Let us choose a volume such that it is in the shape of a rectangular enclosure of volume $V = L_1 L_2 L_3$.  A free-particle solution to the three dimensional Schrödinger equation

$$- \frac{\hbar^2}{2m} \vec{\nabla}^2 u(x) = \varepsilon u(x),$$ (125)

normalized such that

$$\int_V d^3x |u(x)|^2 = 1, \qquad (126)$$

is given by

$$u(x) = \frac{1}{L^{3/2}} e^{i\vec{p}\cdot\vec{x}/\hbar}. \qquad (127)$$

We imagine that the volume V under consideration is actually in contact with identical volumes along its boundaries as shown below:
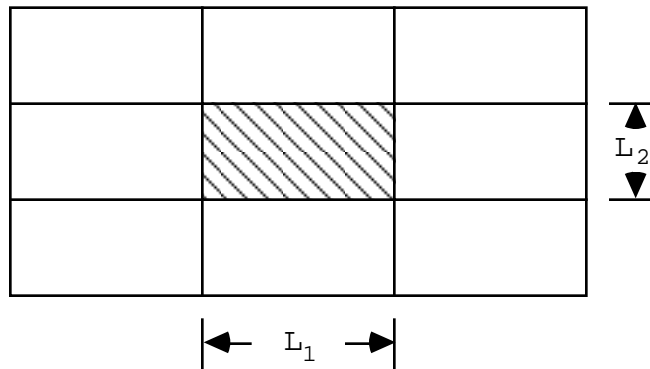


Given this situation, the appropriate boundary conditions on the wavefunctions $u_i(x)$ are

$$\left. \begin{array}{l} u(x_1 + L_1, x_2, x_3) = u(x_1, x_2, x_3) \\[2mm] u(x_1, x_2 + L_2, x_3) = u(x_1, x_2, x_3) \\[2mm] u(x_1, x_2, x_3 + L_3) = u(x_1, x_2, x_3). \end{array} \right\} \qquad (128)$$

These are just periodic boundary conditions. (We saw these also in the Kronig-Penney model of Chapter 3.) They determine the allowed momenta values in the spatial directions. For example, in the one-direction we must have

$$\frac{p_1}{\hbar} (x_1 + L_1) = \frac{p_1}{\hbar} x_1 + 2\pi n_1. \tag{129}$$

where $n_1 = 0, \pm 1, \pm 2, \ldots$ . Therefore

$$p_1 = \left(\frac{h}{L_1}\right) n_1. \tag{130}$$

Similarly in the 2 and 3 directions. Thus, in counting such states, the total number of integers in the range $\Delta n_1 \Delta n_2 \Delta n_3$ is given by

$$\Delta n_1 \Delta n_2 \Delta n_3 = \frac{V}{h^3} \Delta p_1 \Delta p_2 \Delta p_3 . \tag{131}$$

If, in the expressions (123) and (124), the sums are such that the summands vary very little over increments of $n_1$, $n_2$ and $n_3$ by unity, then we may make the replacement

$$\sum_{i = \text{all states}} \rightarrow \frac{V}{h^3} \int d^3 p = \frac{4\pi V}{h^3} \int dp \, p^2 \tag{132}$$

in these expressions. Actually, in the case of particles with spin, the sum over all states must also include a sum over all values of the third component of spin, $S_3' = \hbar m_S$. For noninteracting particles, this just adds a numerical factor:

$$\sum_{i = \text{all states}} \rightarrow \omega \frac{V}{h^3} \int d^3 p = \frac{4\pi \omega V}{h^3} \int dp \, p^2 \tag{133}$$

where $\omega = 2S + 1$ counts the number of components of $S_3'$.

Although the results (132) and (133) have been justified for very particular boundary conditions that might seem

unrealistic, it can be shown that these results are really

insensitive to the shape of the container the noninteracting

particles are placed in and the exact surface boundary

conditions as long as the average deBroglie wavelength of the

particles is small compared to the physical dimensions of the

container. (See the discussion in Reif, <u>Foundations of</u>

<u>Statistical and Thermal Physics</u>, 1st ed., p.362 on this

point.)

We will examine the evaluation of (123) and (124), using

the replacement (133), for the cases of Maxwell-Boltzmann,

Bose-Einstein and Fermi-Dirac statistics. (The following

treatments are, as usual, non-rigorous.)


**(a)   Maxwell-Boltzmann  Statistics**

The expressions for n and $\varepsilon$ are as follows:  (We set

$\omega = 1$, appropriate for a spinless particle)

$$n = \frac{4\pi V}{h^3} \int_0^\infty dp\, p^2 e^{-\alpha} e^{-\beta p^2/2m} \ , \tag{134}$$

$$\varepsilon = \frac{4\pi V}{h^3} \int_0^\infty dp\, p^2 e^{-\alpha} e^{-\beta p^2/2m} \left(\frac{p^2}{2m}\right) \ . \tag{135}$$

Introducing the dimensionless variable

$$x^2 = \frac{\beta p^2}{2m} \tag{136}$$

we can write (134) and (135) as

$$\frac{n}{V} = \frac{4\pi e^{-\alpha}}{h^3} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty dx\, x^2 e^{-x^2}, \tag{137}$$

$$\frac{\varepsilon}{V} = \frac{4\pi e^{-\alpha}}{h^3} \left(\frac{2m}{\beta}\right)^{3/2} \frac{1}{\beta} \int_0^\infty dx\, x^4 e^{-x^2}. \tag{138}$$

We recall from Eq$\underline{^n}$s (53) and (54) of Chapter 2 that

$$\int_0^\infty dx\, e^{-x^2} = \frac{\sqrt{\pi}}{2}. \tag{139}$$

Letting $x^2 \to \lambda x^2$, taking derivatives with respect to $\lambda$ and then setting $\lambda = 0$, we find that

$$\int_0^\infty dx\, x^2 e^{-x^2} = \frac{\sqrt{\pi}}{4}, \tag{140}$$

$$\int_0^\infty dx\, x^4 e^{-x^2} = \frac{3\sqrt{\pi}}{8}. \tag{141}$$

We therefore have that

$$\frac{n}{V} = e^{-\alpha} \left(\frac{2\pi m}{\beta h^2}\right)^{3/2} \tag{142}$$

and

$$\frac{\varepsilon}{V} = \frac{3}{2\beta} e^{-\alpha} \left(\frac{2\pi m}{\beta h^2}\right)^{3/2}. \tag{143}$$

(142) and (143) imply

$$\varepsilon = \frac{3}{2} nkT \tag{144}$$

as we would expect for an ideal gas of n particles from the equipartition theorem. Actually, (142) is just a definition

of $e^\alpha$. Notice that the deBroglie wavelength, $\lambda = \frac{h}{mv}$ , of a particle with energy kT is given by

$$\lambda = \left(\frac{h^2}{2mkT}\right)^{1/2} .$$  (145)

Solving for $e^\alpha$ from (142) allows us to write

$$e^\alpha \approx \frac{1}{\lambda^3 \left(\frac{n}{V}\right)}$$  (146)

Since $\frac{n}{V}$ is just the overall particle density and $\lambda^3$ is a volume, this says in words that

$$e^\alpha \approx \frac{1}{\begin{array}{c}\text{(no. of particles}\\ \text{in a } \lambda^3 \text{ volume)}\end{array}} .$$  (147)

This is the classical meaning of $e^\alpha$. We have already observed that the quantum results, (120) and (121), go over to the classical Maxwell-Boltzmann case when $\alpha \gg 1$. In terms of particle attributes, (147) says, as we might expect, that $\alpha \gg 1$ describes a situation where the deBroglie wavelengths of particles overlap very little. As a numerical example, consider a gas of Helium atoms (m $\approx 10^{-24}$ gm) at room temperature (T = 300°K) and a particle density approximately equal to that of air at one atmosphere $\left(\frac{n}{V} \approx 10^{19} \text{ cm}^{-3}\right)$:

$$e^\alpha \approx 10^5 .$$  (148)

## (b)   Bose-Einstein  Statistics

Apparently, the expressions which we must evaluate in the Bose-Einstein case are (we continue to set $\omega = 1$):

$$n = \frac{4\pi V}{h^3} \int_0^\infty dp\, p^2\, \frac{1}{e^{\alpha+\beta p^2/2m} - 1} \ , \tag{149}$$

$$\varepsilon = \frac{4\pi V}{h^3} \int_0^\infty dp\, p^2\, \frac{\dfrac{p^2}{2m}}{e^{\alpha+\beta p^2/2m} - 1} \ . \tag{150}$$

We expect from the previous classical expression for $e^\alpha$ that we will begin to encounter quantum effects as we increase the deBroglie wavelength overlap by decreasing the temperature. This means lowering the value of $\alpha$.  On the basis of the expression (121), however, it is clear that $\alpha$ can never turn negative.  Consider an $\alpha = -|\alpha|$; then, in a certain mode i we would have for some energies $\varepsilon_i$

$$\varepsilon_i < \frac{|\alpha|}{\beta} \Rightarrow n_i < 0. \tag{151}$$

A negative occupation number is certainly not physical.  Let us therefore consider the extreme situation that $\alpha = 0$.  Then (149) becomes

$$\frac{n}{V} = \frac{4\pi}{h^3} \int_0^\infty dp\, p^2\, \frac{1}{e^{\beta p^2/2m} - 1} \ . \tag{152}$$

Some mathematical steps to reduce the integral in (152) are as follows:

$$\frac{n}{V} = \frac{4\pi}{h^3} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty dx x^2 \frac{1}{e^{x^2} - 1} \, , \tag{153}$$

$$\int_0^\infty dx x^2 \frac{1}{e^{x^2} - 1} = \int_0^\infty dx x^2 \frac{e^{-x^2}}{1 - e^{-x^2}} = \int_0^\infty dx x^2 \sum_{n=1}^\infty e^{-nx^2} . \tag{154}$$

Interchanging the sum and integral in the last quantity and doing the integral (using a scaled version of (140)) then informs us that

$$\sum_{n=1}^\infty \int_0^\infty dx x^2 e^{-nx^2} = \frac{\sqrt{\pi}}{4} \sum_{n=1}^\infty \frac{1}{n^{3/2}} = \frac{\sqrt{\pi}}{4} \zeta\left(\frac{3}{2}\right) \tag{155}$$

where the "Riemann zeta-function" is defined as $(x > 1)$

$$\zeta(x) \equiv \sum_{n=1}^\infty \frac{1}{n^x} . \tag{156}$$

In our case $\zeta\left(\frac{3}{2}\right) = 2.612... $ . One finds that

$$\frac{n}{V} \overset{?}{=} \left(\frac{2\pi m k T}{h^2}\right)^{3/2} \zeta\left(\frac{3}{2}\right) . \tag{157}$$

Eq$^{\underline{n}}$ (157) dictates a particular relationship between the particle density, $\frac{n}{V}$ , and the absolute temperature when $\alpha = 0$. We would expect to be able to approach the absolute zero of temperature to an arbitary extent. However, once we have lowered the temperature (and $\alpha$) to the point where (157) is satisfied, our present equations give us no further guidance as to the behavior of the system. Since n and V are fixed, we may view (157) as predicting a critical temperature, $T_c$,

for the vanishing (or as we will see, the approximate vanishing) of $\alpha$:

$$T_c \equiv \frac{h^2}{2\pi mk} \left( \frac{\frac{n}{V}}{\varsigma\left(\frac{3}{2}\right)} \right)^{2/3} . \tag{158}$$

Since we know that $\alpha \geq 0$, what happens when we try to lower the temperature further?  To understand this, let us reconsider the more correct discrete formula based upon (121):

$$n = \sum_i \frac{1}{e^{\alpha + \beta \varepsilon_i} - 1} . \tag{159}$$

In order to justify the expression (149), it was necessary to assume that the change in the summand in (159) for neighboring discrete momentum states was small.  This assumption breaks down for temperature less than $T_c$.  As $\alpha \to 0$, the single term in (159) for $\vec{p} = 0$ diverges, and we must split this term off before replacing the sum by an integral.  Thus, we replace (149) with the more correct expression

$$n = \frac{4\pi V}{h^3} \int_0^\infty dp\, p^2 \frac{1}{e^{\alpha + \beta p^2/2m} - 1} + n_0 , \tag{160}$$

where

$$n_0 = \frac{1}{e^\alpha - 1} , \tag{161}$$

represents the occupation associated with the $\vec{p}$ state.  What happens for $T < T_c$ is that the Bose particles of the gas begin to collect into this single state.  We may solve approximately for $n_0$ as follows.  From (160) we have

$$n_0 = n - \frac{4\pi V}{h^3} \int dp p^2 \frac{1}{e^{\alpha+\beta p^2/2m} - 1} \; . \qquad (162)$$

To get a first order expression for $n_0$, we make the zero$^{th}$ order approximation that $\alpha = 0$ for $T \le T_c$ in the second term of (162).  Since we already evaluated the resulting integral, we now easily see that

$$n_0 \simeq n - V\left(\frac{2\pi mkT}{h^2}\right)^{3/2} \zeta\left(\frac{3}{2}\right) = n\left(1 - \left(\frac{T}{T_c}\right)^{3/2}\right) \; . \qquad (163)$$

Thus, a macroscopic fraction of the gas occupies the single $\vec{p} = 0$ state for $T \le T_c$.  This phenomenon is known as "Bose-Einstein condensation".  A plot of $n_0(T)$ looks as follows:
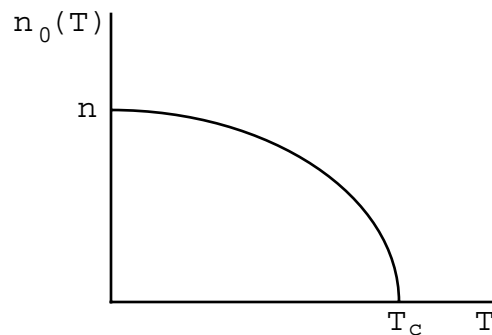


Of course, the non-zero momentum state occupation numbers obey

$$n_{\varepsilon\neq 0} = n\left(\frac{T}{T_c}\right)^{3/2} \qquad (164)$$

below $T_c$.  We do not need to modify the formula (150) for the

total energy $\varepsilon$ since $\varepsilon = 0$ for the mode $n_0$.  The assumption

that $\alpha \ll 1$ when $T \leq T_c$ is consistent since from (161) we have

that

$$\alpha \simeq \frac{1}{n\left(1 - \left(\frac{T}{T_c}\right)^{3/2}\right)} \; . \tag{165}$$

Bose-Einstein condensation is a direct consequence of

the statistics obeyed by integer spin particles whose

available momentum states are, by assumption, changed very

little by interactions between the particles.  This is an

idealistic situation.  The question is:  Is anything like

this seen in nature?  It is known that liquid (not gaseous)

$He^4$, which is a Bose system ($A + Z$ = even), undergoes a

transition, called the "$\lambda$-transition", to a phase called

"Helium II" at a temperature of 2.18°K.  This new phase seems

to be made up of two components called the normal fluid and

the superfluid, similar to the $n_0$ and $n_{\varepsilon \neq 0}$ components

described above.  If one simply substitutes the mass and

density of liquid $He^4$ into the expression (158), one gets $T_c$ =

3.14°K, not too far from the $\lambda$-transition temperature.  In

addition, no such transition is seen in liquid $He^3$, an

isotope of Helium having 2 protons and 1 neutron, which obeys

Fermi-Dirac statistics ($A + Z$ = odd).  Thus, it is tempting

to conjecture that Bose-Einstein condensation is the dominant

(but not exclusive) cause of the $\lambda$-transition in liquid $He^4$.

This conjecture has not yet been thoroughly confirmed however.

Let us also calculate (150) when $\alpha \simeq 0$ (which we now know occurs when $T \leq T_c$). We have

$$\varepsilon = \frac{4\pi V}{h^3} \int dp\, p^2\, \frac{\frac{p^2}{2m}}{e^{\beta p^2/2m} - 1} \quad . \qquad (166)$$

Running through the steps similar to (153) - (157) above, we now find that

$$\varepsilon = \frac{3}{2}\, \zeta\!\left(\frac{5}{2}\right) \left(\frac{2\pi mkT}{h^2}\right)^{3/2} VkT \simeq .770\ n\!\left(\frac{T}{T_c}\right)^{3/2} kT \qquad (167)$$

where $\zeta\!\left(\frac{5}{2}\right) \simeq 1.341...$ . Eq$^n$ (167) has a very different appearance from the Maxwell-Boltzmann result, Eq$^n$ (144). The total specific heat at constant volume is given by

$$C_V \equiv \left.\frac{\partial \varepsilon}{\partial T}\right|_V \simeq 1.925\ \left(\frac{T}{T_c}\right)^{3/2} nk \qquad (168)$$

which is seen to vanish at $T = 0$. Thus, for an ideal gas of weakly interacting BE particles, we expect that the specific heat will initially rise as T increases until reaching a peak at $T = T_c$. After this, the value will level out to the constant value, $\frac{3}{2}$ nk, expected at high temperatures. There is a discontinuity in $C_V$ at $T = T_c$. (The $\lambda$-shape of this graph is the origin for the name of this transition.)

$$\frac{C_V}{nk}$$
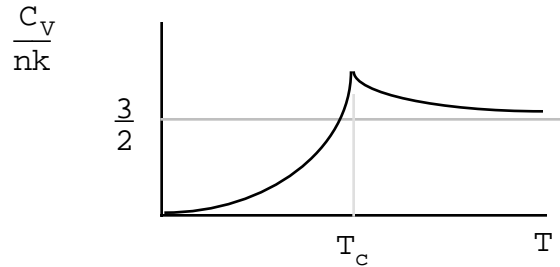


The very imporatant case of Bose-Einstein statistics applied to a gas of photons (dealt with briefly in Chapter 2 in the Compton effect) in an equilibrium temperature black-body cavity is described in this formalism by putting $\alpha = 0$. This makes sense since we have seen that $\alpha$ is determined by the (up to now) fixed number of particles present in the system.   Now the number of particles present is not fixed, but is determined by the temperature and size of the box, and nothing else. Photons have spin 1, and so are described by Eq$\underline{^n}$s (149) and (150) above, but with $\omega = 2$ (two photon polarizations), $\alpha = 0$ and the photon mode energy $\varepsilon_p = pc$, where p is the magnitude of the photon's momentum (remember Eq$\underline{^n}$ (7) of Chapter 2.)   I will leave the treatment of this case to you in a homework problem. We will talk more about the photon, considered as an elementary particle, in the next (unwritten as yet!) Chapter.

## (c)   Fermi-Dirac Statistics

In general, we have (we set $\omega = 2$, appropriate for a spin $\frac{1}{2}$ particle)

$$n = \frac{8\pi V}{h^3} \int_0^\infty dp\, p^2\, \frac{1}{e^{\alpha + \beta p^2/2m} + 1} \quad , \tag{169}$$

$$\varepsilon = \frac{8\pi V}{h^3} \int_0^\infty dp\, p^2\, \frac{\frac{p^2}{2m}}{e^{\alpha + \beta p^2/2m} + 1} \quad . \tag{170}$$

Just as we concentrated on low temperatures in the B-E case, we will also do so here, for this is where we expect to see quantum effects.  Again, $\alpha$ is determined from the expression for n.  Before, we argued that $\alpha$ had to be positive and then evaluated the expressions for n and $\varepsilon$ in the extreme case that $\alpha = 0$.  Here, there is no such restriction since the expression for $n_i$ in the F-D case, Eq$\underline{n}$ (120), is never in any danger of turning negative.  In either the B-E or F-D case, determining $\alpha$ by solving (160) or (169), respectively, is a difficult mathematical problem.  However, the situations simplify at low temperatures.  In the B-E case we found self-consistently that $\alpha \ll 1$ for $T \leq T_c$ and saw that the particles all piled up in the ground state, $\vec{p} = 0$, when $T = 0$.  We expect a completely different situation in the F-D case at low temperatures.  Because of the exclusion principle, no more than a single particle may occupy a given state of the system.  Therefore, at $T = 0$ we expect the identical fermions not to accumulate in the single $\vec{p} = 0$ state, but to fill up the lowest n states of the system with single particles.  The occupation number, n(p), considered as a continuous function of $\frac{p^2}{2m}$ , should appear at $T = 0$ as follows:
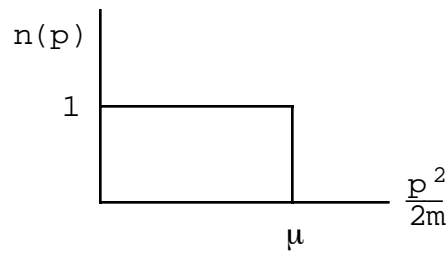
That is, we expect

$$\lim_{T \to 0} \frac{1}{e^{\alpha + \beta p^2/2m} + 1} = \begin{cases} 1, & \dfrac{p^2}{2m} < \mu \\ \\ 0, & \dfrac{p^2}{2m} > \mu. \end{cases} \tag{171}$$

The quantity $\mu$, called the "Fermi energy", is the energy of the highest occupied state and will be determined by the total number of particles present. Eq$^n$ (171) specifies that

$$\alpha(T) = -\frac{\mu}{kT} \tag{172}$$

at low temperatures. Given the expected profile, (171), Eq$^n$s (169) and (170) now read (at <u>exactly</u> T = 0)

$$n = \frac{8\pi V}{h^3} \int_0^{P_F} dp\, p^2, \tag{173}$$

$$\varepsilon_{T=0} = \frac{8\pi V}{h^3} \int_0^{P_F} dp\, p^2 \left(\frac{p^2}{2m}\right), \tag{174}$$

where we have set $\dfrac{p_F^2}{2m} = \mu$. Doing the trivial integrals results in

$$n = \frac{8\pi V}{3} \frac{(2m\mu)^{3/2}}{h^3}, \tag{175}$$

$$\varepsilon_{T=0} = \frac{8\pi V}{10} \frac{(2m\mu)^{5/2}}{mh^3} = \frac{3}{5} \mu \left(\frac{n}{V}\right). \tag{176}$$

Eqⁿ (175) determines $\mu$ (and also the low temperature form for $\alpha(T)$ through (172)). We get

$$\mu = \frac{h^2}{2m} \left(\frac{3}{8\pi} \frac{n}{V}\right)^{2/3}. \tag{177}$$

We define the "Fermi temperature,"

$$kT_F = \mu. \tag{178}$$

Its significance will be discussed shortly.

These results are for T = 0. Let us now look at the first order deviations for low temperatures. In the following, we will be concerned with the approximate evaluation of integrals of the form

$$I = \int_0^\infty d\varepsilon_p \frac{f(\varepsilon_p)}{e^{\alpha + \varepsilon_p/kT} + 1} \tag{179}$$

given that $\alpha(T) \to -\dfrac{\mu}{kT}$ at low temperatures. First, rewrite (179) exactly as

$$I = \int_0^{-\alpha kT} d\varepsilon_p f(\varepsilon_p) - \int_0^{-\alpha kT} d\varepsilon_p \frac{f(\varepsilon_p)}{1 + e^{-\alpha - \varepsilon_p/kT}} + \int_{-\alpha kT}^\infty d\varepsilon_p \frac{f(\varepsilon_p)}{e^{\alpha + \varepsilon_p/kT} + 1}. \tag{180}$$

In the second integral above introduce the variable x through $\varepsilon_p = -\alpha kT - (kT)x$; in the third integral introduce x through $\varepsilon_p = -\alpha kT + (kT)x$. Doing the change of variables, we get (still an exact result)

$$I = \int_0^{-\alpha kT} d\varepsilon_p f(\varepsilon_p) - kT \int_0^{-\alpha} dx \frac{f(-\alpha kT - kTx)}{e^x + 1}$$

$$+ kT \int_0^\infty dx \frac{f(-\alpha kT + kTx)}{e^x + 1}. \tag{181}$$

Now using the fact that $\alpha(T) \rightarrow -\frac{\mu}{kT}$ at low temperatures, we replace the upper limit in the second integral by $+\infty$ and write

$$I \simeq \int_0^{-\alpha kT} d\varepsilon_p f(\varepsilon_p) + kT \int_0^\infty dx \frac{f(-\alpha kT + kTx) - f(-\alpha kT - kTx)}{e^x + 1}. \tag{182}$$

This approximation holds when $\alpha \ll -1$, or equivalently, $T \ll T_F$. We now expand the functions $f(-\alpha kT + kTx)$ and $f(-\alpha kT - kTx)$ about their values at small $x$. This is allowed in (182) because of the suppression of higher order terms by the factor $(e^x + 1)^{-1}$ at large $x$.

$$[f(-\alpha kT + kTx) - f(-\alpha kT - kTx)] \simeq 2kTxf'(-\alpha kT) \tag{183}$$

where

$$f'(-\alpha kT) \equiv \frac{df(\varepsilon_p)}{d\varepsilon_p}\Bigg|_{\varepsilon_p = -\alpha kT}. \tag{184}$$

Our integral now reads

$$I \simeq \int_0^{-\alpha kT} d\varepsilon_p f(\varepsilon_p) + 2(kT)^2 f'(-\alpha kT) \int_0^\infty dx \frac{x}{e^x + 1}. \tag{185}$$

But

$$\int_0^\infty dx \frac{x}{e^x + 1} = -\sum_{n=1}^\infty \frac{(-1)^n}{n^2} = \frac{\pi^2}{12}. \tag{186}$$

The last result can be justified from the theory of Fourier series (see Kaplan, _Advanced Calculus_, 2nd edition, p.487). Therefore, we have arrived at

$$I \underset{T \ll T_F}{\cong} \int_0^{-\alpha kT} d\varepsilon_p f(\varepsilon_p) + \frac{\pi^2}{6} (kT)^2 f'(-\alpha kT). \qquad (187)$$

Let us apply (187) to (169) and (170). First, we rewrite (169) as

$$n = \frac{8\pi\sqrt{2}\ m^{3/2}V}{h^3} \int_0^\infty d\varepsilon_p \frac{\varepsilon_p^{1/2}}{e^{\alpha+\varepsilon_p/kT} + 1}\ , \qquad (188)$$

where $\varepsilon_p = \frac{p^2}{2m}$. In order to use the result (187) we identify $f(\varepsilon_p) = \varepsilon_p^{1/2}$ for which $f'(-\alpha kT) = \frac{1}{2} (-\alpha kT)^{-1/2}$. Plugging in, we find that

$$n \simeq \frac{8\pi V(2m\mu)^{3/2}}{3h^3} \left\{ \left(\frac{-\alpha kT}{\mu}\right)^{3/2} + \frac{\pi^2}{8} \left(\frac{kT}{\mu}\right)^2 \left(-\frac{\mu}{\alpha kT}\right)^{1/2} \right\}. \qquad (189)$$

The first term on the right in (189) just reproduces the result (175) when $T = 0$, as it should. Since n and V are fixed, (189) implies a change in the meaning of $\alpha(T)$. Setting $\alpha(T) = -\frac{\mu}{kT} + z$ where $z \ll \frac{\mu}{kT}$, we find self-consistently from (189) that

$$\alpha(T) \simeq -\frac{\mu}{kT} + \frac{\pi^2}{12} \left(\frac{kT}{\mu}\right). \qquad (190)$$

Raising the value of T is thus seen to increase $\alpha(T)$, as it should, since $\alpha(T) \gg 1$ is the classical limit.

We may employ the same technique on (170) to show that

$$\varepsilon \simeq \frac{8\pi V}{10} \frac{(2m\mu)^{5/2}}{mh^3} \left\{ \left(-\frac{\alpha kT}{\mu}\right)^{5/2} + \frac{5\pi^2}{8} \left(\frac{kT}{\mu}\right)^2 \left(-\frac{\alpha kT}{\mu}\right)^{1/2} \right\} . \quad (191)$$

Since we now know the approximate form for $\alpha(T)$, we may rewrite this to first order as

$$\varepsilon \simeq \frac{8\pi V}{10} \frac{(2m\mu)^{5/2}}{mh^3} \left\{ 1 + \frac{5\pi^2}{12} \left(\frac{kT}{\mu}\right)^2 \right\} . \quad (192)$$

Recognizing from (176) that the overall factor in (192) is just the energy of the gas at zero temperature, we have that

$$\varepsilon \simeq \frac{3}{5} n\mu + \frac{\pi^2}{4} n\mu \left(\frac{kT}{\mu}\right)^2 . \quad (193)$$

We again investigate the specific heat.  We have that

$$C_V = \frac{\pi^2}{2} nk^2 \left(\frac{T}{\mu}\right) = \frac{\pi^2}{2} nk \left(\frac{T}{T_F}\right) . \quad (194)$$

Thus, we see that $T_F$ sets the scale of temperatures for the Fermi-Dirac gas.

The point of this calculation is the following. Remember that the electrons in the Kronig-Penney model are treated as a gas acted upon by the stationary atoms. If we use appropriate electronic concentrations, $\frac{n}{V}$ , in (176) and (177), we find that (see C. Kittel, <u>Introduction to Solid State Physics</u>, 4<u>th</u> edition, p.248 for the values used.)

$$(T_F)_{Cu} = 8.12 \times 10^4 \, ^\circ K \, \Big\}$$

$$(T_F)_{Ag} = 6.36 \times 10^4 \, ^\circ K \, \Big\}$$

(194)

for Copper and Silver.  Thus, the <u>electronic</u> component of $C_V$ at room temperature is very small.  Qualitatively, this effect can be understood from the following diagram which holds at $T \ll T_F$.
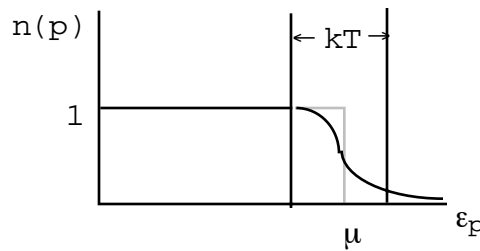


From the approximate form for $n(p)$ when $T \ll T_F$,

$$n(p) \approx \frac{1}{e^{(\varepsilon_p - \mu)/kT} + 1} \, ,$$

(195)

it is easy to see that the width of the region of particles with excited energies is of order kT.  The number of particles excited is therefore of order $n(\frac{kT}{\mu}) = n\left(\frac{T}{T_F}\right)$ .  This means the excitation energy of the gas will be $\left[n\left(\frac{T}{T_F}\right)\right] kT = $ $nk \frac{T^2}{T_F}$ and the specific heat will be $\approx nk\left(\frac{T}{T_F}\right)$, as found above.

This solves the mystery of the missing electronic component of specific heat for these materials since the Fermi temperature is so high compared to room temperature for these materials.

# Problems

1. Show from the commutation relations that

$$N_i \, (a_i^+ \, a_i^+ | \cdots 0_i \cdots >) = 2(a_i^+ \, a_i^+ | \cdots 0_i \cdots >),$$

where $N_i = a_i^+ \, a_i$ for both FD (trivial) and BE statistics.

2. (a) Given

$$a_i \, | \cdots n_i \cdots > = C_i \, | \cdots n_i - 1 \cdots >,$$

where $C_i$ is an unknown constant, and

$$a_i^+ \, a_i \, | \cdots n_i \cdots > = n_i \, | \cdots n_i \cdots >,$$

show that

$$C_i = \sqrt{n_i} \, e^{i\alpha},$$

where $\alpha$ is an arbitrary (real) phase.

   (b)  Deduce from (a) that

$$a_i^+ \, | \cdots n_i - 1 \cdots > = e^{-i\alpha} \sqrt{n_i} \, | \cdots n_i \cdots >.$$

3. For a two-physical property FD system adopt the definition:

$$a_1^+ \, a_2^+ \, |0,0> = a_1^+ \, |0,1> = |1,1>.$$

   (a)  Show that

$$a_1 = |0,0><1,0| + |0,1><1,1|,$$

$$a_2 = |0,0><0,1| - |1,0><1,1|,$$

satisfy (66) – (69) and the above result for $a_1^+ \, a_2^+ \, |0,0>$.

   (b)  Show that

9.52

$$a_i a_i^+ + a_i^+ a_i = \sum_{n_1, n_2 = \{0,1\}} |n_1, n_2\rangle\langle n_1, n_2|,$$

for either i = 1 or 2.

4. From problem 2 above we know for FD statistics that

$$a_i^+ |\cdots 0_i \cdots\rangle = e^{-i\alpha} |\cdots 1_i \cdots\rangle,$$

$$a_i |\cdots 1_i \cdots\rangle = e^{i\alpha} |\cdots 0_i \cdots\rangle.$$

Using the convention that $e^{i\alpha}$ = 1 when the number of occupied one-particle states with index less than i is even, and $e^{i\alpha}$ = -1 when this number is odd, evaluate:

   (a)   $a_3^+$ |1,0,1,0,0>=?
   (b)   $a_5^+ a_1$ |1,0,1,0,0>=?
   (c)   $a_1 a_3 a_5 a_4 a_2$ |1,1,1,1,1>=?
   (d)   $a_3^+ a_1^+ a_4^+ a_5^+ a_2^+$ |0,0,0,0,0>=?

The occupation numbers are given in the order:
$|n_1, n_2, n_3, n_4, n_5\rangle$.


5. (a) Show for both statistics:

$$[a_i, a_j^+ a_k] = \delta_{ij} a_k.$$

Then derive

$$[a_i^+, a_j^+ a_k] = -\delta_{ik} a_j^+.$$

by Hermitian conjugation.
   (b)  Apply part (a) to prove

$$[a_i^+ a_j, a_k^+ a_\ell] = \delta_{jk} a_i^+ a_\ell - \delta_{i\ell} a_k^+ a_j.$$


6. Evaluate:

(a)   $e^{-i\lambda N} a_i e^{i\lambda N} = ?$

(b)   $e^{-i\lambda N} a_i^+ e^{i\lambda N} = ?$

where $N = \sum_i N_i$ is the total number operator.

[Hint:  Differentiate these quantities in $\lambda$ and solve the resulting differential equations.]

7.  $F^{(1)} \equiv \sum_{i,j} a_i^+ a_j \; <i|f|j>,$

$G^{(1)} \equiv \sum_{k,l} a_k^+ a_\ell \; <k|g|l>.$

Show (use the result of problem 1(b))

$$[F^{(1)}, G^{(1)}] = \sum_{i,j} a_i^+ a_j \; <i|[f,g]|j>.$$

8. Use (upper sign is BE, lower sign FD)

$$\frac{1}{e^{\alpha + \beta\varepsilon} \mp 1} \simeq e^{-\alpha - \beta\varepsilon} \pm e^{-2\alpha - 2\beta\varepsilon}$$

when $e^\alpha$ is large to evaluate

$$N = \frac{\omega V}{(2\pi\hbar)^3} \int d^3p \; \frac{1}{e^{\alpha + \beta(\vec{p}^2/2m)} \mp 1}$$

and

$$E = \frac{\omega V}{(2\pi\hbar)^3} \int d^3p \; \frac{\frac{\vec{p}^2}{2m}}{e^{\alpha + \beta(\vec{p}^2/2m)} \mp 1}$$

approximately.  Show that

$$\frac{N}{V} \simeq \omega \left(\frac{mkT}{2\pi\hbar^2}\right)^{3/2} e^{-\alpha}[1 \pm 2^{-3/2} e^{-\alpha}],$$

9.54

$$\frac{E}{V} \simeq \frac{3}{2} \, kT\omega \, (\frac{mkT}{2\pi\hbar^2})^{3/2} \, e^{-\alpha}[1 \pm 2^{-5/2}e^{-\alpha}].$$

Then show that $\frac{E}{V}$ and $\frac{N}{V}$ are related by

$$\frac{E}{V} \simeq \frac{3}{2} \, kT \, (\frac{N}{V}) \, (1 \mp 2^{-5/2}e^{-\alpha}).$$

9. (a) Evaluate the integrals (for photons)

$$n = \frac{8\pi V}{h^3} \int_0^\infty dp \, p^2 \, \frac{1}{e^{\beta pc} - 1} \, ,$$

$$\varepsilon = \frac{8\pi V}{h^3} \int_0^\infty dp \, p^2 \, \frac{pc}{e^{\beta pc} - 1} \, .$$

Note that $\zeta(3) = 1.20206$, $\zeta(4) = \frac{\pi^4}{90}$ .

 (b)  Show that we may also write

$$\frac{\varepsilon}{V} = \int_0^\infty d\nu \, u(\nu), \text{ where}$$

$$u(\nu) = \frac{8\pi h\nu^3}{c^3} \, \frac{1}{e^{h\nu/kT} - 1} \, .$$

Demonstrate that in the $\nu \to 0$ and $\nu \to \infty$ limits we get:

$$\lim_{\nu \to 0} u(\nu) = \frac{8\pi\nu^2}{c^3} \, kT,$$

$$\lim_{\nu \to \infty} u(\nu) = \frac{8\pi h\nu^3}{c^3} \, e^{-h\nu/kT}.$$

**Other  Problems**

10. Consider a gas of n identical nonrelativistic fermions. This gas, in an enclosure of volume V, is placed in a

uniform magnetic field, $B_z$, pointing along the positive z-axis.  An atom with momentum $\vec{p}$ can have only two possible energies:

$$\varepsilon(\pm) = \frac{\vec{p}^2}{2m} \pm \lambda\, B_z,$$

where $\lambda$ is the atom's magnetic moment.  At exactly $T = 0$ (zero temperature), the occupation numbers $n(\pm)$ (associated with the energies $\varepsilon(\pm)$) look like (assuming $\lambda > 0$; "$\mu$" is the common value for the "Fermi energy"):
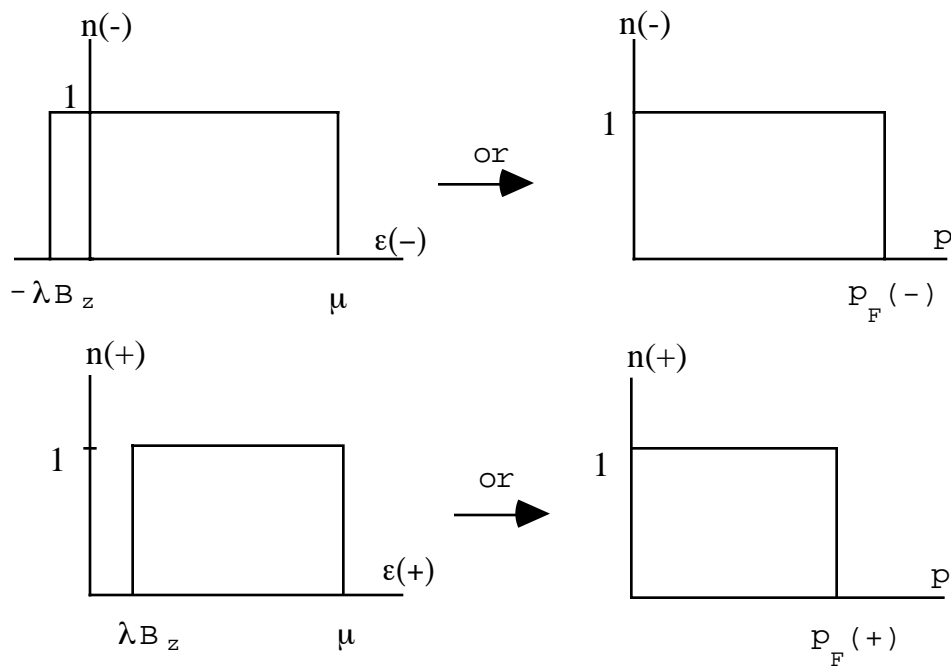


Assuming $n(\pm)$ are given by

$$n(\pm) = \frac{4\pi V}{h^3} \int_0^{p_F(\pm)} dp\, p^2,$$

at T=0, show that $(n = n(-) + n(+))$

$$n(-) - n(+) \simeq \frac{3n\lambda B_z}{2VkT_F} \;,\; n \simeq \frac{8\pi}{3h^3}(2m\mu)^{3/2},$$

where $kT_F = \mu$. (Make the approximation $|\lambda B_z| \ll \mu$.)

11. Consider <u>fermion</u> creation and annihilation operators $a_i^+$ and $a_i$ for some particle property "i".

(a) Find <u>matrix representations</u> for each of the following operators (use the number basis):

(i)  $a_i$

(ii) $a_i^+$

(iii) $a_i^+ a_i$

(b) Evaluate:

(iv) $e^{a_i^\dagger a_i} |\cdots n_i \cdots\rangle$ = ?

(v)  $e^{a_i a_i^\dagger} |\cdots n_i \cdots\rangle$ = ?

12. Consider an ensemble of N one-dimensional harmonic oscillators whose energy levels are quantized but which obey Maxwell-Boltzmann statistics. The energy levels, as usual, are given

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

(a) Given that the occupation number of the $n^{\underline{th}}$ state at temperature T is given by ($\beta = 1/(kT)$)

$$n_n = \exp(-\alpha - \beta E_n),$$

find an expression for the total energy of the gas, E. [Hints: You can eliminate $\alpha$ by using

$$N = \sum_{n=0}^{\infty} n_n.$$

Also, the sum

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$$

may be helpful. You should be able to do (b) and (c) even if you can't do the sums in (a).]

(b) Find the specific heat of this collection of oscillators by taking

$$C_v = \left. \frac{\partial E}{\partial T} \right|_v,$$

(c) Take the limit

$$\lim_{T \to \infty} C_v = ?.$$

[Hint: This limit should give the classical result expected from the equipartition theorem.]

13. Consider a __relativistic__ gas (E = pc) of N identical Fermi-Dirac particles occupying a volume V at zero temperature (T=0). Show that the total energy, $E_{tot}$, of the system is given by

$$E_{tot} = \frac{(9\pi)^{2/3}}{4} \, \hbar c N^{4/3} V^{-1/3}.$$

[Such considerations are important in stellar dynamics of white dwarfs and neutron stars, for example.]

## Chapter 10: Connecting to the Standard Model

   We have came a long way in our understanding and delineation of the quantum aspects of our world.  It is now time to begin to apply our quantum description to nature on the smallest scale that can be determined with present technology. We will find a rich, detailed set of interactions and particles which make up the bricks and mortar of our physical world.

   We have covered some of the important dynamical principles involved, but we have not yet met the particles.  We will do so momentarily, but first we need a way to distinguish or categorize the particles that exist in our world.  We already know about the dynamics associated with some of the properties of particles: mass, parity, and spin.  However, there are also quantum properties which derive from <u>discrete</u> <u>symmetries</u> which characterize a particle and its interactions.  There are three such properties I would like to discuss at this point: parity, time reversal and charge conjugation.  First, we will become reacquainted with the property known as parity, from a transformation, not quantum, point of view.


I. Discrete Symmetries


<div align="center"><u>Parity</u></div>


  We have of course already discussed the parity operation in this text.  The emphasis here will be on the effect of parity on the transformation properties of quantities of quantum fields. Start with the idea of scalar and pseudoscalar fields. Under any orthogonal transformations:


$$\phi'(\vec{x}\,') = \phi(\vec{x}\,). \quad \text{(a scalar \underline{field})} \qquad (10.1)$$

Example: __dot product__. Proof: ($A_i' = \sum_j a_{ij} A_j$, $B_i' = \sum_k a_{ik} B_k$)

$$\vec{A}' \cdot \vec{B}' = \sum_i A_i' B_i' = \sum_{i,j,k} a_{ij} A_j a_{ik} B_k,$$

$$= \sum_{j,k} \left( \sum_i a_{ij} a_{ik} \right) A_j B_k = \vec{A} \cdot \vec{B} \quad . \quad (10.2)$$

$$\underbrace{\phantom{\sum_i a_{ij} a_{ik}}}_{\delta_{ik}}$$

Another example: $\vec{\nabla} \cdot \vec{A}$ (if $\vec{A}$ is a vector). There are also pseudoscalars. They transform as

$$\phi'(\vec{x}') = (\det a) \, \phi(\vec{x}). \qquad (10.3)$$

The transformation of vector and pseudovector fields can be characterized by their behavior under rogations. Component statement:

$$x_i' = \sum_{j=1}^{3} a_{ij} x_j \begin{pmatrix} \text{rotate coord:} & \underline{\text{passive}} \\ \text{rotate system:} & \underline{\text{active}} \end{pmatrix}. \qquad (10.4)$$

__Matrix__ statement:

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \qquad (10.5)$$

$$\text{or} \quad x' = ax. \qquad (10.6)$$

So, anything transforming as

$$V_i'(\vec{x}') = \sum_j a_{ij} V_j(\vec{x}) \left( V(\vec{x}) \text{ a vector } \underline{\text{field}} \right), \quad (10.7)$$

under an active rotation is a vector.

Conditions on the $a_{ij}$? Must preserve the length of vectors.

$$\Rightarrow \quad \sum_1 a_{ij}a_{ik} = \delta_{jk}, \tag{10.8}$$

or

$$a^T a = 1, \tag{10.9}$$

$$\Rightarrow \quad a^T = a^{-1} \quad \text{"orthogonal"}. \tag{10.10}$$

The above condition represents 6 eq$^n$s in 9 unknowns
$\Rightarrow$ 3 $a_{ij}$'s are free (rotations about x,y,z axes).  Determinant:

$$\Downarrow\text{completely antisymmetric}$$
$$\det A = \varepsilon_{ijk}\cdots \quad a_{1i}a_{2j}a_{3k}\cdots \quad ] \tag{10.11}$$

$$\underbrace{\qquad}_{\text{n indices}} \quad \underbrace{\qquad\qquad}_{\text{n factors}}$$

Take determinant of $a^T a = 1$:

$$\Rightarrow \quad (\det a)^2 = 1, \tag{10.12}$$

$$\Rightarrow \quad \det a = \pm 1 \quad \underline{\text{only}}. \tag{10.13}$$

det a = 1 describes pure rotations.  What does det a = −1
describe? Example (makes a new object):
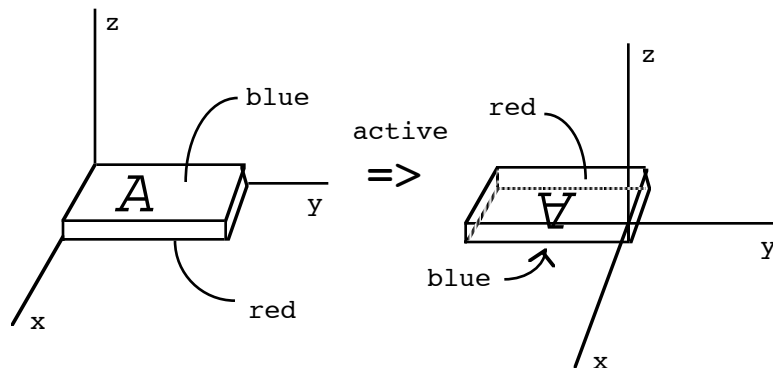


Given by:

$$a = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Can not be done by a rotation. It is an <u>inversion</u> (also orthogonal). Vectors change sign under a complete inversion:
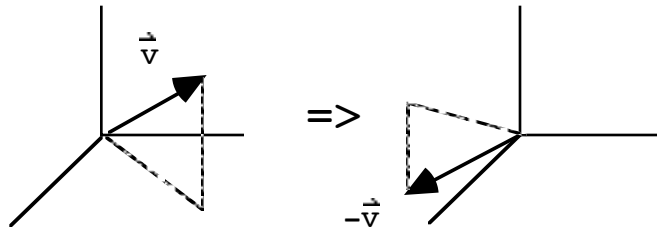


There is another type of vectorial quantity which, however, is distinguished by its behavior under inversion.
Consider ($\vec{A}$, $\vec{B}$ vectors)

$$(\vec{A} \times \vec{B})_i = \sum_{j,k} \varepsilon_{ijk} A_j B_k. \qquad (10.14)$$

Need result (problem):

$$\sum_{j,k} \varepsilon_{ijk} a_{j\ell} a_{km} = (\det a) \sum_n \varepsilon_{n\ell m} a_{in}. \qquad (10.15)$$

Thus

$$(\vec{A}' \times \vec{B}')_i = \sum_{j,k} \varepsilon_{ijk} A'_j B'_k = \sum_{j,k,\ell,m} \varepsilon_{ijk} a_{j\ell} A_\ell a_{km} B_m$$

$$= \sum_{j,k,\ell,m} \varepsilon_{ijk} a_{j\ell} a_{km} A_\ell B_m = (\det a) \sum_{n,\ell,m} \varepsilon_{n\ell m} a_{in} A_\ell B_m$$

$$= (\det a) \sum_n a_{in} (\vec{A} \times \vec{B})_n. \qquad (10.16)$$

This type of vector (<u>pseudovector</u>) does not change sign under an inversion. It's transformation law is simply

$$V'_i(\vec{x}') = (\det a)\sum_j a_{ij}V_j(\vec{x}) \quad \left(V(\vec{x}) \text{ a vector } \underline{\text{field}}\right). \quad (10.17)$$

Example: angular momentum $(\vec{L} = (\vec{x} \times \vec{p}))$.  Choose:

$$a = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Under this a:

$$L'_x = -L_x, \quad L'_y = -L_y, \quad L'_z = +L_z \ .$$
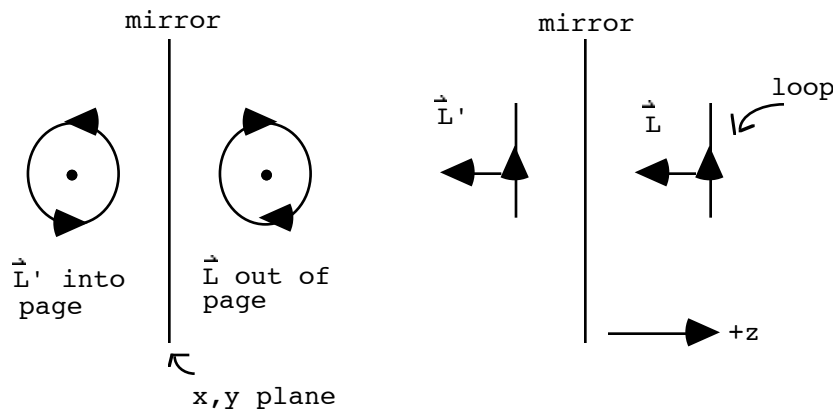


Notice that since $\vec{L}$ is a pseudovector, the <u>lack</u> of change of sign on the z-component under this transformation is what distinguishes it from a vector under the same transformation. Particle <u>spin</u> also behaves as a pseudovector under coordinate transformations.

  A problem at the end of the chapter illustrates the use of parity in the properties and "selection rules" for electromagnetic transitions.  This discrete symmetry is contained in the more general set of relative linear coordinate transformations, called Lorentz transformations, which preserve only the "proper length", $\Delta s^2 \equiv c^2\Delta t^2 - \Delta\vec{x}^2$, between events, rather than space and time intervals separately.  We will learn in an upcoming section that the strong interaction, which conserves parity, is responsible for producing bound states from

several particles.  For example, meson states are bound states of particles known as quarks.  These are classified by their parity, just like atomic states.  However, not all interactions in nature conserve parity. We will see that the weak interactions intrinsically do not conserve parity.  This is not a small violation for this interaction, but in some sense it is "maximally" violated, and was difficult to confirm in the laboratory only because of the extreme weakness of the interaction.  This realization was one of the most profound shocks to particle physics in the 20th century.

## Time Reversal

A better name for this discrete symmetry is reversal of motion.  All non-velocity dependent static forces in classical mechanics are time-reversal invariant.  This can be seen from Newton's law:

$$\left.\begin{array}{l} \vec{F} = m\vec{a}, \\[2mm] \vec{\nabla}V(\vec{x}) = m\ddot{\vec{x}} \end{array}\right\} \text{ invariant under } t \to -t \qquad (10.18)$$

$\Rightarrow$ if $\bar{x}(t)$ is a possible trajectory (solution), then so is $\bar{x}(-t)$

Also, if all one had in the world were electric fields any particle trajectory would be invariant under:

$$\bar{E} \to \bar{E}, \; t \to -t. \qquad (10.19)$$

$$\text{Since } \bar{F} = q\bar{E} = -q\bar{\nabla}\Phi(\bar{x}) \qquad (10.20)$$

However, magnetic fields are a little trickier. Imagine a world where (electron charge, e<0):

Lorentz force
inward on
electron
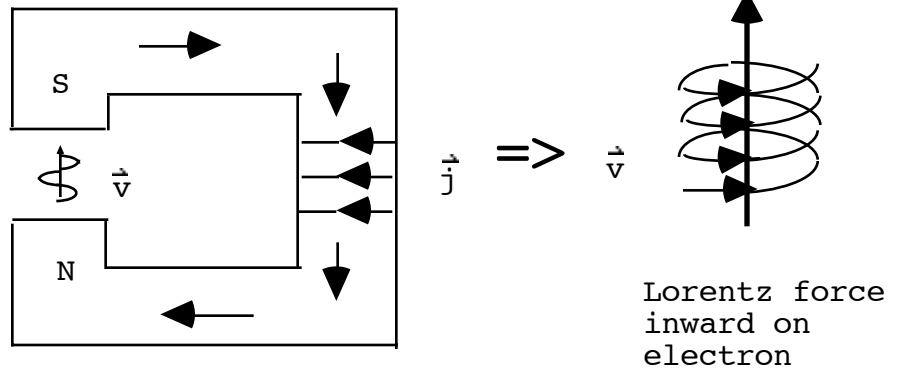
Imagine just letting t → −t, v̄ → −v̄.



Different
axis X

Lorentz force
inward on new
axis
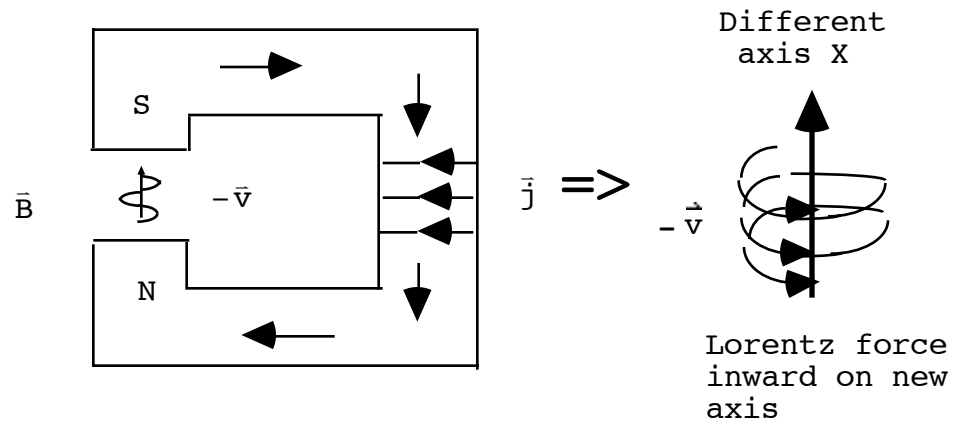
Electron does <u>not</u> retrace former trajectory.  Let

t → −t,  v̄ → −v̄,  j̄ → −j̄,  B̄ → −B̄.

↑ Makes sense from point of view
of reversal of motion.



Lorentz force
inward on same
axis
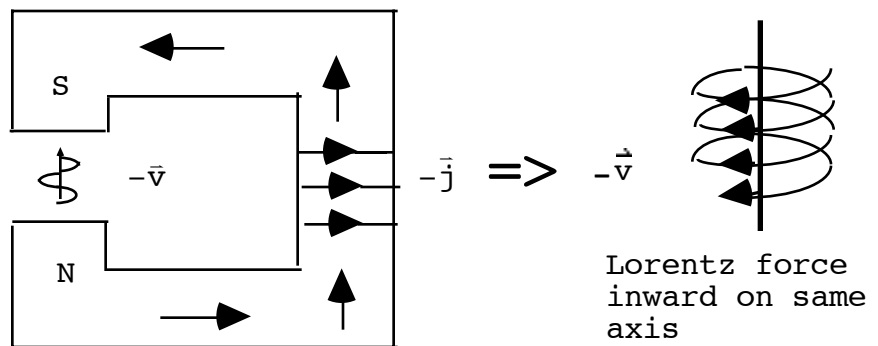
Formally, Maxwell's equations are invariant under

$$t \to -t, \quad \vec{E} \to -\vec{E}, \quad \vec{B} \to -\vec{B}, \quad f \to f, \quad \vec{j} \to -\vec{j}, \quad \vec{v} \to -\vec{v}. \qquad (10.21)$$

What about the Schrödinger equation in this context? It is a first order differential equation in time. We have

$$i\hbar \frac{\partial \psi(\vec{x},t)}{\partial t} = [\frac{-\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{x})]\psi(\vec{x},t). \qquad (10.22)$$

$\psi^*(\vec{x},-t)$ (complex conjugate) is also a solution:

$$i\hbar \frac{\partial \psi^*(\vec{x},-t)}{\partial t} = [\frac{-\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{x})]\psi^*(\vec{x},-t). \qquad (10.23)$$

So we learn that if $\langle \vec{x}|\alpha \rangle$ is the t = 0 wavefunction, the time-reversed (motion-reversed) wavefunction is given by $\langle \vec{x}|\alpha \rangle^* = \langle \alpha|\vec{x}|\rangle$.

$$\langle \vec{x}|\alpha \rangle^* \xrightarrow{\text{time-reversal}} \langle \alpha|\vec{x}\rangle. \qquad (10.24)$$

From this very simple observation, we deduce that if we are to introduce an operator which represents time reversal, it is a very unusual object. The reason is that an operator only takes bras into bras and kets into kets. For example:

$$X = \sum_{i,j} |a_i\rangle\langle a_j|X_{ij}, \qquad (10.25)$$

$$\Rightarrow X|a_k\rangle = \sum_{i} X_{ik}|a_i\rangle, \quad \langle a_k|X = \sum_{i} X_{kj}\langle a_j|. \qquad (10.26a,b)$$

What we want to do now is more akin to an <u>operation</u>, rather than an operator. For example, we had for Hermitian conjugation (see Ch.1, Eq.(188)),

$$(|\alpha\rangle)^+ = \langle\alpha|. \tag{10.27}$$

Under $^+$:

$$(\langle\vec{x}|\alpha\rangle)^+ = \langle\alpha|\vec{x}\rangle. \tag{10.28}$$

However, now consider the momentum space time-reversed wavefunction.  We have (t=0),

$$\tilde{\psi}_\alpha(\vec{p}\,') = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3x \; e^{-i\vec{p}\,'\cdot\vec{x}/\hbar} \; \tilde{\psi}_\alpha(\vec{x}\,'), \tag{10.29}$$

↑ time-reversed
<u>position</u> wavefunction

$$\tilde{\psi}_\alpha(\vec{x}\,') = \tilde{\psi}_\alpha^*(\vec{x}\,'), \tag{10.30}$$

$$\tilde{\phi}_\alpha(\vec{p}\,') = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3x \; e^{-i\vec{p}\,'\cdot\vec{x}/\hbar} \; \tilde{\psi}_\alpha^*(\vec{x}\,'),$$

$$= \tilde{\phi}_\alpha^*(-\vec{p}\,'). \tag{10.31}$$

We have found that

$$\langle\vec{p}\,'|\alpha\rangle^* \xrightarrow[\text{time-reversal}]{\Rightarrow} \langle\alpha|-\vec{p}\rangle. \tag{10.32}$$

$\Rightarrow$ Time-reversal is not equivalent to "$^+$" (dagger).

We will take the following point of view. Define the effect of a <u>anti-unitary</u> operation:

$$\text{base kets} \begin{cases} (\langle\alpha|)^A = \langle\tilde{\alpha}| \; = U|\alpha\rangle, \\ (|\beta\rangle)^A = \langle\tilde{\beta}| = \langle\beta|U^+. \end{cases} \tag{10.33a,b}$$

Hermitian conjugation is one such example with U = 1. However, we also define

No star!
↓       ↓
$$(C_1\langle\alpha| + C_2\langle\beta|)^A = C_1|\tilde{\alpha}\rangle + C_2|\tilde{\beta}\rangle, \tag{10.34}$$

$$\left(\langle\alpha|\beta\rangle\right)^{A} \;=\; \langle\alpha|\,\beta\rangle \;=\; \left\langle\tilde{\beta}|\tilde{\alpha}\right\rangle. \tag{10.35}$$

Notice these rules imply

$$\left(\langle\alpha\,|U\right)^{T} \;=\; \left(\langle\alpha\,|\right)^{T} U^{T}, \tag{10.36}$$

$$\left(U|\,\beta\rangle\right)^{T} \;=\; (U)^{T}\,|\,\beta\rangle^{T}, \tag{10.37}$$

and that

$$U\!\left(|\,\alpha\rangle\right)^{T} \;=\; \left\langle\tilde{\alpha}\,\right|U. \tag{10.38}$$

We now define the effect of the time-reversal <u>operation</u> on $\bar{x}'$, $\bar{p}'$ eigenstates:

$$\left(\langle\bar{x}'\,|\right)^{T}=\;|\,\bar{x}'\rangle, \quad \left(|\,\bar{x}'\rangle\right)^{T} \;=\; \langle\bar{x}'\,| \tag{10.39}$$

$$\left(\langle\bar{p}'\,|\right)^{T} \;=\; |-\bar{p}'\rangle, \quad \left(|\,\bar{p}'\rangle\right)^{T} \;=\; \langle-\bar{p}'\,|. \tag{10.40}$$

Effect on $\bar{x},\bar{p}$?

$$\left(\langle\bar{x}'\,|\bar{x}'\rangle\right)^{T} \;=\; \left(\langle\bar{x}'\,|\right)^{T}\bar{x}' \;=\; \bar{x}'\,|\,\bar{x}'\rangle \;=\; \bar{x}^{T}|\,\bar{x}'\rangle \tag{10.41}$$

$$\Rightarrow\; \bar{x}^{T} \;=\; \bar{x}. \tag{10.42}$$

Similarly $\bar{p}^{T} = -\bar{p}$.

Let us examine the effect of the time reversal operation on particle states using the above formalism. The standard way of time evolving a system for a simple time independent Hamiltonian is

$$|\alpha,\; \delta t\rangle \;=\; \left(1-\frac{iH}{\hbar}\,\delta t\right)\,|\alpha\rangle. \tag{10.43}$$

Visualization (pictures after Sakurai, Fig. 4.11):
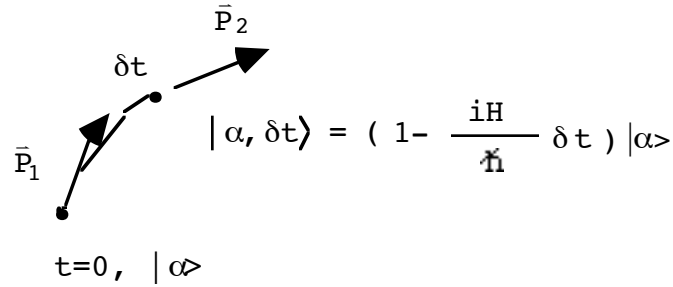
$$|\alpha, \delta t\rangle = (1 - \frac{iH}{\hbar} \delta t) |\alpha\rangle$$

Change the time origin, motion reverse, and evolve backward in time:
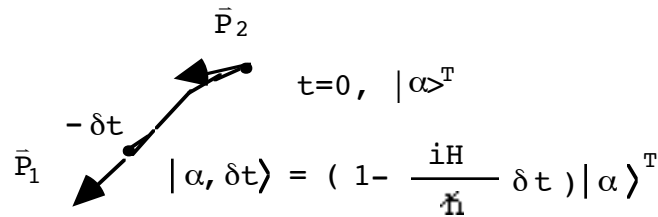


$$|\alpha, \delta t\rangle = (1 - \frac{iH}{\hbar} \delta t) |\alpha\rangle^T$$

Evolve backward, then motion reverse:
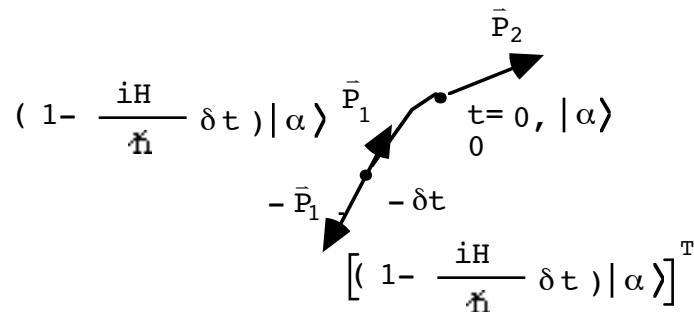


$$(1 - \frac{iH}{\hbar} \delta t) |\alpha\rangle$$

$$\left[ (1 - \frac{iH}{\hbar} \delta t) |\alpha\rangle \right]^T$$

If the system is motion reversal invariant, then we have

$$\langle \tilde{\alpha} | \left( 1 + \frac{iH}{\hbar} \delta t \right)^T = \langle \tilde{\alpha} | \left( 1 + \frac{iH}{\hbar} \delta t \right) . \qquad (10.44)$$

For any $\langle \tilde{\alpha} |$, this implies $(H)^T = H$.


    We are using the special rule that the complex number "i" does not have any special transformational properties under time reversal. There is an immediate consequence of this property, which I will state as a theorem.

Theorem: $H^T = H$ and the energy eigenkets $|n\rangle$ are nondegenerate. Then we may always choose

$$\langle \bar{x}'| n \rangle = \langle \bar{x}'| n \rangle^*. \tag{10.45}$$

Proof:

Given: $H|n\rangle = E_n|n\rangle$. Take $^T$: $(|n\rangle)^T = \langle \tilde{n}|$.

$$\Rightarrow \langle \tilde{n}| H^T = E_n \langle \tilde{n}|, \quad \Rightarrow \langle \tilde{n}| = e^{i\delta}\langle n| \text{ (choose } \delta=0),$$

$$\Rightarrow \langle \bar{x}'| n \rangle = \langle \bar{x}'| n \rangle^T = \left( \langle \bar{x}'| \right)^T \cdot \left( |n\rangle \right)^T = \langle \tilde{n}| \bar{x}' \rangle,$$

$\uparrow$ complex number

$$\Rightarrow \langle \bar{x}'| n \rangle = \langle \bar{x}'| n \rangle^*.$$

My treatment of time reversal as an operation, rather than an operator, is not standard. However, much of the Dirac notation formalism developed in earlier chapters must be modified or abandoned if we take the operator point of view. For example, the operator point of view has $\langle \alpha|\left( \Theta| \beta \rangle \right) \neq \left( \langle \alpha |\Theta| \right) \beta \rangle$ where $\Theta$ is the time reversal <u>operator</u>. In addition, even though $(H)^T = H$ characterizes a time reversal invariant Hamiltonian, unlike parity there is no "conservation of time reversal quantum number." The operation point of view makes this explicit since there is no operator that commutes with the Hamiltonian!

## **Charge Conjugation**

Charge conjugation refers to the act of changing particle into anti-particles. we have not yet talked about anti-particles, which are particle states with all <u>additive</u> quantum numbers reversed in sign. (Additive quantum numbers consist of electric charge, the various quark "flavors," such as strangeness, and the

three types of lepton number.) Calling this operator C, it's
affect on a state with additive quantum numbers A, B, C, would be

$$C|A,B,C\rangle = |-A,-B,-C\rangle. \qquad (10.46)$$

Given that the electric charge, Q, is such a quantum number, one
can immediately see that these operators have the property that
they ant-commute:

$$\{C,Q\} = 0. \qquad (10.47)$$

This last equation implies that no state of non-zero charge can
be an eigenstate of C. In general, this applies for any additive
charge. In addition, applying charge conjugation twice gives back
the same state, which says that

$$C^2 = I. \qquad (10.48)$$

This equation implies that the eigenvalues of C are ±1, for non-
degenerate states, just like parity. It is a <u>multiplicative</u>
quantum number, just like parity.

One particle which has no additive charge is the photon.
Changing the sign of the charge will change the sign of the
photon field, $A_\mu$. We learn in quantum field theory that this
field may be used to create or destroy a photon. Thus, each
additional photon in a given state changes the sign of the charge
conjugation state. This quantum number will be conserved if C
commutes with the electromagnetic Hamiltonian causing the decay,
and thus the charge conjugation of the initial and finial states
are the same. The neutral pion decay,

$$\pi^0 = \gamma + \gamma, \qquad (10.49)$$

implies that $C_{\pi^0} = 1$. This particle is never seen to decay into 3
photons. The neutral pion state is formed from the strong
interaction, and thus this is a very strong hint that the strong

interactions also conserve charge conjugation. Other neutral states for which charge conjugation conservation are applicable are positronium ($e^+e^-$) and so-called quarkonium states, such as $c\bar{c}$ (charm/anti-charm).

The following table summarizes the transformation properties of states under the various discrete operations described above.

---

Table I: Summary of described discrete transformations

<u>Parity:</u> $\vec{J} \rightarrow \vec{J}$ ($\vec{J} = \vec{L}$ or $\vec{S}$), $\vec{p} \rightarrow -\vec{p}$, $\vec{x} \rightarrow -\vec{x}$.

<u>Time Reversal:</u> $\vec{J} \rightarrow -\vec{J}$, $\vec{p} \rightarrow -\vec{p}$, $\vec{x} \rightarrow \vec{x}$.

<u>Charge conjugation:</u> $\vec{J} \rightarrow \vec{J}$, $\vec{p} \rightarrow \vec{p}$, $\vec{x} \rightarrow \vec{x}$,
(All additive quantum numbers such as electric charge, baryon number and the various lepton numbers change sign.)

---

II. Particle Zoo

In this section we will be briefly introduced to the particles of the so-called "Standard Model" of particle physics. Our world is remarkably and intricately made from a collection of 61 particles, some stable from decay into other fundamental particles and some not. All of these 61 particles are now known directly from experiment with one notable exception to be described later. These particles can be classified according to their types of interactions: the quarks feel the EM, strong, and weak interactions, and the leptons participate in EM and weak interactions. In addition, there are the particle mediators of these interactions, the so-called gauge bosons. For a quick picture of the types of interactions in which these particles can

participate, see the pictorial list of allowed particle vertices in Appendix A of the present chapter.

Let us begin first with the quarks. Quarks have spin 1/2 and come in 6 different "flavors" as far as the strong interactions are concerned – these have come to be called "down" (d), "up" (u), "strange" (s), "charmed" (c), and "bottom" (b), and "top" (t). These are presented along with their electric charges (in units of the proton's electric charge) in Table II. (There are also anti-quarks, $\bar{d}$, $\bar{u}$, $\bar{s}$, $\bar{c}$, $\bar{b}$, and $\bar{t}$ with the opposite electric, color charges of their particle partners.) The reason for the grouping of two flavors is that each combination (d,u), (s,c) and (b,t) is considered a different "generation" or "family." Note from the table that the electric charges in each generation are repeated. Each flavor of quark has a unique mass, but specifying their mass values is difficult because, as we will see, quarks are only detected in bound states; individual quarks, such as u or d, are never detected in the laboratory. At high enough energies, one can define the quark masses by analyzing certain experiments. From this method, one finds that the up quark mass is anywhere from about 2 MeV to 4 MeV, the down quark has a mass of from 4 MeV to 8 MeV, and the strange quark has a mass of about 80 MeV to 130 MeV. The first three quarks are considered "light", the last three are "heavy". The mass of the charm quark is about 1.2 GeV, the b–quark has a mass of approximately 4.26 GeV, while the top quark's mass is a whopping 174 GeV. Each quark also has a "baryon number", arbitrarily assigned as 1/3, so that three quark combinations, such as the proton or neutron, have unit baryon number. This quantum number is conserved in all particle interactions as far as is known. Counting up the number of quarks gives a counting of 6 (flavors) x 3 (colors) x 2 (particle/anti-particle) = 36 particles of the 61 standard model particles.

## Table II- Quark Additive Quantum Numbers

| Flavor | Charge | Baryon Number |
|--------|--------|---------------|
| u   (up) | $2/3$ | $1/3$ |
| d   (down) | $-1/3$ | $1/3$ |

| Flavor | Charge | Baryon Number |
|--------|--------|---------------|
| c   (charmed) | $2/3$ | $1/3$ |
| s   (strange) | $-1/3$ | $1/3$ |

| Flavor | Charge | Baryon Number |
|--------|--------|---------------|
| t   (top) | $2/3$ | $1/3$ |
| b   (bottom) | $-1/3$ | $1/3$ |

As I said above, the group of particles known as leptons feel only the EM and weak forces. There are 6 of these particles, just as there are 6 flavors of quarks, and they also have spin 1/2. Also, just as the (d,u), (s,c) and (b,t) combination of quarks forms a different "generation" or "family", the leptons are similarly grouped, but of course their electric charges are different. There are electron and electron neutrinos, $(e, \nu_e)$, the muon and muon neutrono, $(\mu, \nu_\mu)$, and the tau and tau neutrino, $(\tau, \nu_\tau)$. (Again, there are also the anti-particles for each of these.) And just like the quarks, the masses of each generation increase $(m_e > m_\mu > m_\tau)$, with the possible exception of the neutrinos, which for many years have been thought of as being

massless. It is now known that all these particles very likely have small, but non-zero masses. This is a subject of on-going research. Since the neutrinos are neutral, they participate only in weak interactions.

All the leptons in Table III are fundamental (not made up of other particles), but not all of them are stable. For example, the muon and tau leptons (masses of 105.6 MeV and 1.78 GeV, respectively) decay. The primary decay mode for the muon is $\mu \rightarrow e \, \bar{\nu}_e \, \nu_\mu$. The lifetime associated with the muon is $2.2 \times 10^{-6}$ sec. Notice that in this decay the lepton numbers defined in the table below are conserved. The muon has $L_\mu = 1$; this decays into particles having $L_e = 1$ (e), $L_e = -1$ ($\bar{\nu}_e$), and $L_\mu = 1$ ($\nu_\mu$).

The counting of leptons is: 6 (types) x 2 (particle/anti-particle) = 12 of the 61. We are up to 36 + 12 = 48 of 61 particles.

All forces in nature are mediated by other particles, known as gauge bosons. Bosons have integer spin and gauge bosons all have spin 1. The gauge boson in electrodynamics is the photon, the particle of light; the gauge boson in QCD is called the gluon. These two particles are massless and travel at the speed of light in vacuum. There are 8 such particles (see the below discussion under QCD particle interactions). The gauge bosons of the weak interactions are the charged $W^\pm$ and the neutral $Z^0$ bosons. Both of these particles are massive; in fact, the $W^\pm$ has a mass of about 86 times that of a proton (80.4 GeV as opposed to .938 GeV), and the other has a mass of about 97 times times a proton (91 GeV).

The counting of gauge bosons is: 8 (gluons) + 2 ($W^\pm$) + 1 (photon) + 1 (Z) = 12. We are up to 60 of 61 particles.

### Table III: Lepton Additive Quantum Numbers

| Lepton | Charge | $L_e$ | $L_\mu$ | $L_\tau$ |
|---|---|---|---|---|
| e   (electron) | -1 | 1 | 0 | 0 |
| $\nu_e$ (electron neutrino) | 0 | 1 | 0 | 0 |
| $\mu$   (muon) | -1 | 0 | 1 | 0 |
| $\nu_\mu$ (muon neutrino) | 0 | 0 | 1 | 0 |
| $\tau$   (tau) | -1 | 0 | 0 | 1 |
| $\nu_\tau$ (tau neutrino) | 0 | 0 | 0 | 1 |

The only particle we have not mentioned yet in the standard model is the Higgs boson. It is a spin 0 neutral scalar particle. It's mass is not determined by the Standard model. It has not been observed, but it's mass is now constrained by various experiments to lie in the range of from about 65 GeV to about 114 GeV. We will have more to say about the Higgs boson in the weak interaction section of the following particle interactions section. There is supposedly only one Higgs boson.

These are the 61 particles of the Standard Model of particle physics.

III. Particle Interactions

## Quantum Electrodynamics (QED)

The basic interaction vertex in electrodynamics is shown in the Appendix A. Quantum electrodynamics, QED, is the best knows and most studied aspect of the Standard Model. Using Feynman diagrams, the structure and exact properties of many processes are straightforward to compute. For example, the Feynman diagrams for $e^-$ $\gamma$ -> $e^-$ $\gamma$ all given to lowest order as
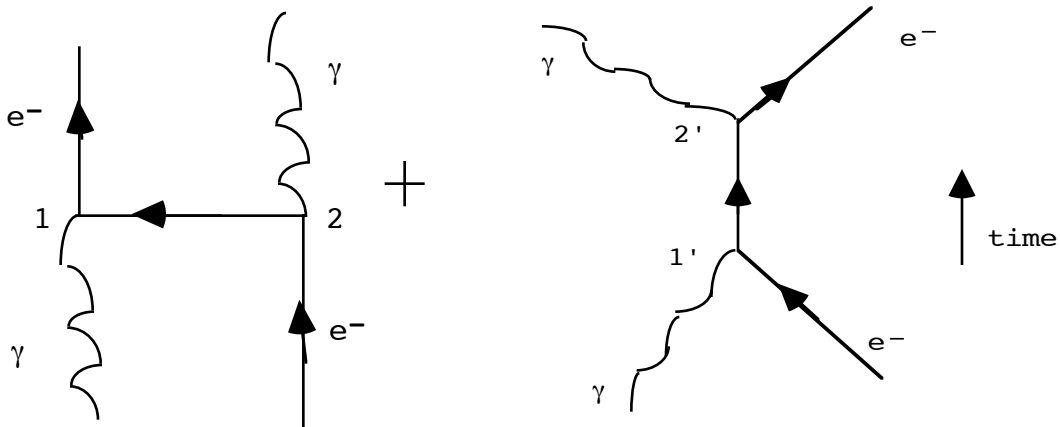


Again, remember that the vertices can come in any time ordering, the only thing invariant is the structure or topology of the diagram. For example,

If "1" is before "2"; e+ line
If "2" is before "1"; e- line

A similar statement can be made about the vertices 1' and 2' in the right diagram. The particles which are exchanged in these diagrams, the internal lines in the Feynman diagrams, are called "virtual" particles. We already know a little about the range and lifetime of such particles from the discussion in chapter 7 regarding the pion. There we learned that the Heisenberg uncertainty principle determined these quantities, see Equations

(7.112) and (7.113). It is similar here with the electron; it is exchanged over an approximate range of $\hbar/(m_e c) = 3.9 \times 10^{-11}$ cm (the so-called electron Compton wavelength) with a time uncertainty of $\hbar/(m_e c^2) = 1.3 \times 10^{-21}$ sec.

The most fundamental and stringently tested conservation law in nature is the conservation of electric charge. This conservation is intimately tied up with the fact that the photon, the gauge boson of QED, is exactly massless, as we have already pointed out. Massless particles have and important property related to spin: the have only two degrees of freedom, rather than the expected. 2s+1=3 m values of s=1. Formally, the range of massless photons is infinite. It turns out that relativistic field theory requires the spin of massless particles to only point along the direct of motion, $\vec{p}$, as in the opposite direction. That is, the photon's <u>helicity</u>, $\dfrac{\vec{s} \cdot \vec{p}}{|\vec{P}|}$, can take only the values + 1 or − 1. This particle property is familiar to us as the 2 polarizations of light we know about from classical electrodynamics. These two polarizations can be pictured a follows. Let us say that the motion of a photon is out of the page. Then, the two independent polarizations may be pictured as
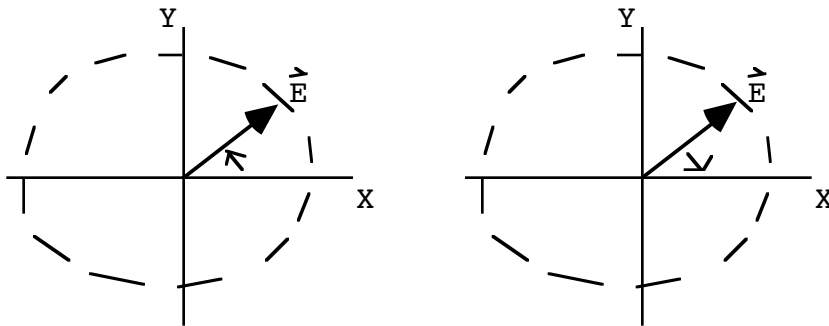


  The left diagram has an $\vec{E}$ field vector which is instantaneously rotating in a circular path in a counterclockwise fashion. This is called left circular polarization and is associated with negative helicity. Of course, we may use a <u>linear</u> polarization basis to describe light beams; these are just linear combinations of these two polarizations. We also learned in the above section on charge conjugation that the photon has no additive quantum

numbers. This implies that there is no such thing as an anti-photon. The photon can be considered it's own anti-particle.

Electrodynamics provides the "glue" which makes atoms possible and essentially all the forces which make up the material around us. QED is so well understood that it can be used to test or understand other forces. For example, in the early 60's a series of experiments were done at the Stanford Linear Acceleration Center (SLAC) involving the scattering of electrons off of protons and neutrons. ($e^-$ N -> $e^-$ X, where "X" is anything. This is called "inclusive" scattering). These scattering just involved exchanging photons, are one would expect from the theory. What physicists discovered from a careful study is that the cross sections behaved as is the electrons were hitting point objects within the nucleus, just as Rutherford long ago discovered the atomic nucleus from the scattering of b-particles (electrons). This type of interaction is called <u>deep inelastic scattering</u>, and by this physicists discovered quarks. The quarks could not be knocked out of their environment (this is called <u>confinement</u>), but the electromagnetic interaction made it clear that there were a number of point objects in photons and neutrons. Although the quarks could hide themselves within a hardon, they could not escape detection by the infinite-ranged photon!

## Quantum Chromodynamics (QCD)

As I said before, the theory of strong interactions is described by QCD.  The force is mediated my the gluons and is extremely strong compared, for example, to electromagnetism. This force is Coulombic (~ $\frac{1}{r^2}$) at small distances and distance independent (~ constant) at large distances.  This constant force, or tension, equivalent to about 15 tons, is what is responsible for quark confinement.  This string can break at large enough distances, but instead of getting two pieces of

string, another quark/anti-quark pair emerges from the vacuum to terminate the string ends, much like what happens to a magnet which is broken in half.  Quarks are only seen confined, in hadrons. The energy scales and sizes of hardons are determined by the strength and range of this force. They are the major building blocks, by mass, of our known physical world. Of these, the up and down quarks dominate; most of hadronic matter is made up of neutrons and protons, which are composed of two down-quarks and an up-quark, or by two up-quarks and a down-quark, respectively*. The small mass differences between up and down quarks, down being slightly more massive than the up quarks, is responsible for neutrons being slightly more massive than protons. (There is also a smaller electromagnetic effect due to the charge on the proton which raises it's energy. This also affects other particles, like charged and neutral pions.)  Just as electrons carry electric charge, which is absolutely conserved, quarks carry three color charges, which are also absolutely conserved (call them green, blue, and red). The name "color" is arbitrary, but suggestive. Just as white light is composed of a mixture of all the colors in the spectrum, one can make white or colorless combinations of quarks. This can be done in two ways: combining a color with its opposite, making a quark-anti-quark pair (mesons) or by combining three different colored quarks (baryons). Hadronic physics is the study of the properties and interactions of these composite particles.

The theory of the strong interactions is called Quantum Chromodynamics, QCD for short. As pointed out above, in all particle theories, forces are mediated by other particles known as gauge bosons. Bosons have integer spin - gauge bosons all have spin 1. The gauge boson in electrodynamics is the photon, the particle of light; the gauge boson in QCD is called the "gluon".

---

* There is also an admixture of all other types of quarks in protons and neutrons. Quantum field theory tells us that quark/anti-quark pairs of all species are continuously appearing and disappearing in the vicinity of hadrons. These are called "sea quarks", and they can have dramatic physical effects, although these flavors are "hidden".

These two particles are massless. There is one all-important different between protons and gluons. Photons are chargeless, whereas gluons also carry color. To see why, consider the diagram:
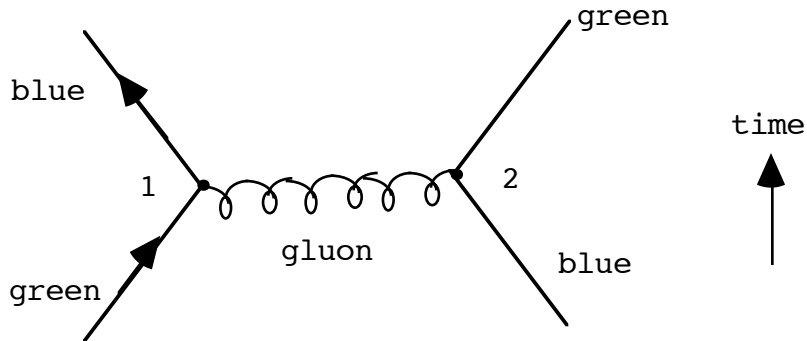


The gluon changes the quark color from one vertex to the other and thus carries color also. This gluon carries the color green/anti-blue ($g\bar{b}$) if vertex 1 comes before vertex 2, and blue/anti-green ($b\bar{g}$) for the opposite ordering. (Remember Feynman diagrams are agnostic on time ordering of vertices, so this specification is really not necessary.) There are 9 such combinations of 3 colors; however, the colorless combination, $g\bar{g}$ + $b\bar{b}$ + $r\bar{r}$, does not correspond to a particle, so there are only 8 gluons. The basic vertices for quarks and gluons are shown in Appendix A, where we see that unlike photons, gluons can couple to themselves. They are also confined inside hadrons, like quarks.

The theory of strong interactions can not be formulated and solved in the usual field theory way of using Feynman diagrams, as can electrodynamics or QED. QED interactions are characterized by a particle/photon coupling strength given by the square root of the fine structure constant, $\alpha = \dfrac{e^2}{\hbar c} \approx \dfrac{1}{137}$ , which is small compared to one and allows the Feynman diagrams to be summed to very high order. For QCD, the quark-gluon vertex has an analogous coupling strength, $\alpha_{strong}$, which is close to one, making the series of Feynman diagrams divergent. This coupling strength is a function of the interaction energy and becomes smaller at high

energies due to sea quark contributions (see the footnote on sea quarks in proton, neutrons). Thus for high energy interactions, Feynman diagrams are again useful. This property of the QCD coupling strength is known as "asymptotic freedom." (This property is used to define quark masses at high energies.)

Let us consider an unphysical world where all quarks are "heavy". Then their motions inside hadrons would be nonrelativistic. This is called the "quark model." This model is an incredibly good guide to overall properties, like magnetic moments and mass orderings. QCD does not change any of the flavors into any of the other flavors, so the u, d, etc. quantum numbers are conserved in strong interactions. Since the u, d quarks are considered degenerate in mass in this model, this gives an effective SU(2) in flavor space which is called "isospin." The mathematics of isospin are exactly the same as the two-valuedness property of Chapter 1, and different isospins may be added exactly as we learned in Chapter 8. The fact that this quantum number is conserved in the strong interactions gives us the ratio of certain matrix elements involved in particle interactions (see the problems at the end of the chapter). Tables IV, V, and VI shows the flavor – spin wavefunctions of the lowest mass particles expected in the quark model, along with their isospin classification. (The various K mesons also have a strangeness quarntum number which distinguishes them.)

There is a numerical method for solving QCD, called "lattice QCD", which solves the theory directly and does not depend on the summation of Feynman diagrams. In fact, the degrees of freedom in quantum field theories are not the number of particles involved (like in nonrelativistic quantum mechanics), but the points of space and time themselves. Thus, the simplest field theory already has an infinite number of degrees of freedom.

## Table IV: Pseudoscalar Meson Table

| | Mass (MeV) | Isospin | Spin$^{\text{Parity}}$ | wave function |
|---|---|---|---|---|
| $\pi^{+}_{-}$ | 139.6 | $I=1$, $I_3=\pm1$ | $0^-$ | $u\bar{d}$ ($d\bar{u}$) x (singlet spin) |
| $\pi^0$ | 135.0 | $I=1$, $I_3=0$ | $0^-$ | $(u\bar{u}-d\bar{d})/\sqrt{2}$ |
| $K^{+}_{-}$ | 493.7 | $I=\frac{1}{2}$, $I_3=\pm\frac{1}{2}$ | $0^-$ | $u\bar{s}$ ($s\bar{u}$) |
| $K^0(\overline{K^0})$ | 497.3 | $I=\frac{1}{2}$, $I_3=\mp\frac{1}{2}$ | $0^-$ | $d\bar{s}$ ($s\bar{d}$) |
| $\eta$ | 548 | $I=0$ | $0^-$ | $(2s\bar{s}-u\bar{u}-d\bar{d})/\sqrt{6}$ |
| $\eta'$ | 958 | $I=0$ | $0^-$ | $(u\bar{u}+d\bar{d}+s\bar{s})/\sqrt{3}$ |

## Table V: Vector Meson Table

| | Mass (MeV) | Isospin | Spin$^{\text{Parity}}$ | wave function |
|---|---|---|---|---|
| $\rho^{+}_{-}$ | 776 | $I=1$, $I_3=\pm1$ | $1^-$ | $u\bar{d}$ ($d\bar{u}$) x (triplet spin) |
| $\rho^0$ | 776 | $I=1$, $I_3=0$ | $1^-$ | $(u\bar{u}-d\bar{d})/\sqrt{2}$ |
| $K^{*+}_{-}$ | 892 | $I=\frac{1}{2}$, $I_3=\pm\frac{1}{2}$ | $1^-$ | $u\bar{s}$ ($s\bar{u}$) |
| $K^{*0}(\overline{K^{*0}})$ | 899 | $I=\frac{1}{2}$, $I_3=\mp\frac{1}{2}$ | $1^-$ | $d\bar{s}$ ($s\bar{d}$) |
| $\omega$ | 782 | $I=0$ | $1^-$ | $(u\bar{u}-d\bar{d})/\sqrt{2}$ |
| $\phi$ | 1020 | $I=0$ | $1^-$ | $s\bar{s}$ |

Table VI: Baryon Octet Table

| | Mass (MeV) | Isospin | wave function(+cyclic permutations) |
|---|---|---|---|
| p | 938.3 | $I=\frac{1}{2}$, $I_3=+\frac{1}{2}$ | $\lvert uud\rangle(2\lvert++-\rangle-\lvert+-+\rangle-\lvert-++\rangle)/3\sqrt{2}$ |
| n | 939.6 | $I=\frac{1}{2}$, $I_3=-\frac{1}{2}$ | $-\lvert ddu\rangle(2\lvert++-\rangle-\lvert+-+\rangle-\lvert-++\rangle)/3\sqrt{2}$ |
| $\Lambda^0$ | 1115 | $I=0$ | $(\lvert uds\rangle-\lvert dus\rangle)(\lvert+-+\rangle-\lvert-++\rangle)/2\sqrt{3}$ |
| $\Sigma^+$ | 1189 | $I=1$, $I_3=1$ | $\lvert uus\rangle(2\lvert++-\rangle-\lvert+-+\rangle-\lvert-++\rangle)/3\sqrt{2}$ |
| $\Sigma^0$ | 1192 | $I=1$, $I_3=0$ | $(\lvert uds\rangle+\lvert dus\rangle)(2\lvert++-\rangle-\lvert+-+\rangle-\lvert-++\rangle)/6$ |
| $\Sigma^-$ | 1197 | $I=1$, $I_3=-1$ | $\lvert dds\rangle(2\lvert++-\rangle-\lvert+-+\rangle-\lvert-++\rangle)3\sqrt{2}$ |
| $\Xi^0$ | 1314 | $I=\frac{1}{2}$, $I_3=+\frac{1}{2}$ | $\lvert uss\rangle(\lvert++-\rangle+\lvert+-+\rangle-2\lvert-++\rangle)3\sqrt{2}$ |
| $\Xi^-$ | 1321 | $I=\frac{1}{2}$, $I_3=-\frac{1}{2}$ | $\lvert dss\rangle(\lvert++-\rangle+\lvert+-+\rangle-2\lvert-++\rangle)3\sqrt{2}$ |

It is this infinitude of degrees of freedom which cause many of the divergences encountered in field theories. In order to control this situation, imagine restricting the number of points to a finite "lattice" of space-time points - the quark degrees of freedom then reside at these points and interact with one another via the gluon fields, which can be considered connections or "links" between the points. It turns out that the entire theory may be formulated in terms of such site-specific quark fields and gluon links.  This lattice theory may be put into computer language and numerical methods used to solve for specific quantities. One important numerical technique used in this program is called <u>Monte Carlo</u>.  Formally, the quantities being measured in the lattice simulations can formally be expressed as an integral over all the degrees of freedom of the lattice system.  The dimensions of this integral are now very large (10's of millions in current simulations) but finite.  The Monte Carlo technique allows an estimation of these integrals by simply averaging over N likely values of the integrand.  It also allows an estimation of the likely variation in this value if the simulation were repeated many times.  Thus, lattice simulations

give values and error bars on physical quantities.  The error bars may be reduced of course for a larger, more computer-time intensive simulation, but they only fall $1/\sqrt{N}$, where N refers to the number of values in the integrand which have been used in the average.

<u>Weak Interactions</u>

We now come to the most interesting and complicated set of particle interactions, the weak interactions. This theory is also a gauge theory, like E&M (QED) and strong interactions (QCD). In fact, it is considered unified with the electromagnetic interaction and is often referred to as the weak-electromagnetic gauge theory. However, because the forces are mediated by massive particles, the $W^{\pm}$ and the $Z^0$, the force is extremely short ranged from Heisenberg's uncertainty principle. Although the weak interactions are a gauge theory, they have no associated exactly conserved quantities like electric charge or color charge. The reason for this is quite subtle. We will get back to this point momentarily.

The weak interactions involve both the quarks and leptons, and cause many types of decays of hadronic states. For example, consider the decay p -> n e $\bar{\nu}_e$.
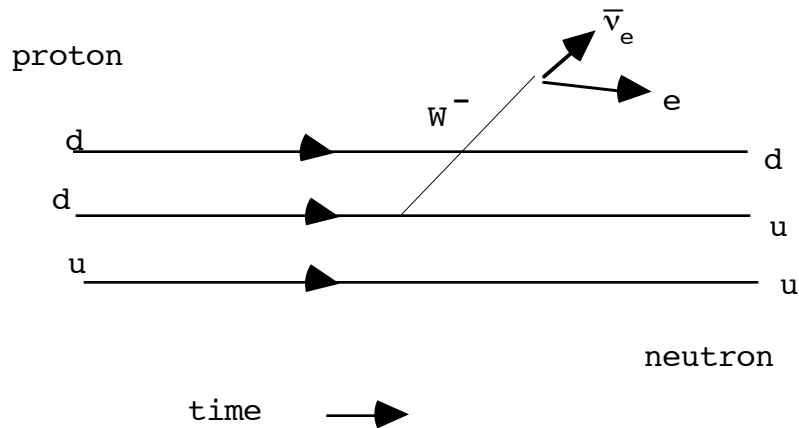
Here we see a proton entering from the left and a neutron, an electron, and an anti-electron neutrino exiting on the right. This occurs because of the coupling of the W⁻ particle to both quarks and leptons. This is called a "charged current" interaction because of the difference in the charge of the two other particles at the two W⁻ vertices. The W⁻ particle in this diagram is virtual, the other particles can be detected in the laboratory. The effects of the $Z^0$ particle, on the other hand, are much more subtle. We learned above that the photon has no additive quantum numbers. The $Z^0$ is also such a particle/anti-particle combination. This means it is present where ever photons are created or destroyed, but because of it's great mass it's range and possible detection are extremely limited.

Although both quarks and leptons participate in the weak interactions, they do so very differently. The quarks are of course all massive, but the neutrinos, for all intents and purposes, are massless. Remember our discussion of particle helicity while discussing photons above. For a massless paricle, whether it is a fermion or a boson (excluding spin 0), the only allowed physical states have helicity values ± 1, representing spin pointed along or anti-parallel to the direction of motion of the particle. This is very different from an electron, which also has only 2 spin degress of freedom, but these can point "up" or "down" relative to *any* coordinate axis. It turns out that all neutrinos (electron, muon, and tau) participate in interactions as if they were completely left-handed, i.e., their spin is pointed anti-parallel to their direction of motion. Anti-neutrinos are right-handed. This association of left-handedness with neutrinos and right-handedness with anti-neutrinos is intrinsically and maximally parity-violating. In fact, the weak interactions are known to violate all of the discrete symmetries we learned about at the beginning of this chapter. The only symmetry that survives is called CPT, a combined charge conjugation, parity and time reversal.

There is a coupling constant in the weak interactions which is analogous to the electromagnetic (proportional to $\sqrt{\alpha}$) and

strong coupling constants ($\sqrt{\alpha_{strong}}$). Let us call it $g_W$. However, because the charged current interactions the W's are all virtual, the *effective* coupling constant is actually ~ $\frac{g_W}{M_W^2}$, where $M_W$ is the mass of the $W^{\pm}$ particles. This is known as Fermi's coupling constant (I have left out some numerical coefficients in the actual value). The reason the mass of the $W^{\pm}$ particle appears squared in the denominator is due to the fact that 1) Feynman diagrams take place in momentum space and 2) that instead of Coulomb's law, ~1/r, which in momentum space is proportional to $1/(\vec{q}^2)$, an extremely massive particle behaves like a point in space, i.e., a Dirac delta function. The Fourier transform of a Dirac delta function is a constant in momentum space. This constant needs to have the same physical dimensions as our $1/(\vec{q}^2)$ function because they represent different limits of the same function*.

How did the $W^{\pm}$ and $Z^0$ particles get to be so massive if the weak interactions are just another gauge theory? It is because of an extremely subtle field theory effect known as <u>spontaneous symmetry breaking</u>. This effect depends upon the existence of the Higgs boson, which is another of the particles without any additive quantum numbers. It is thought that as the temperature of the universe cooled, The Higgs particle, $\phi$, shifts it's value: $\phi \rightarrow \phi' + v$, where "v" is just a number, called the vacuum expectation value (VEV). This is just like the spntaneous formation of little magnetic domains where the direction of the magnetic fields of the atoms are fixed as the temperature of a

---

* This function in field theory is known as the particle propagator, and has is proportional to

$$\frac{1}{q^2 + M_W^2},$$

where |q| represents the momentums allowed in a given interaction). For an infinite ranged interaction the mass $M_W$ would be zero and we would have a Coulomb interaction, whereas in the other limit the particle propagator goes is replaced with $M_W^2$ and represents a point interaction. Actually the $q^2$ above is a four-dimenensional Lorentz dot product, but the idea is the same.

magnet cools. However, not only does the Higgs field (which is originally a doublet of complex fields having a total of 4 real components) pick out a direction in it's isospin space like a magnet, but also picks up a VEV as it cools. This mechanism is thought to give rise to the masses of the $W^{\pm}$ and $Z^0$ particles as well as to all the quarks and leptons, via their original couplings (vertices) to the Higgs. In fact, in one of the most colorful phrases in particle physics, the $W^{\pm}$ and $Z^0$, which were originally massless and therefore had only 2 degrees of freedom (two helicities) are said to "eat" the 3 lost degrees of freedom of the Higgs, acquiring the correct 3 components for massive spin 1 particles. The left over massive particle is the physical Higgs.

One of the most subtle effects in weak interactions is flavor mixing. The weak and strong interactions deal with particle flavor differently. For the quark doublets we saw above that we had the flavor groupings, relevant to the strong interactions:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}. \tag{10.51}$$

For the weak interactions, the flavor groupings are different:

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix}. \tag{10.52}$$

The matrix which connects these particles is called the Cabbibo-Kobayashi-Maskawa, or CKM matrix. It is unitary and three dimensional. That is, one has

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud}, V_{us}, V_{ub} \\ V_{cd}, V_{cs}, V_{cb} \\ V_{td}, V_{ts}, V_{tb} \end{pmatrix} \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}. \tag{10.53}$$

The effects of this mixing on the interactions is too involved to discuss here, but see Appendix B on weak flavor mixing. The Standard Model offers no explanation for the values of the

parameters in this matrix. In fact, it is now thought that the same sort of mixing occurs for neutrinos as well, except there the strong interactions are not involved.  There is strong evidence that neutrinos have small but nonzero masses from the observation of neutrino flavor change, which can only take place if at least one of the neutrinos is massive. The proliferation of parameters in these matrices is one of the reasons most physicists do not think of the Standard Model as the final theory in particle physics, but simply another step toward a more fundamental theory.

That concludes our very short excursion into particle physics. We have learned about both the principles of quantum mechanics and how they apply to the particles around us. I hope you now have a better appreciation for the mathematical beauty and physical structure the natural world, which seems anything but random.

Appendix A: Allowed Standard Model Interactions

The generic allowed interaction vertices in Feynman diagrams for the various Standard model interactions are shown below.  All lines can be considered as "real" (external) or virtual (internal). These are just the primitive "vertices" - one must put them together to make complete Feynman diagrams representing physical processes. Of course, individual quark, gluon lines are confined and can not be external. The Higgs boson couplings are not shown.

Strong:

quark flavor'

quark flavor

g

g
g
g
g

g
g
g

Leptonic

Quarkish

EM

e⁻

e⁻

γ

u

u

γ

(any charged particle):

Leptonic                    Quarkish

Weak (Charged):

d                           e⁻

$W^+$                       $W^+$

u                           v

Leptonic                    Quarkish

Weak (Neutral):

ℓ                           u

$Z^O$                       $Z^O$

ℓ                           u

Other Weak:

W                           $W^+$        $W^-$

$Z^O$

W                           $W^+$        $W^-$

γ                           γ

γ                           $Z^0$

$W^+$        $W^-$          $W^+$        $W^-$
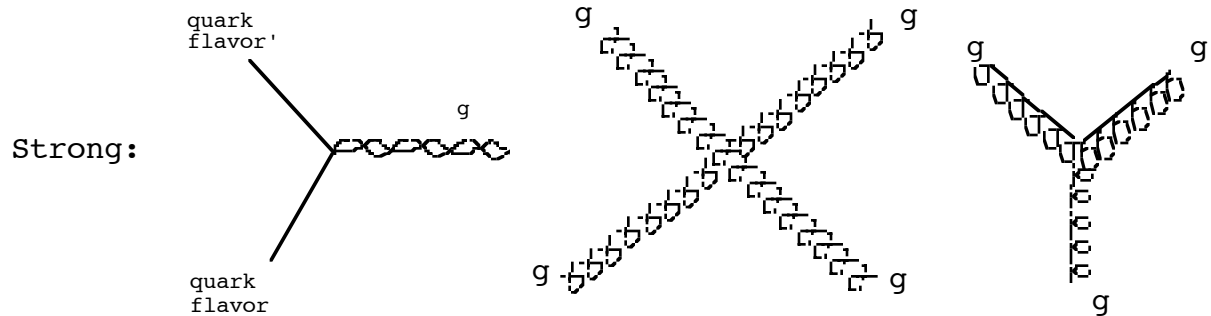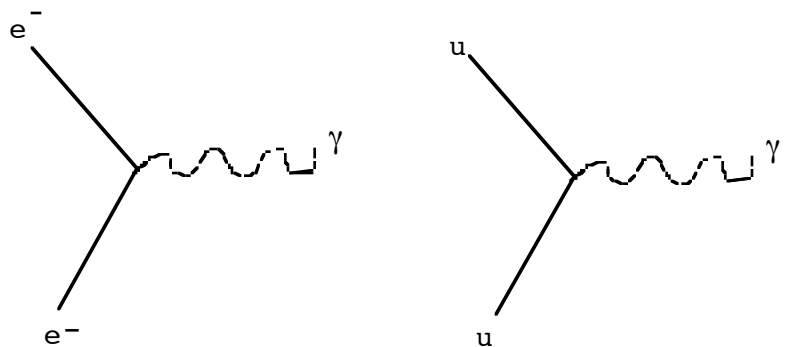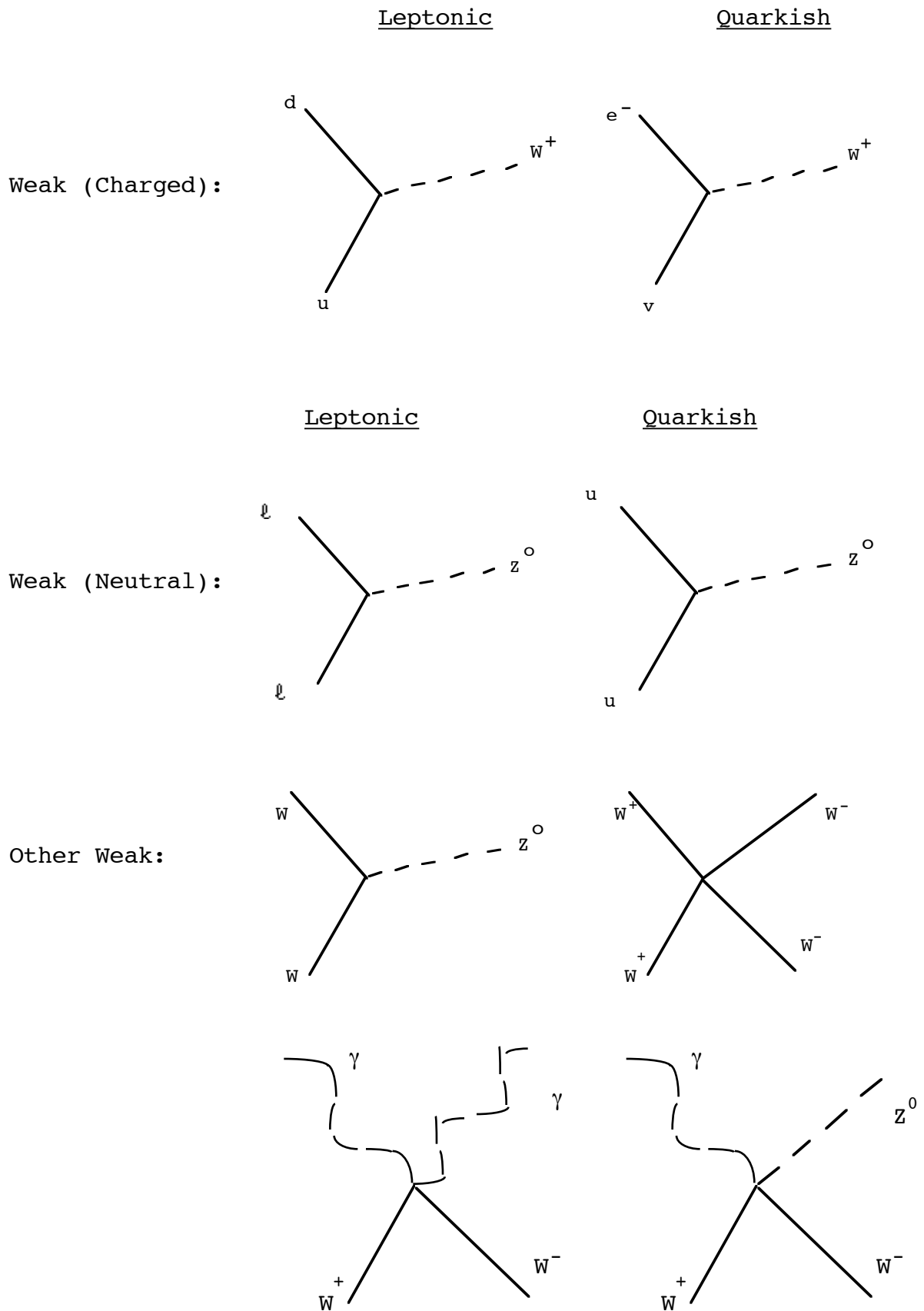
## Appendix B: Weak Flavor Mixing

In this Appendix, I will explain more about the significance of flavor mixing in the weak interactions. I will base this on a very insightful article by Howard Georgi in a Physics Today article which appeared back in April of 1988. The interesting point made here is that the structure of flavor interactions can be understood from a straightforward analogy with coupled harmonic oscillators. I will try to keep my explanation mostly in line with the concepts and ideas already introduced in these notes, so the emphases and presentation will differ a little from Georgi's article. I am presenting this in an Appendix because the text does not require this coverage for the flow of ideas encountered there and also because it is not guaranteed that all the concepts encountered here will have had an appropriate pedagogical introduction. Georgi's article is much more refined than my poor presentation, and I highly recommend the original for a more complete and expert point of view.

Our starting Hamiltonian is just

$$H_0 = \Omega \hbar \sum_j A_j^+ A_j \,, \quad \text{(Drop const. term; called "normal ordering")}$$

which we recognize as Hamiltonian for the three dimensional harmonic oscillator. There are symmetries here. Let

$$A_j \rightarrow \sum_k U_{jk} A_k \,, \tag{B.1}$$

$$\Rightarrow A_j^+ \rightarrow \sum_k A_k^+ U_{jk}^+ = \sum_k A_k^+ U_{kj}^+ \,. \tag{B.2}$$

$$H_0 \rightarrow \Omega \hbar \sum_{j,k,\ell} A_k^+ U_{kj}^+ U_{j\ell} A_\ell = \Omega \hbar \sum_{k,\ell} A_k^+ A_\ell \left( \sum_j U_{kj}^+ U_{j\ell} \right). \tag{B.3}$$

U is unitary:

$$U^{\dagger}U = 1, \quad \Rightarrow \quad \sum_j U^{\dagger}_{kj}U_{j\ell} = \delta_{k\ell}, \qquad (B.4)$$

$$\Rightarrow H_0 \to \Omega\hbar \sum_{k,\ell} A^{\dagger}_k A_{\ell}\delta_{k\ell} = H_0 : \text{unchanged} \qquad (B.5)$$

U is a unitary, 3x3 matrix $\Rightarrow$ the group known as SU(3). Actually, there is a further symmetry here.  Let

$$A_j \to e^{i\alpha}A_j, \quad \Rightarrow A^{+}_j \to e^{-i\alpha}A^{+}_j, \qquad (B.7)$$

$$H_0 \to \sum_j A^{\dagger}_j A_j e^{-i\alpha}e^{i\alpha} = H_0 : \text{unchanged}, \qquad (B.8)$$

Real Symmetry: SU(3)xU(1).  Consequence is that $N_k$ is conserved:

$$[H_0, N_k] = \hbar\Omega \sum_j \left[ A^{\dagger}_j A_j, A^{+}_k A_k \right]. \qquad (B.9)$$

We have shown in the last chapter that

$$\left[ N_j, N_k \right] = \left[ A^{\dagger}_j A_j, A^{+}_k A_k \right] = 0. \qquad (B.10)$$

Ch.7, prob.1 shows that the degeneracy factor for the state $n^{\underline{th}}$ quantum state is $\frac{1}{2}(n + 1)(n + 2)$. Let's go through it here to refresh ourselves.

$$E = \hbar\omega(n_1 + n_2 + n_3 + \tfrac{3}{2}) = \hbar\omega(n + \tfrac{3}{2}),$$
$$n_1 + n_2 + n_3 = \{0, 1, 2 \ldots\}.$$

| | | |
|---|---|---|
| Set $n_1 = 0$ | How many ways? | n+1 |
| Set $n_1 = 1$ | " | n |
| . | . | . |
| . | . | . |
| . | . | . |
| Set $n_1 = 1$ | How many ways? | 1 |

$$\sum_{i=1}^{n+1} i = \tfrac{1}{2}(n + 1)(n + 2). \qquad (\text{Given } \sum_{i=1}^{n} i = \tfrac{1}{2} n(n + 1))$$

First few energies:

$$6\begin{cases} n_1 = 2, n_2 = 2, n_3 = 2 \quad (\text{others} = 0) \\ n_1 = n_2 = 1, n_3 = 0 \\ n_1 = n_3 = 1, n_2 = 0 \\ n_2 = n_3 = 1, n_1 = 0 \quad E = \tfrac{7}{2} \hbar\omega \end{cases}$$

$$3\begin{cases} n_1 = 1, n_2 = n_3 = 0 \\ n_2 = 1, n_1 = n_3 = 0 \\ n_3 = 1, n_1 = n_2 = 0 \quad E = \tfrac{5}{2} \hbar\omega \end{cases}$$

$$1\begin{cases} n_1 = 1, n_2 = n_3 = 0 \quad E = \tfrac{3}{2} \hbar\omega \end{cases}$$

Now cease to think of these as energy states. Consider $n_1$, $n_2$, $n_3$ to be particle occupation numbers of a B.E. system. Remember, I showed in the last chapter that the algebra of

$$a_i \text{ and } a_i^+ \text{ (raising and lowering operators)}$$

for the harmonic oscillator is identical to

$$A \text{ and } A^+ \text{ (creation and annihilation operators)}$$

for the multi-particle bosonic states. If instead $n_1$, $n_2$, $n_3$ were the particle occupation numbers of a Fermi-Dirac system, we would have the Cartesian classification:

|  |  | FD deg. | BE deg. |
|---|---|---|---|
| $E = \frac{9}{2} \hbar\omega$ $\quad |1,1,1\rangle$ | | 1 | 10 |
| $E = \frac{7}{2} \hbar\omega$ $\quad |1,1,0\rangle$ | | | |
| $|0,1,1\rangle$ | | 3 | 6 |
| $|1,0,1\rangle$ | | | |
| $E = \frac{5}{2} \hbar\omega$ $\quad |1,0,0\rangle$ | | | |
| $|0,1,0\rangle$ | | 3 | 3 |
| $|0,0,1\rangle$ | | | |
| $E = \frac{3}{2} \hbar\omega$ $\quad |0,0,0\rangle$ | | 1 | 1 |

Now introduce an interaction:

$$H_{int} = \vec{B}\cdot\left(\vec{p}\times\vec{r}\right) = -\vec{B}\cdot\vec{L} \qquad\qquad (B.11)$$

Why is such a term reasonable?  I argued in Ch.8 (based upon the Ch.1 discussion) that a magnetic field interacts with a magnetic dipole according to

$$H_{int} = \vec{m}\cdot\vec{B}. \qquad\qquad (B.12)$$

Given a gyromagnetic ratio such that

$$\vec{m} = \gamma\,\vec{L}, \qquad\qquad (B.13)$$

$(\gamma = \dfrac{q}{2mc}$ classically) then

$$H_{int} = \gamma\,\vec{B}\cdot\vec{L}. \qquad\qquad (B.14)$$

Given the harmonic oscillator raising and lowering operators $\left(K = m\Omega^2\right)$,

$$A_j = \frac{1}{\sqrt{2m\hbar\Omega}}\, p_j - i\sqrt{\frac{K}{2\hbar\Omega}}\, r_j, \qquad (B.15)$$

$$A_j^+ = \frac{1}{\sqrt{2m\hbar\Omega}}\, p_j + i\sqrt{\frac{K}{2\hbar\Omega}}\, r_j, \qquad (B.16)$$

we can evaluate $H_{int} = \vec{B}\cdot\left(\vec{p}\times\vec{r}\right)$ as (problem):

$$\vec{B}\cdot\left(\vec{p}\times\vec{r}\right) = i\hbar\sum_{i,j,k}\varepsilon_{ijk}B_i A_j A_k. \qquad (B.17)$$

Writing this as

$$H_{int} = \sum_{j,k} A_j^+ M_{j,k} A_k, \qquad (B.18)$$

we identify

$$M_{jh} = ik\sum_i \varepsilon_{ijk}B_i. \qquad (B.19)$$

Notice that

$$M_{jh}^* = M_{kj} \Rightarrow \text{Hermitian.}$$

Also notice (do the transformation again) that under

$$A_j \rightarrow \sum_k U_{jk}A_k, \quad A_j^+ \rightarrow \sum A_k^+ U_{kj}^+, \qquad (B.20)$$

we have, $H_0 \rightarrow H_0$ but

$$H_{int} \rightarrow \sum_{j,h} A_i^+ U_{\ell j}^+ M_{jk} U_{km} A_m = \sum_{\ell,m} A_\ell^+ \tilde{M}_{\ell m} A_m, \qquad (B.21)$$

where

$$\tilde{M}_{\ell m} = \sum_{j,k} U^+_{\ell j} M_{jk} U_{km}. \tag{B.22}$$

As matrices: $\tilde{M} = U^+ M U$.

Since we don't get $H_{int} \rightarrow H_{int}$, we have that the SU(3) symmetry is broken by $H_{int}$. The consequence is that the individual $N_k$ are no longer conserved, just the total $N = \sum_k N_k$.

That is, we have

$$[H_{int}, \ N_j] = \left[ i\hbar \sum_{i,k,\ell} \varepsilon_{ik\ell} B_i A^+_k A_\ell, \ A^+_j A_j \right]$$

$$= i\hbar \sum_{i,k,\ell} \varepsilon_{ik\ell} B_i \left[ A^+_k A_\ell, \ A^+_j A_j \right]. \tag{B.23}$$

But

$$\left[ A^+_i A_j, \ A^+_k A_\ell \right] = \delta_{jk} A^+_i A_\ell - \delta_{i\ell} A^+_k A_j, \tag{B.24}$$

so $(i = j)$

$$\left[ A^+_j A_j, \ A^+_k A_\ell \right] = \delta_{jk} A^+_j A_\ell - \delta_{j\ell} A^+_k A_j.$$

$$\Rightarrow [H_{int} \ N_j] = -i\hbar \sum_{i,k,\ell} \varepsilon_{ik\ell} B_i \left\{ \delta_{jk} A^+_j A_\ell - \delta_{j\ell} A^+_k A_j \right\},$$

$$= -i\hbar \sum_{i,\ell} \varepsilon_{ik\ell} B_i \left( A^+_j A_\ell + A^+_\ell A_j \right) \neq 0. \tag{B.25}$$

But notice, however, that

$$\sum_j [H_{int}, N_j] = -i\hbar \sum_{i,\ell,j} \varepsilon_{ij\ell} B_i \left( A^+_j A_\ell + A^+_\ell A_j \right) = 0. \tag{B.26}$$

Now let's go back to

$$\tilde{M} = U^+ M U. \qquad (B.27)$$

Can show that we can diagonalize any Hermitian matrix, M, in this fashion.  Why do we want to do this?  Because if we can find a U such that

$$\tilde{M}_{ij} = \omega_i \delta_{ij}, \qquad (B.28)$$

then

$$H_{int} = \sum_{\ell,m} A_\ell^+ \tilde{M}_{\ell,m} A_m \rightarrow \sum_{\ell,m} \omega_\ell A_\ell^+ A_\ell. \qquad (B.29)$$

Let's find the w's for our explicit case above.

$$M = i\hbar \begin{pmatrix} 0 & B_3 & -B_2 \\ -B_3 & 0 & B_1 \\ B_2 & -B_1 & 0 \end{pmatrix}. \qquad (B.30)$$

Get the w's from the <u>characteristic</u> <u>equation</u>. (See p.4.11 of this text). We form,

$$\det\left(M - \omega_i\right) = 0, \qquad (B.31)$$

$$\Rightarrow \det \begin{pmatrix} \omega & B_3 & -B_2 \\ -B_3 & \omega & B_1 \\ B_2 & -B_1 & \omega \end{pmatrix} = 0, \qquad (B.32)$$

$$\Rightarrow \omega^3 - \hbar^2 \omega\left(B_1^2 + B_2^2 + B_3^2\right) = 0. \qquad (B.32)$$

Three solutions:

$$\begin{cases} \omega_1 = |\vec{B}|, \\ \omega_2 = -|\vec{B}|, \\ \omega_3 = 0. \end{cases} \qquad (B.33)$$

Therefore

$$H_{int} = \sum_i \omega_i A_i^+ A_i = |\vec{B}|\left(A_1^+ A_1 - A_2^+ A_2\right). \qquad (B.34)$$

Then using 1st order perturbation theory,

$$<H_{int}> = |\vec{B}|\left(n_1 - n_2\right). \qquad (B.35)$$

The energy levels now appear as:



$E = \frac{7}{2}\hbar\omega$    (6)    $n_1=2$ ; $n_1=1$ , $n_3=1$ ; $n_3=2$, $n_1=n_2=1$ ; $n_2=1$ , $n_3=1$ ; $n_2=2$    $2|\vec{B}|$ , $|\vec{B}|$ , $-|$ , $-2|\vec{B}|$

$E = \frac{5}{2}\hbar\omega$    (3)    $n_1=1$ ; $n_3=1$ ; $n_2=2$

$E = \frac{3}{2}\hbar\omega$

before       after
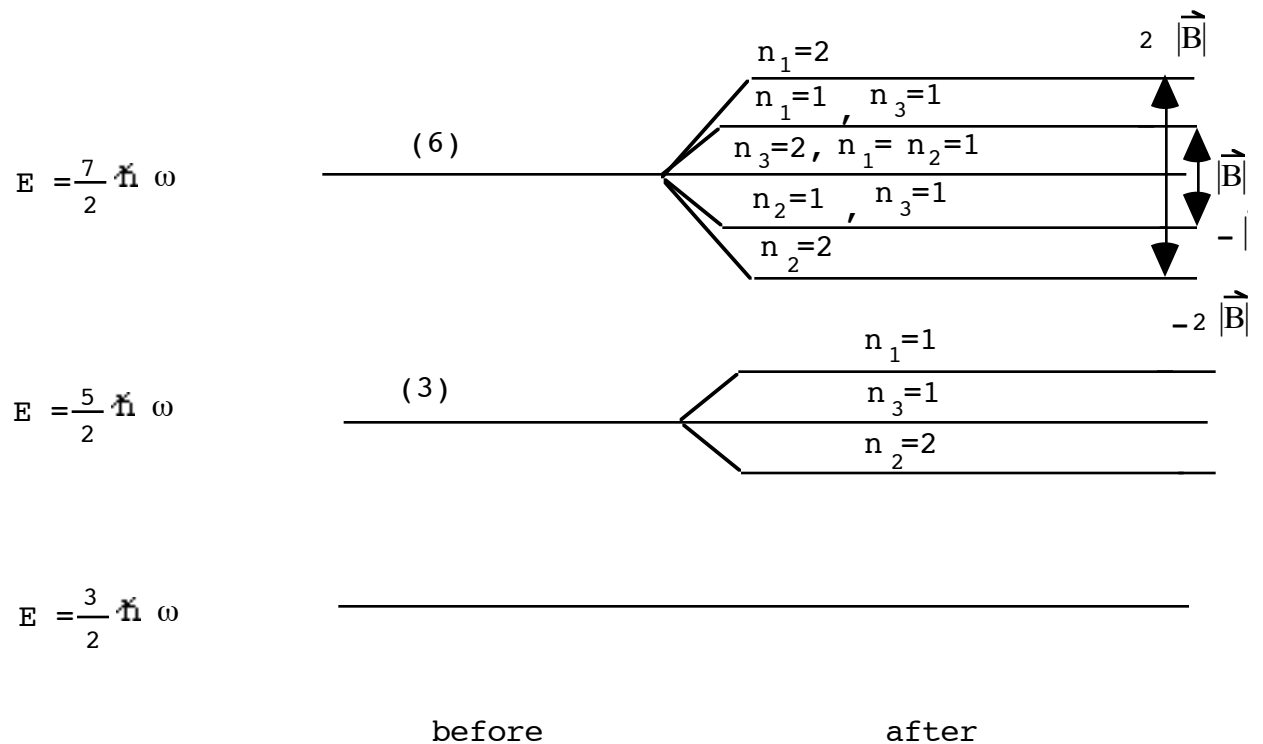
In general, the splitting is

$$<H_{int}> = \left\langle \sum_i \omega_i A_i^+ A_i \right\rangle = \sum_i \hbar\omega_i N_i = \hbar\omega\left(N_\ell\omega_\ell + N_2\omega_2 + N_3\omega_3\right), \quad \text{(B.36)}$$

$( N_1 + N_2 + N_3 = N.)$   For $N = 1$, the splittings are

$$N_1 = 1 \qquad \hbar\,\omega_1$$
$$N_2 = 1 \qquad \hbar\,\omega_2$$
$$N_3 = 1 \qquad \hbar\,\omega_3$$

For $N = 2$, the splittings are

$$N_1 = 2 \qquad 2\hbar\,\omega_1$$
$$N_2 = 2 \qquad 2\hbar\,\omega_2$$
$$N_3 = 2 \qquad 2\hbar\,\omega_3$$
$$N_1 = N_2 = 1 \qquad \hbar\left(\omega_1 + \omega_2\right)$$
$$N_1 = N_3 = 1 \qquad \hbar\left(\omega_1 + \omega_3\right)$$
$$N_2 = N_3 = 1 \qquad \hbar\left(\omega_2 + \omega_3\right)$$

Thus the new Hamiltonian is

$$H_A = \hbar \sum_j \left(\Omega + \omega_j\right)A_j^+ A_j. \qquad \text{(B.37)}$$

Add another interaction, $H_B$, analogous to the first:

$$\left[A_i, B_j\right] = 0, \quad \left[A_i, B_j^+\right] = 0, \qquad \text{(B.38)}$$

$$H_A + H_B = \hbar \sum_j \left[\left(\Omega_A + \omega_{Aj}\right)A_j^+ A_j + \left(\Omega_B + \omega_{Bj}\right)B_j^+ B_j\right]. \qquad \text{(B.39)}$$

We will see that we can view the $H_A$ Hamiltonian as representing the charge 2/3 quarks (u, c, and t), and the $H_B$ Hamiltonian as representing the charge –1/3 quarks (d, s, and b), at least when we are talking about the N=1 states. States are represented by:

$$\left| N_{A1}, N_{A2}, N_{A3}, N_{B1}, N_{B2}, N_{B3} \right\rangle. \tag{B.40}$$

Now also add (we're almost there!)

$$H_w = \hbar W \sum_j \left( A_j^+ B_j + B_j^+ A_j \right), \tag{B.41}$$

Notice $\left( H_w^+ = H_w \right)$. If we were to drop the $\omega_{Aj}$, $\omega_{Bj}$ terms above, we would have a symmetry, as follows:

$$\left. \begin{aligned} A_j &\rightarrow \sum_k U_{jk} A_k \\ B_j &\rightarrow \sum_k U_{jk} B_k \end{aligned} \right\} \quad \text{same } U_{jk}. \tag{B.42}$$

Then

$$H_w \rightarrow \hbar W \sum_{j,k} \left[ A_k^+ U_{kj}^+ U_{j1} B_1 + B_\ell^+ U_{\ell j}^+ A_k \right] = H_w. \tag{B.43}$$

This is a <u>combined</u> $SU_{A+B}(3)$ symmetry. This has the consequence that now only $\left( N_{Ai} + N_{Bi} \right)$ is conserved. That is (problem)

$$\left[ N_{Ai} + N_{Bi}, H_w \right] = 0. \tag{B.44}$$

Now let the $\omega_{Aj}$ and $\omega_{Bj}$ both be non-zero. The resultant full Hamiltonian can be written

$$H = \hbar \sum_j \left( A_j^+ \ B_j^+ \right) \begin{pmatrix} \Omega_A + \omega_{Aj} & W \\ W & \Omega_B + \omega_{Bj} \end{pmatrix} \begin{pmatrix} A_j \\ B_j \end{pmatrix}. \tag{B.45}$$

Now, it is only the <u>total</u> $N_A + N_B = \sum_i \left( N_{Ai} + N_{Bi} \right)$ that is conserved. Let

$$\hat{M} = \begin{pmatrix} \Omega_A + \omega_{Aj} & W \\ W & \Omega_B + \omega_{Bj} \end{pmatrix}.$$  (B.46)

Just as before, diagonalize it to get the energies.  (Call $\Omega_A + \omega_{Aj} = a_j$, $\Omega_B + \omega_{Bj} = b_j$)

$$\det \begin{pmatrix} a_j - \lambda & W \\ W & b_j - \lambda \end{pmatrix} = 0,$$  (B.47)

$$\Rightarrow \lambda^2 - \lambda\left(a_j + b_j\right) - W^2 + a_j b_j = 0,$$  (B.48)

$$\lambda = \frac{a_j + b_j}{2} \pm \frac{1}{2}\sqrt{\left(a_j - b_j\right)^2 + 4W^2}.$$  (eigenvalues)  (B.49)

What about eigenvectors? Can verify that H can be written as

$$H = \frac{\hbar}{2}\sum_j \begin{pmatrix} \alpha_j^+ & \beta_j^+ \end{pmatrix} \begin{pmatrix} a_j + b_j + \sqrt{\left(a_j + b_j\right)^2 + 4W^2} & 0 \\ 0 & a_j + b_j - \sqrt{\left(a_j + b_j\right)^2 + 4W^2} \end{pmatrix} \begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix},$$  (B.50)

where

$$\alpha_j^+ = \cos\Theta_j A_j^+ + \sin\Theta_j B_j^+,$$  (B.51)

$$\beta_j^+ = -\sin\Theta_j A_j^+ + \cos\Theta_j B_j^+.$$  (B.52)

The angle $\Theta$ is given by

$$\tan 2\Theta = \frac{2W}{\left(\beta_j - \alpha_j\right)}.$$  (B.53)

Since

$$\cos^2 2\Theta_j = \frac{1}{1 + \tan^2 2\Theta_j}, \tag{B.54}$$

we have

$$\cos^2 2\Theta_j = \frac{\left(a_j - b_j\right)^2}{\left(a_j - b_j\right)^2 + 4W^2}, \tag{B.55}$$

$$\cos 2\Theta_j = \pm \frac{\left(\alpha_j - b_j\right)}{\sqrt{4\omega^2 + \left(\alpha_j - b_j\right)^2}}, \quad \text{(Choose + sign)} \tag{B.56}$$

$$\Rightarrow \sin 2\Theta_j = \mp \frac{2\omega}{\sqrt{4\omega^2 + \left(\alpha_j - b_j\right)^2}}. \quad \text{(Choose - sign)} \tag{B.57}$$

(Must choose the above signs for the form of H written above. Can see this in the W → 0 limit). Verify this form:

$$\sum_j \left\{ \alpha_j^+ \alpha_j \left( a_j + b_j + \sqrt{\left(a_j - b_j\right)^2 + 4W^2} \right) + B_j^+ B_j \left( a_j + b_j - \sqrt{\left(a_j - b_j\right)^2 + 4W^2} \right) \right\}, \tag{B.58}$$

$$\alpha_j^+ \alpha_j = \cos^2 \Theta_j A_j^+ A_j + \sin^2 \Theta_j B_j^+ B_j + \cos \Theta_j \sin \Theta_j \left( A_j^+ A_j + B_j^+ B_j \right), \tag{B.59}$$

$$\beta_j^+ \beta_j = \sin^2 \Theta_j A_j^+ A_j + \cos^2 \Theta_j B_j^+ B_j - \cos \Theta_j \sin \Theta_j \left( A_j^+ A_j + B_j^+ B_j \right). \tag{B.60}$$

Finish the rest up as a problem. Should get earlier expression (problem):

$$H = \hbar \sum_j \left\{ a_j A_j^+ A_j + b_j B_j^+ B_j + W\left( A_j^+ B_j + B_j^+ A_j \right) \right\}. \tag{B.61}$$

Remember Fermi's Golden Rule. From chapter 9,

$$\text{rate} \sim \left| \langle 2|T|1 \rangle \right|^2, \tag{B.62}$$

for a 2->1 transition.  Let's take an example:

$$\left|1_{\alpha j}\right\rangle = \alpha_j^+|0\rangle, \quad \left|1_{Bj}\right\rangle = \beta_j^+|0\rangle, \tag{B.63}$$

$$\left\langle 1_{\alpha j}|T|1_{Bj}\right\rangle =$$

$$\left\langle 0 \mid \left(\cos \Theta_j A_j + \sin \Theta_j B_j\right)\sum_k \left(\underbrace{A_k^+ A_k}_{N_{Ak}} - \underbrace{B_k^+ B_k}_{N_{Ak}}\right)\left(-\sin \Theta_j A_j^+ + \cos \Theta_j B_j^+\right)\mid 0\right\rangle. \tag{B.64}$$

$$\left(\left[N_{Ak}, A_j^+\right] = A_j^+ \delta_{kj}, \left[N_{Bk}, B_j^+\right] = B_j^+ \delta_{kj}.\right) \text{ Thus}$$

$$\left\langle 1_{\alpha j}|T|1_{Bj}\right\rangle = \left\langle 0 \mid \left(\cos \Theta_j A_j + \sin \Theta_j B_j\right)\left(-\sin \Theta_j A_j^+ - \cos \Theta_j B_j^+\right) \mid 0\right\rangle,$$

$$= -\cos \Theta_j - \sin \Theta_j \left\langle 0 \mid \left(A_j A_j^+ + B_j B_j^+\right) \mid 0\right\rangle,$$

$$= -2 \cos \Theta_j \sin \Theta_j = -\sin\left(2\Theta_j\right). \tag{B.65}$$

Likewise, can show (problem; all i ≠ j):

$$\left\langle 1_{A_i}|T|1_{A_j}\right\rangle = 0, \quad \left\langle 1_{B_i}|T|1_{B_j}\right\rangle = 0, \quad \left\langle 1_{A_i}|T|1_{B_j}\right\rangle = 0. \tag{B.66}$$

This gives rise to the picture of interactions (take the N=1 case again):
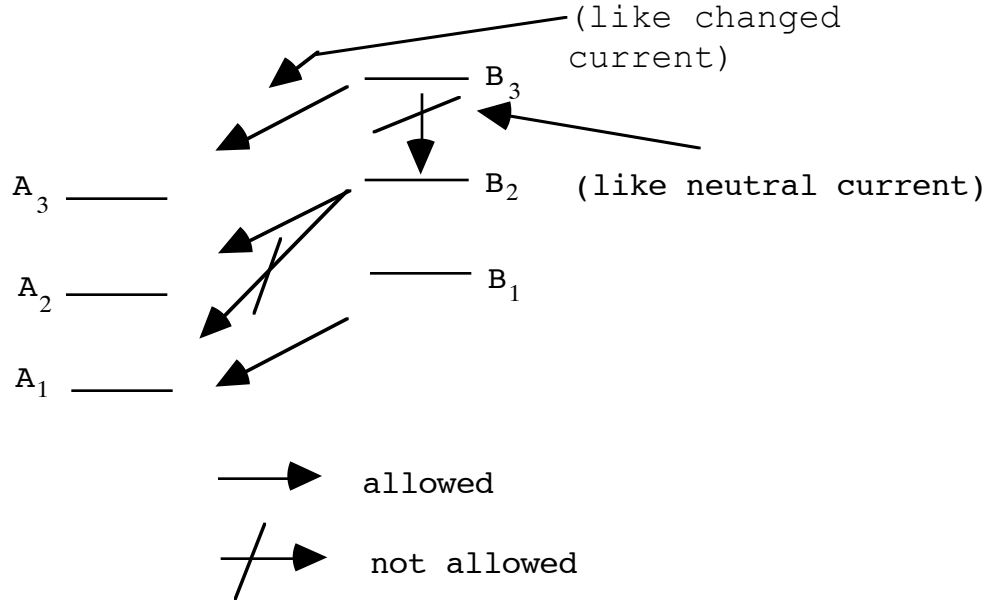
We see that interactions involving the same generation of quark are allowed (the charged current interactions, mediated by the $W^{\pm}$ bosons). However, not all the charged current interactions are possible at this point, but only those with i=j (within each generation). We do not yet have the small neutral current interactions involving different generations. However, they're not hard to model. We need only make one more change to our Hamiltonian.

Here'e the intersting part. Let us simply change the $H_W$ above to

$$H_W = \hbar W \sum_{j,k} \left( A_j^+ V_{jk} B_k + B_j^+ V_{jk}^+ A_k \right), \qquad \left( H_W^+ = H_W \right) \qquad (B.67)$$

where $V_{jk}$ is a unitary matrix. We now have mixings of generations involving all the charged current interactions. This would be completely equivalent to old $H_W$ if $\omega_{Aj}, \omega_{Bj}$ terms in $H_A$ and $H_B$ are dropped. Let

$$A_j \rightarrow \sum_k V_{jk} A_k, \; A_j^+ \rightarrow \sum_k A_k^+ V_{kj}^+, \qquad (B.68)$$

then

$$H_W = \hbar W \sum_{j,k} \left( A_\ell^+ \underbrace{V_{\ell j}^+ V_{jk}}_{\delta_{\ell k}} B_k + B_j^+ \underbrace{V_{jk}^+ V_{k\ell}}_{\delta_{j\ell}} A_\ell \right),$$

$$= \hbar W \sum_{j} \left( A_j^+ B_j + B_j^+ A_j \right) \quad , \text{ old form.} \qquad (B.69)$$

That is, without the $SU_A(3)$, $SU_B(3)$ symmetry breaking terms, this new form is completely equivalent to the old and would not add anything new. However, because of the $\omega_{Aj}, \omega_{Bj}$ terms, we now have

all charged transitions on the previous picture, even those between different generations ($i \neq j$). As I said, the analog of the transitions from the $\beta_i$'s to the $\alpha_j$'s are the charged current interactions, mediated by the $W^\pm$ particles. The $V_{jk}$ matrix is analogous tp the CKM matrix in weak interactions. There are still no direct generation-changing neutral currents (those connecting $\alpha_i$ and $\alpha_j$ or $\beta_i$ and $\beta_j$ for $i \neq j$), but now notice the following. We can now have transitions which look like the following;
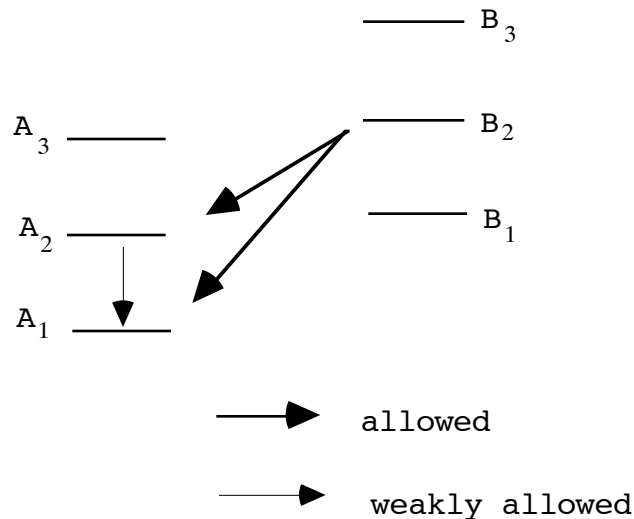


$B_2$ is seen to be mixed into $A_1$ and $A_2$, which means, from the magic of quantum mechanics, that $A_1$ and $A_2$ also are mixed. A "second order transition" would first take $A_2$ into the higher state $B_2$, then back down to $A_1$. However, such transitions are doubly

suppressed. First, this transition only occurs "indirectly" by this two step process, and, as we saw above, the effect is proportional to the presence of the off-diagonal elements of $V_{jk}$. These in fact only contribute because of the presence of the $\omega_{Aj}, \omega_{Bj}$ terms. However, it's not just that the $\omega_{Aj}, \omega_{Bj}$ terms are there, but that they are all *different* that allows these small transitions. An example of such a process in the standard Model is $K^+ \to \pi^+ \nu \bar{\nu}$, which has been observed.

## Problems for appendix B

1. Let's see if we can also understand the $\frac{1}{2}(n+1)(n+2)$ degeneracy factor from the radial classification of states.  Go back to Ch.7, prob.15.  Compared (3-D oscill.)

$$\left[\frac{d^2}{d\rho^2} - \frac{\ell(\ell + 1)}{\rho^2} - \rho^2 + 2(n + \tfrac{3}{2})\right]R_{n\ell}(\rho) = 0,$$

to (2-D oscill.)

$$\left[\frac{d^2}{d\rho^2} - \frac{(m^2 - \frac{1}{4})}{\rho^2} + 2(|m| + 2n_r + 1) - \rho^2\right]\sqrt{\rho}\; P_{n_r m}(\rho) = 0.$$

Should have found:          2-D          3-D

$$m^2 - \tfrac{1}{4} \rightarrow \ell(\ell + 1),$$

$$2n_r + |m| + 1 \rightarrow n + \tfrac{3}{2},$$

or

$$|m| \rightarrow \ell + \tfrac{1}{2}, \quad n_r \rightarrow \frac{n - \ell}{2}. \quad (n_r = 0, 1, 2, 3, \ldots)$$

(a) Considering the n even and n odd cases separately, show that the degeneracy is $\frac{1}{2}(n+1)(n+2)$ .

(b) Write the radial quantum numbers $(n, \ell, m)$ of the $E = \frac{3}{2}\,\hbar\omega$, $\frac{5}{2}\,\hbar\omega$, $\frac{7}{2}\,\hbar\omega$ states out explicitly.

2. Show

$$\vec{B}\cdot\left(\vec{p}\times \vec{r}\right) = i\hbar \sum_{i,j,k} \varepsilon_{i,j,h} B_i A_j A_k.$$

3. Show

$$\left[N_{Ai} + N_{Bi}, H_w\right] = 0.$$

4. After the change of variables described in the text, show that the Hamiltonian can still be written as:

$$H = \hbar \sum_j \left\{ a_j A_j^\dagger A_j + b_j B_j^\dagger B_j + W\left(A_j^\dagger B_j + B_j^\dagger A_j\right) \right\}.$$

5. Explicitly show (before the introduction of the $V_{jk}$ matrix)

$$\left\langle 1_{\alpha i} | T | 1_{\alpha i} \right\rangle = 0 \quad (i \neq j)$$
$$\left\langle 1_{B_i} | T | 1_{Bi} \right\rangle = 0 \quad (i \neq j)$$
$$\left\langle 1_{\alpha i} | T | 1_{Bi} \right\rangle = 0 \quad (i \neq j).$$

6. Show the Feynman diagram for $K^+ \to \pi^+ \nu \bar{\nu}$.

**Problems**

1. Starting with the definition of the determinant of a 3x3 matrix,

$$(\det a) = \sum_{\ell,m,n} \varepsilon_{\ell mn}\, a_{1\ell}\, a_{2m}\, a_{3n}\,.$$

show that,

$$\sum_{j,k} \varepsilon_{ijk}\, a_{j\ell}\, a_{km} = \det(a) \sum_{n} \varepsilon_{n\ell m}\, a_{in}\,.$$

where $a_{ij}$ are elements of a general orthogonal transformation.

2. Considering the matrix element $\langle s'|\vec{J}\cdot\vec{p}|s'\rangle$, show that a nonzero value implies parity nonconservation. (Hints; First consider $|s'\rangle$ a good parity state, then a mixture.)

3. For the time reversal operation, show:

(a) $(AB)^T = B^T A^T$.

(b) $\left(\bar{x}\times\bar{p}\right)^T = -\bar{x}\times\bar{p}$.

4. There is a famous and useful theorem in mathematics called the Wigner-Eckart theorem. It generalizes the spin addition considerations of Ch. 8, and can be applied to spin or isospin matrix elements. (For isospin, see probs. 7, 8, and 9 below.)

$$\langle J',M'|T_{Kq}|J,M\rangle = \langle JK;Mq\,|JK;J'M'\rangle \frac{\langle J'|T_K|J\rangle}{\sqrt{2J+1}},$$

where the $\langle J'|T_K|J\rangle$ is the "reduced matrix element" (independent of M, M'), the $T_{Kq}$ are spherical tensors (see Sakurai's <u>Modern Quantum Mechanics</u>, Ch.3), and the $\langle JK;Mq\,|JK;J'M'\rangle$ are just the

Clebsch–Gordon coefficients of Ch.8. This theorem immediately implies the so-called triangle inequaltiy $|J-K| \leq J' \leq J+K$.

You are given the electromagnetic Hamiltonian density,

$$\mathcal{H}^{em} = \sum_{L,m} \sum_{X=(E,M)} A_{Lm}^{(X)} \; T_{Lm}^{(X)} \, ,$$

where the $A_{Lm}^{(E)}$, $A_{Lm}^{(M)}$ (E=electric, M=magnetic) are expansion coefficients ($A_{00}^{(M)} = 0$) and the $T_{Lm}^{(E)}$, $T_{Lm}^{(M)}$ are spherical tensors. The operators $T_{Lm}^{(E,M)}$ have the properties

$$\pi T_{Lm}^{(E)} \; \pi = (-1)^L \; T_{Lm}^{(E)} \; (L \geq 0) \, ,$$
$$\pi T_{Lm}^{(M)} \; \pi = (-1)^{L+1} \; T_{Lm}^{(M)} \; (L \geq 1) \, ,$$

where $\pi$ is the parity operator ($\pi^\dagger = \pi$). Consider the expectation value (modeling static em moments):

$$<J,m_J,\pi_J \,|\, \mathcal{H}^{em} \,|\, J,m_J,\pi_J>,$$

where $m_J = J$. (The states $|J,m_J,\pi_J>$ have good total angular momentum, $J_z$, and parity.)

(a) For general J ($J=0,\frac{1}{2},1,\frac{3}{2},2,...$), what is the largest L that can contribute?
(b) For general J, how many terms in $\mathcal{H}^{em}$ give nonzero contributions? Specify them.

Consider the more general matrix element (modeling em transitions) where $\mathcal{H}^{rad}$ is the same as $\mathcal{H}^{em}$ but without the L=0 E term:

$$<J_2,m_{J2},\pi_{J2} \,|\, \mathcal{H}^{rad} \,|\, J_1,m_{J1},\pi_{J1}>,$$

(c) For general $J_1$, $J_2$, how many terms in $\mathcal{H}^{rad}$ give nonzero contributions? Specify them.

(d) Assuming the terms with smallest L give the leading contribution, give the transition selection rules for the various $\Delta J = |J_2 - J_1|$ when $\pi_{J1}\pi_{J2}=1$ or $\pi_{J1}\pi_{J2}=-1$.


5. Particle physicists usually express dimensionful quantities in terms of MeV (millions of electron volts), $\hbar$, and c. For example,

$$1 \text{ kg} = 5.61 \times 10^{29} \frac{\text{MeV}}{c^2},$$

$$1 \text{ cm}^{-1} = 1.97 \times 10^{-11} \frac{\text{MeV}}{\hbar c^2},$$

$$1 \text{ sec}^{-1} = 6.58 \times 10^{-22} \frac{\text{MeV}}{\hbar}.$$


Using these values, convert the QCD string tension of 15 tons (1 ton = $8.9 \times 10^3$ Newtons) into $\frac{\text{MeV}}{\text{fm}}$ units.


6. Using the generalized Pauli principle, stating that the total wavefunction, assumed to be a product of space, spin, and isospin parts, must be anti-symmetric under the interchange of the two nucleons, show that pp scattering can only take place in ($^{2S+1}L$ notation)

$$^1S, \; ^3P, \; ^1D, \; ^3F, \ldots$$

scattering states. What is the total isospin of the above states? Note: space interchamge is the same as parity, $(-1)^\ell$, here.


7.(a) Use the Wigner-Eckart theorem from prob. 4 above. Define the isospin states:

$$|\Delta^{++}\rangle = |\tfrac{3}{2}, \tfrac{3}{2}\rangle,$$
$$|\Delta^{+}\rangle = |\tfrac{3}{2}, \tfrac{1}{2}\rangle,$$
$$|p\rangle = |\tfrac{1}{2}, \tfrac{1}{2}\rangle,$$
$$|n\rangle = |\tfrac{1}{2}, -\tfrac{1}{2}\rangle,$$
$$|\pi^{+}\rangle = |1,1\rangle,$$

$$|\pi^0> = |1,0)>.$$

Use equation (8.42) with $I_{\pm}^{\Delta} = I_{\pm}^{\pi} + I_{\pm}^{nuc}$ and

$$|\Delta^{++}> = |p>|\pi^+>$$

to show that

$$\Delta^+ = \sqrt{\frac{1}{3}}\ |n>|\pi^+> + \sqrt{\frac{2}{3}}\ |p>|\pi^0>.$$

(b) Using the given proportionality (good for the isospin parts of wavefunctions in strong interaction decays only)

$$w(\text{initial}\rightarrow \text{final}) \propto |<\text{initial}|\text{final}>|^2,$$

find the decay branching ratio:

$$\frac{w(\Delta^+ \rightarrow p + \pi^0)}{w(\Delta^+ \rightarrow n + \pi^+)}\ .$$

8. From isospin considerations, find the ratio of rates for the decays indicated (see probs. 4, 7):

a) $\dfrac{w(K^{*+} \rightarrow K^0 + \pi^+)}{w(K^{*+} \rightarrow K^+ + \pi^0)} = ?,$ \qquad b) $\dfrac{w(K^{*0} \rightarrow K^+ + \pi^-)}{w(K^{*0} \rightarrow K^0 + \pi^0)} = ?.$

The K mesons have $I = \frac{1}{2}$ ; you can consider $K^+$ the $|\frac{1}{2},\frac{1}{2}>$ state and $K^0$ the $|\frac{1}{2},-\frac{1}{2}>$ .state. The $K^*(892)$ mesons are excited states of the K and also have $I = \frac{1}{2}$ . The pion of course has I=1. I will grade you on well you justify your derivation of the necessary isospin coefficients using the machinery of isospin addition.

9.(a) Find the isospin decomposition of the states:

$$|\pi^0 p>, |\pi^+ n>, \text{ and } |\pi^- n>.$$

[Hints: You need only use the concepts of raising and lowering operators and orthogonality of states. See probs. 4, 7.]

(b) Find the matrix elements

$$<\pi^-n|S|\pi^-n>, \quad <\pi^0p|S|\pi^+n>, \quad \text{and} \quad <\pi^0p|S|\pi^0p>$$

in terms of $<\frac{3}{2}\,|S|\frac{3}{2}> \equiv f_{3/2}$ and $<\frac{1}{2}\,|S|\frac{1}{2}> \equiv f_{1/2}$.

10. Find the magnetic moment of the neutron in the quark model. The normalized "spin up" wavefunction of the neutron can be written as

$$|n,\text{"up"}>=\frac{1}{\sqrt{2}}\left(\frac{(ddu-dud)}{\sqrt{2}}\frac{(|++->-|+-+>)}{\sqrt{2}} + \frac{(ddu+dud-2udd)}{\sqrt{6}}\frac{(|++->+|+-+>-2|-++>)}{\sqrt{6}}\right),$$

where (m is a constituent u,d quark mass)

$$\mu_z = <n,\text{"up"}\,|\,(\mu_z)_{op}\,|n,\text{"up"}>,$$

$$(\mu_z)_{op} = \frac{e\hbar}{2mc}\sum_{i=1,2,3}e_i\,\sigma_{iz},$$

(the $\sigma_{iz}$ are the usual z-component $\vec{\sigma}$ matrices working in the $i\underline{th}$ quark space) and

$$e_i = \begin{cases} \frac{2}{3}\,, & \text{u quark} \\ -\frac{1}{3}\,, & \text{d quark} \end{cases}.$$

11. Draw the lowest order (smallest number of vertices) Feynman diagrams for the following processes. You will have to figure out which interaction is responsible for each. Some of them may have more than one topologically distinct lowest order diagram. Consult the list of allowed vertices passed out early in the class. For hadrons draw the Feynman diagrams at the quark level. There are several ringers in here which can't occur; tell me which ones they are and what conservation law they violate. If you have trouble figuring out the nature of a reaction or decay, looking up the rates should tell you which interaction is responsible or dominant. (Make sure you make clear which lines represent which particles.)

a) $e^+ + \gamma \rightarrow e^+ + \gamma$        e) $\eta^0 \rightarrow \pi^+ + \pi^-$

b) $e^+ + e^- \rightarrow \nu_e + \bar{\nu}_e$        f) $\Sigma^+ \rightarrow n + \pi^+$

c) $p + \pi^- \rightarrow \Lambda^0 + K^0$        g) $\Delta^+ \rightarrow n + \pi^+$

d) $\nu_\mu + p \rightarrow \mu^- + \Delta^{++}$        h) $\mu^+ \rightarrow e^+ + \nu_e + \nu_\mu$

12. The following gives a list of particle decays which do <u>not</u> occur in nature (as far as is known). In each case please give me the physical property or reason why the reaction does not occur. (In some cases there may be more than one reason; in this case give me the "strongest" one.)

a) $n \rightarrow \pi^+ + \pi^-$        d) $\Sigma^- \rightarrow \bar{p} + \pi^0$

b) $\mu^+ \rightarrow e^+ + \nu_e + \nu_\mu$        e) $\Lambda^0 \rightarrow \Xi^0 + \pi^0$

c) $\pi^0 \rightarrow 3\gamma$        f) $\eta^0 \rightarrow \pi^+ + \pi^-$

13. Assuming time reversal invariance, find the ratio of the total (unpolarized) cross sections

$$\frac{\sigma(\bar{p} + p \rightarrow \pi^+ + \pi^-)}{\sigma(\pi^+ + \pi^- \rightarrow \bar{p} + p)} .$$

Write your answer so that it will be valid in any reference frame.

14. Given

$$\begin{cases} \pi^+ = u\bar{d}, \\ \pi^0 = \dfrac{1}{\sqrt{2}}(u\bar{u} - d\bar{d}), \end{cases}$$

(a) Which particle do you think is more massive, and why?

(b) For the decay pseudoscalar $\pi^0 \rightarrow \gamma + \gamma$, which photon amplitude describes the final state and why? ($\hat{\varepsilon}_1$, $\hat{\varepsilon}_2$ are the photon polarization vectars and $\hat{k}$ give the direction of one photon)

(i) $\hat{\varepsilon}_1 \cdot \hat{\varepsilon}_2$

(ii) $\hat{k} \cdot (\hat{\varepsilon}_1 \times \hat{\varepsilon}_2)$

## Index