Quantum Mechanics

Mathematical Structure and Physical Structure Part II

John R. Boccio
Emeritus Professor of Physics
Swarthmore College

February 25, 2021

Contents

Chapter 9

Time-Dependent Perturbation Theory

9.1 Theory

Time-independent or stationary-state perturbation theory, as we developed earlier, allows us to find approximations for the energy eigenvalues and eigenvectors in complex physical systems that are not solvable in closed form and where we could write $\hat{H}$ in two parts as

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{9.1}$$

For these perturbation methods to work, $\hat{V}$ must be *weak* and *time-independent*.

We now turn our attention to the case

$$\hat{H} = \hat{H}_0 + \hat{V}_t \tag{9.2}$$

where $\hat{V}_t$ is *weak* and *time-dependent*.

Examples might be the decays of an atomic system by photon emission or the ionization of an atom by shining light on it.

We assume that at some time t_0 the system has evolved into the state $\left|\psi_t^{(0))}\right\rangle$, i.e., the state $\left|\psi_t^{(0}\right\rangle$ satisfies the time evolution equation

$$i\hbar\frac{\partial}{\partial t}\left|\psi_t^{(0)}\right\rangle = \hat{H}_0\left|\psi_t^{(0)}\right\rangle \qquad t \le t_0 \tag{9.3}$$

It is a solution of the time-dependent Schrodinger equation with no perturbing interactions before t_0 where

$$\hat{H} = \hat{H}_0 \qquad t \le t_0 \tag{9.4}$$

At time t_0 we *turn on* the interaction potential (or perturbation) so that

$$\hat{H} = \hat{H}_0 + \hat{V}_t \qquad t \ge t_0 \tag{9.5}$$

The new state of the system then satisfies

$$i\hbar\frac{\partial}{\partial t}\left|\psi_t\right\rangle = \hat{H}\left|\psi_t\right\rangle = \left(\hat{H}_0 + \hat{V}_t\right)\left|\psi_t\right\rangle \qquad t \geq t_0 \tag{9.6}$$

with the *boundary condition(initial value)*

$$\left|\psi_t\right\rangle = \left|\psi_t^{(0)}\right\rangle \text{ at } t{=}t_0 \tag{9.7}$$

As we said, we assume that the full time-dependent Schrodinger equation cannot be solved in closed form and so we look for approximate solutions.

We let $\hat{V}_t$ be a small perturbation, i.e., we assume there is a natural small parameter in $\hat{V}_t$ (as we saw in time-independent perturbation theory) and we make an expansion of the solution in powers of $\hat{V}_t$ or this small parameter.

Since the effect of $\hat{H}_0$ will be much greater than the effect of $\hat{V}_t$, most of the time dependence comes from $\hat{H}_0$. If we could neglect $\hat{V}_t$, then since $\hat{H}_0$ is independent of time, we would have the simple time dependence

$$\left|\psi_t\right\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t}\left|\psi_t^{(0)}\right\rangle \tag{9.8}$$

Let us assume that this is still approximately true and remove this known time dependence from the solution. This should remove the major portion of the total time dependence from the problem. We do this by assuming a solution of the form

$$\left|\psi_t\right\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t}\left|\psi(t)\right\rangle \tag{9.9}$$

and then determining and solving the equation for the new state vector $|\psi(t)\rangle$.

Substituting this assumption in our original equation, the equation for $|\psi(t)\rangle$ is then given by

$$\begin{aligned}
i\hbar\frac{\partial}{\partial t}\left(e^{-\frac{i}{\hbar}\hat{H}_0 t}\left|\psi(t)\right\rangle\right) &= \left(\hat{H}_0 + \hat{V}_t\right)\left(e^{-\frac{i}{\hbar}\hat{H}_0 t}\left|\psi(t)\right\rangle\right) \\
\hat{H}_0 e^{-\frac{i}{\hbar}\hat{H}_0 t}\left|\psi(t)\right\rangle + i\hbar e^{-\frac{i}{\hbar}\hat{H}_0 t}\frac{\partial}{\partial t}\left|\psi(t)\right\rangle & \\
= \hat{H}_0 e^{-\frac{i}{\hbar}\hat{H}_0 t}\left|\psi(t)\right\rangle + \hat{V}_t e^{-\frac{i}{\hbar}\hat{H}_0 t}\left|\psi(t)\right\rangle & \\
i\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle &= \hat{V}(t)\left|\psi(t)\right\rangle
\end{aligned} \tag{9.10}$$

where

$$\hat{V}(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{V}_t e^{-\frac{i}{\hbar}\hat{H}_0 t} \tag{9.11}$$

The substitution has removed $\hat{H}_0$ from the equation and changed the time dependence of the perturbing potential. We are in the so-called *interaction picture or representation* where *both* the state vectors and the operators depend on time as we discussed earlier.

We develop a *formal solution* by integrating this equation of motion for the state vector to get

$$\begin{aligned} i\hbar \int_{t_0}^{t} \frac{\partial}{\partial t'} |\psi(t')\rangle dt' &= \int_{t_0}^{t} \hat{V}(t') |\psi(t')\rangle dt' \\ &= i\hbar(|\psi(t)\rangle - |\psi(t_0)\rangle) \end{aligned} \tag{9.12}$$

so that the formal solution is given by

$$|\psi(t)\rangle = |\psi(t_0)\rangle + \frac{1}{i\hbar} \int_{t_0}^{t} \hat{V}(t') |\psi(t')\rangle dt' \tag{9.13}$$

This is an *integral equation* for $|\psi(t)\rangle$. We solve it as a power series in $\hat{V}_t$ by the *method of iteration*.

The 0^{th}–order approximation is found by neglecting the perturbing potential. We get

$$|\psi(t)\rangle = |\psi(t_0)\rangle \tag{9.14}$$

The 1^{st}–order approximation is obtained by inserting the 0^{th}–order approximation into the full equation. We get

$$\begin{aligned} |\psi(t)\rangle &= |\psi(t_0)\rangle + \frac{1}{i\hbar} \int_{t_0}^{t} \hat{V}(t') |\psi(t_0)\rangle dt' \\ &= \left(1 + \frac{1}{i\hbar} \int_{t_0}^{t} \hat{V}(t') dt'\right) |\psi(t_0)\rangle \end{aligned} \tag{9.15}$$

The 2^{nd}–order approximation is obtained by inserting the 1^{st}–order approximation into the full equation. We get

$$\begin{aligned} |\psi(t)\rangle = |\psi(t_0)\rangle &+ \frac{1}{i\hbar} \int_{t_0}^{t} \hat{V}(t') |\psi(t_0)\rangle dt' \\ &+ \frac{1}{(i\hbar)^2} \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \hat{V}(t') \hat{V}(t'') |\psi(t_0)\rangle \end{aligned} \tag{9.16}$$

Notice that in all subsequent iterations the operators $\hat{V}(t')$,$\hat{V}(t'')$,..., etc, always occur in order of increasing time from right to left.

We can write the general result as

$$|\psi(t)\rangle = \hat{U}(t, t_0) |\psi(t_0)\rangle \tag{9.17}$$

where

$$\hat{U}(t,t_0) = \hat{I} + \sum_{n=1}^{\infty} \frac{1}{(i\hbar)^n} \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 ..... \int_{t_0}^{t_{n-1}} dt_n \hat{V}(t_1)\hat{V}(t_2) .....\hat{V}(t_n) \quad (9.18)$$

The complete, formal solution to the problem is then given by

$$|\psi_t\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t} |\psi(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}_0 t} \hat{U}(t,t_0) |\psi(t_0)\rangle \quad (9.19)$$

so that

$$e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t,t_0) = \text{ the total time development operator} \quad (9.20)$$

Before developing the detailed techniques of time-dependent perturbation theory, let us spend some time with the operator $\hat{U}(t,t_0)$ and discuss some of its properties.

We first introduce the idea of a *time-ordered product* of operators as follows. The symbol

$$\left(\hat{A}(t)\hat{B}(t_1)\hat{C}(t_2) .....\hat{X}(t_n)\right)_+ \equiv T\left(\hat{A}(t)\hat{B}(t_1)\hat{C}(t_2) .....\hat{X}(t_n)\right) \quad (9.21)$$

means the product of the operators where the operators are written from right to left in order of increasing times, i.e.,

$$\left(\hat{A}(t)\hat{B}(t')\right)_+ = \begin{cases} \hat{A}(t)\hat{B}(t') & t \geq t' \\ \hat{B}(t')\hat{A}(t) & t' \geq t \end{cases} \quad (9.22)$$

Now, we have using the time-ordered product definition

$$\begin{aligned} \left(\left[\int_{t_0}^{t} \hat{V}(t')dt'\right]^2\right)_+ &= \left(\int_{t_0}^{t} \hat{V}(t')dt' \int_{t_0}^{t} \hat{V}(t'')dt''\right)_+ \\ &= \int_{t_0}^{t} dt' \int_{t_0}^{t} dt''(\hat{V}(t')\hat{V}(t''))_+ \\ &= \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\hat{V}(t')\hat{V}(t'') + \int_{t_0}^{t} dt'' \int_{t_0}^{t''} dt'\hat{V}(t'')\hat{V}(t') \\ &= 2\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\hat{V}(t')\hat{V}(t'') \end{aligned} \quad (9.23)$$

and in general

$$\left(\left[\int_{t_0}^{t}\hat{V}(t')dt'\right]^n\right)_+ = \left(\int_{t_0}^{t}\hat{V}(t_1)dt_1\int_{t_0}^{t}\hat{V}(t_2)dt_2.....\int_{t_0}^{t}\hat{V}(t_n)dt_n\right)_+$$
$$= \int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2.....\int_{t_0}^{t}dt_n(\hat{V}(t_1)\hat{V}(t_2).....\hat{V}(t_n))_+$$
$$= n!\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2.....\int_{t_0}^{t}dt_n\hat{V}(t_1)\hat{V}(t_2).....\hat{V}(t_n) \quad (9.24)$$

because there are $n!$ possible orderings of the n terms involved. This last form is identical to the expression for $\hat{U}(t,t_0)$ and thus we have

$$\hat{U}(t,t_0) = \sum_{n=0}^{\infty}\frac{1}{(i\hbar)^n}\frac{1}{n!}\left(\left[\int_{t_0}^{t}\hat{V}(t')dt'\right]^n\right)_+$$
$$= \left(\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t}\hat{V}(t')dt'\right]\right)_+ \quad (9.25)$$

The last expression is just a *convenient shorthand for the infinite sum.* In order to verify that this is in fact a solution of

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{V}(t)|\psi(t)\rangle \quad (9.26)$$

we must prove that

$$i\hbar\frac{\partial}{\partial t}\hat{U}(t,t_0)|\psi(t_0)\rangle = \hat{V}(t)\hat{U}(t,t_0)|\psi(t_0)\rangle \quad (9.27)$$
$$i\hbar\frac{\partial}{\partial t}\hat{U}(t,t_0) = \hat{V}(t)\hat{U}(t,t_0) \quad (9.28)$$

Substituting, we have

$$i\hbar\frac{\partial}{\partial t}\hat{U}(t,t_0) = i\hbar\frac{\partial}{\partial t}\left(\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t}\hat{V}(t')dt'\right]\right)_+$$
$$= \left(\hat{V}(t)\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t}\hat{V}(t')dt'\right]\right)_+ \quad (9.29)$$

In the differentiation we do not have to worry about the non-commutation of the operators inside the time-ordered product since the *order is already specified.* Since t is certainly the latest time in the time-ordered product and therefore

all the other operators will be on the right of $\hat{V}(t)$ we can pull it outside the time-ordered product and write

$$i\hbar\frac{\partial}{\partial t}\hat{U}(t,t_0) = \hat{V}(t)\left(\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t}\hat{V}(t')dt'\right]\right)_+ = \hat{V}(t)\hat{U}(t,t_0)$$

as required.

The *most important question* (really the only question) that is usually asked in quantum mechanics is the following:

> Suppose that the system is initially in an eigenstate $|n\rangle$ of $\hat{H}_0$, i.e., $\hat{H}_0|n\rangle = \epsilon_n|n\rangle$. What is the probability that the system will be observed, after the perturbation has had time to act, in a different (and thus orthogonal) eigenstate of $\hat{H}_0$, say $|m\rangle$?

Alternatively, the question is sometimes posed this way:

> What is the probability that the interaction causes the system to make a *transition* from the state $|n\rangle$ to the state $|m\rangle$?

The probability amplitude for observing the system in the state $|m\rangle$ at time t is given by

$$\begin{aligned}\langle m \mid \psi_t\rangle &= \langle m|\, e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t,t_0)\,|\psi(t_0)\rangle \\ &= \langle m|\, e^{-\frac{i}{\hbar}\hat{H}_0 t}\hat{U}(t,t_0)\,|n\rangle \end{aligned} \tag{9.30}$$

where

$$|\psi(t_0)\rangle = |n\rangle \tag{9.31}$$

is the initial state.

Setting $t_0 = 0$ for simplicity and using the 1^{st}–order approximation for $\hat{U}(t,0)$ and also using

$$\langle m|\, e^{-\frac{i}{\hbar}\hat{H}_0 t} = \left(e^{\frac{i}{\hbar}\hat{H}_0 t}\,|m\rangle\right)^+ = \left(e^{\frac{i}{\hbar}\varepsilon_m t}\,|m\rangle\right)^+ = \langle m|\, e^{-\frac{i}{\hbar}\varepsilon_m t} \tag{9.32}$$

we get

$$\begin{aligned}\langle m \mid \psi_t \rangle &= \frac{1}{i\hbar} e^{-\frac{i}{\hbar}\varepsilon_m t} \int_0^t dt' \, \langle m | \hat{V}(t') | n \rangle \\ &= \frac{1}{i\hbar} e^{-\frac{i}{\hbar}\varepsilon_m t} \int_0^t dt' \, \langle m | e^{\frac{i}{\hbar}\hat{H}_0 t} \hat{V}_{t'} e^{-\frac{i}{\hbar}\hat{H}_0 t} | n \rangle \\ &= \frac{1}{i\hbar} e^{-\frac{i}{\hbar}\varepsilon_m t} \int_0^t dt' e^{\frac{i}{\hbar}(\varepsilon_m - \varepsilon_n)t} \langle m | \hat{V}_{t'} | n \rangle \end{aligned} \tag{9.33}$$

The probability of the transition is then

$$P_{n\to m}(t) = |\langle m \mid \psi_t \rangle|^2 = \frac{1}{\hbar^2} \left| \int_0^t dt' e^{\frac{i}{\hbar}(\varepsilon_m - \varepsilon_n)t} \langle m | \hat{V}_{t'} | n \rangle \right|^2 \tag{9.34}$$

The simplest example is when $\hat{V}_t$ is not a function of t, or $\hat{V}_t = \hat{V}$. We then have

$$P_{n\to m}(t) = |\langle m \mid \psi_t \rangle|^2 = \frac{\left|\langle m | \hat{V}_{t'} | n \rangle\right|^2}{\hbar^2} \left| \int_0^t dt' e^{\frac{i}{\hbar}(\varepsilon_m - \varepsilon_n)t} \right|^2 \tag{9.35}$$

If we define $\Delta = \epsilon_m - \epsilon_n$, then we have

$$P_{n\to m}(t) = \left|\langle m | \hat{V}_{t'} | n \rangle\right|^2 \left| \frac{1 - e^{-\frac{i}{\hbar}\Delta t}}{\Delta} \right|^2 = \left|\langle m | \hat{V}_{t'} | n \rangle\right|^2 \left(\frac{\sin \frac{\Delta t}{2\hbar}}{\Delta/2} \right)^2 \tag{9.36}$$

for the transition probability.

9.1.1 What is the physical meaning of this result?

We must be *very careful* when we use the words

> the perturbation causes a transition
> between eigenstates of $\hat{H}_0$

What this means physically is that the system has absorbed from the perturbing field (or emitted to it) the energy difference $\Delta = \epsilon_m - \epsilon_n$ and therefore the system has changed its energy.

Does the statement also mean that the state vector has changed from an initial value $|\psi(0)\rangle = |n\rangle$ to a final value $|\psi(t)\rangle = |m\rangle$?

We can get a better feeling for the correct answer to this question by deriving the result in a different manner.

We have

$$i\hbar\frac{\partial}{\partial t}\left|\psi_t\right\rangle = \hat{H}\left|\psi_t\right\rangle = \left(\hat{H}_0 + \hat{V}_t\right)\left|\psi_t\right\rangle \tag{9.37}$$

and

$$\hat{H}_0\left|n\right\rangle = \varepsilon_n\left|n\right\rangle \tag{9.38}$$

As in our development of time-independent perturbation theory, we let

$$\hat{V}_t = g\hat{U}_t \tag{9.39}$$

where g is a small parameter.

The set of eigenvectors $\{|n\rangle\}$ is a complete set and therefore we can use it as a basis for the space and, in particular, we can write

$$\left|\psi_t\right\rangle = \sum_n a_n(t)e^{-\frac{i}{\hbar}\varepsilon_n t}\left|n\right\rangle \tag{9.40}$$

The reason for pulling out the phase factors will be clear shortly.

It is clear that if $g = 0$, then this is the correct general solution with

$$a_n(t) = a_n(0) = \text{constant} \tag{9.41}$$

The phase factors we pulled out represent the time dependence due to $\hat{H}_0$ and this is, by assumption, the major time dependence in the system.

If g is small we expect the time dependence of $a_n(t)$, which is due to the perturbation to be weak or that

$$\frac{da_n(t)}{dt} \text{ is small} \tag{9.42}$$

It is in this sense that we can propose to use perturbation theory on the system.

Using the eigenbasis expansion we have

$$\begin{aligned}\sum_n\left(i\hbar\frac{da_n(t)}{dt} + \varepsilon_n a_n(t)\right)e^{-\frac{i}{\hbar}\varepsilon_n t}\left|n\right\rangle \\ = \sum_n\left(\varepsilon_n a_n(t) + g\hat{U}_t a_n(t)\right)e^{-\frac{i}{\hbar}\varepsilon_n t}\left|n\right\rangle\end{aligned} \tag{9.43}$$

Applying the linear functional $\langle m|$ from the left and using the orthonormality relation

$$\left\langle m \mid n\right\rangle = \delta_{mn} \tag{9.44}$$

we get

$$\begin{aligned}i\hbar\frac{da_m(t)}{dt} &= g\sum_n\left\langle m\right|\hat{U}_t\left|n\right\rangle e^{i\omega_{mn}t}a_n(t) \\ &= \sum_n V_{mn}(t)e^{i\omega_{mn}t}a_n(t)\end{aligned} \tag{9.45}$$

where

$$\omega_{mn} = \frac{\varepsilon_m - \varepsilon_n}{\hbar} \tag{9.46}$$

This is an exact equation. It implies that the time dependence of $a_n(t)$ is due entirely to $\hat{V}_t$ (because we explicitly extracted out the dependence due to $\hat{H}_0$). This is the interaction picture that we had earlier.

Exactly Solvable 2-State Example

Consider a 2-state system with

$$H_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \quad , \quad V(t) = \begin{pmatrix} 0 & \delta e^{i\omega t} \\ \delta e^{-i\omega t} & 0 \end{pmatrix} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \tag{9.47}$$

In the interaction picture, as derived above, we have

$$i\hbar \frac{dc_m(t)}{dt} = \sum_n V_{mn}(t) e^{i\omega_{mn} t} c_n(t) \quad , \quad |\psi_t\rangle = \sum_n c_n(t) e^{-\frac{i}{\hbar}\varepsilon_n t} |n\rangle \tag{9.48}$$

or

$$i\hbar \frac{dc_1(t)}{dt} = \delta e^{i\left[\omega + \frac{E_1 - E_2}{\hbar}\right]t} c_2(t) \tag{9.49}$$

$$i\hbar \frac{dc_2(t)}{dt} = \delta e^{i\left[-\omega - \frac{E_1 - E_2}{\hbar}\right]t} c_1(t) \tag{9.50}$$

$$|\psi_t\rangle = c_1(t) e^{-\frac{i}{\hbar}E_1 t} |1\rangle + c_2(t) e^{-\frac{i}{\hbar}E_2 t} |2\rangle \tag{9.51}$$

We can write these equations as

$$\frac{dC(t)}{dt} = -\frac{i\delta}{\hbar} \begin{pmatrix} 0 & e^{i[\omega - \omega_{21}]t} \\ e^{-i[\omega - \omega_{21}]t} & 0 \end{pmatrix} C(t) \tag{9.52}$$

where

$$C(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} \quad , \quad \omega_{21} = \frac{E_2 - E_1}{\hbar} \tag{9.53}$$

We can find an exact solution. With initial conditions $c_1(0) = 1$ and $c_2(0) = 0$ we get

$$|c_1(t)|^2 = \frac{\delta^2}{\delta^2 + \frac{\hbar^2(\omega - \omega_{21})^2}{4}} \sin^2 \Omega t = 1 - |c_2(t)|^2 \tag{9.54}$$

where

$$\Omega^2 = \frac{\delta^2}{\hbar^2} + \frac{(\omega - \omega_{21})^2}{4} \tag{9.55}$$

A graph of these functions is shown in Figure 9.1 below.

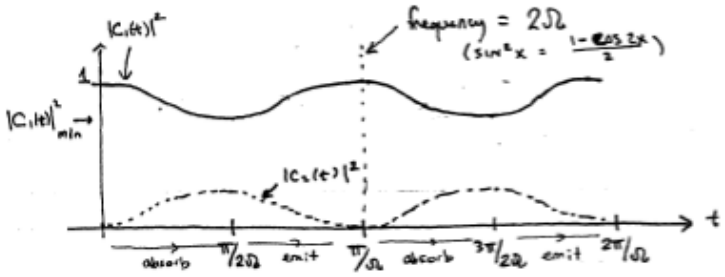


Figure 9.1: Exact Solution

A straightforward calculation gives

$$|c_1(t)|^2_{\min} = \frac{(\omega - \omega_{21})^2}{(\omega - \omega_{21})^2 + \frac{4\delta^2}{\hbar^2}} \tag{9.56}$$

At resonance, $\omega = \omega_{21}$, we have

$$\Omega = \frac{\delta}{\hbar} \quad , \quad |c_1(t)|^2_{\min} = 0 \tag{9.57}$$

as shown in Figure 9.2 below.

Figure 9.2: At Resonance

The amplitude as a function of ω is shown in Figure 9.3 below.

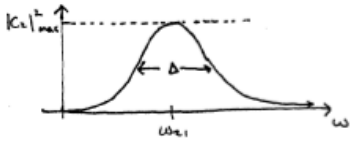


Figure 9.3: Amplitude versus ω

where Δ = full width at half maximum = $4\delta/\hbar$. The amplitude is peaked at resonance and the width is proportional to δ (the strength of the perturbation).

This periodically forced 2−state system is a basic problem - it demonstrates the fundamental features of absorption and emission.

We now return to the full, general equations and look for a perturbation solution. Now we assume (power series)

$$a_n(t) = a_n^{(0)} + g a_n^{(1)} + g^2 a_n^{(2)} + ... \tag{9.58}$$

Substituting and arranging the terms in a power series in g we have

$$\left(\frac{da_n^{(0)}}{dt}\right) g^0 + \left(i\hbar \frac{da_m^{(1)}}{dt} - \sum_n \langle m|\hat{U}_t|n\rangle e^{\frac{i}{\hbar}\omega_{mn}t} a_n^{(0)}\right) g^1 + ...+$$
$$\left(i\hbar \frac{da_m^{(r+1)}}{dt} - \sum_n \langle m|\hat{U}_t|n\rangle e^{\frac{i}{\hbar}\omega_{mn}t} a_n^{(r)}\right) g^r + ... = 0$$

or looking at each order separately we have

$$0^{th} - \text{order} \quad \frac{da_n^{(0)}}{dt} = 0$$
$$1^{st} - \text{order} \quad i\hbar \frac{da_m^{(1)}}{dt} = \sum_n \langle m|\hat{U}_t|n\rangle e^{\frac{i}{\hbar}\omega_{mn}t} a_n^{(0)}$$
.....
.....
$$(r+1)^{st} - \text{order} \quad i\hbar \frac{da_m^{(r+1)}}{dt} = \sum_n \langle m|\hat{U}_t|n\rangle e^{\frac{i}{\hbar}\omega_{mn}t} a_n^{(r)}$$
.....

Note that the coefficients $a_n^{(0)}$ follow from the initial condition

$$|\psi(0)\rangle = \sum_n a_n^{(0)} |n\rangle \tag{9.59}$$

The solution proceeds as follows:

initial condition $\rightarrow a_n^{(0)}$
$a_n^{(0)} \rightarrow a_n^{(1)}$ using the 1^{st} − order equation
.....
$a_n^{(r)} \rightarrow a_n^{(r+1)}$ using the $(r+1)^{st}$ − order equation

Now consider the following example. We assume that

$$\hat{H} = \hat{H}_0 \qquad t \le 0 \tag{9.60}$$

where

$$\hat{H}_0 |n\rangle = \varepsilon_n |n\rangle \tag{9.61}$$

and during the time interval $0 \le t \le T$ a perturbation $\hat{V}_t$ is applied to the system and the $a_n(t)$ change with time.

Finally, for $t \ge T$ the perturbation is turned off and $a_n(t) = a_n(T)$.

The probability that, as a result of the perturbation, the energy of the system becomes ϵ_r, is given by

$$|\langle r \mid \psi_t \rangle|^2 = \left| \sum_i a_i(t) e^{-\frac{i}{\hbar}\varepsilon_i t} \langle r \mid i \rangle \right|^2 = |a_r(t)|^2 \tag{9.62}$$

and as $t \to \infty$ we get

$$|\langle r \mid \psi_t \rangle|^2 = |a_r(T)|^2 \tag{9.63}$$

Now to 1^{st}–order we have

$$i\hbar \frac{da_r^{(1)}}{dt} = \sum_n \langle r | \hat{U}_t | n \rangle \, e^{\frac{i}{\hbar}\omega_{rn} t} a_n^{(0)} \tag{9.64}$$

If $|\psi(0)\rangle = |i\rangle$, then

$$a_n^{(0)} = \begin{cases} 1 & n = i \\ 0 & n \ne i \end{cases} \tag{9.65}$$

This gives

$$i\hbar \frac{da_r^{(1)}}{dt} = \langle r | \hat{U}_t | i \rangle \, e^{\frac{i}{\hbar}\omega_{ri} t} \tag{9.66}$$

Integrating we have

$$a_r^{(1)}(T) = \frac{1}{i\hbar} \int_0^T \langle r | \hat{U}_t | i \rangle \, e^{\frac{i}{\hbar}(\varepsilon_r - \varepsilon_i)t} dt \tag{9.67}$$

and

$$a_r(T) = a_r^{(0)}(T) + g a_r^{(1)}(T) \tag{9.68}$$

which is identical to our earlier result as $t \to \infty$.

Now let return to our question. Has the state changed also?

In the example we found that the perturbation produces a final state $|\psi_t\rangle$ for $t \ge T$ which to 1^{st}–order is

$$|\psi_t\rangle = \sum_n a_n(t) e^{-\frac{i}{\hbar}\varepsilon_n t} |n\rangle \tag{9.69}$$

This is a coherent (definite relative phases) superposition of eigenvectors of $\hat{H}_0$. This is *NOT* a stationary state. Interference effects between the terms in the sum are detectable. They do not, however, affect

$$|a_r(T)|^2 = \text{ probability that the energy changes to } \varepsilon_r \tag{9.70}$$

Thus, the perturbation does not cause a *jump* from one stationary state $|i\rangle$ of $\hat{H}_0$ to another $|r\rangle$, but instead it produces a non-stationary state.

The conventional language of quantum mechanics produces this ambiguity between the two statements

the energy is ϵ_r and *the state is* $|r\rangle$

For the state

$$|\psi_t\rangle = \sum_n a_n(t) e^{-\frac{i}{\hbar}\varepsilon_n t} |n\rangle$$

it is correct to say

the probability of the energy being ϵ_r is $|a_r(T)|^2$

or

$Prob(E = \epsilon_r | |\psi\rangle) = |a_r(T)|^2$

The state, however, is $|\psi_t\rangle$ and *NOT* $|r\rangle$.

An example

Suppose we perturb an oscillator with a decaying electric field of the form

$$\hat{V}_t = -q\mathcal{E}_0 \hat{x} e^{-\frac{t}{\tau}} \qquad t \geq 0 \tag{9.71}$$

To 1^{st}–order, starting with the initial state $|n\rangle$ with energy

$$\varepsilon_n = \hbar\omega(n + \frac{1}{2}) \tag{9.72}$$

we have

$$|\psi(t)\rangle = |n\rangle + \frac{1}{i\hbar}\int_0^t dt' \hat{V}(t) |n\rangle \tag{9.73}$$

where

$$\hat{V}(t) = e^{\frac{i}{\hbar}\hat{H}_0 t} \hat{V}_t e^{-\frac{i}{\hbar}\hat{H}_0 t} \tag{9.74}$$

We let $n = 0$ (the ground state) for this example. We then have

$$|\psi(t)\rangle = |0\rangle + \frac{1}{i\hbar}\int_0^t dt' e^{\frac{i}{\hbar}\hat{H}_0 t}(-q\mathcal{E}_0\hat{x}) e^{-\frac{t}{\tau}} e^{-\frac{i}{\hbar}\hat{H}_0 t} |0\rangle \tag{9.75}$$

Using

$$e^{-\frac{i}{\hbar}\hat{H}_0 t}\,|0\rangle = e^{-i\frac{\omega}{2}t}\,|0\rangle \tag{9.76}$$

$$\hat{x}\,|0\rangle = \sqrt{\frac{\hbar}{2m\omega}}\,|1\rangle \tag{9.77}$$

$$e^{\frac{i}{\hbar}\hat{H}_0 t}\,|1\rangle = e^{i\frac{3\omega}{2}t}\,|0\rangle \tag{9.78}$$

we get (letting $t \to \infty$)

$$|\psi(t)\rangle = |0\rangle + \frac{1}{i\hbar}(-q\mathcal{E}_0)\sqrt{\frac{\hbar}{2m\omega}}\int_0^\infty dt' e^{i\omega t' - \frac{t}{\tau}}\,|1\rangle \tag{9.79}$$

and finally,

$$\begin{aligned} P_{0\to 1} = |\langle 1 \mid \psi(t)\rangle|^2 &= \frac{q^2\mathcal{E}_0^2}{2m\hbar\omega}\left|\int_0^\infty dt' e^{i\omega t' - \frac{t}{\tau}}\right|^2 \\ &= \frac{q^2\mathcal{E}_0^2}{2m\hbar\omega}\frac{\tau^2}{\tau^2\omega^2+1} \end{aligned} \tag{9.80}$$

We now return to the earlier general result (11.36) we derived for the probability, namely,

$$P_{0\to n}(t) = \left|\langle n|\,\hat{V}_{t'}\,|0\rangle\right|^2\left(\frac{\sin\frac{\Delta t}{2\hbar}}{\Delta/2}\right)^2 \tag{9.81}$$

In Figure 9.4 below we plot this function.

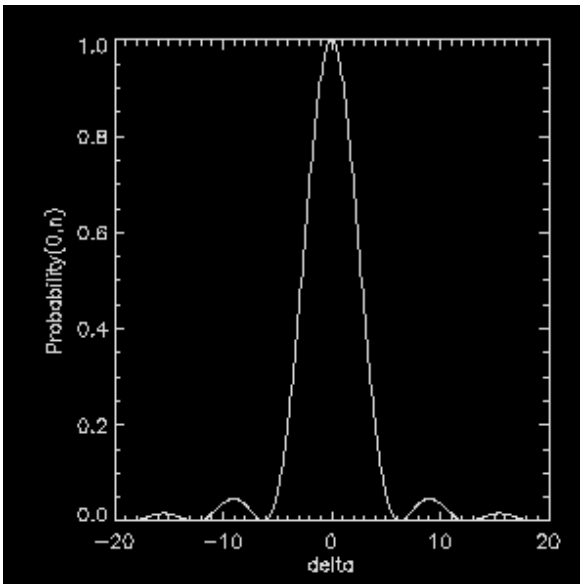


Figure 9.4: Probability(0,n) versus Delta

The height of the central peak is proportional to t^2 and the location of the first zero is at

$$\Delta = \frac{2\pi\hbar}{t} \tag{9.82}$$

so that the width of the peak decreases as $1/t$.

The formula implies that for very short times

$$P_{0\to n} \propto t^2 \text{ for all } \varepsilon_n \tag{9.83}$$

As $t \to \infty$, however, the probability is largest for those states whose energy lies under the sharp bump near $\Delta = 0$ or those states with whose energy lies under the peak around ϵ_0. Now the energy $\epsilon_n \approx \epsilon_0$ lies under the sharp bump when

$$|\Delta| = |\varepsilon_n - \varepsilon_0| < \frac{2\pi\hbar}{t} \tag{9.84}$$

The area under the bump is proportional to t and the rest of the area oscillates in time around zero. This latter feature means that if $\epsilon_n \neq \epsilon_0$, the transition probability oscillates in time with a repetition time of

$$\frac{2\pi\hbar}{|\varepsilon_n - \varepsilon_0|} \tag{9.85}$$

The case, where we are looking for a transition to a single state, is, thus, only valid in perturbation theory for very small time t. Otherwise the condition that the

$$P_{0\to 0}(t) \approx 1 \tag{9.86}$$

will not be true and perturbation theory breaks down. We also note that the probability cannot grow larger than one or that, after a while, the higher-order effects of the perturbation which we have neglected so far must become important and prevent the probability from exceeding one.

The condition that tells us whether a transition probability to a state with an energy appreciably different than the original energy is the same condition in time-independent perturbation theory that tells whether the state vector changes appreciably from the unperturbed state, namely

$$\left| \frac{\langle n| \hat{V} |0\rangle}{\varepsilon_n - \varepsilon_0} \right| << 1 \tag{9.87}$$

Physically, a more interesting case occurs when the state $|n\rangle$ is one of a continuum of energy states, or it lies in a group of very closely spaced levels.

In this case we ask a different experimental question, namely,

> What is the probability that the system makes
> a transition to a small group of states near
> $|n\rangle$ (or has energy near ϵ_n)?

Since the area under the bump near $\Delta = 0$ or $\epsilon_n \approx \epsilon_0$ is proportional to t, we expect that the transition probability to a small group of states near ϵ_0 will grow linearly with t and thus

$$\frac{P_{0\to n}(t)}{t} = \text{ transition rate } = \Gamma = \text{ constant as } t \to \infty \tag{9.88}$$

Quantities that we measure are related to the transition rate and this result says that these measurements will make sense.

Let us now carry out this derivation in detail.

To calculate this transition rate we must sum $P_{0\to n}$ over the group of *final* states. We assume that $|\langle b| \hat{V}_{t'} |0\rangle|^2$ is relatively constant over the small group of states near $|n\rangle$ (has a weak energy dependence).

We then have

$$\sum_{\substack{n \\ in group}} P_{0\to n}(t) = \left|\langle n| \hat{V}_{t'} |0\rangle\right|^2 \int_{group} d\varepsilon_n \rho(\varepsilon_n) \left[\frac{\sin\left[\frac{(\varepsilon_n - \varepsilon_0)t}{\hbar}\right]}{\left[\frac{(\varepsilon_n - \varepsilon_0)}{2}\right]} \right]^2 \tag{9.89}$$

where

$$\rho(\varepsilon_n) = \text{ number of states per unit energy}$$
$$\rho(\varepsilon_n) d\varepsilon_n = \text{ number of states in the interval } d\varepsilon_n$$

Now in the limit $t \to \infty$

$$\left[\frac{\sin\left[\frac{(\varepsilon_n-\varepsilon_0)t}{\hbar}\right]}{\left[\frac{(\varepsilon_n-\varepsilon_0)}{2}\right]}\right]^2 \to \frac{2\pi t}{\hbar}\delta(\varepsilon_0-\varepsilon_n) \tag{9.90}$$

i.e., in general, for a sequence of functions

$$\delta_t(\alpha) = \frac{\sin^2 \alpha t}{\pi \alpha^2 t} \tag{9.91}$$

we have that

$$\delta_t(\alpha) = \begin{cases} \frac{t}{\pi} & \alpha = 0 \\ \leq \frac{1}{\pi\alpha^2 t} & \alpha \neq 0 \end{cases} \tag{9.92}$$

and

$$\lim_{t\to\infty} \int_{-\infty}^{\infty} d\alpha \delta_t(\alpha) F(\alpha) = F(0) \tag{9.93}$$

Therefore,

$$\lim_{t\to\infty} \delta_t(\alpha) = \delta(\alpha) \tag{9.94}$$

Using this result, we have

$$\sum_n P_{0\to n}(t) = \Gamma t \tag{9.95}$$

and thus

$$\Gamma = \text{ transition rate } = \frac{2\pi}{\hbar}\left|\langle n|\hat{V}|0\rangle\right|^2 \rho(\varepsilon_n)_{\varepsilon_n=\varepsilon_0} \tag{9.96}$$

which is called *Fermi's Golden Rule*.

We now consider a perturbation that depends explicitly on time. In particular, suppose we have a harmonic perturbation of the form

$$\hat{V}_t = \hat{V} e^{\eta t}\cos\omega t = \frac{\hat{V}}{2} e^{\eta t}\left(e^{-i\omega t} + e^{i\omega t}\right) \tag{9.97}$$

and $|\psi(t_0)\rangle = |0\rangle$, where we let $t_0 \to -\infty$. The $e^{\eta t}$ factor is necessary to make the mathematical operations valid in the limit. It is equivalent for small η to turning the perturbation on slowly. In the end we will let $\eta \to 0$.

We have

$$\begin{aligned} \langle n \mid \psi(t)\rangle &= \langle n \mid 0\rangle + \frac{1}{i\hbar} e^{-\frac{i}{\hbar}\varepsilon_n t} \int_{-\infty}^{0} dt' \langle n|\hat{V}(t')|0\rangle \\ &= \frac{e^{\eta t}}{2}\left[\frac{e^{i(\varepsilon_n-\varepsilon_0-\hbar\omega)\frac{t}{\hbar}}}{\varepsilon_0-\varepsilon_n+\hbar\omega+i\hbar\eta} + \frac{e^{i(\varepsilon_n-\varepsilon_0+\hbar\omega)\frac{t}{\hbar}}}{\varepsilon_0-\varepsilon_n-\hbar\omega+i\hbar\eta}\right]\langle n|\hat{V}|0\rangle \end{aligned} \tag{9.98}$$

Thus, the probability of the energy being ϵ_n at the time t is

$$|\langle n \mid \psi(t)\rangle|^2$$
$$= \frac{e^{2\eta t}}{4}\left|\langle n|\hat{V}|0\rangle\right|^2 \left\{ \begin{array}{c} \frac{1}{(\varepsilon_0-\varepsilon_n+\hbar\omega)^2+(\hbar\eta)^2} + \frac{1}{(\varepsilon_0-\varepsilon_n-\hbar\omega)^2+(\hbar\eta)^2} \\ +2\mathrm{Re}\frac{e^{-2i\omega t}}{(\varepsilon_0-\varepsilon_n+\hbar\omega+i\hbar\eta)(\varepsilon_0-\varepsilon_n-\hbar\omega+i\hbar\eta)} \end{array} \right\} \tag{9.99}$$

The first term comes from the $e^{-i\omega t}$ part of $\hat{V}_t$ (positive frequency) and the second term comes from the $e^{i\omega t}$ part of $\hat{V}_t$ (negative frequency). The last term represents *interference effects*.

Since

$$P_{0\to n}(t) = \Gamma_{0\to n} t \tag{9.100}$$

we have

$$\Gamma_{0\to n} = \frac{dP_{0\to n}(t)}{dt} \tag{9.101}$$

and thus

$$\Gamma_{0\to n} \tag{9.102}$$
$$= \frac{e^{2\eta t}}{4}\left|\langle n|\hat{V}|0\rangle\right|^2 \left\{ \begin{array}{c} \left[\frac{1}{(\varepsilon_0-\varepsilon_n+\hbar\omega)^2+(\hbar\eta)^2} + \frac{1}{(\varepsilon_0-\varepsilon_n-\hbar\omega)^2+(\hbar\eta)^2}\right](1-\cos 2\omega t) \\ +2\sin 2\omega t\left[\frac{e^{-2i\omega t}}{(\varepsilon_0-\varepsilon_n+\hbar\omega+i\hbar\eta)(\varepsilon_0-\varepsilon_n-\hbar\omega+i\hbar\eta)}\right] \end{array} \right\}$$

The $\sin\omega t$ and $\cos\omega t$ terms arise from the interference term. In the limit $\eta \to 0$ and assuming that $|n\rangle$ is in the continuum part of spectrum, we have

1. the first two terms are not equal to zero only if $\varepsilon_n - \varepsilon_0 = \pm\hbar\omega$
2. the $\sin\omega t$ and $\cos\omega t$ terms average to zero if we assume that $\Gamma_{0\to n}$ is $dP_{0\to n}(t)/dt$ averaged over a few cycles of $\hat{V}_t$

which gives the result

$$\Gamma_{0\to n} = \frac{2\pi}{\hbar}\frac{\left|\langle n|\hat{V}|0\rangle\right|^2}{4}\left[\delta(\varepsilon_n - \varepsilon_0 - \hbar\omega) + \delta(\varepsilon_n - \varepsilon_0 + \hbar\omega)\right] \tag{9.103}$$

The positive and negative frequency parts act independently and the interference averages to zero.

Thus, the $e^{-i\omega t}$ part produced a $\Delta E > 0$ process (absorption), while the $e^{i\omega t}$ part produced a $\Delta E < 0$ process (emission).

To enhance our understanding of time-dependent perturbation theory, we look at a variation of this harmonic perturbation. Suppose we have a harmonic perturbation of the form

$$\hat{V}_t = \hat{V}e^{-i\omega t} + \hat{V}^+ e^{i\omega t} \tag{9.104}$$

which is only applied for a finite time interval $0 \le t \le T$. If we start with energy ϵ_i in the state $|i\rangle$, then at any time $t \ge T$

$$a_f^{(1)}(t) = 1^{st} - \text{order amplitude for the state } |\psi(t)\rangle$$
$$\text{to have energy } \varepsilon_f \text{ (be in state } |f\rangle ?) \quad f \neq i$$

is given by

$$a_f^{(1)}(T) = \frac{1}{i\hbar} \langle f|\hat{V}|i\rangle \int_0^T e^{i(\omega_{fi}-\omega)t} dt + \frac{1}{i\hbar} \langle f|\hat{V}^+|i\rangle \int_0^T e^{i(\omega_{fi}+\omega)t} dt \qquad (9.105)$$

and

$$\left|a_f^{(1)}(T)\right|^2 = \text{ probability that the final energy will be } \epsilon_f \qquad (9.106)$$

As an example we consider spin resonance (we solved this problem exactly earlier).

We consider a spin = 1/2 particle in a static magnetic field B_0 (in the z–direction). This says that the unperturbed Hamiltonian is

$$\hat{H}_0 = -\frac{1}{2}\hbar\gamma B_0 \hat{\sigma}_z \qquad (9.107)$$

This operator has the eigenvectors and eigenvalues

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad , \quad \epsilon_\pm = \mp\frac{1}{2}\hbar\gamma B_0 \qquad (9.108)$$

We now perturb the system with another magnetic field B_1, which is rotating in the $x-y$ plane with angular velocity ω. This implies that

$$\begin{aligned} \hat{V}_t &= -\frac{1}{2}\hbar\gamma B_1 (\cos\omega t \hat{i} + \sin\omega t \hat{j}) \cdot \hat{\sigma} \\ &= -\frac{1}{2}\hbar\gamma B_1 \left[\hat{\sigma}_x \cos\omega t + \hat{\sigma}_y \sin\omega t\right] = -\frac{1}{2}\hbar\gamma B_1 \begin{pmatrix} 0 & e^{-i\omega t} \\ e^{i\omega t} & 0 \end{pmatrix} \\ &= -\frac{1}{2}\hbar\gamma B_1 \left(\hat{V} e^{-i\omega t} + \hat{V}^+ e^{i\omega t}\right) \end{aligned} \qquad (9.109)$$

where

$$\hat{V} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad , \quad \hat{V}^+ = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad (9.110)$$

We choose the initial state to be

$$|i\rangle = |n\rangle \ \text{(spin up in the z-direction)} \qquad (9.111)$$

The first order perturbation theory approximation for this result is

$$\left|a_f^{(1)}(T)\right|^2 = \left(\frac{\omega_1}{\omega_0 + \omega}\right)^2 \sin^2 \frac{1}{2}(\omega_0 + \omega)T \qquad (9.112)$$

where

$$\omega_0 = \gamma B_0 \quad , \quad \omega_1 = \gamma B_1 \tag{9.113}$$

and we have used

$$\langle - | \hat{V} | + \rangle = 0, \langle - | \hat{V}^+ | + \rangle = -\frac{1}{2}\hbar\gamma B_1 \tag{9.114}$$

$$\omega_{fi} = \frac{\varepsilon_f - \varepsilon_i}{\hbar} = \omega_0 \tag{9.115}$$

When is the first order perturbation theory result valid?

If we compare the exact result with perturbation theory by expanding the exact result in a power series, we find that the two results agree exactly if

$$\left| \frac{\omega_1}{\omega_0 + \omega} \right| << 1 \tag{9.116}$$

which corresponds $B_1 \ll B_0$ (as long as $\omega_0 + \omega \neq 0$). When $\omega_0 + \omega = 0$ we have a phenomenon called resonance. The exact solution gives

$$|a_f(T)|^2 = \sin^2 \frac{1}{2}\alpha T \tag{9.117}$$

where $\alpha^2 = (\omega_0 + \omega)^2 + \omega_1^2$ and perturbation theory gives

$$\left| a_f^{(1)}(T) \right|^2 = \left(\frac{\omega_1 T}{2} \right)^2 \tag{9.118}$$

Thus, the results agree only if $|\omega_1 T| \ll 1$ or if the perturbation only acts for a short time.

9.2 Atomic Radiation and Selection Rules

We now apply time-independent perturbation theory to the absorption and emission electromagnetic radiation by matter.

The Hamiltonian for an electron in an atom interacting with and electromagnetic field is

$$\hat{H} = \frac{\left(\vec{p}_{op} - \frac{q}{c}\vec{A} \right)^2}{2m_e} + q\phi + U \tag{9.119}$$

where $q = -e$, U = the potential energy function that binds the electrons in the atom, and $\vec{A}$ and φ are the vector and scalar potentials associated with the electromagnetic field.

These potentials imply the electric and magnetic fields

$$\vec{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}, \vec{B} = \nabla \times \vec{A} \tag{9.120}$$

Note that if $\vec{A} = 0$, the $\vec{B} = 0$ and

$$\phi = -\int_0^{\vec{r}} \vec{E}(\vec{r}, t) \cdot d\vec{r} \tag{9.121}$$

We rewrite $\hat{H}$ as

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{9.122}$$

where

$$\hat{H}_0 = \frac{\vec{p}_{op}^{\,2}}{2m_e} + U \;=\; \text{Hamiltonian for the atom with no electromagnetic field} \tag{9.123}$$

and

$$\hat{V} = \frac{q}{2m_e c}\left(\vec{p}_{op} \cdot \vec{A} + \vec{A} \cdot \vec{p}_{op}\right) + \frac{q^2}{2m_e c^2}\left(\vec{A} \cdot \vec{A}\right) + q\phi \tag{9.124}$$

is the perturbation due to the presence of the electromagnetic field, i.e., the term $\hat{V}$ tells us how the atom interacts with the electromagnetic field.

9.2.1 The Electric Dipole Approximation

The typical wavelength of visible electromagnetic radiation is ≈ 5000 and the typical dimension of an atom is $\approx$ a few . This implies that the electromagnetic fields are approximately constant over the volume of the atom.

In Gaussian units $|\vec{E}| \approx |\vec{B}|$, but the force due to $\vec{B} \approx (v/c)\times$ the force due to $\vec{E}$. Thus, magnetic effects are negligible in most atoms compared to electric effects. We therefore assume

1. $\vec{E} \approx$ constant over the volume of the atom
2. $\vec{B}$ can be neglected

This is the so-called electric dipole approximation.

In this approximation, we have $\vec{A} = 0$ and as we said above (11.121), $\vec{B} = 0$ and

$$\phi = -\int_0^{\vec{r}} \vec{E}(\vec{r}, t) \cdot d\vec{r} \tag{9.125}$$

This last integral is independent of path since

$$\nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} = 0 \tag{9.126}$$

Since we are assuming that $\vec{E} \approx$ constant over the volume of the atom, we get

$$\phi = -\vec{r} \cdot \vec{E}(t) \tag{9.127}$$

This adds a perturbation of the form

$$\hat{V} = -q\vec{r} \cdot \vec{E}(t) \tag{9.128}$$

to the Hamiltonian.

In some derivations the quantity

$$\hat{V} = \frac{q}{2m_e c}\left(\vec{p}_{op} \cdot \vec{A} + \vec{A} \cdot \vec{p}_{op}\right) + \frac{q^2}{2m_e c^2}\left(\vec{A} \cdot \vec{A}\right) + q\phi \tag{9.129}$$

is chosen as the perturbation. The next step would be to expand in powers of the potentials. Since $\nabla \cdot \vec{E} = 0$ for the radiation field, we can choose $\varphi = 0$ and $\nabla \cdot \vec{A} = 0$ (if these relations were not true we could make them so with a gauge transformation).

Therefore, we end up with an expansion in powers of $\vec{A}$. The first term in the expansion is

$$\hat{V}_{Ap} = -\frac{q}{2m_e c}\left(\vec{p}_{op} \cdot \vec{A} + \vec{A} \cdot \vec{p}_{op}\right) = -\frac{q}{m_e c}\vec{A} \cdot \vec{p}_{op} \tag{9.130}$$

since

$$\vec{p}_{op} \cdot \vec{A} - \vec{A} \cdot \vec{p}_{op} \propto \nabla \cdot \vec{A} = 0 \tag{9.131}$$

In this case, in the electric dipole approximation

$$\begin{aligned} &\vec{E} \approx \text{ constant} \quad , \quad \vec{B} = 0 \\ &\vec{A}(\vec{r}, t) = -c\int_0^{\vec{r}} \vec{E}(\vec{r}, t)dt \\ &\phi(\vec{r}, t) = 0 \end{aligned}$$

If $\vec{E}(\vec{r}, t) = \vec{E}(\vec{r})e^{-i\omega t}$, then $\vec{A} = e\vec{E}/i\omega$. Using

$$\frac{\vec{p}_{op}}{m} = \frac{i}{\hbar}\left[\hat{H}_0, \vec{r}_{op}\right] \tag{9.132}$$

we get

$$\hat{V}_{Ap} = -\frac{q}{\hbar\omega}\left[\hat{H}_0, \vec{r}_{op}\right] \cdot \vec{E} \tag{9.133}$$

We can then calculate matrix elements

$$\begin{aligned} \langle m|\hat{V}_{Ap}|n\rangle &= -\frac{q}{\hbar\omega}\langle m|\hat{H}_0\vec{r}_{op} - \vec{r}_{op}\hat{H}_0|n\rangle \cdot \vec{E} \\ &= -\frac{q}{\hbar\omega}\left(\varepsilon_m - \varepsilon_n\right)\langle m|\vec{r}_{op} \cdot \vec{E}|n\rangle \\ &= \frac{\omega_{mn}}{\omega}\langle m|\hat{V}|n\rangle \end{aligned} \tag{9.134}$$

where

$$\omega_{mn} = \frac{\varepsilon_m - \varepsilon_n}{\hbar} \tag{9.135}$$

Thus, the matrix elements of $\hat{V}_{Ap}$ (from $\vec{A}\cdot\vec{p}_{op}$) differ from the matrix elements of $\hat{V}$ (from φ) by the factor ω_{mn}/ω. This implies different transition probabilities in first order except at resonance where $\omega_{mn}/\omega = 1$. The reason for the differences is as follows:

1. we assumed that the perturbation = 0 for $t \le 0\,,\ t \ge T$
2. there is no physical problem with $V = -q\vec{r}\cdot\vec{E}$ changing discontinuously
3. however, if $\vec{A}\cdot\vec{p}_{op}$ changes discontinuously, then the relation

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} \tag{9.136}$$

   generates spurious $\delta-$function type $\vec{E}$ fields

It is clear that one must exercise great care in choosing a starting point for perturbation theory.

9.2.2 Induced Emission and Absorption

We now look at the physics connected with the harmonic perturbation. Physically, it represents an electromagnetic wave interacting with the atom.

We consider the perturbing potential

$$\hat{V} = \begin{cases} -q\vec{r}\cdot\vec{E}_0\left(e^{-i\omega t} + e^{i\omega t}\right) & 0 < t < T \\ 0 & t < 0, t > T \end{cases} \tag{9.137}$$

where $\vec{E}_0$ = a constant vector which tells us the strength and polarization of the electromagnetic field.

This perturbing potential corresponds to monochromatic(single wavelength) electromagnetic radiation.

For an initial state $|\psi(0)\rangle = |i\rangle$ where $\hat{H}_0\,|i\rangle = \epsilon_i\,|i\rangle$ the probability, at any time $t \ge T$, that the atom will have a final energy ϵ_f is

$$P_{i\to f}(T) = \left|a_f^{(1)}(T)\right|^2 \tag{9.138}$$

to first order, where we have from our earlier derivation

$$a_f^{(1)}(T) = \frac{\langle f|\left(-q\vec{r}\cdot\vec{E}_0\right)|i\rangle}{\hbar}\left[\frac{1-e^{i(\omega_{fi}-\omega)T}}{\omega_{fi}-\omega} + \frac{1-e^{i(\omega_{fi}+\omega)T}}{\omega_{fi}+\omega}\right] \tag{9.139}$$

If $\epsilon_f > \epsilon_i$ this gives the probability amplitude for absorbing radiation and if $\epsilon_f < \epsilon_i$ this gives the probability amplitude for emitting radiation.

In the limit $T \to \infty$, as we saw earlier, $P_{i\to f}(t) = \Gamma_{i\to f} t$ which gives

$$\Gamma_{i\to f} = \frac{2\pi q^2}{\hbar} \frac{\left|\langle f|\, \vec{r}\cdot\vec{E}_0\, |i\rangle\right|^2}{4} \left[\delta(\varepsilon_i - \varepsilon_f - \hbar\omega) + \delta(\varepsilon_i - \varepsilon_f + \hbar\omega)\right] \tag{9.140}$$

In this expression

$$\begin{aligned} \delta(\varepsilon_i - \varepsilon_f - \hbar\omega) &\to \text{ absorption} \\ \delta(\varepsilon_i - \varepsilon_f + \hbar\omega) &\to \text{ emission} \end{aligned}$$

We then have the transition rates

$$\Gamma^a_{i\to f} = \Gamma^e_{i\to f} = \frac{2\pi q^2}{\hbar} \frac{\left|\langle f|\, \vec{r}\cdot\vec{E}_0\, |i\rangle\right|^2}{4} \tag{9.141}$$

These expressions are *zero/nonzero* or transitions are *not-allowed/allowed* depending on the matrix element of the perturbation, i.e.,

if for $i \to f \quad \langle f|\, \vec{r}\cdot\vec{E}_0\, |i\rangle = 0$, then the transition $i \to f$ is not allowed

These relationships between the quantum numbers of the initial and final states that tell us whether or not a transition is allowed are called *selection rules*.

To determine the selection rules for one-electron atoms we only need to consider matrix elements of the form

$$\langle n'\ell' m_{\ell'} m_{s'}|\, \vec{r}\cdot\vec{E}_0\, |n\ell m_\ell m_s\rangle \tag{9.142}$$

or we need to look at three matrix elements, namely,

$$\langle n'\ell' m_{\ell'} m_{s'}|\, \hat{x}\, |n\ell m_\ell m_s\rangle, \langle n'\ell' m_{\ell'} m_{s'}|\, \hat{y}\, |n\ell m_\ell m_s\rangle \text{ and } \langle n'\ell' m_{\ell'} m_{s'}|\, \hat{z}\, |n\ell m_\ell m_s\rangle$$

Now

$$z = r\cos\theta \propto rY_{10} = x_3 \tag{9.143}$$
$$x = r\sin\theta\cos\varphi \propto r(Y_{11} + Y_{1,-1}) = x_1 \tag{9.144}$$
$$y = r\sin\theta\sin\varphi \propto r(Y_{11} - Y_{1,-1}) = x_2 \tag{9.145}$$

A typical matrix element, therefore, will have a term like the following:

$$\langle sm_{s'} \mid sm_s\rangle \int R_{n'\ell'}(r) R_{n\ell}(r) r^3 dr \int Y^*_{\ell' m_{\ell'}} Y_{\ell m_\ell} Y_{1m} d\Omega \tag{9.146}$$

where $m = \pm 1, 0$.

The radial integral would equal zero only by accident implying that it is not part of the general selection rules, which must come from the other terms.

The term $\langle sm_{s'} \mid sm_s \rangle = \delta_{m_{s'} m_s}$ gives us a simple selection rule (this simple rule arises here because the interaction does not depend on spin). We have

$$m_{s'} = m_s \rightarrow \text{ the } \Delta m_s = 0 \qquad \textbf{SELECTION RULE} \tag{9.147}$$

The rest of the selection rules come from the angular integration terms

$$\int Y^*_{\ell' m_{\ell'}} Y_{\ell m_\ell} Y_{1m} d\Omega \tag{9.148}$$

This integral equals 0 unless $\ell + \ell' + 1 =$ an even integer. This rule follows from parity considerations. For any angular integration over all angles to be nonzero, the integrand must be even under the parity operation.

Now

$$Y_{\ell m} \rightarrow (-1)^\ell Y_{\ell m} \tag{9.149}$$

under the parity operation. Therefore

$$Y^*_{\ell' m \ell'} Y_{\ell m_\ell} Y_{1m} \rightarrow (-1)^{\ell + \ell' + 1} Y^*_{\ell' m \ell'} Y_{\ell m_\ell} Y_{1m} \tag{9.150}$$

which gives the stated rule.

The $\theta-$integration says that we must have

$$|\ell' - \ell| \leq 1 \leq \ell' + \ell$$

This corresponds to thinking of the integrand as made up of two states, namely,

$$\langle \ell' m_{\ell'} | \left(|\ell m_\ell\rangle \otimes |1m\rangle \right)$$

Our angular momentum addition rules say that

$$|\ell m_\ell\rangle \otimes |1m\rangle = |\ell + 1\rangle \oplus |\ell\rangle \oplus |\ell - 1\rangle$$

and the selection rule then follows from the orthogonality condition.

The two $\ell-$rules when combined imply the selection rule

$$\Delta m_\ell = m_{\ell'} - m_\ell = \pm 1, 0 \qquad \textbf{SELECTION RULE} \tag{9.151}$$

Thus, for transitions *within the electric dipole approximation*, as defined above, we have the SELECTION RULES

$$\Delta m_\ell = m_{\ell'} - m_\ell = \pm 1 \, , \, 0 \tag{9.152}$$

$$\Delta \ell = \ell' - \ell = \pm 1 \tag{9.153}$$

$$\Delta m_s = 0 = \Delta s \tag{9.154}$$

The derivation is more complex for multi-electron atoms due the complexity of the wave function (see next chapter), but it can be shown that, in general, the *SELECTION RULES* in the electric dipole approximation are

1. parity changes
2. $\Delta(\sum \ell_i) = \pm 1$
3. $\Delta S = 0$
4. $\Delta M_S = 0$
5. $\Delta L = \pm 1, 0$ (possibility of 0 which is not allowed for 1-electron atoms)
6. $\Delta M_L = \pm 1, 0$
7. $\Delta J = \pm 1, 0$
8. $\Delta M_J = \pm 1, 0$
9. $J = 0 \to J = 0$ is strictly forbidden

9.3 A Real Physical Process - Ionization

We now carry out the calculations of the transition rate for the ionization(transition to the continuum) of hydrogen by electromagnetic radiation.

The initial state of the electron is the ground state of hydrogen

$$|i\rangle = |100\rangle \text{ with energy } \varepsilon_i = \varepsilon_{100} = -\frac{e^2}{2a_0} \tag{9.155}$$

The final state of the electron in the ionization process is a free particle state(ionized electron)

$$|f\rangle = \left|\vec{k}\right\rangle \text{ with energy } \varepsilon_f = \varepsilon_k = \frac{\hbar^2 k^2}{2m} \tag{9.156}$$

We then have

$$\vec{p}_{op}\left|\vec{k}\right\rangle = \hbar\vec{k}\left|\vec{k}\right\rangle \tag{9.157}$$

which says this is a momentum eigenstate also (for free particles $[\hat{H}, \vec{p}_{op}] = 0$ and momentum and energy have the same eigenstates). The momentum is given by $\vec{p} = \hbar\vec{k}$. Since $\varepsilon_f > \varepsilon_i$ this is an absorption process and thus the transition rate is given by

$$\begin{aligned}\Gamma_{0\to\vec{k}} &= \text{transitionrate} = \lim_{t\to\infty} \frac{P_{0\to\vec{k}}(t)}{t} \\ &= \frac{2\pi}{\hbar}\left|\left\langle\vec{k}\right|\hat{V}\left|100\right\rangle\right|^2 \delta(\varepsilon_k - \varepsilon_{100} - \hbar\omega)\end{aligned} \tag{9.158}$$

where ω = frequency of the electromagnetic radiation and

$\hat{V} = -e\vec{r}\cdot\vec{\mathcal{E}}$ (electric dipole approximation)

$\vec{\mathcal{E}}$ = the electric field vector

We define

$$\begin{aligned} d\Gamma &= \text{ rate of transition into a small solid angle } d\Omega \\ &= \sum_{\vec{k} \text{ in } d\Omega} \Gamma_{0\to\vec{k}} \end{aligned} \tag{9.159}$$

For convenience we use a common trick and assume that the universe is a large box (side = L, volume = L^3). This allows us to normalize the plane wave states associated with the free electron. We have

$$\left\langle \vec{r} \mid \vec{k} \right\rangle = Ae^{i\vec{k}\cdot\vec{r}} \tag{9.160}$$

$$\begin{aligned} \left\langle \vec{k} \mid \vec{k} \right\rangle = 1 &= \int d^3\vec{r} \left\langle \vec{k} \mid \vec{r} \right\rangle \left\langle \vec{r} \mid \vec{k} \right\rangle \\ &= A^2 \int d^3\vec{r} e^{-i\vec{k}\cdot\vec{r}} e^{i\vec{k}\cdot\vec{r}} = A^2 \int d^3\vec{r} = A^2 L^3 \end{aligned} \tag{9.161}$$

or

$$A = \frac{1}{L^{3/2}} \tag{9.162}$$

Now there are

$$L^3 \frac{d^3\vec{k}}{(2\pi)^3} = \frac{L^3}{(2\pi)^3} d\Omega k^2 dk = \frac{L^3 mk}{(2\pi)^3\hbar^2} d\Omega d\varepsilon_k \tag{9.163}$$

states in the volume $d^3\vec{k}$ of phase space. This implies that there are

$$\frac{L^3 mk}{(2\pi)^3\hbar^2} \tag{9.164}$$

states per unit energy per unit solid angle.

Therefore,

$$d\Gamma = \sum_{\vec{k} \text{in} d\Omega} \Gamma_{0\to\vec{k}} \to d\Omega \int_0^\infty \frac{L^3 mk}{(2\pi)^3\hbar^2} d\varepsilon_k \frac{\pi e^2}{2\hbar} \left| \left\langle \vec{k} \right| \vec{r}\cdot\vec{\mathcal{E}} \left| 100 \right\rangle \right|^2 \delta(\varepsilon_k - \varepsilon_{100} - \hbar\omega)$$

Doing the integration (using the delta function) we get

$$d\Gamma = d\Omega \frac{L^3 mke^2}{16\pi^2\hbar^3} \left| \left\langle \vec{k} \right| \vec{r}\cdot\vec{\mathcal{E}} \left| 100 \right\rangle \right|^2 \tag{9.165}$$

where

$$\varepsilon_k = \varepsilon_{100} + \hbar\omega = \frac{\hbar^2 k^2}{2m} \tag{9.166}$$

$$k = \left(\frac{2m\omega}{\hbar} - \frac{1}{a_0^2} \right)^{1/2} \tag{9.167}$$

9.3.1 Evaluation of the Matrix Element

We have

$$\begin{aligned}\left\langle \vec{k}\right| \vec{r}\cdot\vec{\mathcal{E}}\left|100\right\rangle &= \int d^3\vec{r}\left\langle \vec{k}\mid\vec{r}\right\rangle \left\langle \vec{r}\right| \vec{r}\cdot\vec{\mathcal{E}}\left|100\right\rangle \\ &= \frac{1}{L^{3/2}}\int d^3\vec{r}e^{-i\vec{k}\cdot\vec{r}}\vec{r}\cdot\vec{\mathcal{E}}\left\langle \vec{r}\mid 100\right\rangle \\ &= \frac{1}{L^{3/2}}\int d^3\vec{r}e^{-i\vec{k}\cdot\vec{r}}\vec{r}\cdot\vec{\mathcal{E}}R_{10}(r)Y_{00} \\ &= \frac{2}{a_0^{3/2}L^{3/2}}\int d^3\vec{r}e^{-i\vec{k}\cdot\vec{r}}e^{-\frac{r}{a_0}}\vec{r}\cdot\vec{\mathcal{E}} \end{aligned} \tag{9.168}$$

We can arbitrarily choose $\vec{k}=k\hat{z}$ which gives

$$\left\langle \vec{k}\right| \vec{r}\cdot\vec{\mathcal{E}}\left|100\right\rangle = \frac{2}{a_0^{3/2}L^{3/2}}\int d\Omega dr r^2 e^{-ikr\cos\theta-\frac{r}{a_0}}\vec{r}\cdot\vec{\mathcal{E}} \tag{9.169}$$

using $\vec{k}\cdot\vec{R}=kz=kr\cos\theta$. Now in spherical-polar coordinates we have

$$\vec{\mathcal{E}} = \mathcal{E}\left(\sin\theta_\varepsilon\cos\varphi_\varepsilon\hat{e}_x+\sin\theta_\varepsilon\sin\varphi_\varepsilon\hat{e}_z+\cos\theta_\varepsilon\hat{e}_z\right) \tag{9.170}$$

$$\vec{r} = r\left(\sin\theta\cos\varphi\hat{e}_x+\sin\theta\sin\varphi\hat{e}_y+\cos\theta\hat{e}_z\right) \tag{9.171}$$

so that

$$\vec{\mathcal{E}}\cdot\vec{r} = \mathcal{E}r\left(\cos\theta\cos\theta_\varepsilon+\sin\theta\sin\theta_\varepsilon\cos(\varphi-\varphi_\varepsilon)\right) \tag{9.172}$$

We then have

$$\begin{aligned}&\left\langle \vec{k}\right| \vec{r}\cdot\vec{\mathcal{E}}\left|100\right\rangle \\ &= \frac{2\varepsilon}{a_0^{3/2}L^{3/2}}\int_0^{2\pi}d\varphi\int_0^{\pi}\sin\theta d\theta\int drr^3e^{-ikr\cos\theta-\frac{r}{a_0}}\left[\cos\theta\cos\theta_\varepsilon+\sin\theta\sin\theta_\varepsilon\cos(\varphi-\varphi_\varepsilon)\right]\end{aligned}$$

Since

$$\int_0^{2\pi}d\varphi\cos\varphi = \int_0^{2\pi}d\varphi\sin\varphi = 0 \tag{9.173}$$

the $\varphi-$ integration wipes out the $\cos(\varphi-\varphi_\varepsilon)$ term. Letting $x=\cos\theta$ we then have

$$\left\langle \vec{k}\right| \vec{r}\cdot\vec{\mathcal{E}}\left|100\right\rangle = \frac{4\pi\mathcal{E}\cos\theta_\varepsilon}{a_0^{3/2}L^{3/2}}\int_{-1}^{1}\left[\int_0^{\infty}drr^3e^{-ikrx-\frac{r}{a_0}}\right]xdx \tag{9.174}$$

Now

$$\int_0^{\infty}drr^3e^{-ikrx-\frac{r}{a_0}} = \int_0^{\infty}drr^3e^{-\alpha r} = \frac{3!}{\alpha^4} = \frac{6}{\left(ikx+\frac{1}{a_0}\right)^4} \tag{9.175}$$

Therefore, we have

$$\left\langle \vec{k} \right| \vec{r} \cdot \vec{\mathcal{E}} \left| 100 \right\rangle = \frac{4\pi \mathcal{E} \cos\theta_\varepsilon}{a_0^{3/2} L^{3/2}} \int_{-1}^{1} \frac{6x}{\left(ikx + \frac{1}{a_0} \right)^4} dx$$
$$= \frac{4\pi \mathcal{E} \cos\theta_\varepsilon}{a_0^{3/2} L^{3/2}} \frac{16 k a_0^5}{(1 + k^2 a_0^2)^3} \tag{9.176}$$

and

$$\left| \left\langle \vec{k} \right| \vec{r} \cdot \vec{\mathcal{E}} \left| 100 \right\rangle \right|^2 = \frac{4096 \pi^2 \mathcal{E}^2 \cos^2\theta_\varepsilon a_0^7 k^2}{L^3 (1 + k^2 a_0^2)^6} \tag{9.177}$$

where

$$k^2 = \frac{2m\omega}{\hbar} - \frac{1}{a_0^2} \tag{9.178}$$

Finally, we get

$$\Gamma_{ionization} = \int d\Gamma = \frac{L^3 m k e^2}{16\pi^2 \hbar^3} \int d\Omega_{\vec{k}} \left| \left\langle \vec{k} \right| \vec{r} \cdot \vec{\mathcal{E}} \left| 100 \right\rangle \right|^2 \tag{9.179}$$

where

$$d\Omega_{\vec{k}} = \text{ integration over the angles of } \vec{k}$$
$$\text{(varies direction of arbitrary z - axis)}$$

However, varying the z-direction is the same as keeping $\vec{\mathcal{E}}$ fixed and integrating over $d\Omega_{\vec{\mathcal{E}}}$. Therefore, we have

$$\Gamma_{ionization} = \frac{L^3 m k e^2}{16\pi^2 \hbar^3} \frac{4096 \pi^2 \mathcal{E}^2 a_0^7 k^2}{L^3 (1 + k^2 a_0^2)^6} \int d\Omega_{\vec{\mathcal{E}}} \cos^2\theta_\varepsilon$$
$$= \frac{64 \pi m e^2 \mathcal{E}^2 a_0^7}{3\hbar^3} \frac{k^3}{(1 + k^2 a_0^2)^6} \tag{9.180}$$

Now

$$1 + k^2 a_0^2 = \frac{2m\omega a_0^2}{\hbar} \tag{9.181}$$

and letting

$$\omega_0 = \frac{\hbar}{2m a_0^2} \tag{9.182}$$

we find that

$$1 + k^2 a_0^2 = \frac{\omega}{\omega_0} \text{ and } k = \frac{1}{a_0} \left(\frac{\omega}{\omega_0} - 1 \right)^{1/2} \tag{9.183}$$

and we get

$$\Gamma_{ionization} = \frac{64 \pi e^2 \mathcal{E}^2 a_0^3}{3\hbar} \left(\frac{\omega_0}{\omega} \right)^6 \left(\frac{\omega}{\omega_0} - 1 \right)^{3/2} \tag{9.184}$$

Thus, there exists a threshold energy for this process, i.e., it cannot occur unless the energy of the photon is greater than a minimum amount, which makes physical sense. In particular, we must have $\omega \geq \omega_0$ so that

$$\hbar\omega_0 = |\varepsilon_{100}| = E_{ionization} = \text{ minimum} = \text{threshold} \tag{9.185}$$

We also get the correct $6^{th}-$power term in the answer which agrees with experiment.

9.4 Adiabatic and Sudden Approximations

In standard time-dependent perturbation theory, we assume that the time-dependent perturbation is *weak*. An alternative approach, where we assume the time-dependence is *slow*, is called the *adiabatic approximation*.

Suppose that $\hat{H} = \hat{H}(g(t))$, where $g(t)$ tells us the dependence on time. This might correspond to a variation in time of some parameters. We still have

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}(g(t))\,|\psi(t)\rangle \tag{9.186}$$

and at *any instant of time* we have

$$\hat{H}(g(t))\,|n(g(t))\rangle = E_n(g(t))\,|n(g(t))\rangle \tag{9.187}$$

where $n(g(t))$ represents the quantum numbers describing the instantaneous state vector.

Let us assume that the instantaneous eigenvectors always form a complete set so that we can write

$$|\psi(t)\rangle = \sum_n \alpha_n(t)e^{i\beta_n(t)}\,|n(g(t))\rangle \tag{9.188}$$

where we have generalized the phase factor

$$e^{-\frac{i}{\hbar}\varepsilon_n t} \tag{9.189}$$

that appeared in a similar expression in our earlier derivations to include the term

$$\beta_n(t) = -\frac{1}{\hbar}\int_0^t E_n(g(t'))dt' \tag{9.190}$$

which is called the *dynamical phase*.

Inserting this expression for the state vector $|\psi(t)\rangle$ into the time-dependent

Schrodinger equation we get

$$
\begin{aligned}
&ih\frac{d}{dt}\sum_n \alpha_n(t)e^{i\beta_n(t)}\left|n(g(t))\right\rangle = \hat{H}(g(t))\sum_n \alpha_n(t)e^{i\beta_n(t)}\left|n(g(t))\right\rangle \\
&i\hbar\sum_n \frac{d\alpha_n(t)}{dt}e^{i\beta_n(t)}\left|n(g(t))\right\rangle + ih\sum_n \alpha_n(t)\frac{de^{i\beta_n(t)}}{dt}\left|n(g(t))\right\rangle \\
&\qquad + ih\sum_n \alpha_n(t)e^{i\beta_n(t)}\frac{d}{dt}\left|n(g(t))\right\rangle \\
&\quad = \sum_n \alpha_n(t)e^{i\beta_n(t)}\hat{H}(g(t))\left|n(g(t))\right\rangle = \sum_n \alpha_n(t)e^{i\beta_n(t)}E_n(g(t))\left|n(g(t))\right\rangle \\
&i\hbar\sum_n \frac{d\alpha_n(t)}{dt}e^{i\beta_n(t)}\left|n(g(t))\right\rangle + ih\sum_n \alpha_n(t)i\frac{d\beta_n(t)}{dt}e^{i\beta_n(t)}\left|n(g(t))\right\rangle \\
&\qquad + ih\sum_n \alpha_n(t)e^{i\beta_n(t)}\frac{d}{dt}\left|n(g(t))\right\rangle = \sum_n \alpha_n(t)e^{i\beta_n(t)}E_n(g(t))\left|n(g(t))\right\rangle
\end{aligned}
$$

Now

$$\frac{d\beta_n(t)}{dt} = -\frac{1}{\hbar}\frac{d}{dt}\int_0^t E_n(g(t'))dt' = -\frac{1}{\hbar}E_n(g(t)) \tag{9.191}$$

Therefore, we get

$$\sum_n \frac{d\alpha_n(t)}{dt}e^{i\beta_n(t)}\left|n(g(t))\right\rangle + \sum_n \alpha_n(t)e^{i\beta_n(t)}\frac{d}{dt}\left|n(g(t))\right\rangle = 0 \tag{9.192}$$

Applying the linear functional

$$\langle m| = \langle m(g(t))| \tag{9.193}$$

from the left we get

$$\sum_n \frac{d\alpha_n}{dt}e^{i\beta_n}\left\langle m \mid n\right\rangle + \sum_n \alpha_n e^{i\beta_n}\left\langle m\right|\frac{d}{dt}\left|n\right\rangle = 0 \tag{9.194}$$

Using

$$\left\langle m \mid n\right\rangle = \delta_{mn} \tag{9.195}$$

we have

$$\frac{d\alpha_m}{dt}e^{i\beta_m} = -\sum_n \alpha_n e^{i\beta_n}\left\langle m\right|\frac{d}{dt}\left|n\right\rangle \tag{9.196}$$

$$\frac{d\alpha_m}{dt} = -\sum_n \alpha_n e^{i(\beta_n-\beta_m)}\left\langle m\right|\frac{d}{dt}\left|n\right\rangle \tag{9.197}$$

Now taking the time derivative of the eigenvalue equation we have

$$\frac{d\hat{H}}{dt}\left|n\right\rangle + \hat{H}\frac{d}{dt}\left|n\right\rangle = \frac{dE_n}{dt}\left|n\right\rangle + E_n\frac{d}{dt}\left|n\right\rangle \tag{9.198}$$

Again, applying the linear functional $\langle m|$ from the left we get

$$\langle m| \frac{d\hat{H}}{dt} |n\rangle + \langle m| \hat{H} \frac{d}{dt} |n\rangle = \frac{dE_n}{dt} \langle m \mid n\rangle + E_n \langle m| \frac{d}{dt} |n\rangle \tag{9.199}$$

For $m \neq n$, using $\langle m| \hat{H} = \langle m| E_m$, we have

$$\langle m| \frac{d\hat{H}}{dt} |n\rangle + E_m \langle m| \frac{d}{dt} |n\rangle = E_n \langle m| \frac{d}{dt} |n\rangle \tag{9.200}$$

$$\langle m| \frac{d}{dt} |n\rangle = \frac{\langle m| \frac{d\hat{H}}{dt} |n\rangle}{E_n - E_m} \tag{9.201}$$

Thus, we finally have

$$\frac{d\alpha_m}{dt} = \sum_n \alpha_n e^{i(\beta_n - \beta_m)} \frac{\langle m| \frac{d\hat{H}}{dt} |n\rangle}{E_m - E_n} \tag{9.202}$$

We choose the *initial* state to be one of the instantaneous eigenvectors

$$|\psi(0)\rangle = |n(g(0))\rangle \tag{9.203}$$

which implies that

$$\begin{aligned} \alpha_n(0) &= 1 \\ \alpha_m(0) &= 0 \quad m \neq n \end{aligned}$$

Therefore, for $m \neq n$ at small t we have

$$\frac{d\alpha_m}{dt} \approx e^{i(\beta_n - \beta_m)t} \frac{\langle m| \frac{d\hat{H}}{dt} |n\rangle}{E_m - E_n} \tag{9.204}$$

We now assume that

$$\langle m| \frac{d\hat{H}}{dt} |n\rangle \text{ and } E_m - E_n \tag{9.205}$$

have slow time dependence and that to this order of approximation we can write

$$e^{i(\beta_n - \beta_m)t} = e^{i(E_m - E_n)\frac{t}{\hbar}} \tag{9.206}$$

which is what we would have if there was *no* extra time dependence.

We then get

$$\alpha_m(t) \approx -i\hbar \frac{\langle m| \frac{d\hat{H}}{dt} |n\rangle}{(E_m - E_n)^2} \left[e^{i(E_m - E_n)t} - 1 \right] \tag{9.207}$$

This implies that

$$\alpha_m(t) \text{ remains small for } m \neq n \tag{9.208}$$

The adiabatic theorem assumes that in the case where the system starts in an eigenstate $|n\rangle$ at $t = 0$, i.e.,

$$\alpha_m(t) = 0 \quad m \neq n \tag{9.209}$$

and that

$$|\psi(t)\rangle = e^{-i\beta_n t}\,|n(g(t))\rangle \tag{9.210}$$

which says that if the system was in the eigenstate $|n\rangle$ at $t = 0$, i.e.,

$$\hat{H}(g(0))\,|n\rangle = \hat{H}_0\,|n\rangle = \varepsilon_n\,|n\rangle \tag{9.211}$$

then at a later time t, it is still in the *same eigenstate* $|n(g(t))\rangle$ of the new Hamiltonian $\hat{H}(g(t))$, i.e.,

$$\hat{H}(g(t))\,|n(t)\rangle = E_n(t)\,|n(t)\rangle \tag{9.212}$$

This result is *independent* of the size of the perturbation. It depends only on the change in time being *slow*.

This means that if we start with a particle in the ground state of a harmonic oscillator potential

$$V = \frac{1}{2}k(0)x^2 \rightarrow \psi_0(k(0), x) \tag{9.213}$$

and assume that

$$k(0) \rightarrow k(T) \tag{9.214}$$

slowly, the particle ends up in the ground state of the harmonic oscillator potential

$$V = \frac{1}{2}k(T)x^2 \rightarrow \psi_0(k(T), x) \tag{9.215}$$

to within a phase factor.

The opposite result comes from the so-called *sudden approximation* , where the change occurs so fast that no changes of the state vector are possible.

Since the state vector does not change at all, if you are in the ground state and a sudden change in the parameters occurs, then you remain in the ground state for the *old parameters*. This is *not* the ground state with new parameters. It is some *linear combination* of the new states.

Let us look at the adiabatic approximation in another way. We consider a time dependent part of the Hamiltonian of the form

$$\hat{H}(t) = \hat{H}_0 + \hat{H}' \quad , \quad \hat{H}' = \hat{V}\,f(t) \tag{9.216}$$

where $f(t)$ has the form shown in Figure 9.5 below.

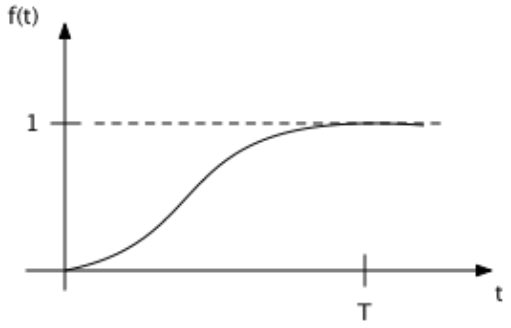


Figure 9.5: Time Dependence

We assume that the particle starts out in the n^{th} eigenstate of $\hat{H}_0 = \hat{H}(0)$

$$|\psi(0)\rangle = |n\rangle_i \tag{9.217}$$

where the subscripts are

$$\begin{aligned} i &\rightarrow \text{ initial parameters} \\ f &\rightarrow \text{ final parameters} \end{aligned}$$

The state vectors is assumed to change in time to $|\psi(t)\rangle$.

If $\hat{V}$ is small, we can write, using first order time-independent perturbation theory

$$|m\rangle_f = |m\rangle_i + \sum_{k \neq m} \frac{V_{km}}{E_m - E_k} |k\rangle_i \tag{9.218}$$

where

$$V_{km} = {}_i \langle k| \hat{V} |m\rangle_i \tag{9.219}$$

On the other hand, first order time-dependent perturbation theory implies that

$$|\psi(t)\rangle = \sum_n \alpha_n(t) e^{-\frac{i}{\hbar} E_n t} |n\rangle_i \tag{9.220}$$

with

$$\alpha_n(t) = 1 - \frac{i}{\hbar} V_{nn} \int_0^t f(t')dt' \tag{9.221}$$

$$\alpha_m(t) = -\frac{i}{\hbar} V_{mn} \int_0^t f(t') e^{i(E_m - E_n)\frac{t'}{\hbar}} dt' \quad , \quad m \neq n \tag{9.222}$$

Now we are assuming that df/dt is small, we integrate by parts to get

$$\begin{aligned}
\alpha_m(t) &= -\frac{i}{\hbar} V_{mn} \int_0^t f(t') e^{i(E_m - E_n)\frac{t'}{\hbar}} dt' \\
&= -\frac{V_{mn}}{E_m - E_n} \int_0^t f(t') \frac{d}{dt'} \left[e^{i(E_m - E_n)\frac{t'}{\hbar}} \right] dt' \\
&= -\frac{V_{mn}}{E_m - E_n} \left[f(t') e^{i(E_m - E_n)\frac{t'}{\hbar}} \right]_0^t \\
&\quad + \frac{V_{mn}}{E_m - E_n} \int_0^t \frac{df(t')}{dt'} e^{i(E_m - E_n)\frac{t'}{\hbar}} dt'
\end{aligned} \tag{9.223}$$

Using $f(0) = 0$ and neglecting the last term because df/dt is small gives

$$|\psi(T)\rangle = \left[(1 - i\frac{V_{nn}}{\hbar}\gamma) |n\rangle_i - \sum_{q \neq n} \frac{V_{qn}}{E_q - E_n} |q\rangle_i \right] e^{-\frac{i}{\hbar} E_n T} \tag{9.224}$$

where

$$\gamma = \int_0^T f(t) dt \tag{9.225}$$

Therefore

$$\begin{aligned}
{}_f\langle n \,|\, \psi(T)\rangle &= \left[{}_i\langle n| + \sum_{k \neq n} \frac{V_{kn}}{E_n - E_k} {}_i\langle k| \right] \left[(1 - i\frac{V_{nn}}{\hbar}\gamma) |n\rangle_i - \sum_{q \neq n} \frac{V_{qn}}{E_q - E_n} |q\rangle_i \right] e^{-\frac{i}{\hbar} E_n T} \\
&= \left[1 - i\frac{V_{nn}}{\hbar}\gamma + \sum_{q \neq n} \frac{|V_{qn}|^2}{(E_n - E_q)^2} \right] e^{-\frac{i}{\hbar} E_n T}
\end{aligned} \tag{9.226}$$

$$\begin{aligned}
{}_f\langle m \,|\, \psi(T)\rangle &= \left[{}_i\langle m| + \sum_{k \neq m} \frac{V_{km}}{E_m - E_k} {}_i\langle k| \right] \left[(1 - i\frac{V_{nn}}{\hbar}\gamma) |n\rangle_i - \sum_{q \neq n} \frac{V_{qn}}{E_q - E_n} |q\rangle_i \right] e^{-\frac{i}{\hbar} E_n T} \\
&= \left[-i\gamma \frac{V_{nn} V_{nm}}{\hbar(E_m - E_n)} + \sum_{m \neq k \neq n} \frac{V_{nk} V_{km}}{(E_n - E_k)(E_m - E_k)} \right] e^{-\frac{i}{\hbar} E_n T}
\end{aligned} \tag{9.227}$$

Note that all the first order terms cancel in the last expression. If we only keep terms to first order (which is consistent with the derivation) we then have

$${}_f\langle k \,|\, \psi(T)\rangle = \begin{cases} \left[1 - i\frac{V_{nn}}{\hbar}\gamma \right] & k = n \\ 0 & k \neq n \end{cases} \tag{9.228}$$

which implies that

$$\begin{aligned}
&|{}_f\langle n \,|\, \psi(T)\rangle|^2 = 1 + \frac{|V_{nn}|^2 \gamma^2}{\hbar^2} \to 1 \text{ to first order} \\
&|{}_f\langle m \,|\, \psi(T)\rangle|^2 = 0 \text{ to first order } m \neq n
\end{aligned}$$

which is the *adiabatic approximation*.

Therefore, the adiabatic approximation says:

> if the Hamiltonian $\hat{H}(t)$ changes slowly in time,
> then there will be no transitions from the
> eigenstate $|n\rangle_i$ of $\hat{H}(0)$ to a *different*
> eigenstate $|m\rangle_f$ of $\hat{H}(t)$

In time-dependent perturbation theory, we assume the perturbation is turned on and off and thus $\hat{H}(0) = \hat{H}(T)$, i.e., we have the *same* unperturbed Hamiltonian at the end.

The transitions in that case are from one eigenvector of the unperturbed Hamiltonian to another eigenvector of the same unperturbed Hamiltonian.

This implies a first-order transition amplitude and hence a second-order transition probability.

In the adiabatic approximation, however, we have a second-order transition amplitude ($m \neq n$) and hence a fourth-order transition probability. That is why we can assume that the transition probability for $m \neq n$ is equal to zero.

The first derivation gives the adiabatic approximation for any size perturbation. In the second derivation,however, we not only assumed a slow change in time, but also assumes a small perturbation so that we could use first order perturbation theory.

What happens in the second derivation if the perturbation is not small?

The way to handle this is to divide the time interval $(0, T)$ into N subintervals such that the perturbation ΔV is small within any subinterval. In fact, it is of $O(V/N)$. Thus, if N is large, ΔV is small.

We then apply our arguments to each subinterval. If the transition amplitude is first-order in the perturbation, then the total transition amplitude behaves like

$$N\left(\frac{V}{N}\right) \to V \tag{9.229}$$

with each of the N steps giving a contribution proportional to ΔV. This says that the net result is of order V, and thus, if V is large, the transition amplitude will be large.

However, the transition amplitude is second-order and thus the total transition

amplitude behaves like

$$N\left(\frac{V}{N}\right)^2 \to \frac{V^2}{N} \to 0 \text{as} N \to \infty \tag{9.230}$$

Therefore, the transition amplitude ($m \neq n$) is zero *independentofthesize* of V.

An Example

Let us consider a 1–dimensional square well where

$$V(x) = \begin{cases} 0 & |x| \leq \frac{a}{2} \\ \infty & |x| > \frac{a}{2} \end{cases} \tag{9.231}$$

The eigenfunctions and energies are

$$\psi_n(x) = \begin{cases} \cos\frac{n\pi x}{2a} & n = 1, 3, 5, ... \\ \sin\frac{n\pi x}{2a} & n = 2, 4, 6, ... \end{cases} \tag{9.232}$$

for $|x| \leq \frac{a}{2}$ and zero otherwise and

$$E_n = \frac{\pi^2\hbar^2 n^2}{8ma^2} \quad n = 1, 2, 3, ... \tag{9.233}$$

Suppose that we change the size of the well and ask what happens to the ground state in the sudden and adiabatic approximations.

Sudden

$$\psi_1(x) = \cos\frac{\pi x}{2a} \text{ before} \tag{9.234}$$

leads to

$$\psi(x) = \cos\frac{\pi x}{2a} \text{ after (no change in the wave function)} \tag{9.235}$$

However, after the change we have new eigenfunctions and energies

$$\psi'_n(x) = \begin{cases} \cos\frac{n\pi x}{4\bar{a}} & n = 1, 3, 5, ... \\ \sin\frac{n\pi x}{4\bar{a}} & n = 2, 4, 6, ... \end{cases} \tag{9.236}$$

for $|x| \leq \frac{a}{2}$ and zero otherwise and

$$E'_n = \frac{\pi^2\hbar^2 n^2}{8m\bar{a}^2} \quad n = 1, 2, 3, ... \tag{9.237}$$

The state of the system is still an eigenstate of the old well and, thus, is *not* an eigenstate of the new well. In fact, we have

$$\psi(x) = \cos\frac{\pi x}{2a} = \sum_n b_n \psi'_n(x) \tag{9.238}$$

Adiabatic

$$\psi_1(x) = \cos\frac{\pi x}{2a} \text{ before (ground state of old well)} \tag{9.239}$$

$$\psi_1'(x) = \cos\frac{\pi x}{4a} \text{ after (ground state of new well)} \tag{9.240}$$

The state of the system is an eigenstate of the new well and, thus, is *not* an eigenstate of the old well any longer. In fact, we have

$$\psi_n'(x) = \cos\frac{\pi x}{4a} = \sum_n b_n \psi_n(x) \tag{9.241}$$

It is a superposition of the old energy eigenstates.

9.5 Problems

9.5.1 Square Well Perturbed by an Electric Field

At time $t = 0$, an electron is known to be in the $n = 1$ eigenstate of a 1–dimensional infinite square well potential

$$V(x) = \begin{cases} \infty & \text{for } |x| > a/2 \\ 0 & \text{for } |x| < a/2 \end{cases}$$

At time $t = 0$, a uniform electric field of magnitude $\mathcal{E}$ is applied in the direction of increasing x. This electric field is left on for a short time τ and then removed. Use time-dependent perturbation theory to calculate the probability that the electron will be in the $n = 2, 3$ eigenstates at some time $t > \tau$.

9.5.2 3-Dimensional Oscillator in an electric field

A particle of mass M, charge e, and spin zero moves in an attractive potential

$$V(x, y, z) = k\left(x^2 + y^2 + z^2\right) \tag{9.242}$$

(a) Find the three lowest energy levels E_0, E_1, E_2 and their associated degeneracy.

(b) Suppose a small perturbing potential $Ax\cos\bar{\omega}t$ causes transitions among the various states in (a). Using a convenient basis for degenerate states, specify in detail the allowed transitions neglecting effects proportional to A^2 or higher.

(c) In (b) suppose the particle is in the ground state at time $t = 0$. Find the probability that the energy is E_1 at time t.

9.5.3 Hydrogen in decaying potential

A hydrogen atom (assume spinless electron and proton) in its ground state is placed between parallel plates and subjected to a uniform weak electric field

$$\vec{\mathcal{E}} = \begin{cases} 0 & \text{for } t < 0 \\ \vec{\mathcal{E}}_0 e^{-\alpha t} & \text{for } t > 0 \end{cases}$$

Find the 1^{st}order probability for the atom to be in any of the $n = 2$ states after a long time.

9.5.4 2 spins in a time-dependent potential

Consider a composite system made up of two spin = 1/2 objects. For $t < 0$, the Hamiltonian does not depend on spin and can be taken to be zero by suitably adjusting the energy scale. For $t > 0$, the Hamiltonian is given by

$$\hat{H} = \left(\frac{4\Delta}{\hbar^2}\right) \vec{S}_1 \cdot \vec{S}_2$$

Suppose the system is in the state $|+-\rangle$ for $t \leq 0$. Find, as a function of time, the probability for being found in each of the following states $|++\rangle$, $|-+\rangle$ and $|--\rangle$.

(a) by solving the problem exactly.

(b) by solving the problem assuming the validity of 1^{st}–order time-dependent perturbation theory with $\hat{H}$ as a perturbation switched on at $t = 0$. Under what conditions does this calculation give the correct results?

9.5.5 A Variational Calculation of the Deuteron Ground State Energy

Use the empirical potential energy function

$$V(r) = -Ae^{-r/a}$$

where $A = 32.7\,MeV$, $a = 2.18 \times 10^{-13}\,cm$, to obtain a variational approximation to the energy of the ground state energy of the deuteron ($\ell = 0$).

Try a simple variational function of the form

$$\phi(r) = e^{-\alpha r/2a}$$

where α is the variational parameter to be determined. Calculate the energy in terms of α and minimize it. Give your results for α and E in MeV. The experimental value of E is $-2.23\,MeV$ (your answer should be VERY close! Is your answer above this? [HINT: do not forget about the *reduced mass* in this problem]

9.5.6 Sudden Change - Don't Sneeze

An experimenter has carefully prepared a particle of mass m in the first excited state of a one dimensional harmonic oscillator, when he sneezes and knocks the center of the potential well a small distance a to one side. It takes him a time T to blow his nose, and when he has done so, he immediately puts the center back where it was. Find, to lowest order in a, the probabilities P_0 and P_2 that the oscillator will now be in its ground state and its second excited state.

9.5.7 Another Sudden Change - Cutting the spring

A particle is allowed to move in one dimension. It is initially coupled to two identical harmonic springs, each with spring constant K. The springs are symmetrically fixed to the points $\pm a$ so that when the particle is at $x = 0$ the classical force on it is zero.

(a) What are the energy eigenvalues of the particle when it is connected to both springs?

(b) What is the wave function in the ground state?

(c) One spring is suddenly cut, leaving the particle bound to only the other one. If the particle is in the ground state before the spring is cut, what is the probability that it is still in the ground state after the spring is cut?

9.5.8 Another perturbed oscillator

Consider a particle bound in a simple harmonic oscillator potential. Initially($t < 0$), it is in the ground state. At $t = 0$ a perturbation of the form

$$H'(x,t) = Ax^2 e^{-t/\tau}$$

is switched on. Using time-dependent perturbation theory, calculate the probability that, after a sufficiently long time ($t \gg \tau$), the system will have made a transition to a given excited state. Consider all final states.

9.5.9 Nuclear Decay

Nuclei sometimes decay from excited states to the ground state by internal conversion, a process in which an atomic electron is emitted instead of a photon. Let the initial and final nuclear states have wave functions $\varphi_i(\vec{r}_1, \vec{r}_2, ..., \vec{r}_Z)$ and $\varphi_f(\vec{r}_1, \vec{r}_2, ..., \vec{r}_Z)$, respectively, where $\vec{r}_i$ describes the protons. The perturbation giving rise to the transition is the proton-electron interaction,

$$W = -\sum_{j=1}^{Z} \frac{e^2}{|\vec{r} - \vec{r}_j|}$$

where $\vec{r}$ is the electron coordinate.

(a) Write down the matrix element for the process in lowest-order perturbation theory, assuming that the electron is initially in a state characterized by the quantum numbers $(n\ell m)$, and that its energy, after it is emitted, is large enough so that its final state may be described by a plane wave, Neglect spin.

(b) Write down an expression for the internal conversion rate.

(c) For light nuclei, the nuclear radius is much smaller than the Bohr radius for a give Z, and we can use the expansion

$$\frac{1}{|\vec{r}-\vec{r}_j|} \approx \frac{1}{r} + \frac{\vec{r}\cdot\vec{r}_j}{r^3}$$

Use this expression to express the transition rate in terms of the dipole matrix element

$$\vec{d} = \langle\varphi_f| \sum_{j=1}^{Z} \vec{r}_j \, |\varphi_i\rangle$$

9.5.10 Time Evolution Operator

A one-dimensional anharmonic oscillator is given by the Hamiltonian

$$H = \hbar\omega \left(a^\dagger a + 1/2\right) + \lambda a^\dagger a a$$

where λ is a constant. First compute a^+ and a in the interaction picture and then calculate the time evolution operator $U(t, t_0)$ to lowest order in the perturbation.

9.5.11 Two-Level System

Consider a two-level system $|\psi_a\rangle$, $|\psi_b\rangle$ with energies E_a , E_b perturbed by a jolt $H'(t) = \hat{U}\delta(t)$ where the operator $\hat{U}$ has only off-diagonal matrix elements (call them U). If the system is initially in the state a, find the probability $P_{a\to b}$ that a transition occurs. Use only the lowest order of perturbation theory that gives a nonzero result.

9.5.12 Instantaneous Force

Consider a simple harmonic oscillator in its ground state. An instantaneous force imparts momentum p_0 to the system. What is the probability that the system will stay in its ground state?

9.5.13 Hydrogen beam between parallel plates

A beam of excited hydrogen atoms in the $2s$ state passes between the plates of a capacitor in which a uniform electric field exists over a distance L. The hydrogen atoms have a velocity v along the $x-$axis and the electric field $\vec{\mathcal{E}}$ is directed along the $z-$axis as shown in the figure.

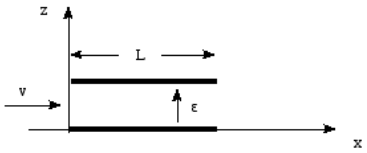


Figure 9.6: Hydrogen beam between parallel plates

All of the $n = 2$ states of hydrogen are degenerate in the absence of the field $\vec{\mathcal{E}}$, but certain of them mix (Stark effect) when the field is present.

(a) Which of the $n = 2$ states are connected (mixed) in first order via the electric field perturbation?

(b) Find the linear combination of the $n = 2$ states which removes the degeneracy as much as possible.

(c) For a system which starts out in the $2s$ state at $t = 0$, express the wave function at time $t \leq L/v$. No perturbation theory needed.

(d) Find the probability that the emergent beam contains hydrogen in the various $n = 2$ states.

9.5.14 Particle in a Delta Function and an Electric Field

A particle of charge q moving in one dimension is initially bound to a delta function potential at the origin. From time $t = 0$ to $t = \tau$ it is exposed to a constant electric field $\mathcal{E}_0$ in the $x-$direction as shown in the figure below:

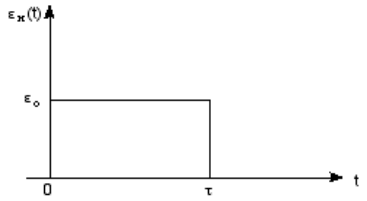


Figure 9.7: Electric Field

The object of this problem is to find the probability that for $t > \tau$ the particle will be found in an unbound state with energy between E_k and $E_k + dE_k$.

(a) Find the normalized bound-state energy eigenfunction corresponding to the delta function potential $V(x) = -A\delta(x)$.

(b) Assume that the unbound states may be approximated by free particle states with periodic boundary conditions in a box of length L. Find the normalized wave function of wave vector k, $\psi_k(x)$, the density of states as

a function of k, $D(k)$ and the density of states as a function of free-particle energy E_k, $D(E_k)$.

(c) Assume that the electric field may be treated as a perturbation. Write down the perturbation term in the Hamiltonian, $\hat{H}_1$, and find the matrix element of $\hat{H}_1$ between the initial and the final state $\langle 0|\hat{H}_1|k\rangle$.

(d) The probability of a transition between an initially occupied state $|I\rangle$ and a final state $|F\rangle$ due to a weak perturbation $\hat{H}_1(t)$ is given by

$$P_{I\to F}(t) = \frac{1}{\hbar^2}\left|\int_{-\infty}^{t}\langle F|\hat{H}_1(t')|I\rangle\, e^{i\omega_{FI}t'}dt'\right|^2$$

where $\omega_{FI} = (E_F - E_I)/\hbar$. Find an expression for the probability $P(E_k)dE_k$ that the particle will be in an unbound state with energy between E_k and $E_k + dE_k$ for $t > \tau$.

9.5.15 Nasty time-dependent potential [complex integration needed]

A one-dimensional simple harmonic oscillator of frequency ω is acted upon by a time-dependent, but spatially uniform force (not potential!)

$$F(t) = \frac{(F_0\tau/m)}{\tau^2 + t^2} \quad , \quad -\infty < t < \infty$$

At $t = -\infty$, the oscillator is known to be in the ground state. Using time-dependent perturbation theory to 1^{st}–order, calculate the probability that the oscillator is found in the 1^{st} excited state at $t = +\infty$.

Challenge: $F(t)$ is so normalized that the impulse

$$\int F(t)dt$$

imparted to the oscillator is always the same, that is, independent of τ; yet for $\tau >> 1/\omega$, the probability for excitation is essentially negligible. Is this reasonable?

9.5.16 Natural Lifetime of Hydrogen

Though in the absence of any perturbation, an atom in an excited state will stay there forever(it is a stationary state), in reality, it will *spontaneously decay* to the ground state. Fundamentally, this occurs because the atom is always perturbed by vacuum fluctuations in the electromagnetic field. The spontaneous emission rate on a dipole allowed transition from the initial excited state $|\psi_e\rangle$ to all allowed ground states $|\psi_g\rangle$ is,

$$\Gamma = \frac{4}{3\hbar}k^3\sum_g \left|\langle\psi_g|\hat{\vec{d}}\,|\psi_e\rangle\right|^2$$

where $k = \omega_{eg}/c = (E_e - E_g)/\hbar c$ is the emitted photon's wave number.

Consider now hydrogen including fine structure. For a given sublevel, the spontaneous emission rate is

$$\Gamma_{(nLJM_J)\to(n'L'J')} = \frac{4}{3\hbar}k^3 \sum_{M'_J} \left|\langle n'L'J'M'_J|\,\vec{d}\,|nLJM_J\rangle\right|^2$$

(a) Show that the spontaneous emission rate is independent of the initial M_J. Explain this result physically.

(b) Calculate the lifetime ($\tau = 1/\Gamma$) of the $2P_{1/2}$ state in seconds.

9.5.17 Oscillator in electric field

Consider a simple harmonic oscillator in one dimension with the usual Hamiltonian

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

Assume that the system is in its ground state at $t = 0$. At $t = 0$ an electric field $\vec{\mathcal{E}} = \mathcal{E}\hat{x}$ is switched on, adding a term to the Hamiltonian of the form

$$\hat{H}' = e\mathcal{E}\hat{x}$$

(a) What is the new ground state energy?

(b) Assuming that the field is switched on in a time much faster than $1/\omega$, what is the probability that the particle stays in the unperturbed ground state?

9.5.18 Spin Dependent Transitions

Consider a spin= 1/2 particle of mass m moving in three kinetic dimensions, subject to the spin dependent potential

$$\hat{V}_1 = \frac{1}{2}k\,|-\rangle\langle -| \otimes |\vec{r}|^2$$

where k is a real positive constant, $\vec{r}$ is the three-dimensional position operator, and $\{|-\rangle, |+\rangle\}$ span the spin part of the Hilbert space. Let the initial state of the particle be prepared as

$$|\Psi_0\rangle = |-\rangle \otimes |0\rangle$$

where $|0\rangle$ corresponds to the ground state of the harmonic (motional) potential.

(a) Suppose that a perturbation

$$\hat{W} = \hbar\Omega\left(|-\rangle\langle +| + |+\rangle\langle -|\right) \otimes \hat{I}^{CM}$$

where $\hat{I}^{CM}$ denotes the identity operator on the motional Hilbert space, is switched on at time $t = 0$.

Using Fermi's Golden Rule compute the rate of transitions out of $|\Psi_0\rangle$.

(b) Describe qualitatively the evolution induced by $\hat{W}$, in the limits $\Omega \gg \sqrt{k/m}$ and $\Omega \ll \sqrt{k/m}$. HINT: Make sure you understand part(c).

(c) Consider a different spin-dependent potential

$$\hat{V}_2 = |+\rangle\langle+| \otimes \Sigma_+(\vec{x}) + |-\rangle\langle-| \otimes \Sigma_-(\vec{x})$$

where $\Sigma_\pm(\vec{x})$ denote the motional potentials

$$\Sigma_+(\vec{x}) = \begin{cases} +\infty & |x| < a \\ 0 & |x| \geq a \end{cases}$$

$$\Sigma_-(\vec{x}) = \begin{cases} 0 & |x| < a \\ +\infty & |x| \geq a \end{cases}$$

and a is a positive real constant. Let the initial state of the system be prepared as

$$|\Psi_0\rangle = |-\rangle \otimes |0'\rangle$$

where $|0'\rangle$ corresponds to the ground state of $\Sigma_-(\vec{x})$. Explain why Fermi's Golden Rule predicts a vanishing transition rate for the perturbation $\hat{W}$ specified in part (a) above.

9.5.19 The Driven Harmonic Oscillator

At $t = 0$ a 1–dimensional harmonic oscillator with natural frequency ω is driven by the perturbation

$$H_1(t) = -Fxe^{-i\Omega t}$$

The oscillator is initially in its ground state at $t = 0$.

(a) Using the lowest order perturbation theory to get a nonzero result, find the probability that the oscillator will be in the $2nd$ excited state $n = 2$ at time $t > 0$. Assume $\omega \neq \Omega$.

(b) Now begin again and do the simpler case, $\omega = \Omega$. Again, find the probability that the oscillator will be in the $2nd$ excited state $n = 2$ at time $t > 0$

(c) Expand the result of part (a) for small times t, compare with the results of part (b), and interpret what you find.

In discussing the results see if you detect any parallels with the driven classical oscillator.

9.5.20 A Novel One-Dimensional Well

Using tremendous skill, physicists in a molecular beam epitaxy lab, use a graded semiconductor growth technique to make a GaAs(Gallium Arsenide) wafer containing a single 1-dimensional (Al,Ga)As quantum well in which an electron is confined by the potential $V = kx^2/2$.

(a) What is the Hamiltonian for an electron in this quantum well? Show that $\psi_0(x) = N_0 e^{-\alpha x^2/2}$ is a solution of the time-independent Schrodinger equation with this Hamiltonian and find the corresponding eigenvalue. Assume here that $\alpha = m\omega/\hbar$, $\omega = \sqrt{k/m}$ and m is the mass of the electron. Also assume that the mass of the electron in the quantum well is the same as the free electron mass (not always true in solids).

(b) Let us define the raising and lowering operators $\hat{a}$ and $\hat{a}^+$ as

$$\hat{a}^+ = \frac{1}{\sqrt{2}}\left(\frac{d}{dy} - y\right) \quad , \quad \hat{a} = \frac{1}{\sqrt{2}}\left(\frac{d}{dy} + y\right)$$

where $y = \sqrt{m\omega/\hbar}x$. Find the wavefunction which results from operating on ψ_0 with $\hat{a}^+$ (call it $\psi_1(x)$). What is the eigenvalue of ψ_1 in this quantum well? You can just state the eigenvalue based on your knowledge - there is no need to derive it.

(c) Write down the Fermi's Golden Rule expression for the rate of a transition (induced by an oscillating perturbation from electromagnetic radiation) occuring between the lowest energy eigenstate and the first excited state. State the assumptions that go into the derivation of the expression.

(d) Given that $k = 3.0\,kg/s^2$, what photon wavelength is required to excite the electron from state ψ_0 to state ψ_1? Use symmetry arguments to decide whether this is an allowed transition (explain your reasoning); you might want to sketch $\psi_0(x)$ and $\psi_1(x)$ to help your explanation.

(e) Given that

$$\hat{a}\,|\nu\rangle = \sqrt{\nu}\,|\nu - 1\rangle \quad , \quad \hat{a}^+\,|\nu\rangle = -\sqrt{\nu + 1}\,|\nu + 1\rangle$$

evaluate the transition matrix element $\langle 0|\,x\,|1\rangle$. (HINT: rewrite x in terms of $\hat{a}$ and $\hat{a}^+$). Use your result to simplify your expression for the transition rate.

9.5.21 The Sudden Approximation

Suppose we specify a three-dimensional Hilbert space $\mathcal{H}_A$ and a time-dependent Hamiltonian operator

$$H(t) = \alpha \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} + \beta(t) \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -2 \end{pmatrix}$$

where α and $\beta(t)$ are real-valued parameters (with units of energy). Let $\beta(t)$ be given by a step function

$$\beta(t) = \begin{cases} \alpha & t \leq 0 \\ 0 & t > 0 \end{cases}$$

The Schrodinger equation can clearly be solved by standard methods in the intervals $t = [-\infty, 0]$ and $t = (0, +\infty]$, within each of which H remains constant. We can use the so-called *sudden approximation* to deal with the discontinuity in H at $t = 0$, which simply amounts to assuming that

$$|\Psi(0_+)\rangle = |\Psi(0_-)\rangle$$

Suppose the system is initially prepared in the ground state of the Hamiltonian at $t = -1$. Use the Schrodinger equation and the sudden approximation to compute the subsequent evolution of $|\Psi(t)\rangle$ and determine the function

$$f(t) = \langle |\Psi(0)\rangle \, | \, |\Psi(t)\rangle \rangle \quad , \quad t \geq 0$$

Show that $|f(t)|^2$ is periodic. What is the frequency? How is it related to the Hamiltonian?

9.5.22 The Rabi Formula

Suppose the total Hamiltonian for a spin$-1/2$ particle is

$$H = -\gamma \left[B_0 S_z + b_1 \left(\cos(\omega t) S_x + \sin(\omega t) S_y \right) \right]$$

which includes a static field B_0 in the z direction plus a rotating field in the $x - y$ plane. Let the state of the particle be written

$$|\Psi(t)\rangle = a(t) \, |+_z\rangle + b(t) \, |-_z\rangle$$

with normalization $|a|^2 + |b|^2 = 1$ and initial conditions

$$a(0) = 0 \quad , \quad b(0) = 1$$

Show that

$$|a(t)|^2 = \frac{(\gamma b_1)^2}{\Delta^2 + (\gamma b_1)^2} \sin^2 \left(\frac{t}{2} \sqrt{\Delta^2 + (\gamma b_1)^2} \right)$$

where $\Delta = -\gamma B_0 - \omega$. This expression is known as the *Rabi Formula*.

9.5.23 Rabi Frequencies in Cavity QED

Consider a two-level atom whose pure states can be represented by vectors in a two-dimensional Hilbert space $\mathcal{H}_A$. Let $|g\rangle$ and $|e\rangle$ be a pair of orthonormal basis states of $\mathcal{H}_A$ representing the ground and excited states of the atom, respectively. Consider also a microwave cavity whose lowest energy pure states can be

described by vectors in a three-dimensional Hilbert space $\mathcal{H}_C$. Let $\{|0\rangle, |1\rangle, |2\rangle\}$ be orthonormal basis states representing zero, one and two microwave photons in the cavity.

The experiment is performed by sending a stream of atoms through the microwave cavity. The atoms pass through the cavity one-by-one. Each atom spends a total time t inside the cavity (which can be varied by adjusting the velocities of the atoms). Immediately upon exiting the cavity each atom hits a detector that measures the atomic projection operator $P_e = |e\rangle\langle e|$.

Just before each atom enters the cavity, we can assume that the joint state of that atom and the microwave cavity is given by the factorizable pure state

$$|\Psi(0)\rangle = |g\rangle \otimes (c_0|0\rangle + c_1|1\rangle + c_2|2\rangle)$$

where $|c_0|^2 + |c_1|^2 + |c_2|^2 = 1$

(a) Suppose the Hamiltonian for the joint atom-cavity system vanishes when the atom is not inside the cavity and when the atom is inside the cavity the Hamiltonian is given by

$$H_{AC} = \hbar\nu\,|e\rangle\langle g| \otimes \left(|0\rangle\langle 1| + \sqrt{2}\,|1\rangle\langle 2|\right) + \hbar\nu\,|g\rangle\langle e| \otimes \left(|1\rangle\langle 0| + \sqrt{2}\,|2\rangle\langle 1|\right)$$

Show that while the atom is inside the cavity, the following joint states are eigenstates of H_{AC} and determine the eigenvalues:

$$|E_0\rangle = |g\rangle \otimes |0\rangle$$
$$|E_{1+}\rangle = \frac{1}{\sqrt{2}}(|g\rangle \otimes |1\rangle + |e\rangle \otimes |0\rangle)$$
$$|E_{1-}\rangle = \frac{1}{\sqrt{2}}(|g\rangle \otimes |1\rangle - |e\rangle \otimes |0\rangle)$$
$$|E_{2+}\rangle = \frac{1}{\sqrt{2}}(|g\rangle \otimes |2\rangle + |e\rangle \otimes |1\rangle)$$
$$|E_{2-}\rangle = \frac{1}{\sqrt{2}}(|g\rangle \otimes |2\rangle - |e\rangle \otimes |1\rangle)$$

Then rewrite $|\Psi(0)\rangle$ as a superposition of energy eigenstates.

(b) Use part (a) to compute the expectation value

$$\langle P_e\rangle = \langle\Psi(t)|\,P_e \otimes I^C\,|\Psi(t)\rangle$$

as a function of atomic transit time t. You should find your answer is of the form

$$\langle P_e\rangle = \sum_n P(n)\sin^2[\Omega_n t]$$

where $P(n)$ is the probability of having n photons in the cavity and Ω_n is the $n-$photon Rabi frequency.

Chapter 10

Quantum Measurement

In this chapter, the solution to the so-called "quantum measurement" problem is fully developed within the standard structure of quantum mechanics (4 postulates). I have collected together and organized the thoughts of many other physicists (especially Hobson), filled in many details and derivations, added my own thought along the way, and generally produced a unified picture of the solution.

10.1 Basic Quantum Mechanics Reviewed

I am assuming that you have already studied the basics of quantum mechanics prior to this discussion.

We can then state the standard formulation of quantum mechanics based on the 4 postulates listed below(with some embellishments):

- •• All physical systems are represented by ket vectors $|\psi\rangle$ normalized to 1, i.e., $\langle\psi\,|\,\psi\rangle = 1$. The ket labels represent **everything that we know about the system**.

- •• Measurable properties of physical systems are represented by linear operators called **observables**.

  So restating part of the first postulate, the ket labels represent **the values of all observables of the system that have been measured**.

  If a vector associated with particular physical state $|\psi\rangle$ is an eigenvector, with eigenvalue α, of operator $\hat{A}$ associated with particular measurable property of system, i.e., if $\hat{A}|\psi\rangle = \alpha\,|\psi\rangle$, then the system in that state **definitely has** the value α of that measurable property.

  This implies that if one performs a measurement corresponding to the ob-

servable represented by $\hat{A}$ on a system in the state $|\psi\rangle$, then with **certainty** (probability =1) the measurement yields the value α for that measurable property.

Observables are represented by Hermitian operators (real eigenvalues). Since the eigenvectors of any Hermitian operator form a complete, orthonormal set, they can be used as a basis for the Hilbert space of the system.

Finally, if the system is in the state $|\psi\rangle$ and one measures an observable $\hat{B}$, where $|\psi\rangle$ is **not** an eigenvector of $\hat{B}$, the the only possible results of the measurement are one of the eigenvalues $\{b_k\}$ of $\hat{B}$.

•• Dynamics of state vectors

The state vectors of any system change with time via **deterministic** laws (similar to classical rules).

We define the **time evolution or development** operator that governs how a state vector changes in time by the relationship

$$|A, t + \Delta t\rangle = \hat{U}(\Delta t)\,|A, t\rangle \tag{10.1}$$

or the state vector at time $t + \Delta t$ is given by the time evolution operator(Unitary) $\hat{U}$ operating on the state vector at time t.

In general, the ket labels(which contain whatever we know(have measured) about state) are the only thing that changes.

The time evolution operator is a unitary operator since the state vector must remain normalized to 1, i.e, the vector length cannot change, which is guaranteed by the use of a unitary operator. The only changes to state vectors in quantum mechanics are changes in direction(phase).

The time evolution operator is related to the energy operator (this follows from time translation invariance)

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar} \tag{10.2}$$

•• Connection with Experiment/Measurements

We have specified above what happens when one measures a certain property of physical system at a moment when state vector of system is an eigenvector of the operator representing that property - probability equals 1 that we get the corresponding eigenvalue.

What if one measures a certain property of physical system at a moment

when state vector of system does not happen to be an eigenvector of the operator representing that property (which is most of the time) - what are the probabilities?

We need a new assumption.

Suppose the system is in the state $|\psi\rangle$, and one carries out a measurement of a property (observable) associated with the operator $\hat{B}$. We assume the eigenvectors of $\hat{B}$ are the vectors(states) $|b_i\rangle$, which means that $\hat{B}\,|b_i\rangle = b_i\,|b_i\rangle$, $i = 1, 2, , 3...$ where the b_i are the corresponding eigenvalues.

Quantum theory now assumes that the outcome of measurement is strictly matter of ***probability****.*

Quantum theory stipulates that the probability that the outcome of measurement of $\hat{B}$ on the state $|\psi\rangle$ (not an eigenvector) will yield the result b_i (remember the only possible results of measurement are the eigenvalues of $\hat{B}$ no matter what state the system is in), is equal to $|\langle b_i\,|\,\psi\rangle|^2$ (the Born rule). **The probability is given by the absolute square of of the corresponding component!**

The quantum mechanics formalism based on these postulates + embellishments correctly predicts experimental results for all known experiments.

Some ideas implied by these rules are:

These rules imply that one cannot say anything definite about the value of the observable represented by $\hat{B}$ when system is in a state $|\psi\rangle$, which is NOT an eigenvector of $\hat{B}$.

One can only make probability statements.

Before one measures the observable represented by $\hat{B}$ when the system is in a state $|\psi\rangle$, which is NOT an eigenvector of $\hat{B}$, the system does not have a value of that observable, according to quantum theory!

Our information about any state is only set of probabilities.

But all of your experience says that objects have values for measured quantities before they are measured, i.e., your experience tells you that the observable represented by $\hat{B}$ has a value even if we do not measure it.

That is your view (the standard classical view) about what is real and

what is not real.

Quantum theory implies you are wrong in both cases!!

10.1.1 Where is the "collapse" postulate?

Since the system has a definite value of the observable represented by $\hat{B}$ after the measurement, i.e., a pointer points to a value or a counter clicks or a mark is registered on a piece of paper (note that these are all **irreversible** occurrences which must be the end result of any measurement) and there is no mechanism to produce a single value in the rules as presented so far, most presentations add another rule at this point called **collapse of the state vector**.

It is usually proposed that the effect of a measurement is to irreversibly change(collapse) the state vector (which was not an eigenvector of $\hat{B}$) into an eigenvector of $\hat{B}$(corresponding to the eigenvalue jsut measured) so that it would be observed to have a definite (probability = 1) value for a subsequent measurement of the operator $\hat{B}$.

This extra rule says that state vector changes(discontinuously) during measurement from representing range of possibilities (superposition of all possible states) to definite state or only one possible outcome.

Which particular eigenvector it gets changed into is determined by outcome of measurement and cannot be known until then!. It cannot be predicted! It is at this point that randomness enters quantum mechanics.

I believe that this last rule is not needed and should not be added.

First of all, no real mechanism is ever given for "how" this process actually takes place, and second no specifications are given as to exactly "when" it occurs.

That suggests to me that it does not exist!

I will now proceed to develop a proposal for "definite outcomes" without using any "collapse" rule.

10.2 The Measurement Process

We consider a system consisting of a quantum system (Q-system) and a measurement system (M-system).

If the meter, which we assume is initially OFF (state $|0\rangle_M$) was turned ON when system was in $|+\rangle_Q$ state, then according to the above rules the combined

system evolves to

$$|+\rangle_Q |0\rangle_M \rightarrow |+\rangle_Q |+1\rangle_M \quad \text{i.e., meter (a good working meter) reads } +1 \quad (10.3)$$

Similarly, if the meter turned ON when the system is in the $|-\rangle_Q$ state, then combined system evolves to

$$|-\rangle_Q |0\rangle_M \rightarrow |-\rangle_Q |-1\rangle_M \quad \text{i.e., meter (a good working meter) reads -1} \quad (10.4)$$

This indicates that measurement, within framework of our rules, CORRELATES or ENTANGLES the dynamical variables (Q-system) being measured and the macroscopic (M-system) indicator of the meter, which we assume can be directly (macroscopically) observed.

Let us expand the discussion a bit. We have supposed above that the meter has eigenvectors (labelled by the corresponding eigenvalues)

$$|+1\rangle_M \rightarrow \text{meter on: reading } +1 \quad (10.5)$$

$$|-1\rangle_M \rightarrow \text{meter on: reading -1} \quad (10.6)$$

$$|0\rangle_M \rightarrow \text{meter off} \quad (10.7)$$

and the system has eigenvectors (labelled by eigenvalues)

$$|+\rangle_Q \rightarrow \text{value} = +1 \quad (10.8)$$

$$|-\rangle_Q \rightarrow \text{value} = \text{-1} \quad (10.9)$$

Now suppose that the initial state of the system is a superposition

$$|\psi\rangle = a|+\rangle_Q + b|-\rangle_Q \quad (10.10)$$

and thus the initial state of the **combined** system is given by

$$|initial\rangle = (a|+\rangle_Q + b|-\rangle_Q)|0\rangle_M \quad (10.11)$$

which represents the system in a superposition state **and** the meter OFF. We are interested in the evolution of this state according to QM. We note as we stated above that, if, instead of above initial state, we started with the initial states

$$|A\rangle = |+\rangle_Q |0\rangle_M \quad \text{OR} \quad |B\rangle = |-\rangle_Q |0\rangle_M \quad (10.12)$$

and then turn on the meter, these states must evolve as

$$|A\rangle = |+\rangle_Q |0\rangle_M \rightarrow |A'\rangle = |+\rangle_Q |+\rangle_M \quad (10.13)$$

$$|B\rangle = |-\rangle_Q |0\rangle_M \rightarrow |B'\rangle = |-\rangle_Q |-\rangle_M \quad (10.14)$$

respectively, indicating that the meter measured the appropriate value (the definition of good meter) since system is in eigenstate and has definite value

with certainty.

If system is in initial state corresponding to a superposition, however, then the **linearity** of quantum mechanics says it must evolve

$$|initial\rangle = (a\,|+\rangle_Q + b\,|-\rangle_Q)\,|0\rangle_M \rightarrow |final\rangle = a\,|+\rangle_Q\,|+\rangle_M + b\,|-\rangle_Q\,|-\rangle_M \quad (10.15)$$

We note the problem immediately, i.e., the meter has not ended up in a state with a definite value - it remains in a superposition of two macroscopically different pointer readings, which is never observed in the real world.

Hence, if as most physicists assume, the state vector represents a complete description of the Q-system, there seems to be a need for the "collapse" rule to fix the result and obtain "definite" values!

Since we will not be incorporating the "collapse" rule, we must proceed in a different way.

10.2.1 The Density Operator

The problem, as we will now see, lies with assuming that the state vector is the proper way to represent the Q-system during the measurement process.

The **expectation value** is the average of set of measurement results taken from a collection of systems in the same state. A straightforward calculation of the expectation value in a specific state takes the following form, with $\hat{O}$ being an operator representing the measurement of the specific physical variable and $|\phi\rangle$ is a state vector of some system in the collection:

$$\langle\hat{O}\rangle = \langle\phi|\,\hat{O}\,|\phi\rangle = \text{ expectation value of } \hat{O} \text{ in state } |\phi\rangle \quad (10.16)$$

If we choose **any** set of basis vectors $\{|i\rangle\}$, $i = 1, 2, ...$ for our vector space, we can expand $|\phi\rangle$ and $\langle\phi|$ as

$$|\phi\rangle = \sum_i a_i\,|i\rangle \quad , \quad \langle\phi| = \sum_j a_j^*\,\langle j| \quad (10.17)$$

where

$$a_i = \langle i\,|\,\phi\rangle \quad , \quad a_j^* = \langle\phi\,|\,j\rangle \quad (10.18)$$

Plugging these expansions into the expression for the expectation value:

$$\begin{aligned}\langle\hat{O}\rangle &= \sum_j a_j^*\,\langle j|\,\hat{O}\sum_i a_i\,|i\rangle = \sum_j\left(\sum_i a_j^* a_i\,\langle j|\,\hat{O}\,|i\rangle\right)\\ &= \sum_j\left(\sum_i \langle\phi\,|\,j\rangle\,\langle i\,|\,\phi\rangle\,\langle j|\,\hat{O}\,|i\rangle\right) = \sum_j\left(\sum_i\left[\langle i\,|\,\phi\rangle\,\langle\phi\,|\,j\rangle\,\langle j|\,\hat{O}\,|i\rangle\right]\right)\end{aligned} \quad (10.19)$$

Now we define a new operator $\hat{\rho} = |\phi\rangle\langle\phi|$, which is just the **projection operator** onto the state $|\phi\rangle$. We can then write the expectation value as

$$\langle\hat{O}\rangle = \sum_i \left(\sum_j \langle i|\hat{\rho}|j\rangle\langle j|\hat{O}|i\rangle \right) \tag{10.20}$$

Using the relation

$$\sum_j |j\rangle\langle j| = \hat{I}$$

one finds that the expectation value can be written in a very interesting form

$$\langle\hat{O}\rangle = \sum_i \langle i|\hat{\rho}\hat{O}|i\rangle = \mathrm{Tr}(\hat{\rho}\hat{O}) \tag{10.21}$$

i.e., the expectation value is given by the sum over the diagonal matrix elements of the operator product $\hat{\rho}\hat{O}$ (the symbol Tr = Trace is just shorthand for the diagonal sum).

The new operator $\hat{\rho}$ is called the **density operator**.

Why bother?

The real power of the density operator approach to QM comes when we have to deal with a situation in which we **cannot be sure what state system is in**(as in the measurement problem).

Imagine we have a whole collection of identical systems, some in $|\phi_1\rangle$, some in $|\phi_2\rangle$, etc. We might not know which system in which state, and might not even know how many systems are in any one state.

Example: Think about a beam of electrons that has not passed through any Stern-Gerlach(S-G) magnets. Chances are that the spin states of the electrons are completely random. Perhaps the best one can know is the probability of finding an electron in any state.

$$P_1 = \mathrm{Prob}(|\phi_1\rangle) \quad , \quad P_2 = \mathrm{Prob}(|\phi_2\rangle) \quad , \quad P_3 = \mathrm{Prob}(|\phi_3\rangle) \quad , \quad \ldots\ldots\ldots \tag{10.22}$$

These probabilities have nothing to do with quantum theory; they simply represent our ignorance of the details of what is happening. Thus, they are **not** related to any quantum amplitudes.

Given a situation like this, one should be able to do some useful calculations.

For example, one could work out expectation value of any measurement as follows.

If one can calculate the expectation value for each individual state, then the

overall expectation value is simply given by

$$\langle \hat{O} \rangle = P_1 \langle \phi_1 | \hat{O} | \phi_1 \rangle + P_2 \langle \phi_2 | \hat{O} | \phi_2 \rangle + P_3 \langle \phi_3 | \hat{O} | \phi_3 \rangle + \cdots + P_n \langle \phi_n | \hat{O} | \phi_n \rangle \quad (10.23)$$

Think back to original definition of expectation value which just represents the average value of measurement. What we have done here is put together the weighted average of the average value for each state, which is just the definition of the overall average value. Now if one constructs a density operator that is given by

$$\hat{\rho} = P_1 |\phi_1\rangle \langle \phi_1| + P_2 |\phi_2\rangle \langle \phi_2| + P_3 |\phi_3\rangle \langle \phi_3| + \cdots + P_n |\phi_n\rangle \langle \phi_n| \quad (10.24)$$

then the expectation value is still given by

$$\langle \hat{O} \rangle = \mathrm{Tr}(\hat{\rho}\hat{O}) \quad (10.25)$$

as expected.

Some notation: when the density operator takes the form $\hat{\rho} = |\phi\rangle \langle \phi|$ it is said to represent a **pure state** and when the density operator takes the form $P_1 |\phi_1\rangle \langle \phi_1| + P_2 |\phi_2\rangle \langle \phi_2| + P_3 |\phi_3\rangle \langle \phi_3| + \cdots + P_n |\phi_n\rangle$ it is said to represent a **mixed state**

10.2.2 A Crucial Example

Consider a box containing a very large number of electrons, each having spin $= 1/2$. This means that each electron spin can have measurable component $\pm 1/2$ along any direction. Now, suppose the box has a hole so that the electrons can get out and go into a Stern-Gerlach device oriented to measure the z-components of spin(arbitrary choice).

In order to proceed, we need to know how the box of electrons was prepared.

Let us consider two very different cases:

1. In the first case, we fill the box with electrons that have been prepared in a superposition state

$$|\psi\rangle_1 = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle + |\downarrow_z\rangle) = |\uparrow_x\rangle \quad (10.26)$$

   This preparation can be done by sending the electrons through an x-oriented magnet and choosing one of the resulting beams. Thus, in this case, each electron is in the indicated state - each electron is in a superposition of "up" and "down" in the z-direction. We then fill the box with these electrons.

2. In the second case, we send electrons through a z-oriented magnet and collect electron from both beams "z-up" and "z-down". In this case we

know that the electrons are either "z-up" OR "z-down", i.e., they each have a definite value - they are not in a superposition. We then fill the box with the two collections of electrons. Remember the electrons in the box in this case are EITHER in the state $|\uparrow_z\rangle$ OR in the state $|\downarrow_z\rangle$.

Now we proceed with the experiment. In case (1) we observe "z-up" 50% of the time and "z-down" 50% of the time and we know that in order to describe this system by a state vector we must say

$$|\psi\rangle_{BOX(1)} = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle + |\downarrow_z\rangle) \rightarrow \text{50-50 up-down} \tag{10.27}$$

i.e., every electron in the box is in this superposition state.

In case (2)) we observe "z-up" 50% of the time and "z-down" 50% of the time. But we now have a problem. If I did not know that each electron in the box in this case had a definite value, I would be tempted to describe the this system by the same state vector as in case (1). However, we know that is not true! The electrons in case (2) are not each in a superposition but they have definite values!

So, if I measure z-components I cannot tell whether I have case (1) or case (2) and I do not know how to write the state vector for the box in case (2),

But remember how I created the electron in case (1). They all have a definite value of the x-component, namely, "x-up'. So if I subject the electrons coming out of the box in case (1) to a x-measurement instead of a z-measurement, I will end up with only one beam!

However, in case (2), the electrons coming out of the box are either "z-up" OR "z-down" each of which is 50-50 in the x-direction and thus I would end up with two beams after the extra measurement!

The different measurement results mean that their states must be described differently in QM.

State vectors do not give us the freedom to do this unless we want to monkey around with relative phases, i.e., we would need to write

$$|\psi\rangle_{BOX(1)} = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle + |\downarrow_z\rangle) \quad , \quad |\psi\rangle_{BOX(2)} = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle + e^{i\alpha}\,|\downarrow_z\rangle) \tag{10.28}$$

where α is a completely unknown relative phase factor, which must be averaged over during any calculations since it is different (and random) for each separate measurement (each member of ensemble). I actually do not think such an object is a legitimate state vector!

If we use density matrices, we have a very different story. For pure state a

density operator (or matrix) is defined as

$$\hat{\rho} = |\psi\rangle\langle\psi| \tag{10.29}$$

for some state vector $|\psi\rangle$, i.e., it is the pure state projection operator.

In case (1) this gives

$$\hat{\rho} = \frac{1}{2}(|1/2\rangle\langle 1/2| + |-1/2\rangle\langle 1/2| + |1/2\rangle\langle -1/2| + |-1/2\rangle\langle -1/2|) \tag{10.30}$$

Derivation of the $\hat{\rho}$ matrix in the $(+1/2, -1/2)$ basis:

$$\rho = \begin{pmatrix} \langle 1/2|\hat{\rho}|1/2\rangle & \langle 1/2|\hat{\rho}|-1/2\rangle \\ \langle -1/2|\hat{\rho}|1/2\rangle & \langle -1/2|\hat{\rho}|-1/2\rangle \end{pmatrix} \tag{10.31}$$

Now

$$\begin{aligned} &\langle 1/2|\hat{\rho}|1/2\rangle \\ &\quad = \langle 1/2|\frac{1}{2}(|1/2\rangle\langle 1/2| + |-1/2\rangle\langle 1/2| + |1/2\rangle\langle -1/2| + |-1/2\rangle\langle -1/2|)|1/2\rangle \\ &\quad = \frac{1}{2} \end{aligned} \tag{10.32}$$

and so on, so that

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{10.33}$$

where the diagonal matrix elements represent probabilities and the non-zero off-diagonal matrix elements imply that one will observe quantum interference effects in this system. It is clear that any pure state density operator cannot be written as the sum of pure state projection operators.

In case (2), however, have

$$\hat{\rho} = \frac{1}{2}(|1/2\rangle\langle 1/2| + |-1/2\rangle\langle -1/2|) \tag{10.34}$$

and

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{10.35}$$

which clearly is sum of pure state projection operators. This corresponds to a mixed state. Note that off-diagonals are zero so that this density operator cannot lead to any quantum interference effects.

Remember, this system(case (2)) is such that electrons have values so that the density operator take this form (sum of projection operators)!!

Note that when electrons DO NOT HAVE VALUES (case (1)) the density operator has interference terms and cannot be written as a sum of projection operators - A DIFFERENCE that does not show up when using state vectors!!

We note that if we treat case (2) as pure state with the extra relative phase factor we would obtain:

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & e^{i\alpha} \\ e^{i\alpha} & 1 \end{pmatrix} \tag{10.36}$$

which becomes

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{(same result as before)} \tag{10.37}$$

when we average over α.

Now, let us digress to see what happens in a real classical system.

Consider rolling a die which has possible values = 1,2,3,4,5,6 where the probability of occurrence of each value = 1/6. In this case, the density operator representing the die will be for a mixed state (no interference effects)and the die **has** a value before/after each roll so that we have

$$\hat{\rho} = \frac{1}{6}(|1\rangle\langle 1| + |2\rangle\langle 2| + |3\rangle\langle 3| + |4\rangle\langle 4| + |5\rangle\langle 5| + |6\rangle\langle 6|) \quad \text{mixed state} \tag{10.38}$$

and the expectation value of the $\widehat{ROLL}$ operator is

$$\langle\widehat{ROLL}\rangle = \frac{1}{6}(1+2+3+4+5+6) = 3.5 \quad \text{standard definition} \tag{10.39}$$

More formally, we know that any operator can be written as sum of eigenvalues × projections operators(called the **spectral decomposition**), i.e., for the $\hat{B}$ operator introduced earlier we can write

$$\hat{B} = \sum_i b_i |b_i\rangle\langle b_i|$$

so that we have for the $\widehat{ROLL}$ operator

$$\widehat{ROLL} = 1|1\rangle\langle 1| + 2|2\rangle\langle 2| + 3|3\rangle\langle 3| + 4|4\rangle\langle 4| + 5|5\rangle\langle 5| + 6|6\rangle\langle 6| \tag{10.40}$$

The operator product of the density operator and the $\widehat{ROLL}$ operator can be written(using the orthonomality of the basis state vectors) as

$$\hat{\rho}\,\widehat{ROLL} = \frac{1}{6}(1|1\rangle\langle 1| + 2|2\rangle\langle 2| + 3|3\rangle\langle 3| + 4|4\rangle\langle 4| + 5|5\rangle\langle 5| + 6|6\rangle\langle 6|) \tag{10.41}$$

Thus, the expectation or average value is

$$\langle\widehat{ROLL}\rangle = \mathrm{Tr}(\hat{\rho}\widehat{ROLL}) = \sum_k \langle k|\,\hat{\rho}\widehat{ROLL}\,|k\rangle = \frac{1}{6}(1+2+3+4+5+6) = 3.5 \tag{10.42}$$

But, in this case, we know the values are real before the measurement, i.e., we are using a macroscopic die with numbers we can see!

Thus, this particular form(sum of pure state projection operators) of a density operator represents that case. The same was true earlier for the electrons in the box when we knew they were either "up" or "down"! In both cases the system had known values before measurements! **REMEMBER this fact for later.**

If we were to add the "collapse" rule it raises a host of questions: What exactly do we mean, physically and mathematically, by a "collapse during measurement" of quantum system?

Does collapse occur all at one instant?

Wouldn't instantaneous collapse contradict special relativity?

If collapse occurs during a time interval, then what equation describes its time-evolution during that interval?

Quantum states are presumed to follow the Schrödinger equation, which prescribes continuous time evolution; how can instantaneous state collapse be reconciled with smooth evolution?

How can we resolve "problem of outcomes" that appears to arise when a superposed quantum's state is measured by "which-state" detector, creating a so-called entangled state of the quantum and the detector that **appears** to be an indefinite superposition of two macroscopically distinct states of a composite system?

Such questions comprise the **quantum measurement problem**.

With all these thoughts/ideas in hand, we will now continue this discussion of the measurement problem and finally suggest a resolution of problem of definite outcomes that lies entirely within standard quantum physics.

A "quantum measurement" means any quantum process that results in a macroscopic effect, regardless of whether humans or laboratories are involved.

Thus not only is an electron striking a laboratory viewing screen and creating a visible flash a measurement, but also a cosmic-ray muon striking and macroscopically moving a sand grain on a planet in some other galaxy is a measurement.

To analyze a measurement, we look at a specific experiment: suppose an electron beam passes through a pair of double slits and then impacts a viewing screen. Just as in Thomas Young's similar double-slit experiment using light, performed in 1801, a pattern is formed on the viewing screen that seems to

shows interference between the two portions of the electron beam which are seemingly(classically) coming through the two slits: a broad dark-and-bright striped pattern spreads out widely on the screen - much wider than slits - indicating regions of destructive (dark) and constructive (bright) interference.

On closer inspection, the bright lines are formed by a very large number of tiny individual electron impacts, each one making a small flash on the screen.

According to above definition, each flash is a measurement of the position of an electron as it hits the screen.

Each electron's flash on screen is a measurement!

For the purposes of this analysis, however, it is better to consider a related example of measurement, still based on the double-slit experiment.

Suppose an electron detector is installed at the slits and assume that the detector can detect the electron's position as it passes through slits while disturbing each electron only minimally (in the precise sense described below). As it turns out, measurement, even by such a minimally-disturbing "which-path detector", changes everything.

Exactly when the detector turns on, the pattern on screen changes from a striped interference pattern to a smoothly-spread-out sum of two single-slit patterns, each showing diffraction but no interference.

The interference pattern abruptly vanishes.

An analogous experiment has been done using light (photons) instead of electrons, and using an interferometer rather than double-slit interference setup.

A which-path detector was randomly switched on or off as each photon passed through this experiment; photons for which the detector was "off" formed an interference pattern while photons for which the detector was "on" formed the expected no-interference pattern.

Reminder:

If the system is in a "pure" superposition state, then the density operator takes the form

$$\hat{\rho} = \frac{1}{2}(|1/2\rangle\langle 1/2| + |-1/2\rangle\langle 1/2| + |1/2\rangle\langle -1/2| + |-1/2\rangle\langle -1/2|) \qquad (10.43)$$

i.e., sum of projection operators and "cross-terms" (interference terms) or the matrix form

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{10.44}$$

where the diagonal matrix elements represent probabilities and the off-diagonal matrix elements imply that one will observe quantum interference effects in this system. Clearly, any pure state density operator cannot be written as the sum of pure state projection operators.

If in a "mixed state", then the density operator takes the form

$$\hat{\rho} = \frac{1}{2}(|1/2\rangle\langle 1/2| + |-1/2\rangle\langle -1/2|) \tag{10.45}$$

or the matrix form

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{10.46}$$

i.e., sum of projection operators (no interference terms).

As we found in the dice example:

$$\hat{\rho} = \frac{1}{6}(|1\rangle\langle 1| + |2\rangle\langle 2| + |3\rangle\langle 3| + |4\rangle\langle 4| + |5\rangle\langle 5| + |6\rangle\langle 6|) \quad \text{mixed state} \tag{10.47}$$

When we observe a quantum system evolve into such a "mixed state" density operator, then the quantum system can be interpreted "classically", i.e., a measurement has taken place and it has been irreversibly recorded.

This is the important point!! A full "collapse" is not necessary!! Before we go deeper to see how this all works, we take two important digression to introduce a topics that will become central to the discussion of quantum measurement - these subjects will be greatly expanded in next book revision.

10.3 The so-called Gambler's Ruin problem - a possible way to get to the irreversible recording

10.3.1 Mathematical Problem of the Points

A sequence of fair coins is flipped. Player A gets a point for every head and player B gets a point for every tail. Player A wins if there are a heads before b tails, otherwise B wins. Find the probability that A wins. Let $\alpha(a,b)$ be the probability that A wins and $\beta(a,b)$ the probability that B wins, which implies that

$$\alpha(a,b) + \beta(a,b) = 1$$

Solution (due to Pascal and Huygens -1500's)

$$\alpha(a,b) = \frac{1}{2^{a+b-1}} \sum_{k=0}^{b-1} \binom{a+b-1}{k} \text{ where } \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

An example is the **Gambler?s Ruin problem**.

Just as in the problem of points, suppose that at some stage A has a counters, and B has $m+n-a$ counters so that total number of counters = $m+n$, and let A's chances of victory at that point be $v(a)$.

The solution when A starts with m counters is given by

$$p_A = v(m) = \frac{1-(\beta/\alpha)^m}{1-(\beta/\alpha)^{m+n}}$$

This works for all cases except $\alpha = \beta = 1/2$ (the 50-50 probability case). For that special case, the solution is

$$v(a) = \frac{a}{m+n}$$

Thus, there is always a winner (hence the the name **ruin** for the loser) - even when $\alpha = \beta = 1/2$. Note that this relation is linear in a in this case!

Pearle's Theorem What does this have to do with quantum collapse? I now present an idea from Philip Pearle:

I soon found a charming analogy for collapse dynamics, useful for providing an intuitive and non-technical explanation of how it works. I happened to be browsing in Feller?s book on probability (a favorite textbook, from an undergraduate course taught by Stanislaus Ulam) when I encountered the gambler?s ruin game.

Two gamblers, initially possessing, respectively, a fraction $x_1(0)$, $x_2(0)$ of their combined wealth (so $x_1(0)+x_2(0) = 1$) repeatedly toss a fair coin, and the result, heads or tails, determines which one gives one dollar to the other.

They play until one gambler loses all his money, and the game ends.

The analogy is that the amount of money possessed by one gambler at any time is proportional to the squared amplitude of one of two states whose sum is the state vector representing the physical system undergoing collapse.

Just as one gambler loses all his money, so one of the states loses all its amplitude, and as the other gambler wins all the money, so the state vector ends up as totally described by the other state (?> collapse!!).

Let $Q(x)$ be the conditional probability that a gambler wins the game, given that he has the fraction x of the total wealth.If Δ is the fraction of the total wealth they exchange at each toss (i.e., Δ = 1 \$1/total dollars), the difference equation

$$Q(x) = \frac{1}{2}Q(x - \Delta) + \frac{1}{2}Q(x + \Delta)$$

expresses that there are two routes to win if one has fractional wealth x, namely lose the next toss and drop to $x - \Delta$ but win thereafter, or win the next toss and rise to $x + \Delta$ and win thereafter.

The solution of the difference equation is $Q(x) = Ax + B$, where A and B are constants. Now we have boundary conditions:

Since $Q(0) = 0$ (because you can't win if you have no money) and $Q(1) = 1$ (because you have won if you have all the money), then $Q(x) = x$, i.e., $Q(1) = 1 = A + B$, $Q(0) = B = 0 \rightarrow A = 1, B = 0 \rightarrow Q(x) = x$

That is, if one starts with the fraction $x = x(0)$ of the money, one has the probability x(0) of attaining all the money ?> you get to $x = 1$, which is just collapse behavior.

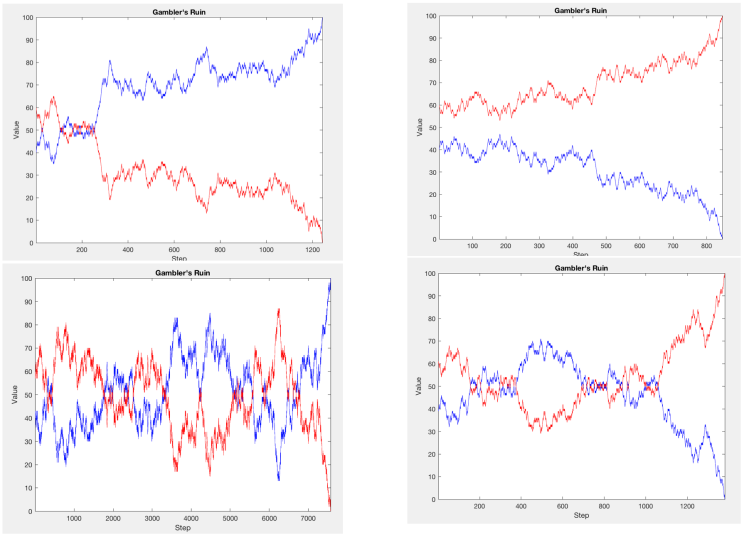
The game can be modified to have many players, to have Δ change as the game progresses (e.g., to get smaller as one gambler gets closer to losing), etc.

So, one may think of quantum collapse as a gambler's ruin competition among the states in a superposition, to see which final state wins the game (remains at the end).

I ran a simulation of the Gambler?s Ruin problem using OCTAVE on a computer.OCTAVE code - gambruin.m:

```
x1=[];
y1=[];
x=40;x1=[x1,x];
y=100-x;y1=[y1,y];
count=1;
while (x ~= 0) && (x ~= 100)
r=rand;
x=x+(r <= 0.5)-(r > 0.5);x1=[x1,x];
y=100-x;y1=[y1,y];
count=count+1;
endwhile
figure
plot(1:count,x1,'-b');
axis([1,count,0,100]);
xlabel('Step');
ylabel('Value');
title('Gambler''s Ruin')
hold on;
plot(1:count,y1,'-r');
hold off
count
```

Some sample runs are shown below - note there is always a winner(loser)



One state always wins out!! Collapse occurs even though there is no direct collapse mechanism!! An irreversible mark always appears!! Could this be the way collapse works without a postulate or mechanism?

10.4 Another short digression - another way to get to the irreversible recording - Decoherence

If we have a quantum system is the state

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

then this states exhibits quantum interference effects.

Decoherence says that as a state interacts with macroworld in its environment (which it has to do) off diagonal elements go to zero, i.e., the state makes transition to

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

As we already said, this state has zero quantum interference effects and **a definite measurement value has been irreversibly recorded somewhere**.

If we had quantum dice the state vector would be

$$|Dice\rangle = \frac{1}{\sqrt{6}}(|1\rangle + |2\rangle + |3\rangle + |4\rangle + |5\rangle + |6\rangle) \text{ and } \rho = |Dice\rangle\langle Dice|$$

giving

$$\begin{aligned}\rho = \frac{1}{6}\,[&|1\rangle\langle 1| + |2\rangle\langle 1| + |3\rangle\langle 1| + |4\rangle\langle 1| + |5\rangle\langle 1| + |6\rangle\langle 1| \\ &+ |1\rangle\langle 1| + |2\rangle\langle 1| + |3\rangle\langle 1| + |4\rangle\langle 1| + |5\rangle\langle 1| + |6\rangle\langle 1| \\ &+ |1\rangle\langle 2| + |2\rangle\langle 2| + |3\rangle\langle 2| + |4\rangle\langle 2| + |5\rangle\langle 2| + |6\rangle\langle 2| \\ &+ |1\rangle\langle 3| + |2\rangle\langle 3| + |3\rangle\langle 3| + |4\rangle\langle 3| + |5\rangle\langle 3| + |6\rangle\langle 3| \\ &+ |1\rangle\langle 4| + |2\rangle\langle 4| + |3\rangle\langle 4| + |4\rangle\langle 4| + |5\rangle\langle 4| + |6\rangle\langle 4| \\ &+ |1\rangle\langle 5| + |2\rangle\langle 5| + |3\rangle\langle 5| + |4\rangle\langle 5| + |5\rangle\langle 5| + |6\rangle\langle 5| \\ &+ |1\rangle\langle 6| + |2\rangle\langle 6| + |3\rangle\langle 6| + |4\rangle\langle 6| + |5\rangle\langle 6| + |6\rangle\langle 6|]\end{aligned}$$

Thus, the quantum dice generally has a lot of quantum interference terms. This corresponds to a density matrix of the form

$$\rho = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{pmatrix}$$

It has lots of off-diagonal terms. Rolling the dice (a measurement) or decoherence if dice left alone, produces

$$\rho = \begin{pmatrix} a_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{66} \end{pmatrix}$$

where

$$a_{11} = a_{22} = a_{33} = a_{44} = a_{55} = a_{66} = \frac{1}{6}$$

or

$$\rho = \frac{1}{6}(|1\rangle\langle 1| + |2\rangle\langle 2| + |3\rangle\langle 3| + |4\rangle\langle 4| + |5\rangle\langle 5| + |6\rangle\langle 6|)$$

which is a mixed state implying as we said earlier that a measurement has occurred!! Maybe that is what Gambler?s Ruin is doing!

The next revision of this text will have lots more to say about Gambler's Ruin and Decoherence in relation to the measurement question.

Now let us continue on earlier path.

10.5 Look closely at "which-path" experiments

In the so-called "delayed-choice" two-slit experiment mentioned above, collapses were instantaneous to within the accuracy of fast switching between the two states; in addition, each collapse is executed entirely while the photon was inside the interferometer.

We can gain considerable insight by studying how quantum theory describes a which-path measurement. Note: it is a measurement as defined earlier, because the detector registers "slit 1" or "slit 2" macroscopically for each electron.

We denote the state of one electron passing through slit 1 as $|\psi_1\rangle$ and the state of one electron passing through slit 2 as $|\psi_2\rangle$. John von Neumann, who was the first to carefully analyze measurement in purely quantum-theoretical terms, insisted on treating not only the measured quantum but also the macroscopic detector as quantum systems because, after all, detectors are made of atoms and they perform a quantum function by detecting individual quanta. I agree with that view of the macroscopic detector.

Now let us repeat the earlier discussion again, filling in any remaining gaps and confusions.

Accordingly, one represents the "ready to detect" quantum state of detector by $|ready\rangle$, and the state of detector after detecting an electron by $|1\rangle$ if $|\psi_1\rangle$ was detected, and by $|2\rangle$ if $|\psi_2\rangle$ was detected. A properly operating detector will surely transition from $|ready\rangle$ to $|1\rangle$ upon measurement of an electron that has been prepared (perhaps by simply shutting slit 2) in state $|\psi_1\rangle$.

As a limiting idealization, we assume, with von Neumann, that the measurement of an electron prepared in state $|\psi_1\rangle$ leaves the electron still in state $|\psi_1\rangle$ after detection. Such a minimally-disturbing measurement would cause the electron-plus-detector composite system, initially in composite state $|\psi_1\rangle\,|ready\rangle$, to transition into the final state $|\psi_1\rangle\,|1\rangle$.

We summarize the process as

$$|\psi_1\rangle\,|ready\rangle \rightarrow |\psi_1\rangle\,|1\rangle \tag{10.48}$$

Similarly, a minimally-disturbing measurement of and electron initially prepared in $|\psi_2\rangle$ is described mathematically by

$$|\psi_2\rangle\,|ready\rangle \rightarrow |\psi_2\rangle\,|2\rangle \tag{10.49}$$

Now suppose that both slits are open so each electron can pass through either slit, and suppose the preparation and experiment (e.g. slit widths) is symmetric with respect to two slits. Then the state of each electron as it approaches the

slits prior to detection must be described by a symmetric superposition

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\psi_1\rangle + |\psi_2\rangle) \tag{10.50}$$

But quantum physics, including its time dependence, is **linear**. This imply that $|\psi\rangle\,|ready\rangle$ evolves according to

$$\begin{aligned} |\psi\rangle\,|ready\rangle &= \frac{1}{\sqrt{2}}(|\psi_1\rangle + |\psi_2\rangle)\,|ready\rangle \\ &= \frac{1}{\sqrt{2}}(|\psi_1\rangle\,|ready\rangle + |\psi_2\rangle\,|ready\rangle) \\ &\to \frac{1}{\sqrt{2}}(|\psi_1\rangle\,|1\rangle + |\psi_2\rangle\,|2\rangle) \end{aligned} \tag{10.51}$$

The final state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\psi_1\rangle\,|1\rangle + |\psi_2\rangle\,|2\rangle) \tag{10.52}$$

following detection said to be "entangled" because it cannot be factored into simple product of states of two sub-systems. As indicated in Figure 1

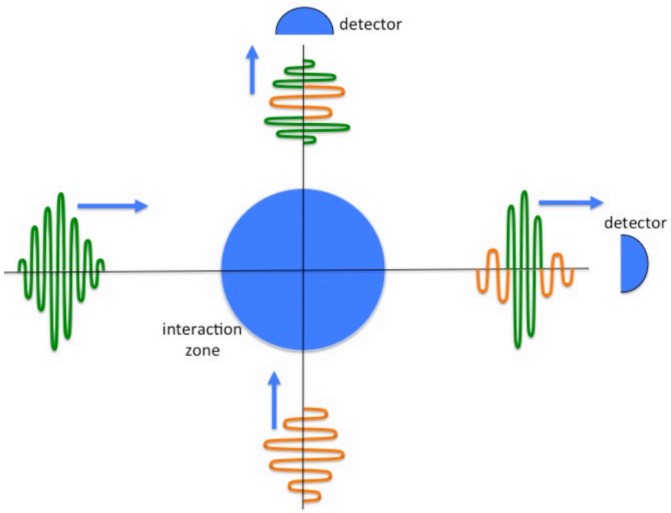


Figure 10.1:

when two independent quanta pass near each other, interact, and subsequently separate, the interaction generally entangles the two quanta and entanglement then persists after interaction regardless of how far apart the two quanta might eventually travel, provided only that the two quanta experience no further interactions.

Despite possibly wide spatial separation, entangled quanta have a

unity not possessed by non-entangled quanta.

This unity is source of quantum non-locality.

Entanglement is ubiquitous in nature.

The entangled "measurement state" (52) that is at the heart of quantum measurement is remarkably subtle.

To fully understand "entanglement", we first need to understand "superposition".

The quantum principles says that **any** linear combination of possible quantum states of a system, as in (50) and (52) for example, is also a possible quantum state of that system. Figure 2 pictures an experiment that demonstrates such a superposition of states.

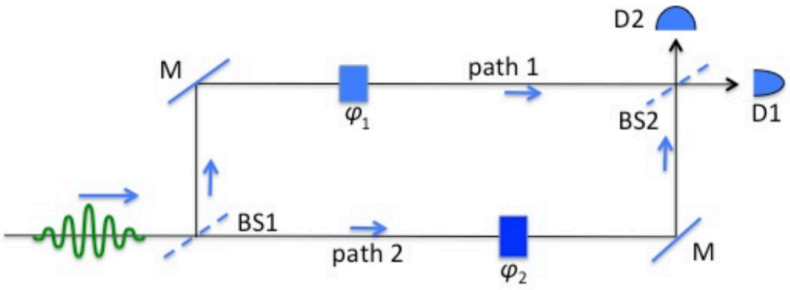


Figure 10.2:

This represents a layout of optical paths called a "Mach-Zehnder interferometer.".

A light beam enters at the lower left passing through the "beam splitter" BS1; the reflected beam makes a right angle with the incoming direction, while the transmitted beam passes straight through. So the beam splits and each "half" traverses one of two paths; mirrors M bring the paths back to crossing point as shown.

Devices called "phase shifters", denoted by φ_1 and φ_2, are placed into each path. The phase shifter can add a short variable length to a path. A second beam splitter BS2 can be placed at crossing point. Without BS2, each "half"-beam moves straight ahead along one path to detector on that path.

Things get more interesting with BS2 in place.

Because 50% of each of two beams then goes to each detector, BS2 mixes two

beams together so one can show interference.

The interferometer is constructed so that, when the phase shifters are set to zero, the two "optical paths" (number of wavelengths, after accounting for phase changes upon reflection and refraction) from entry point to D1 are equal while two optical paths to D2 differ by half a wavelength.

It is thus found that the light interferes constructively at D1 and destructively at D2, i.e., all light goes to D1. If one then uses φ_1 or φ_2 to add a half wavelength to either path, the light then interferes constructively at D2 and destructively at D1, i.e., all light goes to D2.

As one continuously varies the length of one or other path by varying one or other phase shifter, one finds that the amount of light arriving at D1 varies continuously from 100% down to 0%, while amount arriving at D2 varies from 0% to 100%.

The two paths are clearly interfering!

The experiment is an interferometer-based analog of Young's double-slit interference experiment demonstrating the wave nature of light.

But, as we know, light is really just photons, and photons are indivisible.

How does nature explain this experiment when we dim the light to point where only one photon at a time traverses the interferometer?

After all, the photon still traverses BS1, yet it cannot split in two because a quantum is unified and cannot be split!

With BS2 removed, one finds either D1 or D2 registers a single entire photon, randomly, with 50-50 probabilities, regardless of how the phase shifters are set.

The randomness is absolute. It is more random than any human macroscopic game, such as coin flips, that only mimics randomness. Nature invents quantum randomness to deal with obstacles such as beam splitters while preserving unity of quantum.

The detectors never register half of a photon. They get either a whole photon or no photon.

What happens in the single-photon experiment with BS2 present?

As discussed earlier, beginning from equal path lengths, which gives constructive interference at D1 and destructive interference at D2, as phase shifters vary, probabilities of detecting photon at D1 and D2 vary as Figure 3 which gives the

percentage of photons impacting D1.

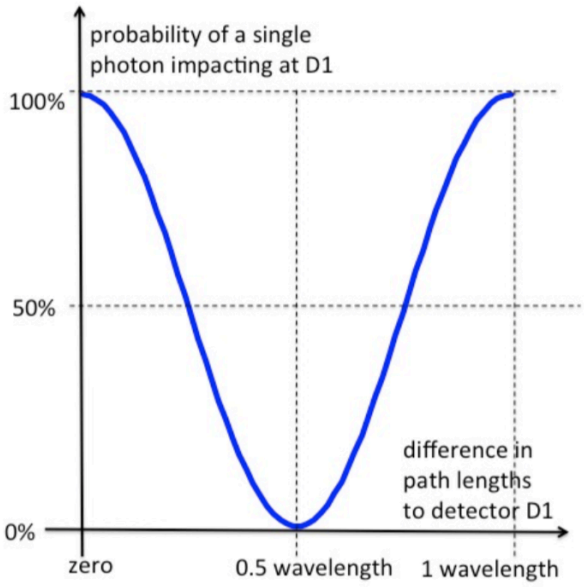


Figure 10.3:

Importantly, the results do not depend on which phase shifter the experimenter chooses to vary. Since each photon responds to changes in either path length, each photon must "follow both paths"! This verifies the superposition principle and shows that quanta can be in "two places at same time"!

This seems paradoxical if we assume photons are tiny particles, but if we assume the photons are waves it is not paradoxical, i.e., each photon simply spreads along both paths, interfering with itself at D1 and D2. One must conclude that each photon travels both paths even when BS2 is not present because once a photon enters the interferometer it must behave in same manner regardless of whether BS2 is placed or not placed at the far end.

A delayed-choice experiment provides further evidence for conclusion:

Since photons "do not know" whether BS2 will be inserted, they must travel both paths on all trials including those for which BS2 not inserted - this is connected with entanglement.

With BS2 removed, the situation is like the double-slit experiment with a which-slit detector present:

Each photon is entangled with the macroscopic detectors D1 or D2 as in (52).

With BS2 present, the two paths mix and we have a situation like double-slit experiment with no which-slit detector:

Each photon follows two paths to each detector where it interferes with itself, and one detects the interference state (50).

All of this suggests that measurements affect superposed quantum states via entanglement of the superposed quantum with a detector.

10.6 Resolving Paradoxes and Understanding Measurement

10.6.1 The apparent paradox of Schrödinger's cat

A cat is penned up in a steel chamber, along with the following device:

In a Geiger counter there is tiny bit of radioactive substance, so small, that perhaps in course of an hour one of its atoms decays, but also, with equal probability, perhaps none decay; if it happens, the counter tube discharges and through a relay releases a hammer which shatters a small flask of hydrocyanic acid(poisonous).

If one left the entire system to itself for hour, we would say that the cat still lives if no atom has decayed!

If one uses a state vector approach to understanding the entire system, one would express this by having a living and a dead cat mixed or smeared out in equal parts - a superposition of "dead" and "alive"! It is typical of these cases, that the indeterminacy originally restricted to the atomic domain becomes transformed into macroscopic indeterminacy, which can only then be resolved by direct observation, which prevents us from naively accepting as valid the "blurred model" for representing reality.

Mathematically, the nucleus and the cat have become entangled in measurement state (52), with $|\psi_1\rangle$ and $|\psi_2\rangle$ representing undecayed and decayed states of nucleus and $|1\rangle$ and $|2\rangle$ representing the alive and dead cat.

According to Schrödinger's understanding of the situation, the indeterminacy of the nuclear state "becomes transformed into macroscopic indeterminacy" of cat, and since he could not comfortably accept this as a "blurred" state", i.e., a cat that is in a superposition of being both alive and dead, he hoped this would say something is wrong with QM.

As we will show, according to standard quantum physics, Schrödinger's 1937 understanding was incorrect:

The composite system (cat-plus-nucleus) is not predicted to be in superposition of two states of a cat and a nucleus.

Instead, the composite system is predicted to be in a superposition of two **correlations** between cat and nucleus, one in which a live cat is 100% correlated with an undecayed nucleus, and a second in which a dead cat is 100% correlated with a decayed nucleus.

Entanglement will have transformed a pure state superposition of nuclear states to pure state superposition of correlations between subsystem states.

We will see that this is precisely what one expects from quantum mechanics, and it is not paradoxical!!

This so-called "problem of definite outcomes" applies of course to more than Schrödinger's dramatized example.

Regardless of whether the measuring instrument is a which-slit detector, a Geiger counter, or a cat, the entangled state (52) applies. This state appears at first glance to represent a quantum superposition in which the detector is in two macroscopically different states simultaneously. If so, then there is an inconsistency within quantum physics, because obviously it cannot be this easy to create a macroscopic superposition.

The question is: Is it true that (52) really represents a macroscopic superposition?

Turns out, there is more to this entangled state than meets the eye.

If one assumes the detector to be in superposed state $a\,|1\rangle + b\,|2\rangle$, one finds that (52) necessitates either $a = 0$ or $b = 0$, implying that detector is not in an individually superposed state within its own Hilbert space.

The same applies to the detected quantum:

It is not in a superposed state $a\,|\psi_1\rangle + b\,|\psi_2\rangle$ with both $a \neq 0$ and $b \neq 0$.

The entanglement process leaves neither sub-system superposed!

So far as I know, this simple fact has long been ignored by physicists studying the measurement problem.

10.6.2 Density Operator to the Rescue

The density operator formalism(discussed earlier) for quantum physics provides a stronger version of this conclusion.

The density operator for a quantum system whose state is $|\psi\rangle$ (a pure state) is simply the projection operator

$$\hat{\rho} = |\psi\rangle\langle\psi| \tag{10.53}$$

As we saw earlier, if the system is in state whose density operator is $\hat{\rho}$, then standard quantum expectation value $\langle\hat{O}\rangle$ of arbitrary observable $\hat{O}$ is found from

$$\langle\hat{O}\rangle = \mathrm{Tr}(\hat{\rho}\hat{O}) \tag{10.54}$$

This approach is especially useful if the quantum system is a composite of two subsystems A and B.

We define the density operator $\hat{\rho}_A$ for subsystem A alone by by

$$\hat{\rho}_A = \mathrm{Tr}_B(\hat{\rho}) \tag{10.55}$$

where Tr_B means that trace taken only over states of subsystem B.

It is then easy to show(see later) that standard quantum expectation values for subsystem A alone (values obtained by an observer of A without any knowledge of B) are

$$\langle\hat{O}_A\rangle = \mathrm{Tr}(\hat{\rho}_A\hat{O}_A) \tag{10.56}$$

where $\hat{O}_A$ means any observable operating on system A alone (i.e., operating within A's Hilbert space). Applying this formulation to measurement state (52), the reduced density operators for the quantum system (call it A) and its detector (call it B), respectively, are

$$\hat{\rho}_A = \frac{1}{2}(|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|) \tag{10.57}$$

$$\hat{\rho}_B = \frac{1}{2}(|1\rangle\langle 1| + |2\rangle\langle 2|) \tag{10.58}$$

The plus signs in (57) and (58) make one think of superpositions such as (50), but these are not superpositions. The density operator for the superposition (50) has cross-terms:

$$\hat{\rho} = (|\psi\rangle\langle\psi|) = \frac{1}{2}(|\psi_1\rangle\langle\psi_1| + |\psi_1\rangle\langle\psi_2| + |\psi_2\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|) \tag{10.59}$$

The two cross-terms, involving both $|\psi_1\rangle$ and $|\psi_2\rangle$, are missing from (57). So (57) does not describe system in superposition of two quantum states.

However, (57) is precisely the density operator one should use if one **knows** the

quantum system is either in state $|\psi_1\rangle$ or in state $|\psi_2\rangle$ but one didn't know which and so, due your own to lack of information, you simply assign probability of $1/2$ to each of two possibilities.

The same goes for (58).

(57) and (58) are "classical" probabilistic states - analogous to the "states of knowledge" one would assign to coin flip when you know the outcome to be either heads or tails with equal probability but don't know which has occurred. Remember the dice example!

The situation described by a density operator such as (57) is known as "mixture" of states $|\psi_1\rangle$ and $|\psi_2\rangle$, as distinct from "superposition" of states as observed in the Mach-Zehnder experiment and represented by (50).

Equation (56) tells us that all correct statistics for subsystem A alone can be found from standard formula (54) applied to subsystem A alone.

But we have just seen that (57) is the density operator one should use if one knows A to be in either $|\psi_1\rangle$ or $|\psi_2\rangle$ without knowing which.

The same goes for subsystem B and (58).

In the case of Schrödinger's cat, it follows that the observer of the cat alone sees outcomes appropriate to a cat that is either alive or dead, **not** both.

For subsystems, the interference terms are missing, and an "ensemble" of repeated trials must exhibit a nonsuperposed mixture rather than superposition.

This is the clear prediction of quantum physics for the entangled state (52).

But one must be careful, because (57) and (58) are not complete descriptions of quantum states of the nucleus or the cat.

In fact, (57) and (58) are not quantum states at all but merely "reduced states" arising from actual state (52) of composite system when one part of the composite system is removed from the equations.

In case of Schrödinger's cat, (57) and (58) give the correct predictions for observations of either the nucleus alone or the cat alone, but do not represent the state of either subsystem because this is given by (52).

In fact, when two quanta entangled, neither one has a quantum state of its own!

Physicists, philosophers and mathematicians who specialize in quantum foundations have in past objected to the argument that reduced density operators

can be adduced in this manner to clarify the measurement problem.

They offer two key objections:

First, "basis ambiguity", charges that 'the 'basis set" (set of orthogonal eigenvectors) for operator (58) (for example) is entirely ambiguous, so (58) cannot represent the true quantum state.

It is true that (58) doesn't represent the true state of the subsystem, because (58) is actually just identity operator $|1\rangle\langle 1| + |2\rangle\langle 2|$ in B's subspace, divided by 2, so that any other orthogonal basis set could be used instead. Given only the description (58), subsystem B could just as well be described by any other pair of orthonormal vectors in B's subspace, for example

$$\frac{1}{\sqrt{2}}(|1\rangle \pm |2\rangle) \tag{10.60}$$

But B's state of affairs is certainly not entirely described by (58). Rather, it is described by composite state (52).

Equation (58) merely tells us the following:

If the cat and the nucleus are in state (52) then, when one looks at the cat, one is going to see cat that is either alive or dead. There is no claim that (58) represents complete quantum state of the cat.

That is, there is no claim that cat is really in either state $|1\rangle$ or state $|2\rangle$, because the state it's really in is admittedly (52).

Thus, the basis ambiguity objection to our conclusion (namely, that Schrödinger's cat is either alive or dead, not both) fails.

The second key objection is that (57) and (58) are "improper density operators" because they arise not from insufficient knowledge (as classical probabilities arise) but from reductions of the full density operator for state (52) to Hilbert subspaces of each subsystem. It's true that these reduced density operators do not arise from insufficient knowledge about an actual state.

In fact, we do have complete knowledge of the state of both A and B, namely, the measurement state (52). So this objection fails not because it is false but because it is irrelevant:

Reduced operators admittedly do not represent the state of the composite system. They tell us only what we will observe at the nucleus and at the cat and they tell us nothing about the correlations between these observations, so these density operators do not tell us the real state of system.

And so plot thickens.

The entangled state (52) properly describes both individual subsystems.

However, the plus sign in (52) signifies superposition of two terms. We know, however, that neither subsystem A nor subsystem B is superposed.

What then is meaning of plus sign?

This superposition arose from the superposition represented by (50).

We cannot logically ignore this fact - a strategy known as the "shut up and calculate" approach to quantum measurement.

Instead, we must ask: Exactly what is superposed when the two subsystems are in this entangled state?

Superpositions preserve the all-important **unity of quantum**.

When Max Planck proposed in 1900 that electromagnetic radiation occurs in energy steps of magnitude $E = h\nu$, he tacitly implied the central quantum principle:

> **The unity of an individual quantum.**

Energy (electromagnetic energy in case of radiation) comes in spatially extended bundles, each having a definite and identical quantity of energy. One cannot have half a quantum, or 2.7 quanta!

You must have either 0 or 1 or 2 etc., quanta. In its own way, this is a fairly natural notion - apparently nature prefers to sub-divide the universe into a countable or even a finite set of entities as opposed to an uncountable continuum!

The spatial extension of these bundles then implies nonlocality:

If we have one quantum and destroy it (by transforming it to something else), we must destroy all of it everywhere simultaneously, because we cannot, at any time, have just part of quantum in existence.

Louis de Broglie put it perfectly in 1924, regarding another kind of quantum namely the electron:

> **The energy of an electron is spread over all space with a strong concentration in a very small region....That which makes an electron an atom of energy is not the small volume that it occupies in space - I repeat it occupies all space - but the fact that it is indivisible, that it constitutes a unit.**

When one transforms the state of a quantum, one must transform the entire extended quantum all at once. Hence there are quantum jumps. Furthermore, composite entangled systems such as atoms also behave in a unified fashion.

This unity is the source of the nonlocality seen in experiments involving entangled pairs of photons.

Nonlocality is exactly what one would expect, given the unity and spatial extension of the quantum and the unitary (i.e. unity-preserving) nature of the entanglement process.

Standard nonrelativistic quantum theory prescribes two kinds of time evolution: collapse upon measurement, and Schrödinger equation between measurements.

The key feature of Schrödinger equation is that it prescribes a so-called "unitary" time evolution, meaning time evolution that preserves pure states, i.e., transforms unit Hilbert space vectors into other unit vectors.

Some ideas required physically by unity of quantum are expressed as follows.

If quantum is described by a pure quantum state at $t = 0$, it should remain pure at later times. This notion prompts us to ask whether measurement process also preserves pure states.

At least in case of the idealized process described in (51), the answer is “yes” because both “before” and “after” states are pure.

Measurement state (52), since it is pure, represents a highly unified state of affairs, even though one of its subsystems is a macroscopic detector. Thus one suspects that this state, like its progenitor (50), is truly a superposition in which the superposed terms represent two situations or states of the same object.

But precisely what is that object, i.e., what is superposed?

We have seen that the states of subsystem A are not superposed, nor are states of subsystem B.

The conventional interpretation (which, as we will see, is subtly incorrect) of a product state such as $|\psi_1\rangle|1\rangle$ is that it represents a state of a composite system AB in which subsystem A is in state $|\psi_1\rangle$ while B is in state $|1\rangle$.

In this case, (52) would represent a superposition in which AB is simultaneously in state $|\psi_1\rangle|1\rangle$ and also in state $|\psi_2\rangle|2\rangle$.

The situation of Schrödinger’s cat would be: live cat and undecayed nucleus superposed with dead cat and decayed nucleus.

This is at least as physically outrageous as a live cat superposed with a dead cat, and it contradicts the physical implications (a cat that is either alive or dead) of reduced states (57) and (58) as described earlier.

Something is still wrong!

Let us repeat some of our discussion and provide even more details.

10.7 Some Repetition and More Intricate Details

Proposal:

The solution to the so-called quantum measurement problem is completely contained within standard quantum mechanics and needs no elaborate new structures and interpretations.

10.7.1 Remember the standard discussion from earlier:

Total system T = quantum system S (states $|s_i\rangle$) + measuring device A (states $|a_i\rangle$) , where $|a_0\rangle$ = state of measuring device "OFF". The unitary time evolution rule says that $|s_1\rangle|a_0\rangle \to |s_1\rangle|a_1\rangle$ and $|s_2\rangle|a_0\rangle \to |s_2\rangle|a_2\rangle$. The linearity of QM the says that

$$|\psi\rangle_{SA} = (c_1|s_1\rangle + c_2|s_2\rangle)|a_0\rangle = c_1|s_1\rangle|a_0\rangle + c_2|s_2\rangle|a_0\rangle \to c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$$

In this derivation

$$|\psi\rangle_S = (c_1|s_1\rangle + c_2|s_2\rangle$$

is state of the quantum system(superposition), while

$$|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$$

is state of the combined system(superposition).

The quantum measurement problem is the observation that the state

$$|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$$

is not observed as the outcome of the measurement!

What is seen is not a so-called superposition, but either

$$|s_1\rangle|a_1\rangle \quad \text{or} \quad |s_2\rangle|a_2\rangle$$

That is "problem of outcomes".

The state

$$|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$$

is usually referred to as a superposition, which is misleading.

Entanglement outweighs superposition as the defining feature of this state. Without entanglement-correlations, we would not have a measurement problem!

Careful investigation of this state in a 2007 experiment(Roch) (Wheeler delayed choice experiment) demonstrates its strikingly non-local character:

A photon jumps from state $|\psi\rangle_S = (c_1|s_1\rangle + c_2|s_2\rangle$ to state $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ precisely when A(the measuring device) switches on and while photon is still inside the interferometer, and jumps from $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ to $|\psi\rangle_S = (c_1|s_1\rangle + c_2|s_2\rangle$ when A switches off.

Quantum jumps, removal of interferences, and the observed non-locality are due to entanglement.

Of course, an entangled state = superposition, but it is a very special superposition.

To call $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ simply a "superposition" misses the crucial physics of entanglement and makes all the difference in the understanding of the state.

Entanglement is a characteristic trait of quantum mechanics; one that enforces the entire departure from classical lines of thought.

It is well known that, for 2-part systems, all non-product states exhibit non-locality (Bell inequality).

Thus when S and A are entangled $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$, they share a non-local channel.

The measurement state $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ is subtle. Although $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ is called a superposition of S and/or superposition of A, neither is true.

When S and A are in the measurement state $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ neither S nor A in superposition. $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ = superposition, but neither a superposition of S nor of A and also not superposition of states of the composite system SA.

In $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$, S is in both states $|s_1\rangle$ and $|s_2\rangle$ simultaneously, as we know from the observed interference between 2 states.

However, in case of $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ experiment shows that SA is

not in both two-part states $|s_1\rangle|a_1\rangle$ and $|s_2\rangle|a_2\rangle$ simultaneously, **but only in two correlations simultaneously.**

The entanglement of two systems is quite different from superposition of one system.

Experiments(Rarity et al) have demonstrated the precise sense in which $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ represents superposition.

Experiments answer the question "given that S and A are in state $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$, what (if any) entities interfere and what is nature of interference?"

10.7.2 The Local State Solution of the Problem of Definite Outcomes

Consider a single quantum S(electron or photon), passing through double-slit experiment, with a "downstream" viewing screen.

Suppose an ideal "which-slit detector" A is present so that, upon detection, S and A become entangled in measurement state with orthogonal detection states $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$. Imagine S and A separated by meters or kilometers.

Complete observation of the experiment requires two "local observers".The 1st observer is S and the 2nd observer is A.

Such a non-local setup has been carried out experimentally.

What did the local observers observe?

A well-known prediction of quantum physics says that first observer observes implications of "local state of S", represented by "reduced" density operator, where the degrees of freedom 2nd system are averaged over by the "trace" (Tr) operation (see derivation below)

$$\rho_S = \mathrm{Tr}_A(\rho_{SA}) = |c_1|^2 |s_1\rangle\langle s_1| + |c_2|^2 |s_2\rangle\langle s_2|$$

and 2nd observer would observe implications of "local state of A", represented by the reduced density operator

$$\rho_A = \mathrm{Tr}_S(\rho_{SA}) = |c_1|^2 |a_1\rangle\langle a_1| + |c_2|^2 |a_2\rangle\langle a_2|$$

where the density operator gives probabilities via the relation

$$P(q) = \mathrm{Tr}(\rho P_q) = \mathrm{Tr}(\rho|q\rangle\langle q|)$$

The "local state of S" is found by completely removing from the density operator any effects of A and vice versa.

This is the important idea!

Derivation of reduced density operator

$$\rho_{SA} = |\psi_{SA}\rangle\langle\psi_{SA}| = (c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle)(c_1^*\langle s_1|\langle a_1| + c_2^*\langle s_2|\langle a_2|)$$

$$\begin{aligned}
\rho_S &= \mathrm{Tr}_A(\rho_{SA}) = \sum_{k=1}^{2}\langle a_k|\rho_{SA}|a_k\rangle \\
&= \sum_{k=1}^{2}\langle a_k|(|\psi\rangle_{SA\ SA}\langle\psi|)|a_k\rangle \\
&= \sum_{k=1}^{2}\langle a_k|((c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle)(c_1^*\langle s_1|\langle a_1| + c_2^*\langle s_2|\langle a_2|))|a_k\rangle \\
&= \sum_{k=1}^{2}(c_1|s_1\rangle\delta_{k1} + c_2|s_2\rangle\delta_{k2})(c_1^*\langle s_1|\delta_{k1} + c_2^*\langle s_2|\delta_{k2}) \\
&= |c_1|^2|s_1\rangle\langle s_1| + |c_2|^2|s_2\rangle\langle s_2| \qquad (10.61)
\end{aligned}$$

where we have used orthonormality via $\langle a_i|a_j\rangle = \delta_{ij}$. We note that the Tr (trace) operation = sum over all designated states removes all knowledge of designated system from equation. This is clearly useful if we do not know much about reduced system. A similar derivation holds for ρ_A.

Continuing our discussion.

Reduced states are mixtures, not superpositions. QM predicts both LOCAL observers find mixtures not superpositions. An ensemble of experimental trials verifies this via the mixed-state pattern in agreement with assertion made earlier that neither S nor A is in a superposition.

For a different example, $\rho_A = |c_1|^2|a_1\rangle\langle a_1| + |c_2|^2|a_2\rangle\langle a_2|$ predicts Schrödinger's cat is in a mixture of either dead or alive, not a superposition of both dead and alive.

The local states $\rho_S = \mathrm{Tr}_A(\rho_{SA}) = |c_1|^2|s_1\rangle\langle s_1| + |c_2|^2|s_2\rangle\langle s_2|$ and $\rho_A = \mathrm{Tr}_S(\rho_{SA}) = |c_1|^2|a_1\rangle\langle a_1| + |c_2|^2|a_2\rangle\langle a_2|$ must be taken seriously as implying the outcomes predicted to be observed at the two sites.

The local states cannot be dismissed simply by the argument that the only "real" state is the "global state" $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$.

There is no contradiction between the predicted local mixtures $\rho_S = \mathrm{Tr}_A(\rho_{SA}) = |c_1|^2|s_1\rangle\langle s_1| + |c_2|^2|s_2\rangle\langle s_2|$ and $\rho_A = \mathrm{Tr}_S(\rho_{SA}) = |c_1|^2|a_1\rangle\langle a_1| + |c_2|^2|a_2\rangle\langle a_2|$ and

the unitarily-evolving global pure state $|\psi\rangle_{SA} = c_1 |s_1\rangle |a_1\rangle + c_2 |s_2\rangle |a_2\rangle$.

Important: To consider the combined system SA as single system only evolving unitarily misses the essential physics of nonlocality.

Haroche and Raimond weighed in on question of when a two-part system should be considered as a composite of two subsystems, versus when it should be considered as a single system B = SA.

The composite system should be considered a single system whenever the binding between the parts is much stronger than the interactions involved in the dynamics, so that internal structure of composite system is left unchanged as it travels through the experiment.

By this criterion, SA is not single system. Not only does the relation between S and A change during experiment, the relation is entangled and thus nonlocal.

The implication is that $|\psi\rangle_{SA} = c_1 |s_1\rangle |a_1\rangle + c_2 |s_2\rangle |a_2\rangle$ must be considered an entanglement of two separate systems S and A, not superposition of a single composite system SA.

That's the basics of the measurement theory! It follows completely from standard QM.

10.7.3 Now for even more details

Roch's "delayed choice" experiments used a Mach-Zehnder interferometer(Figure 4) rather than the logically equivalent double-slit setup, to observe photons.

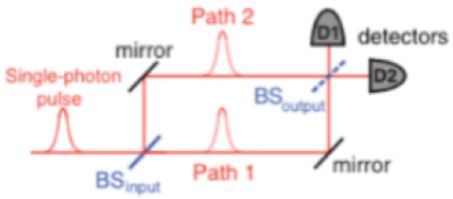


Figure 10.4:

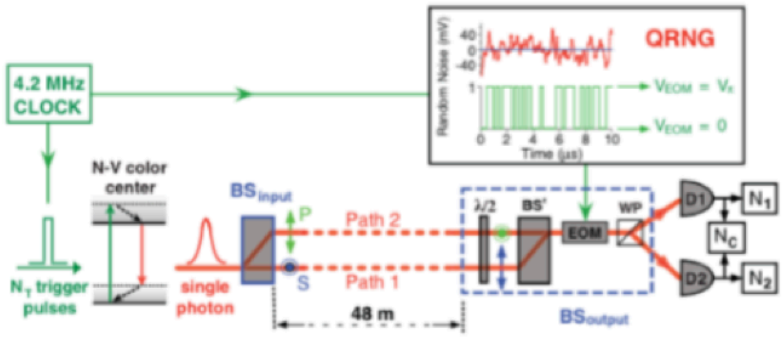


Figure 10.5:

As shown in detail in Figure 5, while a photon is on the 48-meter-long interferometer paths, a quantum-based random number generator "decided" whether the second beam splitter (positioned at the end of paths - note the spatial separation) would be incorporated or omitted, i.e., it decided whether the detectors would not or would (respectively) determine "which path".

In trials incorporating the second beam splitter, an interference pattern is observed, indicating the photon passed through the device as a superposition along both paths.

In trials omitting the second beam splitter, no interference is observed, indicating the photon passed through the device as a mixture along one or the other path.

The two parallel paths were 5 millimeters apart.

Precisely (so far as the experiment could determine) when the second beam-splitter switched from "on" to "off", the photon changed in mid-flight from being on both paths to being on one or the other path.

The quantum jump was correlated with and coincident with the incorporation or the omission of the second beam splitter (i.e., with the decision to not entangle or to entangle the detectors with the photon).

All experimental results are just as predicted by the reduced states $\rho_S = \mathrm{Tr}_A(\rho_{SA}) = |c_1|^2 |s_1\rangle \langle s_1| + |c_2|^2 |s_2\rangle \langle s_2|$ and $\rho_A = \mathrm{Tr}_S(\rho_{SA}) = |c_1|^2 |a_1\rangle \langle a_1| + |c_2|^2 |a_2\rangle \langle a_2|$.

Switching the second beam-splitter off entangles the photon and the detector in the measurement state $|\psi\rangle_{SA} = c_1 |s_1\rangle |a_1\rangle + c_2 |s_2\rangle |a_2\rangle$ causing photon

to collapse from the superposition $|\psi\rangle_S = (c_1|s_1\rangle + c_2|s_2\rangle$ into the mixture $\rho_S = \mathrm{Tr}_A(\rho_{SA}) = |c_1|^2|s_1\rangle\langle s_1| + |c_2|^2|s_2\rangle\langle s_2|$ that is observed at the detector.

At this point we are in the diagonal density matrix stage. Thus, we are doing a classical measurement and it should be interpreted as such.

This resolves problem of definite outcomes!

Quantum theory predicts and experiment verifies that, with the detector in operation, observers of S and of A find them to be in definite mixtures, not indefinite superpositions!

Any strategy of imagining the quantum and the detector to be widely separated obviously changes nothing - it does not matter whether the quantum and the detector are close together or far apart.

The key to understanding quantum measurements comes from understanding the nonlocal relationship that develops between S and A when they evolve unitarily into entangled measurement state $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$.

In experiments, the global state $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ violates Bell's inequality, implying an instantaneous non-local transfer of **correlations** across arbitrarily large distances. Without entanglement one would observe different results - the interferences would not disappear as in the experiments. But entanglement "decoheres-collapses" coherent states so that S and A impact their detectors randomly. Locally, entanglement decoheres each photon so that they exhibit definite outcomes.

But quantum dynamics is unitary, implying that the global state $|\psi\rangle_{SA} = c_1|s_1\rangle|a_1\rangle + c_2|s_2\rangle|a_2\rangle$ remains coherent despite the incoherence of subsystems.

Since the individual photons are now incoherently mixed, what has happened to coherence?

Answer: it is the experimentally observed global coincidence measurements which compare the impact points of entangled pairs.

In a double-slit experiment with two screens, each photon "knows" the impact point (i.e., phase shift) of the other photon and instantly adjusts its own impact point in order to form an interference pattern as a function of the difference between the two photons' phase shifts!

This is strikingly non-local, and the experimental results violate Bell's inequality.

Thus the coherence of the entangled state resides in the correlations between subsystems, rather than in the subsystems themselves!!

Entanglement transforms the coherence of states of S into coherence of correlations between states of S and A, allowing S and A to exhibit definite outcomes while preserving the global coherence as demanded by unitary evolution.

We can now answer the question: Precisely what is superposed and what interferes in the measurement state?

The answer is surprisingly simple: Only the correlations between S and A are superposed. Thus the measurement state $|\psi\rangle_{SA} = c_1 |s_1\rangle |a_1\rangle + c_2 |s_2\rangle |a_2\rangle$ should be read as:

> The state $|s_1\rangle$ is positively correlated with the state $|a_1\rangle$, and the state $|s_2\rangle$ is positively correlated with the state $|a_2\rangle$
>
> **Only correlations are superposed, not states.**

When the superposition $|\psi\rangle_S = (c_1 |s_1\rangle + c_2 |s_2\rangle$ of S entangles with the states of A, the superposition shifts from a superposition of states of S to a superposition of correlations between S and A, so S can be in an incoherent mixture while maintaining unitary global dynamics!

This is how nature resolves problem of definite outcomes.

So the coherence exhibited by the measurement state $|\psi\rangle_{SA} = c_1 |s_1\rangle |a_1\rangle + c_2 |s_2\rangle |a_2\rangle$ must be invisible to local observers, and yet show up in the global measurement state in order to preserve unitary dynamics.

One could regard this as the underlying reason why entanglement (i.e., measurement) must shift the coherence from the states of S and A to the correlations between S and A. "Collapse" can be viewed as consequence of the measurement state's nonlocality plus special relativity's ban on instant signaling. We note that the global measurement state is a very different animal from local states. While the local mixed states are immediately observed at both local sites, the global state can be "observed" only **at some time after** measurement by traveling to both local sites, gathering data from both, and then assembling information and noting the correlations between two sets of data.

Another difficulty(as mentioned earlier) often raised in conjunction with measurement problem: "basis ambiguity" or "the preferred basis problem". The argument is that the local state solution based on reduced states is mathematically ambiguous in the special case that $|c_1|^2 = |c_2|^2 = 1/2$ because the reduced density operators become $\rho_S = I_S/2$ and $\rho_A = I_A/2$, where I_S and I_A are identity operators in the two reduced Hilbert spaces. Measurement does not determine a unique pair of subsystem basis vectors; any orthonormal basis can be used for

each subsystem implying an ambiguous situation.

This is specious criticism because the measurement device is not constructed to pick out particular basis set in the subsystem's Hilbert space. It is designed instead to correlate with the particular physical state of measured quantum system. Regardless of what basis is used, the detector is designed to transition into $|a_i\rangle$ upon detecting S to be in $|s_i\rangle$, $(i = 1, 2)$. It's not designed to transition into some other states such as $(|s_1\rangle + |s_2\rangle)/\sqrt{2}$. The particular physical states $|s_1\rangle$ and $|s_2\rangle$ determine, unambiguously, the "natural" basis of the measurement. In the double-slit experiment, these states are for example "quantum comes through slit 1" and "quantum comes through slit 2". The ambiguity of basis sets for the reduced density operators isn't important. What's important is the specific correlations established by detector!

Our conclusions so far:

When the detector measures the superposed quantum, quantum physics predicts that the states actually observed are local (i.e., mixed or reduced) states of subsystems, not the superposed global state that follows from Schrödinger's equation.

The local states directly observed in the measurement must contain no hint of nonlocal correlations between two subsystems since this would violate relativity's prohibition on instant signaling.

Thus the local states describe what actually happens at both subsystems.

The global state predicts these local states(can be derived from), and also predicts indirectly-observable (by gathering global data at later time) nonlocal correlations between these states.

10.7.4 An even more dramatic experiment

The unity of the quantum suggests that the measurement state (52) represents a unified, hence superposed, and pure, quantum state of composite system. We asked the question: precisely what is superposed?

We answered this question by studying the simple (i.e., non-composite) superposition (50) via interference exhibited in a Mach-Zehnder experiment.

Varying the lengths of either path 1 or path 2 created varying interference effects in detectors, demonstrating each photon really must travel both paths to its detector.

Quantum theory agrees entirely with these conclusions, as can be shown by using photon wavelengths to show that path differences correctly predict the

interferences observed at each detector.

This implies that to understand the measurement state, one needs to find and analyze entanglement experiments that demonstrate interference. This has been done for several decades in connection with quantum non-locality. The key theoretical analysis was done by John Bell. Many nonlocal interference experiments have been done beginning with Clauser and Freedman culminating in experiments demonstrating nonlocality across great distances and that simultaneously closed all possible loopholes in all previous experiments.

By now, it is well known that the entangled state (52) predicts nonlocal effects between two subsystems, and that phase variations of either subsystem cause instantaneous, i.e., non-local, re-adjustments (correlations) of the possibly-distant other subsystem.

When macroscopic systems are involved (i.e., cats) we have a problem.

It is not easy to vary the phase of cat, and, as we saw in the Mach-Zehnder experiment, one cannot understand superposition without varying the phases of superposed parts.

Thus, all nonlocality experiments are carried out with pairs of simpler quanta such as photons.

The most recent nonlocal entanglement experiments most appropriate for investigating measurement were conducted nearly simultaneously by Rarity and Tapster and Ou. Figure 6 shows the layout for these "RTO" experiments.

The "source" creates entangled photon pairs by "parametric down-conversion".

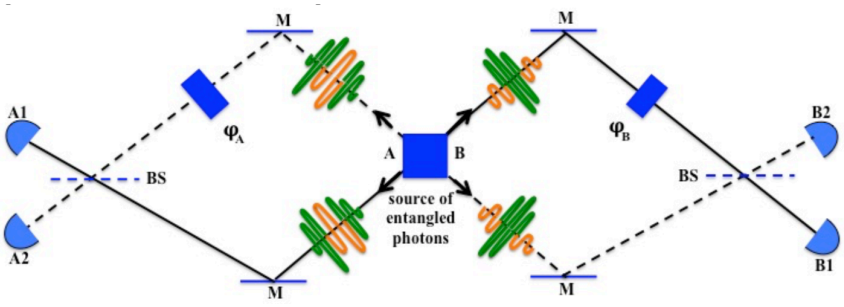


Figure 10.6:

The RTO experiment is two back-to-back interferometer experiments but with the first beam splitter for each photon located inside the source of entangled photons.

Without entanglement, each single photon (either A or B) would interfere with itself at own detectors according to its own phase shift φ_A or φ_A.

Two entangled photons are emitted into a superposition of the solid paths connecting detectors A1 and B1, and the dashed paths connecting detectors A2 and B2.

Note that the two photons are already entangled when emitted.

Entanglement changes everything.

No longer does either photon interfere with itself at its own detectors.

Instead, photons are entangled in measurement state (52) with $|\psi_1\rangle$ and $|\psi_2\rangle$ representing (say) the solid-line and the dashed-line states of A and $|1\rangle$ and $|2\rangle$ representing the solid-line and the dashed-line states of B, although in the RTO experiments neither subsystem is macroscopic.

Each photon now acts like which-path detector for the other photon.

Recall the double-slit experiment:

When the which-slit detector is switched on, the pattern on screen switches abruptly from striped interference pattern indicating the pure state nature of each electron across both slits, to the phase-independent sum of two non-interfering single-slit patterns.

Entanglement between the electron and the which-slit detector breaks the pure state into two single-slit parts, so that the measured electron comes through either slit 1 or slit 2.

This suggests that in the RTO experiment, the entanglement should break the pure-state superposition into two non-interfering parts.

This is exactly what is observed. Both photons impact their detectors as random 50-50 mixtures, just like a flipped coin.

Entanglement breaks the single-photon pure state (50) observed in the Mach-Zehnder experiment, causing each photon to behave "incoherently" with no dependence on its phase setting.

But (52) is pure state.

Where has phase dependence gone?

The answer lies in the phase-dependent but nonlocal relationship observed be-

tween the solid and the dashed branches. This phase dependence is observed experimentally in coincidence (or correlation) measurements comparing the detections of entangled pairs.

The "flipped coins" mentioned above turn out to be correlated with each other.

This phase dependence across two separated subsystems is essential to preserve the unity of the (now entangled) quantum.

This is not an easy experiment to perform:

The source creates a stream of photon pairs, and one must compare the impact of single photon A at detectors A1, A2 with the impact of corresponding entangled photon B at detectors B1, B2. RTO figured out how to do this, with the result shown in Figure 7.

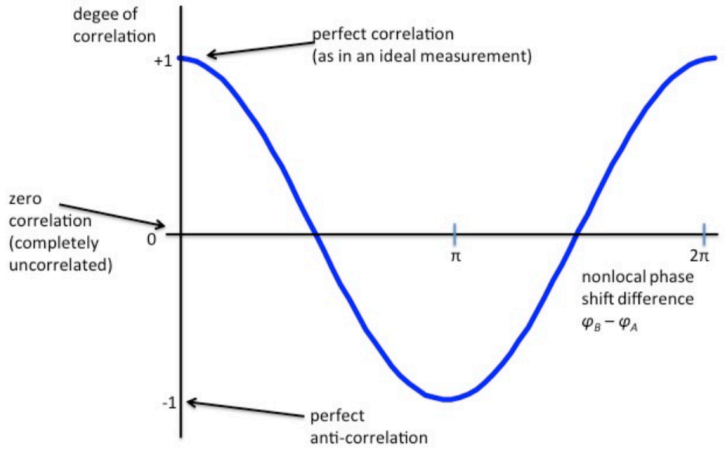


Figure 10.7:

The figure graphs the degree of correlation between A and B. This is a measure of the agreement between the outcomes at A's detectors and B's detectors.

A correlation of +1 means perfect, or 100%, agreement: Either both sets of detectors register outcome 1 (i.e., A1 and B1 click) or both register outcome 2.

The opposite extreme is a correlation of -1, meaning 100% disagreement: If one detector registers 1, the other registers 2.

Either correlation, +1 or -1, implies that either photon's outcome is predictable from the other photon's outcome.

A correlation of zero means one photon's outcome does not at all determine

other's outcome: Each photon has random 50-50 chance of either outcome regardless of the other photon.

Correlations between 0 and +1 mean the outcomes are more likely to agree than to disagree, with larger correlations denoting higher probability of agreement; for example, correlation of +0.5 means 75% probability of agreement.

Similarly, correlations between 0 and -1 mean outcomes are more likely to disagree than to agree; a correlation of -0.5 means a 75% probability of disagreement.

The RTO experiment agrees entirely with predictions of standard quantum physics.

When accounting is made of the optical paths for both photons, they obtain the following result:

$$P(correlated) = P(A1\,and\,B1) + P(A2\,and\,B2) = \frac{1}{2}[1 + \cos(\varphi_B - \varphi_A)] \quad (10.62)$$

$$P(anticorrelated) = P(A1\,and\,B2) + P(A2\,and\,B1) = \frac{1}{2}[1 - \cos(\varphi_B - \varphi_A)] \quad (10.63)$$

where P(correlated) is a single-trial probability that A's and B's detectors will agree, and P(anticorrelated) is a single-trial probability that A's and B's detectors will disagree.

The degree of correlation, defined as P(correlated) - P(anticorrelated), is then simply $\cos(\varphi_B - \varphi_A)$, as graphed in Figure 7.

In 1964, John Bell published a ground-breaking article stating the sufficient condition for a statistical theory such as quantum physics to meet condition known as "locality". He defined locality to mean "that result of measurement on one system be unaffected by operations on distant system with which it has interacted in past".

Bell expressed a sufficient condition in form of inequality that any local theory must obey.

He demonstrated that certain statistical predictions of quantum physics violate Bell's inequality, i.e., quantum physics makes nonlocal predictions.

The results in Figure 7 implies the case in point: Figure 7 violates Bell's inequality at all phase differences $\varphi_B - \varphi_A$ other than 0, π, and 2π.

Let me underline meaning of this:

Violation of Bell's inequality means that the statistics of the measurements on photon A - photon A's "statistical behavior" - is necessarily affected by the setting of photon B's phase shifter.

In fact, even without Bell's condition, the nonlocality of the experiment intuitively obvious.

Here's why:

Suppose we set the phase shifters to zero and that all four optical paths (two solid, two dashed) are then equal; thus $\varphi_B - \varphi_A$ is zero.

Without the two beam splitters BS, the two photons emitted into the solid pair and the dashed pairs of paths would impact either detectors A1 and B1 or A2 and B2 because of symmetry of experiment and conservation of momentum. This is neither surprising nor nonlocal, and would happen even if the photons were not entangled.

But the beam splitter is a randomizing device that mixes the solid and dashed paths; any photon passing through it has a 50-50 chance of reflection or transmission.

With non-entangled photons and both beams splitters in place, there would then be no correlation between photon A's outcome and B's outcome because the two photons are independent of each other.

With entanglement, correlation is perfect.

How does one photon "know" which path the other photon took at the other photon's beam splitter?

Each photon is now "detecting" the quantum state of the other photon, from a distance that could be large.

Perfect correlation certainly "feels" nonlocal even though (as mentioned above) this perfect correlation at $\varphi_B - \varphi_A = 0$ does not violate Bell's inequality. Note that such a violation is a sufficient but not a necessary condition for nonlocality.

Non-locality is written all over the RTO experiment.

Each photon "knows" which direction the other photon takes at its beam splitter and adjusts its selection accordingly.

The key nonlocal feature of graph, which is simply a cosine function, has $(\varphi_B - \varphi_A)$ as its independent variable. Thus any desired shift in correlations can be made by an observer at either of possibly-widely-separated phase shifters.

Bell suspected that this situation meant that observer A (call her Alice) could use her phase shifter to alter outcomes that would have occurred at both her own and observer B's (call him Bob) detector and, following up on this hypothesis, derived his inequality involving probabilities at both Alice's and Bob's detectors which, if violated, implied that both photons must have readjusted their states.

Such a readjustment is just what we expect, given the unity of quantum and thus the unity of atoms and other entangled systems such as our two photons.

Two photons form a single "bi-quantum", an "atom of light", in the pure state (52).

When Alice varies her phase shifter, both photons "know" both path lengths and readjust their behavior accordingly to produce the proper correlations.

Analogously, a single photon "knows" both path lengths in single-photon interferometer experiment.

Finally, we come to the central question of the discussion:

What is actually superposed in entangled superposition (52)?

A Mach-Zehnder experiment tests the simple superposition (50), while the RTO experiment tests the entangled superposition (52).

We know what is superposed in Mach-Zehnder, namely quantum states $|\psi_1\rangle$ (path 1) and $|\psi_2\rangle$ (path 2). This is deduced from the effect that either phase shifter has on both states.

Now consider the RTO experiment. What is the effect of shifting either phase shifter?

One thing that does not change is the state ("local state" would be a better term, as discussed earlier) of either photon A or photon B:

As we know, both photons remain in 50-50 mixtures regardless of either phase setting.

What does change with variations in either phase shifter is the correlations between A and B.

With $\varphi_B - \varphi_A = 0$ we have perfect correlation: Either A1 and B1 (which we denote (11)) or A2 and B2 (denoted (22)).

As we vary either φ_B or φ_A we obtain non-zero probabilities of anti-correlated individual trials, denoted (12) (outcomes A1 and B2) and (21) (A2 and B1).

When non-local phase angle difference $(\varphi_B - \varphi_A)$ reaches $\pi/2$, we have zero correlation, and when it eaches π have perfect anti-correlation.

Table 1 summarizes the crucial points in more detail.

The column titled "simple superposition" shows how the superposition state of single photon (M-Z) varies from "100% state 1" to "100% state 2" as the phase angle between two states varies.

The column titled "entangled superposition of two subsystems" shows that the state of each photon remains unchanged throughout the entire range of both phase settings, while the nonlocal correlation between states of two photons varies from "100% correlated" to "zero correlation" and then to "100% anticorrelated" as either of the two local phase angles varies.

Simple superposition:		**Entangled superposition of two sub-systems:**		
φ	State of photon	φ_B-φ_A	State of each photon	Correlation between the two photons
0	100% "1", 0% "2"	0	50-50 "1" or "2"	100% corr, 0% anti
$\pi/4$	71% "1", 29% "2"	$\pi/4$	50-50 "1" or "2"	71% corr, 29% anti
$\pi/2$	50% "1", 50% "2"	$\pi/2$	50-50 "1" or "2"	50% corr, 50% anti
$3\pi/4$	29% "1", 71% "2"	$3\pi/4$	50-50 "1" or "2"	29% corr, 71% anti
π	0% "1", 100% "2"	π	50-50 "1" or "2"	0% corr, 100% anti

Table 1. In a simple superposition, the photon's state varies with phase angle. In an entangled superposition, the relationship between states of the two photons varies, while individual states of both photons are phase-independent (or "mixed").

So once again, what is superposed in RTO experiment?

The hallmark of a superposition is the dependence on phase difference between the objects superposed.

But Table 1 exhibits no such phase dependence of states of two photons.

Each photon remains in unchanging 50-50 mixture of their own "path 1" and "path 2" states - a situation that is radically at odds with the true superposition of path 1 and path 2 exhibited by M-Z experiment.

Thus, in the entangled RTO state, neither photon is superposed.

We see here the source of the "classical" or non-superposed nature of reduced density operators (57) and (58)), not to mention the non-superposed and hence non-paradoxical nature of Schrödinger's cat.

Examination of the phase-dependence of the measurement state (52), as demonstrated by nonlocality experiments such as RTO experiment, reveals the true nature of Schrödinger's cat. The last column of Table 1 shows us what actually is superposed when two subsystems are entangled in measurement state (52).

Since the correlations between two photons vary sinusoidally as the non-local phase angle between the two photons varies, clearly these are correlations between the states of two photons, and not the states themselves, that are interfering.

Entanglement has shifted superposition, from states of one photon A ((50), M-Z) to correlations between photon A and photon B ((52), RTO).

More Repetition, More Details and Some Conclusions

In order to resolve problem of definite outcomes of measurements, aka Schrödinger's cat, our discussion analyzed the entangled state (52) of a microscopic quantum and its macroscopic measuring apparatus.

This state is a "superposition" of two composite entities $|\psi_1\rangle|1\rangle$ and $|\psi_2\rangle|2\rangle$, with a phase angle between these entities that can range over 2π radians.

To resolve problem of definite outcomes we must ask(as we have already done several times):

Precisely what does the composite superposition (52) actually superpose, physically?

In order to understand a simple non-composite superposition (50), we looked at

the effect of varying the phase angle between superposed entities $|\psi_1\rangle$ and $|\psi_2\rangle$ in experimental setting such as M-Z interferometer.

The theoretically predicted and experimentally observed results then made it obvious that the quantum whose state is (50) flows simultaneously along two separate paths described by $|\psi_1\rangle$ and $|\psi_2\rangle$.

To understand the superposition (52), one should proceed similarly by studying situations in which phase angle between superposed entities $|\psi_1\rangle|1\rangle$ and $|\psi_2\rangle|2\rangle$ varies.

One lesson of this analysis is that, in order to understand the measurement problem, one must understand the significance of nonlocality.

This is because the key measurement state (52) that caused Schrödinger and decades of experts so much concern has nonlocal characteristics.

It must be understood as a superposition of correlations, rather than a superposition of states, but this cannot become apparent until one considers the effect of variations in phase angle between its superposed terms.

Experimental or theoretical studies of such phase variations will have nonlocal ramifications, because such variations are inherently nonlocal!!

It's worth emphasizing that, when two subsystems are entangled in measurement state (52), neither subsystem is superposed - only the correlations between subsystems are superposed.

In RTO experiments, the two correlations in question are represented by the solid and the dashed paths connecting pairs of outcomes. A pair of photons entangled in state (52) follows both of these paths simultaneously.

The subsystems themselves, however, are not in superpositions but instead in indeterminate mixtures of definite states. Thus observers of either subsystem will observe only definite outcomes, as predicted by the local mixtures (57) and (58).

RTO experiments are the entangled analog of the M-Z interferometer experiment: a pair of back-to-back interferometer experiments, with entangled pair of quanta of which one quantum passes through each interferometer.

As we said earlier, the experiment and its theoretical analysis shows that, when a superposed photon A becomes entangled with second photon B to form state (52), the nonlocal aspect of A's superposition is transferred to the correlations between A and B.

Thus, an entangled state such as (52) is neither a superposition of states of A nor of states of B, but instead superposition of correlations between states of A and states of B.

To see this most clearly, we compare the simple superposition (50) with the entangled superposition (52).

In the simple superposition, the state observed by a "which-state" detector varies smoothly from 100% $|\psi_1\rangle$, through 50% $|\psi_1\rangle$ and 50% $|\psi_2\rangle$, and finally to 100% $|\psi_1\rangle$ as the phase angle φ between $|\psi_1\rangle$ and $|\psi_2\rangle$ varies from 0 to π.

In the entangled superposition, neither the state of A nor the state of B varies as φ_A or φ_B varies; both A and B remain in 50-50 mixtures throughout. What does vary is the correlation between A and B.

A non-local "correlation detector" (i.e., an RTO-type of experiment!) would find that the relation between the two subsystems varies from 100% positively correlated (either pair state 11 or 22, pictured by solid and dashed paths in Figure 6), to 50% positively correlated and 50% anti-correlated, and finally to 100% anti-correlated (12 or 21), as the nonlocal phase difference $\varphi_B - \varphi_A$ varies from 0 to π.

This is a superposition of correlations, not a superposition of composite states or of non-composite (single-system) states.

At least in the idealized case of a minimally-disturbing von Neumann measurement, the initial stage of measurement process (through formation of measurement state (52)) can be described as follows:

A quantum in a simple superposition such as (50) entangles with a macroscopic which-path detector.

At the instant of entanglement, the local states of both the quantum and the detector undergo a radical change, a quantum jump.

Locally, the detector and the quantum jump into mixtures (57) and (58).

Simultaneously, the global state (52) continues evolving smoothly according to Schrödinger equation.

Entanglement causes the superposed single quantum to be instantly transformed into superposed correlations between the quantum and the detector.

This stage of measurement process is entirely describable in terms of a pure global states following Schrödinger equation.

"Collapse" from the local superposition to the local mixtures occurs because of the formation of the entangled state (52) and the resulting formation of subsystems whose local states ((57) and (58)) have definite outcomes.

Note that the phenomenon of nonlocality is essential to preserving the pure-state nature (the unity) of the composite system.

To put this more intuitively, a reorganization throughout the entire extent of the composite entangled system is required in order to preserve the unity of the (now entangled) quantum.

According to Table 1, when two systems entangle to form the state (52), both "collapse" into phase-independent local mixtures.

Relativity requires this phase independence:

If any phase-dependent aspect of entangled state were locally observable, instant information-containing messages could be sent, violating special relativity. Local states of entangled subsystems must be invariant to phase changes.

Thus, only the relationship - the correlations - between A and B, but not A or B themselves, can vary with phase angle. Since local observers cannot detect these correlations, the entangled state cannot be used to send superluminal signals.

This is, ultimately, the reason Schrödinger's cat must be either alive or dead rather than superposition of both. A phase-dependent superposition involving both local states would permit nonlocal signaling, violating relativity.

This conclusion implies that standard physical description of composite non-entangled (i.e., factorable) product state such as $|\psi_1\rangle|1\rangle$ has been long mistaken.

Usually we regard $|\psi_1\rangle|1\rangle$ as state of the composite system AB, where subsystem A is in state $|\psi_1\rangle$ and subsystem B is in state $|1\rangle$. But this leads us into the paradox of Schrödinger's cat, where (52) represents a state in which two macroscopically different composite states exist simultaneously as a superposition.

According to present discussion, quantum theory and quantum experiments imply this entangled state to be a superposition of correlations between states rather than a superposition of composite states.

Thus $|\psi_1\rangle|1\rangle$ is not a state of composite system, but instead a correlation between two subsystems. That is, $|\psi_1\rangle|1\rangle$ means "subsystem A is in state $|\psi_1\rangle$ if and only if subsystem B is in state $|1\rangle$", an important departure from usual description.

Even if one of two subsystems happens to be a macroscopic detector, the en-

tangled state (52) is simply a non-paradoxical superposition of correlations. It says merely that state $|\psi_1\rangle$ of A is correlated with state $|1\rangle$ of B, and state $|\psi_2\rangle$ of A is correlated with the state $|2\rangle$ of B, with non-local phase angle $\varphi_B - \varphi_A$ determining the degree of each correlation.

Regardless of the phase angle, neither subsystem is in a superposition.

The entangled measurement state (52) is best described as a "macroscopic correlation": a pair of superposed (i.e., phase-dependent) quantum correlations in which one subsystem happens to be macroscopic. It is technically very difficult to create a macroscopic superposition, but macroscopic which-path detectors routinely achieve the state (52). It's not paradoxical, even though many analyses have puzzled over it.

In entanglement, nature employs an ingenious tactic. She must not violate relativistic causality, yet she must be nonlocal in order to maintain the pure-state nature of original single-quantum superposition over composite objects such as bi-photons.

Thus, she accomplishes nonlocality entirely via the superposition of correlations, because correlations cannot be locally detected and thus their superposition cannot violate relativity. This tactic lies behind the nonlocal spread of phase-dependence over large spatial distances.

By means of the superposition of correlations - entanglement - nature creates a phase-dependent pure-state quantum structure across extended quantum systems such as bi-photons.

I've frequently used the term "local" as contrasted with "global". For composite systems, and especially the entangled measurement state, it's a crucial distinction.

Entangled states such as (52) have distinct local and global (nonlocal) aspects.

The local description corresponds to two observers, each observing only one subsystem. In the case of (52), this "local description" is fully captured by the reduced density operators (57) and (58) - each local observer detects a mixture, not a superposition, of one subsystem.

The "global description" means the evolving pure state of the entire composite system, in our case (52). It is a superposition of nonlocal correlations that can only be detected by observing both subsystems and, via an ensemble of trials that individually record corresponding outcomes at both subsystems, determining the state of the correlations between them.

Although the global state implies the local description, the local description

cannot hint at the global correlations because any such hint would violate Einstein causality.

Thus, when an electron shows up in your lab, neither an examination of the electron nor an examination of an ensemble of identically-created electrons can give you the least hint of whether or how this electron is entangled with other quanta elsewhere in the universe.

This clarification of entanglement resolves problem of definite outcomes, aka Schrödinger's cat.

An ideal measurement of a superposed microscopic system A by a macroscopic detector B establishes the measurement state (52) at 100% positive correlation. This state is equivalent to the logical conjunction "A is in local state $|\psi_1\rangle$ if and only if B is in local state $|1\rangle$, AND A is in local state $|\psi_2\rangle$ if and only if B is in local state $|1\rangle$", where AND indicates the superposition. This conjunction is precisely what we want following a measurement.

Schrödinger's cat is not in the least paradoxical.

This analysis does not entirely resolve the quantum measurement problem.

It resolves the problem of definite outcomes associated with the measurement state (52), but this state continues to obey Schrödinger's equation and, hence, is reversible.

In fact, the entangled state between a quantum and its which-path detector can actually be reversed in the Stern-Gerlach experiment.

In my view, a quantum measurement must result in a macroscopic indication such as a recorded mark, and a mark is irreversible.

The above analysis shows the entangled state (52) describes a mixture of definite, not superposed, outcomes of measurements, but these outcomes remain indeterminate and the global state remains reversible.

The irreversibility problem is the question of how this nonlocal superposition of correlations then further "collapses" irreversibly to just one of its possible outcomes, a "collapse" that occurs in the RTO experiment only when one photon impacts a detector.

In the case of the RTO experiment, however, it seems fairly clear that the nonlocal superposition described by Eq. (52) must irreversibly decohere(register a value) when either of its subsystems A or B interacts with a detector.

The RTO experiment furnishes a particularly good setting for this question, because the two photons remain in the reversible entangled state (52) throughout

their flights from the source to detectors, and thus the two key questions of the measurement problem (the problem of definite outcomes and the problem of irreversibility) can be analyzed individually.

Let us now attempt to resolve the seeming existence of an "irreversiblity" problem.

10.8 The Environment as Monitor

I hope you are by now convinced that quantum physics describes the microscopic world entirely consistently and with unparalleled experimental accuracy.

But still, there's one slight problem:

> Quantum physics seems to fail utterly in describing the world around us!

We never see tables or teapots, not to mention ice cream, as wavy, space-filling fields that are possibly here and possibly there and possibly both here and there. Tables and teapots don't quantum jump, nor does ice cream. Classical physics may be inaccurate at the microscopic level, but it does explain how such ordinary objects move in response to forces.

If quantum physics describes the microscopic world correctly, and if the macroscopic world is made of microscopic objects, then the quantum principles should lead ultimately to tables and teapots.

Let us now explain, at least in part, how.

As we have seen, quantum physics began with Max Planck's hypothesis that eventually led to the quantum. The crucial quantum principle is the universe is made of these highly unified, extended bundles of field energy. To the extent that any given phenomenon depends on the spatially extended field nature of quanta, the phenomenon should be considered as "quantum". To the extent that quanta can be represented by pointlike objects, the phenomenon can usually be considered "classical".

But this classical-quantum boundary is a useful metaphor, not a law of nature. Most physicists, me included, regard the world as fully quantum. The clearest expression of the extended field nature of quanta is the principle of superposition.

Because electrons, photons, atoms, and so on are simply disturbances in fields, these objects superpose just as disturbances in the surface of a tub of water superpose. Several states can all be present at the same time.

We will describe how the process called **decoherence** converts these wavy superpositions into particle-like mixtures later.

The quantum measurement problem must be part and parcel of any discussion of how quantum physics explains our normal world. After all, quantum measurements - quantum phenomena that cause macroscopic changes - are the bridge from the micro- to the macroworld. Schrodinger's cat is a dramatic example.

The microscopic decay of a radioactive atom triggers a device that can kill a macroscopic cat. Schrodinger launched the measurement problem when he noted that the quantum rules seem to imply something we never see: a macroscopic superposition - namely, a cat that is both alive and dead.

Earlier, we suggested a resolution of this problem of definite outcomes. But there's more than that to the measurement problem.

Macroscopic processes are irreversible. The moving finger moves on. "Things run down". The second law of thermodynamics demands it.

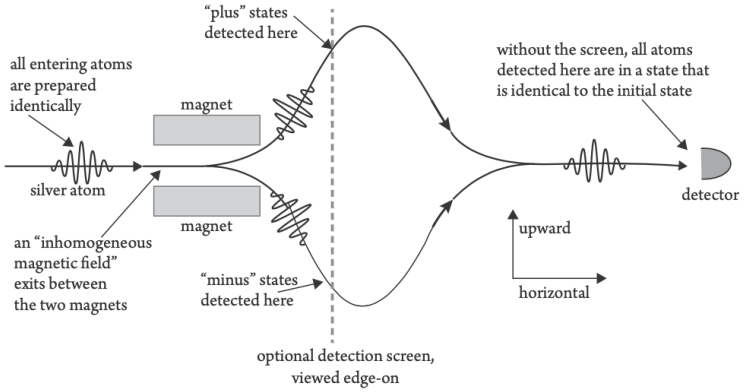
So quantum measurements must be irreversible, even though there appears to be no trace of irreversibility in the microscopic world. As we'll see, decoherence solves this mystery.

10.8.1 The Problem of Irreversibility

Every measurement involves a macroscopic change of some sort, and such changes must obey the second law of thermodynamics, so entropy increases. Furthermore, we can observe this entropy increase as a permanent (i.e., irreversible) mark made by the measurement.

But when a quantum is measured, the quantum and its detector obey the time evolution equation as they entangle to form the measurement state, so it seems that entropy doesn't increase. How can we resolve this conundrum?

A variation on the Stern-Gerlach experiment demonstrates the connection between quantum measurements and the second law.I'll focus only on the features of this experiment relevant to the irreversibility of measurements.Here's the experiment.



A horizontal stream of silver atoms passes between a pair of magnets (see figure). The entering atoms, at the left, have all been prepared previously in identical states that I'll call the "zero state". There's no need for us to concern ourselves with the precise nature of this initial state. The magnets are shaped to create a so-called "inhomogeneous magnetic field" in the space between the magnets, as shown in the figure. This field separates the stream in such a way that if a detection screen is placed "downstream" from the magnets as shown, atoms make visible impacts on the screen at two different spots, one above and the other below the original direction. Individual atoms impact randomly at one or the other spot with a 50% probability - a striking example of quantum indeterminacy - because the atoms were all prepared identically.

Examination of the two beams shows that atoms striking the upper spot are no longer in the zero state but are in a different state called the "plus state", whereas the atoms striking the lower spot are in a new state called the "minus state".

So the magnet-plus-screen combination acts as a detector to determine which atoms are in the plus state and which are in the minus state. This is entirely analogous to a double-slit experiment with a which-slit detector putting quanta into the "slit 1 state" or the "slit 2 state" before the quanta strike a viewing screen.

There's more.

If one removes the detection screen and instead installs some appropriately chosen magnets (not shown in figure) at certain points along both paths, one can bend each stream back onto its original horizontal path, as shown. On studying the atoms in the converged stream, a perhaps surprising result emerges; none of these atoms are in the plus state and none are in the minus state. Instead, every one is in the zero state from which it started!

We created a "do-nothing" box!

So this experiment, without the detection screen, is reversible. But with the screen, the experiment is obviously irreversible because the atoms make an irreversible mark when they strike the screen. It follows that any irreversibility and corresponding entropy increase in this experiment are entirely due to the impact on the screen, because we have seen that, without the screen, the experiment is reversible.

Thus the macroscopic detection makes all the difference, and it is here that we must search for all the irreversible effects of measurement. Specifically, mere entanglement, such as occurs when a which-slit detector operates in the double-slit experiment, is not responsible for irreversibility. The detection at the screen works as follows:

The detection screen interacts with each atom to form the entangled measurement state:

> atom is in the plus state, flash appears at the upper spot

entangled(superposed) with

> atom is in the minus state, flash appears at the lower spot

This is the controversial Schrodinger's cat state that we analyzed earlier (the radioactive nucleus is now the silver atom, and the cat is now the screen), where we argued that this state is only a superposition of correlations.

This experiment demonstrates that the macroscopic recording of a measurement is no small detail. No measurement is complete until it makes a mark on the macroscopic world, and making such a mark must, because of the second law, involve an entropy increase, implying irreversibility.

As John Wheeler repeatedly stressed,

> No elementary quantum phenomenon is a phenomenon until
> it is ... brought to a close by an irreversible act of amplification

If the measurement is recorded by the audible click of a detector, for example, the irreversible mark is a sound wave spreading out in all directions into the air around the detector. This sound wave warms the air a little and eventually disperses (vanishes, for all practical purposes) into a large volume of air. There is no way nature is going to spontaneously gather up every last bit of this wave's dispersed energy (while necessarily cooling the air), reverse the entire dispersal process, and use this energy to restore the detector and the air to their states before the click.

In fact, the reversed process is prohibited by the second law because it would

reduce the total entropy of the universe! We conclude that microscopic quantum processes, including entanglement, remain reversible as long as they remain microscopic, and that the irreversibility of measurements must be rooted in the macroscopic recording process. The next section looks at a macroscopic recording process that goes on all over the universe all the time.

10.8.2 How Environmental Decoherence Collapses Superpositions

A pebble lies on a sunny beach, immersed in an environment that includes atmospheric molecules, photons from the sun, cosmic rays from stars, and even photons from the Big Bang. During every second, many such quanta interact with the pebble, reflecting or otherwise scattering off the pebble in every possible direction. The scattered photons must carry away data about the pebble's orientation, structure, and color, because otherwise the pebble could not be seen. Such natural "measurement" processes occur all the time, regardless of the presence or absence of humans to consciously observe the gathered data.

To study the quantum features of such a natural measurement, instead of a pebble let's consider an atmospheric atom that is in a highly nonclassical state of being superposed at two or more macroscopically separated locations. We know that such a superposition can be created in the laboratory, for instance by passing the atom through a double-slit apparatus, but it could also occur naturally, for example if the atom passes through an opening that is sufficiently narrow to cause the atom to diffract widely.What happens when such an atom interacts with quanta in the surrounding environment?

The superposed atom interacts with, say, an environmental photon in a manner analogous to the way a superposed electron coming through double slits interacts with a which-path detector:

If the interaction is significantly different at the atom's two superposed locations, a photon can "measure" the atom by entangling with it in a measurement-like state. Such an entangling interaction is just like a which-path measurement as discussed earlier! The scattered photon carries away which-path data about the atom, just as a which-path detector carries away data about an electron coming through the slits.

As discussed earlier(reduced density operator), this measurement collapses or decoheres (removes, via a series of small environmental interactions, the interference pattern from) the superposed atom, converting its state from a superposition of being at both locations to a local mixture of being either at the first location or at the second.

In the natural environment, a single scattered photon is not likely to entirely measure (entirely decohere) the atom, because environmental quanta are not

specifically organized, the way a laboratory detector is organized, to measure superpositions.

A which-path detector in a laboratory is carefully constructed to respond in detectably different ways at different locations of the measured quantum, quite unlike environmental quanta, which interact randomly in all kinds of ways. Although a good which-path detector requires only one interaction to reliably distinguish between - reliably decohere - the superposed states of the detected quantum, a large number of environmental quanta must scatter from a typical superposed atom to completely decohere it and turn it into a local mixture.

Careful analysis shows that decoherence of typical naturally occurring superpositions requires many environmental interactions. So decoherence by the environment involves a series of partial measurements and partial collapses, each of them instantaneous, nonlocal, and similar to the single-step measurement that occurs at the slits in a double-slit experiment with a which-path detector.

It's through this environmental decoherence process that everyday objects such as pebbles and this slide lose their extended quantum field nature and behave classically, with no obvious trace of superposition, interference, or nonlocality. Small objects, of atomic dimensions, are less susceptible to environmental decoherence simply because fewer environmental quanta scatter from them as a result of their smaller size, and because each scattering event can cause only a tiny amount of decoherence because the entanglement involves nearly indistinguishable locations. Thus, a superposed photon from the Big Bang might travel the universe for 13.8 billion years without decohering, while a grain of sand on Earth, should it happen to show any signs of superposition, is decohered environmentally and essentially instantly because of the myriad environmen-tal interactions it experiences. Widely extended superpositions of small objects, such as a fine dust grain superposed in two locations separated by a millimeter, are also decohered nearly instantly because the branches of such a superposition are so distinct that a single photon reflecting from one branch but not the other can turn the entire superposition into a mixture.

In a similar way, a single detector at only one of two parallel slits is sufficient to turn a quantum that's initially in a superposition of coming through both slits into a mixture that comes through either one or the other slit. For mesoscopic objects such as dust grains, environmental decoherence turns superpositions quickly into mixtures.

How quick?

This has been calculated theoretically and measured in a few experiments.

A fine dust grain is some 10^{-5} m across. If it's in a superposition of being in two places with its two superposed branches separated by this same distance so that

the two branches are right next to each other, as though the grain had come through two closely adjacent slits, it would be decohered entirely by normal air on Earth in only 10^{-31} second. Even the best laboratory vacuum (which still contains plenty of air molecules) would decohere it within 10^{-14} second - a hundredth of a trillionth of a second. If this grain were in deep outer space, cosmic background radiation from the Big Bang would decohere it in 1 second.

For a smaller object, such as a large molecule with a diameter of 10^{-8} m, these decoherence times are longer (still assuming it's in a superposition with two branches that are separated by the molecule's diameter): 10^{-19} second if the molecule is in normal air, 0.01 second if it's in the best laboratory vacuum, but much longer than the age of the universe if it's in deep space.

The message is that sufficiently small superpositions can survive awhile, but meso- or macroscopic superpositions are fragile and are decohered quickly by tiny environmental interactions.

This is why the "quantum world" is usually identified with the microworld.

When typical quantum features become meso- or macroscopic, they generally vanish quickly. So the quantum universe appears classical at the macroscopic level because the enveloping quantum environment "monitors" every object constantly, and macroscopic objects are especially susceptible to this decoherence process. Nature is full of which-branch detectors! Note that humans aren't required in any of this - no physicists, no laboratories. Nature has been collapsing superpositions, and quanta have been losing their wavy field nature, at least since the Big Bang.

The role of the environment as an ever-present which-path monitor that turns mesoscopic and macroscopic superpositions into mixtures was first clearly recognized by Zurek during the early 1980s. The work of Zurek, his colleagues, and others has, in large part, explained how the quantum world leads to the classical world of our experience.

A wide variety of experiments have demonstrated environmental decoherence and convinced physicists that decoherence really is the mechanism that converts the quantum world into non-wavy tables and teapots. One beautiful example is an experiment by the University of Vienna group under Zeilinger and Arndt.They use the Talbot?Lau interferometer technique that demonstrated, in 2002, interference in large molecules such as C70 and certain biological molecules. The interference showed each of these large molecules to be in a superposition of following more than one path on its journey through the interferometer. Using the same technique in 2004, this group was able to demonstrate convincingly the "environmental" decoherence of superpositions of C70 molecules.

I put environmental in quotation marks because in this experiment the environment came from within the molecules themselves rather than from an external environment. Individual C70 molecules passed through the interferometer, as in the 2002 experiment. But there was a new twist; the experimenters heated the molecules just before sending them through the interferometer. They expected that, with sufficient heating, the molecules would themselves emit thermal radiation in the form of visible and infrared "thermal photons", just as an electric hot plate emits thermal radiation (you can feel its warmth at a distance, and it might glow red) when heated. The radiation comes from the random thermal motion of the many atoms and other quanta within the molecule. According to decoherence theory, each of these radiated photons should act as a partial which-path detector, collapsing the molecule's superposition state partially by carrying a certain amount of which-path data from the molecule into its surroundings. If enough such data are transferred, this which-path measurement should decohere the molecule, causing the interference pattern to dim or vanish - the signature of a superposition evolving into a mixture.Their results demonstrated decoherence in action. With the molecules only slightly heated, the emitted photons had low energies and thus low frequencies and long wavelengths - too long to distinguish between the possible paths, which for such a massive molecule are separated by extremely small distances. But when the molecules were heated to a few thousand kelvins, the emitted photons' wavelengths became short enough to distinguish between the possible paths, so the which-path data transmitted to the surroundings was sufficient to partially decohere the superposed molecules, causing the interference pattern to partially vanish. At sufficiently high temperatures, theory and experiment showed that a mere two or three emitted photons sufficed to decohere each molecule. There was quantitative agreement between the predictions and observations. The interference pattern began decohering at just the predicted temperature, and the degree of visibility of the remaining interference was just as predicted. The experiment revealed the step-by-step action of decoherence as data leaked, photon by photon, into the environment.

It also demonstrated the extreme sensitivity of superpositions to decoherence. Just a few high-temperature photons turn a massive superposed molecule into an incoherent mixture. Superposed quanta are delicate.

10.8.3 Decoherence and the Measurement Problem

Decoherence provides a solution of the irreversibility problem for natural environmental measurements.

Decoherence was first introduced during the early 1980s to explain how our apparently classical surroundings arise from measurement-like interactions with the environment.
12pt] The local state argument, presented earlier, shows that such measurements turn superpositions into entangled nonlocal measurement states with definite outcomes, resolving the problem of definite outcomes for environmental mea-

surements.

But just as for laboratory measurements, we must ask if this also resolves the irreversibility problem associated with natural measurements.The answer is yes.

Here's why.

As we've seen, when a small superposed dust grain is decohered by the surrounding environment, the object undergoes a series of small entanglement-caused partial collapses. These environmental measurements are ?recorded? by the many environmental photons and air molecules that interact with the grain and then disperse widely. Data about the superposition state of the grain before it decohered are now scattered randomly far and wide. Just as one cannot unscramble an egg, one cannot reversibly gather these pieces back together and reconstruct the global state of the air plus the superposed grain.

This argument is especially compelling in the case of environmental decoherence, because of the environment's enormous size. This is an obvious example of the second law in action; for all practical purposes, the process is irreversible and entropy has increased.

Because there is famous opposition to introducing for-all-practical-purposes arguments into physics, it needs to be noted that the second law itself is inherently a for-all-practical-purposes principle. We know that a box full of gas could evolve spontaneously into separate regions of hot gas and cold gas simply by chance, with no external assistance. The chances of this are ridiculously small, but they are not zero, and if we consider boxes containing smaller and smaller amounts of gas, these odds increase. Such for-all-practical-purposes arguments that trace back to the second law are entirely in keeping with the principles of physics.

Here's a tale about that. Joe, deep in philosophical conversation with Schmo, points out, "Entropy never decreases."

"Never?" asks Schmo.

"No, never," responds Joe.

"What? NEVER?" shouts Schmo.

Looking up, Joe shrugs. "Well, hardly ever."

So the local state resolution plus decoherence combine to resolve the measurement problem entirely for the case of natural environmental measurements. The environment measures a superposed object (molecule, dust grain, and so on) when myriad environmental quanta convert the superposition into a measurement state that now takes the form of a nonlocal entanglement between the

object and a very dispersed state of the many environmental quanta.

The local state argument implies this measurement state represents a mixture of definite properties, whereas the highly dispersed nature of the environmental quanta guarantees the process is irreversible.

There has, for years, been a question about what is the exact relationship between decoherence and the solution of the measurement problem. Some accounts appear to imply that decoherence alone resolves the measurement problem, but this is not true. An especially clear proof of this appears in an article provocatively titled "Why Decoherence Has Not Solved the Measurement Problem: A Response to P. W. Anderson," by Stephen Adler. Anderson, a Nobel laureate, claimed in 2001 that decoherence does solve the measurement problem. Adler is convincingly correct; decoherence alone does not solve the problem of definite outcomes, so it does not by itself solve the measurement problem.

Maximilian Schlosshauer, a quantum foundations expert who has written widely about decoherence, also states that

> decoherence cannot solve the problem of definite outcomes in quantum measurement

The connection of decoherence to the measurement problem is that it resolves the irreversibility problem in the case of natural environmental measurements, and it shows how environmental interactions transform a superposed quantum into the measurement state, but it does not solve the problem of definite outcomes associated with this measurement state. The problem of definite outcomes is, however, resolved by the local state analysis, as explained earlier.

So the local state solution combines with environmental decoherence to solve the measurement problem for the case of environmental measurements.

What about the case of laboratory measurements?

Here, the preceding discussion resolved the definite outcomes problem, but not the irreversibility problem. Irreversibility poses slightly different problems in the two cases, because lab measurements are sufficiently controlled that the natural environment has little effect. In fact, much of an experimentalist's efforts go precisely into ensuring that random environmental interactions have no significant effect on an experiment's outcome.

Our discussion of the Stern-Gerlach experiment suggested that the answer lies in the irreversible nature of the macroscopic detection process. This suggestion resembles the resolution of the irreversibility problem for environmental measurements:

Environmental measurements are "recorded" by innumerable environmental quanta whereas lab measurements are recorded by detection screens, electronic clicks, or, perhaps, cats.

Let's focus on a single impact made on a detection screen by one electron in a double-slit experiment - one of the small spots.

How was it recorded?

This mark was initiated by a single electron, but to observe it macroscopically the original impact had to be detected and then amplified sufficiently for humans to see it. This is the purpose of every laboratory detector of microscopic events, and it's why an ?irreversible act of amplification? (as Wheeler puts it) is essential to quantum measurements.

In the 1989 experiment that produced the patterns of dots on a screen, each electron was emitted in a single coherent state, the same state for each electron. EM fields accelerated these electrons to high speeds, about 180,000 kilometers per second, or 60% of light speed - much faster than the electron in a normal hydrogen atom, which orbits at an average speed of "only" 2000 kilometers per second. High energies were needed to make a detectable impact. Each electron went through a pair of parallel slits, then spread out into an interference state that interacted indeterminately with the screen. The interaction entangled the electron with the screen, decohering the electron and collapsing it instantaneously from its earlier superposition state over the entire screen to a locally observed mixture. In other words, the electron collapsed randomly into one or the other of many small atom-size regions in the screen. For the fluorescent screen used in the experiment, this interaction created some 500 photons that marked and amplified the location of the impact - already an irreversible process because the photons are emitted randomly in all directions. In a process called the photoelectric effect, each photon then struck a metal plate and dislodged an electron from it, converting each of the 500 photons into a low-energy electron. These 500 low-energy electrons were then focused into a point image that could be displayed on a TV monitor in much the same way that old (before flat screens) TV tubes operated. Every step of this process - creation of 500 photons, conversion to photoelectrons, and amplification to create the final display - creates entropy and is irreversible thermodynamically. It's the second law of thermodynamics in action.

This illustrates the general case:

Laboratory detection and amplification of a quantum event necessarily involves irreversibility and entropy production. The laboratory detector, a macroscopic device made of many microscopic quanta, plays the decohering role that the environment plays in natural measurements.

To summarize: The quantum measurement problem, in both its laboratory and environmental senses, is resolved entirely by the local state solution of the problem of definite outcomes and by the irreversibility of the decoherence process. You have passed through the looking glass and explored the land of the quantum. I hope the journey has been meaningful and fun. For more than a century, much has been made of the odd and supposedly paradoxical nature of the quantum. This presumed quantum spookiness has led to an excess of attempted fixes and interpretations. Many experts have even declared the theory to be not a description of reality at all, but only a mathematical recipe that helps humans predict the results of experiments.

As Niels Bohr put it,

> There is no quantum world. There is only an abstract quantum description.

According to this hypothesis, quantum theory doesn't describe anything real at all, so there's no cause for concern about collapse of the quantum state and other odd quantum behaviors. Such an easy resolution of the quantum quandaries amounts to giving up on science's project of understanding the realities of the natural world. It's an extraordinary claim, requiring extraordinary proof. But there is no such proof, and there are no grounds for regarding quanta as any less real than rocks. Indeed, rocks are made of quanta.

Although it has long been a conceit of humankind to imagine the universe to be centered around us, there is no reason to think that reality comprises only the kinds of things we experience in our own daily lives. The real world does not fade from existence, nor does it become incomprehensible, at distances that happen to be several powers of ten smaller than teapots. Atomic and subatomic processes are just as real as teapots and, with the help of technology, accessible to human experimentation and understanding.

From the viewpoint of the macroscopic and classical world that we are pleased to call "normal", there is certainly oddness in wave-particle duality, indeterminacy, quantum states, superposition, nonlocality, measurement, and quantum jumps. But there are no logical contradictions here, no disagreements with experiment, and nothing that should persuade us that quantum physics is about anything other than the real world.

Quantum physics is either charmingly counterintuitive or maddeningly puzzling, depending on your taste, but it is entirely self-consistent and experimentally accurate. It's time to accept it with all its charms and puzzles, and stop trying to repair or reinterpret it. It's time, in other words, to relax and admit the world is not as we had thought.

Nature is far more creative than we could have conceived. Our most fundamental theory is in better shape than its detractors suppose. Quantum physics is a

remarkable treasure trove of far-reaching phenomena and ideas whose surface we have probably only begun to scratch. It's time to fully embrace these ideas, incorporating them into our ways of thinking about the universe, about our planet, and about ourselves. This is a process that will engage our minds and stretch our imaginations far into the future, for quantum physics is, indeed, not what anybody could have imagined.

10.9 Closing Story #1

Let us look at EPR again now that we understand what is happening in entangled quantum systems.

Depending on basis used (language chosen) we have one of two states for a pair of electrons:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle|\downarrow_z\rangle - |\downarrow_z\rangle|\uparrow_z\rangle) \quad \text{or} \quad |\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow_x\rangle|\downarrow_x\rangle - |\downarrow_x\rangle|\uparrow_x\rangle) \tag{10.64}$$

Let us once again consider the EPR reasoning to see whether it leads to the conclusion of incompleteness of quantum mechanics. If Alice performs a measurement on spin along the z direction and the outcome is spin up, the state vector after measurement is updated to $|\uparrow_z\rangle|\downarrow_z\rangle$.

This just means that if Bob were to measure the z-component of spin, the spin value would be DOWN according to Alice. This description of physical reality is true only **relative to Alice**.

From Bob's perspective, before he knows the Alice's measurement result, he still views the composite system in the original state, i.e., no quantum event happened yet. In other words, Bob still predicts that future measurement spin will find the UP state with 50% chance.

Since the spins of the two electrons are entangled, Alice's spin is a measuring apparatus for Bob's spin.

Since Alice performed a spin measurement, she effectively reads the measuring apparatus. Therefore, she is the intrinsic observer, and Bob is an external observer. At this point, both observers are out of synchronization on the relational information of the two particles, thus they give different predictions of Bob's particle spin value.

To verify the physical description Alice obtained for Bob's spin value after measuring the Alice particle spin value, Alice can travel to Bob's location to perform a measurement, or can send the measurement result to Bob and ask Bob to perform a measurement.

Suppose Alice sends the measurement outcome to Bob. Bob updates state to $|\uparrow_z\rangle|\downarrow_z\rangle$, the same as state relative to Alice.

He now can confirm the physical reality that the Bob spin value is DOWN with unit probability.

However, in this state, he cannot predict deterministically that Bob particle spin value is would be in the x-direction, since $|\downarrow_z\rangle$ is superposition of x-components.

Similarly, if Alice performs a measurement on the Alice particle spin along the x-direction and the outcome is spin UP, then the Bob particle spin value is deterministically DOWN(x-direction) relative to Alice, but nothing happened from Bob's view.

If Alice sends the measurement result to Bob, Bob updates the state vector accordingly to $|\uparrow_x\rangle|\downarrow_x\rangle$. He now can confirm the physical reality that the Bob particle spin value is the right state $|\downarrow_x\rangle$.

However, in this state, he cannot predict deterministically that the Bob's particle spin value would be in the z-direction, since $|\downarrow_x\rangle$ is superposition of z-components.

Since Alice cannot perform measurement on the Alice spin value along z and x directions at the same time, Bob cannot confirm the Bob spin value has spin values in both z and x directions simultaneously. The reality that the Bob spin value simultaneously has definite values for S_z and S_x cannot be verified.

This is consistent with the Heisenberg Uncertainty Principle. There is no incompleteness issue for quantum mechanics. Hence the original EPR argument.

However, there is still a puzzle here, namely, a non-causal correlation.

It appears Bob's measurement outcome on the Bob spin value "depends" on which direction Alice chooses to measure the Alice spin value.

Since Alice's measurement does not impact the physical property of particle Bob spin value, we can ask exactly what spin state is Bob spin value in before Alice's measurement?

To answer this subtle question, first note that it is Alice's new knowledge of the Bob spin value, not the physical reality of the Bob spin value, that depends on the axis along which the measurement is performed.

One cannot assume there exists an absolute reality for Bob spin value.

To confirm the new found reality of the Bob spin value relative to Alice, Alice

sends the measurement result to Bob who performs a subsequent measurement.

There is no faster-than-light action here.

Secondly, it is true that Bob's measurement outcome correlates to the Alice's measurement result. But this is an informational correlation, not a causal relation.

This correlation is encoded in the entangled state of the composite system $|\psi\rangle$. Since the entanglement is preserved even when both particles are space-like separated, the correlation is preserved.

Such an entangled quantum state contains not only the classical correlation, but also the coherence information of the composite system.

When Alice measures the Alice particle spin value, she effectively measures the composite system, because she obtains information not only about the Alice spin value, but also about the correlation between the Alice spin value and the Bob spin value.

In addition, the measurement induces the decoherence of the composite system.

Before Alice performs the measurement, it is meaningless to speculate what spin state the Bob particle spin is in.

When Alice measures the Alice spin value along the z direction and obtains result of spin up, she knows that in this condition, the Bob spin value is in spin down and later this is confirmed by Bob.

If instead, she measures the Alice spin value along the x direction and obtains result of spin left, she knows that in this new condition, the Bob spin value is in spin right and is later confirmed by Bob.

To better understand this non-causal relation, suppose we have many identical copies of the entangled pairs described by $|\psi\rangle$.

Alice measures the Alice particle spin values sequentially along the z direction and she does not send measurement results to Bob. Bob independently measures the Bob particle spin values along z direction as well.

Both of them observe their own measurement results for S_z as randomly spin up or spin down, but with fifty percent of chance for each. When later they meet, compare measurement results.They find two sequences of S_z values that are exactly opposite.

They can even choose a random sequence of z or x directions but both follow

exact sequence in their independent measurements. When later they meet and compare measurement results, they still find their measured values are the opposite sequentially.

What does this all mean?

Special Relativity forces us to abandon the concept of absolute time.

Measurement of time is observer-dependent.

Similarly, in relativistic QM, the idea of an observer independent quantum state must be abandoned.

Space-like separated observers, however, can reconcile the different descriptions of the same quantum system through classical communication of information obtained from local measurements.

10.10 Closing Story #2 - Some Provocative Thoughts - A possible future direction.....

Quantum mechanics began with Heisenberg's "Umdeutung(Reinterpretation)" paper, i.e., his proposed "reinterpretation" of physical quantities at the fundamental level as non-commutative.

To say that the algebra of physical quantities is commutative is equivalent to saying that the projection operators form a Boolean algebra (Boolean algebra is the branch of algebra in which the values of the variables are the truth values true and false, usually denoted 1 and 0 respectively).

They represent yes-no observables, or properties (for example, the property that the energy of the system lies in a certain range of values), or propositions (the proposition asserting that the value of the energy lies in this range), with the two eigenvalues corresponding to the truth values, true and false.

Heisenberg's insight amounts to the proposal that certain phenomena in our Boolean macroworld that defy a classical physical explanation can be explained probabilistically as a manifestation of collective behavior at a non-Boolean micro-level.

The Boolean algebra of physical properties of classical mechanics is replaced by a family of "intertwined" Boolean algebras, one for each set of commuting observables.

The intertwinement precludes the possibility of embedding the whole collection

into one inclusive Boolean algebra, so you can't assign truth values consistently to the propositions about observable values in all these Boolean algebras.

Putting it differently: there are Boolean algebras in the family of Boolean algebras of a quantum system, notably the Boolean algebras for position and momentum, or for spin components in different directions, that don't fit together into a single Boolean algebra, unlike the corresponding family for a classical system.

The intertwinement of commuting and noncommuting observables in Hilbert space imposes objective pre-dynamic probabilistic constraints on correlations between events.

The probabilistic constraints encoded in the geometry of Hilbert space provide the framework for the physics of a genuinely indeterministic universe.

They characterize the way probabilities fit together in a world in which there are nonlocal probabilistic correlations that violate Bell's inequality.

Quantum probabilities don't quantify incomplete knowledge about a state, but reflect the irreducibly probabilistic relation between the non-Boolean microlevel and the Boolean macrolevel.

This means that quantum mechanics is quite unlike any theory we have dealt with before in the history of physics, and there is no reason, apart from tradition, to assume that the theory can provide the sort of explanation we are familiar with in a theory that is commutative or Boolean at the fundamental level.

Quantum probabilities can't be understood in the Boolean sense as quantifying ignorance about the pre-measurement value of an observable, but give results in terms of what you'll find if you "measure", which involves considering the outcome, at the Boolean macrolevel, of manipulating a quantum system in a certain way.

A quantum "measurement" is a bit of a misnomer and not really the same sort of thing as a measurement of a physical quantity of a classical system.

It involves putting a microsystem, like a photon, in a situation, say a beamsplitter or an analyzing filter, where the photon is forced to make an intrinsically random transition recorded as one of two macroscopically distinct alternatives in a device like a photon detector.

The registration of the measurement outcome at the Boolean macrolevel is crucial, because it is only with respect to a suitable structure of alternative possibilities that it makes sense to talk about an event as definitely occurring or not

occurring, and this structure is a Boolean algebra.

From this perspective, Heisenberg's theory provides a way of deriving probabilities and probabilistic correlations with no causal explanation.

They are "uniquely given from the start" as a feature of the non-Boolean structure, related to the angles in Hilbert space, not measures over states as they are in a classical or Boolean theory.

The really significant thing about a noncommutative mechanics is the novel possibility of correlated events that are intrinsically random, not merely apparently random like coin tosses, where the probabilities of "heads" and "tails" represent an averaging over differences among individual coin tosses that we don't keep track of for practical reasons.

This intrinsic randomness allows new sorts of nonlocal probabilistic correlations for "entangled" quantum states of separated systems.

The view that Hilbert space is fundamentally a theory of probabilistic correlations that are structurally different from correlations that arise in Boolean theories is, in effect, an information-theoretic interpretation of quantum mechanics.

On this way of understanding quantum mechanics, as a non-classical theory of information or a new way of generating probabilities and probabilistic correlations between intrinsically random events, probabilities are defined with respect to a single Boolean frame, the Boolean algebra generated by the "pointer-readings" - the "ultimate measuring instruments", which are "kept outside the system subject to quantum mechanical treatment"

It's not that unitarity is suppressed at a certain level of complexity, where non-Booleanity becomes Booleanity and quantum becomes classical.

Rather, there is a macrolevel, which is Boolean, and there are actual events at the macrolevel.

Any system, of any complexity, is fundamentally a quantum system and can be treated as such, in principle, which is to say that a unitary dynamical analysis can be applied to whatever level of precision you like.

The crucial assumption in this probabilistic interpretation of the theory is that the outcome of a measurement is an intrinsically random event at the macrolevel, something that actually happens, not described by the deterministic unitary dynamics, so outside the theory, or "irrational".

Putting it differently, the "collapse", as a conditionalization of the quantum

state, is something you put in by hand after recording the actual outcome.

The physics doesn't give it to you.

Special relativity, as a theory about the structure of space-time, provides an explanation for length contraction and time dilation through the geometry of Minkowski space-time, but that's as far as it goes.

This explanation didn't satisfy Lorentz, who wanted a dynamical explanation in terms of forces acting on physical systems used as rods and clocks.

Quantum mechanics, as a theory about randomness and nonlocality, provides an explanation for probabilistic constraints on events through the geometry of Hilbert space, but that's as far as it goes.

10.11 Last Thoughts

When a detector measures a superposed quantum, quantum physics predicts that states actually observed are local (i.e., mixed or reduced) states of subsystems, not the superposed global state that follows from Schrödinger's equation. The local states that are directly observed in a measurement must contain no hint of the nonlocal correlations between the two subsystems, lest relativity's prohibition on instant signaling be violated.

Thus the local states describe what actually happens in both subsystems.

The global state predicts these local states, and also predicts the indirectly-observable (by gathering global data at a later time) nonlocal correlations between these states.

Experiments confirm both the local and the global predictions, namely that outcomes are definite but unpredictable, and correlations between these definite outcomes are as described by the entangled measurement state.

This resolves problem of definite outcomes of measurements.

Nonlocality experiments demonstrate the precise nature of superposition inherent in the global measurement state:

This state is a superposition only of correlations between the detector and its observed quantum, not a superposition of states of the detector or quantum. It should be read as "first the quantum state is correlated with the first detector state AND the second quantum state is correlated with the second detector state".

The word AND indicates superposition. By this shifting of coherence from states of subsystems to correlations between subsystems means that evolution of global state can remain unitary, and global state can remain coherent, while both subsystems collapse into incoherent mixtures of unpredictable but definite outcomes. This is the way nature resolves problem of definite outcomes.

Analyses such as this are sometimes called "no-collapse" solutions of measurement problem, but this is a misnomer. There is a physically real, instantaneous, and observable collapse - a quantum jump - of observed local states, as verified quite explicitly by Roch's delayed-choice experiment.

The global state cannot collapse because it obeys unitary dynamics; it accordingly entangles rather than collapses upon measurement. Observed phenomena collapse, while the global state continues smoothly evolving as predicted by Schrödinger's equation.

The notion of "collapse of the quantum state" must be replaced by "collapse of the local state, and unitary evolution of the global state".

Both aspects, local and global, are correctly predicted by standard quantum theory and verified by experiments.

If we add decoherence into the mix, then the problem of irreversibility is also removed!

So let me repeat.....

There are no problems with standard quantum mechanics! We have now finished our coverage of basic Quantum Mechasnics.

In the next(last) chapter we again discuss topics brought up earlier, such as EPR and theBell Inequality in more detail and more mathematically, for completeness.

But before doing that I am going to insert Chapter 11 of Jauch's book on Foundations of Quantum Mechanics(1968) where he presents the local-state idea and solves the measurement problem in his own way. I hope that, after you read Jauch's work you will agree that it is basically the same idea as I have presented and thus confirms my ideas.

Chapter 11

The EPR Argument and Bell Inequality

11.1 Hidden variables and Bell's Inequalities-1st Try

As we have seen, when a quantum system possesses more than one degree of freedom, the associated Hilbert space is a tensor product of the spaces associated with each degree of freedom. This structure leads to specific properties of quantum mechanics, whose *paradoxical* nature has been discussed in earlier chapters. In this section we will study an example of such a situation by considering entangled states for the spins of two particles.

The system under consideration is a hydrogen atom which is dissociated into an electron and a proton. We consider the spin states of these two particles when they have left the dissociation region and are located in spatially distinct regions, e.g. a few meters from one another. They are then considered to be free particles whose spin states do not evolve.

11.1.1 The Electron Spin

Consider a unit vector $\hat{u}_\phi$ n the (z,x) plane given by $\hat{u}_\phi = \cos\phi\,\hat{u}_z + \sin\phi\,\hat{u}_x$ where $\hat{u}_x$ and $\hat{u}_z$ are unit vectors in the x and z directions. We note that $\hat{S}_{e\phi} = \hat{S}_e \cdot \hat{u}_\phi$ is the component of the electron spin in the $\hat{u}_\phi$ direction.

In the eigenbasis $|e:\pm\rangle$ of $\hat{S}_{ez}$, the matrix representing $\hat{S}_{e\phi}$ is

$$\begin{pmatrix} \langle +|\hat{S}_{e\phi}|+\rangle & \langle +|\hat{S}_{e\phi}|-\rangle \\ \langle -|\hat{S}_{e\phi}|+\rangle & \langle -|\hat{S}_{e\phi}|-\rangle \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix} \tag{11.1}$$

which has eigenvalues $\pm\hbar/2$ (true for any direction).

The corresponding eigenvectors are

$$|e:+\phi\rangle = \cos\frac{\phi}{2}\,|e:+\rangle + \sin\frac{\phi}{2}\,|e:-\rangle \tag{11.2}$$

$$|e:-\phi\rangle = -\sin\frac{\phi}{2}\,|e:+\rangle + \cos\frac{\phi}{2}\,|e:-\rangle \tag{11.3}$$

If the electron is emitted in the state $|e:+\phi\rangle$, the probability $P_+(\alpha)$ of finding the electron in the state $|e:+\alpha\rangle$ is given by

$$\begin{aligned} P_+(\alpha) &= |\langle e:+\alpha \mid e:+\phi\rangle|^2 \\ &= \left|\left(\cos\frac{\alpha}{2}\,\langle e:+| + \sin\frac{\alpha}{2}\,\langle e:-|\right)\left(\cos\frac{\phi}{2}\,|e:+\rangle + \sin\frac{\phi}{2}\,|e:-\rangle\right)\right|^2 \\ &= \left|\left(\cos\frac{\alpha}{2}\cos\frac{\phi}{2} + \sin\frac{\alpha}{2}\sin\frac{\phi}{2}\right)\right|^2 = \cos^2\frac{\phi-\alpha}{2} \end{aligned} \tag{11.4}$$

and similarly,

$$\begin{aligned} P_-(\alpha) &= |\langle e:-\alpha \mid e:+\phi\rangle|^2 \\ &= \left|\left(-\sin\frac{\alpha}{2}\,\langle e:+| + \cos\frac{\alpha}{2}\,\langle e:-|\right)\left(\cos\frac{\phi}{2}\,|e:+\rangle + \sin\frac{\phi}{2}\,|e:-\rangle\right)\right|^2 \\ &= \left|\left(-\sin\frac{\alpha}{2}\cos\frac{\phi}{2} + \cos\frac{\alpha}{2}\sin\frac{\phi}{2}\right)\right|^2 = \sin^2\frac{\phi-\alpha}{2} \end{aligned} \tag{11.5}$$

Using these results, the expectation value of $\hat{S}_{e\alpha}$ in the $|e:+\phi\rangle$ state is then

$$\begin{aligned} \langle \hat{S}_{e\alpha}\rangle &= \left(+\frac{\hbar}{2}\right)P_+(\alpha) + \left(-\frac{\hbar}{2}\right)P_-(\alpha) \\ &= \frac{\hbar}{2}\left(\cos^2\frac{\phi-\alpha}{2} - \sin^2\frac{\phi-\alpha}{2}\right) = \frac{\hbar}{2}\cos(\phi-\alpha) \end{aligned} \tag{11.6}$$

11.1.2 Correlations Between the Two Spins

We assume that after the dissociation, the electron-proton system is in the factorized spin state $|e:+\phi\rangle \otimes |p:-\phi\rangle$. Now if $|u_1\rangle \in E$ and $|u_2\rangle \in E$ and $|v_1\rangle \in F$ and $|v_2\rangle \in F$, then $|u\rangle \otimes |v\rangle \in G = E \otimes F$ and if $\hat{A}$ and $\hat{B}$ act respectively in E and F, then $\hat{C} = \hat{A} \otimes \hat{B}$ acts in G. Then one has $\langle u_2| \otimes \langle v_2|\,\hat{C}\,|u_1\rangle \otimes |v_1\rangle = \langle u_2|\,\hat{A}\,|u_1\rangle\,\langle v_2|\,\hat{B}\,|v_1\rangle$.

We now determine the probability $P_+(\alpha)$ of finding $+\hbar/2$ when measuring the component $\hat{S}_{e\alpha}$ of the electron spin in this state.

The projector on the eigenstate $|e:+\alpha\rangle$, corresponding to the measured value, is $|e:+\alpha\rangle\,\langle e:+\alpha| \otimes \hat{I}_p$, where $\hat{I}_p$ is the identity operator on the proton states.

Therefore,

$$\begin{aligned} P_+(\alpha) &= \langle p:-\phi| \otimes \langle e:+\phi| \left(|e:+\alpha\rangle \langle e:+\alpha| \otimes \hat{I}_p\right) |e:+\phi\rangle \otimes |p:-\phi\rangle \\ &= \langle p:-\phi| \hat{I}_p |p:-\phi\rangle \langle e:+\phi \mid e:+\alpha\rangle \langle e:+\alpha \mid e:+\phi\rangle \\ &= |\langle e:+\alpha \mid e:+\phi\rangle|^2 = \cos^2 \frac{\phi-\alpha}{2} \end{aligned} \tag{11.7}$$

and the state after the measurement is $|e:+\alpha\rangle \otimes |p:-\phi\rangle$. The proton spin is not affected, because the initial state is factorized (and all probability laws are factorized).

For $\hat{u}_\phi = \cos\phi\hat{u}_z + \sin\phi\hat{u}_x$ and $\hat{u}_\beta = \cos\beta\hat{u}_z + \sin\beta\hat{u}_x$ we can calculate the expectation values (as earlier)

$$\left\langle \hat{S}_{e\alpha} \right\rangle = \frac{\hbar}{2}\cos(\phi-\alpha) \quad , \quad \left\langle \hat{S}_{p\beta} \right\rangle = -\frac{\hbar}{2}\cos(\phi-\beta) \tag{11.8}$$

The correlation coefficient between the two spins $E(\alpha,\beta)$ is defined by

$$E(\alpha,\beta) = \frac{\left\langle \hat{S}_{e\alpha} \otimes \hat{S}_{p\beta} \right\rangle - \left\langle \hat{S}_{e\alpha} \right\rangle \left\langle \hat{S}_{p\beta} \right\rangle}{\left(\left\langle \hat{S}^2_{e\alpha} \right\rangle \left\langle \hat{S}^2_{p\beta} \right\rangle \right)^{1/2}} \tag{11.9}$$

Now

$$\hat{S}^2_{e\alpha} = \frac{\hbar^2}{4}\hat{I}_e \quad , \quad \hat{S}^2_{p\beta} = \frac{\hbar^2}{4}\hat{I}_p \tag{11.10}$$

and

$$\begin{aligned} \left\langle \hat{S}_{e\alpha} \otimes \hat{S}_{p\beta} \right\rangle &= \langle e:+\alpha| \hat{S}_e |e:+\alpha\rangle \langle p:+\beta| \hat{S}_p |p:+\beta\rangle \\ &= -\frac{\hbar^2}{4}\cos(\phi-\alpha)\cos(\phi-\beta) \end{aligned} \tag{11.11}$$

Thus,

$$E(\alpha,\beta) = \frac{-\frac{\hbar^2}{4}\cos(\phi-\alpha)\cos(\phi-\beta) + \frac{\hbar^2}{4}\cos(\phi-\alpha)\cos(\phi-\beta)}{\frac{\hbar^2}{4}} = 0$$

This just reflects the fact that in a factorized state, the two spin variables are independent.

Correlations in the Singlet State

Now assume that, after the dissociation, the two particles are in the singlet spin state

$$|\psi_s\rangle = \frac{1}{\sqrt{2}}\left(|e:+\rangle \otimes |p:-\rangle - |e:-\rangle \otimes |p:+\rangle\right) \tag{11.12}$$

If we measure the component $\hat{S}_{e\alpha}$ of the electron spin along the direction $\hat{u}_\alpha = \cos\alpha\hat{u}_z + \sin\alpha\hat{u}_x$, we find the following results and corresponding probabilities: there are two possible values

$$+\frac{\hbar}{2} \Leftrightarrow \textit{projector} \quad |e:+\alpha\rangle\langle e:+\alpha| \otimes \hat{I}_p$$
$$-\frac{\hbar}{2} \Leftrightarrow \textit{projector} \quad |e:-\alpha\rangle\langle e:-\alpha| \otimes \hat{I}_p$$

with probabilities

$$P_+(\alpha) = \left(|\langle e:+\alpha \mid e:+\rangle|^2 + |\langle e:+\alpha \mid e:-\rangle|^2\right) = \frac{1}{2}$$
$$P_-(\alpha) = \left(|\langle e:-\alpha \mid e:+\rangle|^2 + |\langle e:-\alpha \mid e:-\rangle|^2\right) = \frac{1}{2}$$

This result is a consequence of the rotational invariance of the singlet state.

Now suppose the result of this measurement is $+\hbar/2$ and then later on, one measures the component $\hat{S}_{p\beta}$ of the proton spin along the direction $\hat{u}_\beta = \cos\beta\hat{u}_z + \sin\beta\hat{u}_x$.

Since the electron spin is measured to be $+\hbar/2$, the state after that measurement is

$$\langle e:+\alpha \mid \psi_s\rangle |e:+\alpha\rangle = \frac{1}{\sqrt{2}}(\langle e:+\alpha \mid e:+\rangle |e:+\alpha\rangle \otimes |p:-\rangle - \langle e:+\alpha \mid e:-\rangle |e:+\alpha\rangle \otimes |p:+\rangle)$$
$$\Rightarrow \cos\frac{\alpha}{2}|e:+\alpha\rangle \otimes |p:-\rangle - \sin\frac{\alpha}{2}|e:+\alpha\rangle \otimes |p:+\rangle \tag{11.13}$$

The probabilities for the two possible results of measurement of the proton spin, $\pm\hbar/2$, are

$$P_+(\beta) = \sin^2\frac{\alpha-\beta}{2} \quad , \quad P_-(\beta) = \cos^2\frac{\alpha-\beta}{2} \tag{11.14}$$

What would have happened if we had measured the proton spin first?

If the proton had been measured first then we would have

$$P_+(\beta) = \frac{1}{2} \quad , \quad P_-(\beta) = \frac{1}{2} \tag{11.15}$$

as we found for the electron when the electron was measured first.

The fact that the measurement on the electron affects the probabilities of the results of a measurement on the proton, although the two particles are spatially separated, is in contradiction to Einstein's assertion or belief that *the real states of two spatially separated objects must be independent of one another.* This is the starting point of the EPR paradox. Quantum mechanics is *not a local theory* as far as measurements are concerned.

Note, however, that this non-locality does not allow the instantaneous transmission of information, From a measurement of the proton spin, one can not determine whether the electron spin has been previously measured. It is only when, for a series of experiments, the results of the measurements on the electron and the proton are later compared, that one can find this non-local character of quantum mechanics.

We now recalculate the expectations values $\left\langle \hat{S}_{e\alpha} \right\rangle$ and $\left\langle \hat{S}_{p\beta} \right\rangle$in the singlet state. We get (using the same process as above) $\left\langle \hat{S}_{e\alpha} \right\rangle = 0 = \left\langle \hat{S}_{p\beta} \right\rangle$. This is so because one does not worry about the other variable.

Finally, we can calculate the correlation coefficient in the singlet state. We have, since the spins are correlated now, that

$$\left\langle \hat{S}_{e\alpha} \otimes \hat{S}_{p\beta} \right\rangle = \frac{\hbar^2}{4} \left(\sin^2 \frac{\alpha - \beta}{2} - \cos^2 \frac{\alpha - \beta}{2} \right) = -\frac{\hbar^2}{4} \cos(\alpha - \beta) \tag{11.16}$$

and therefore

$$E(\alpha, \beta) = \frac{-\frac{\hbar^2}{4} \cos(\alpha - \beta) + 0}{\frac{\hbar^2}{4}} = -\cos(\alpha - \beta) \tag{11.17}$$

in the singlet state.

11.1.3 A Simple Hidden Variable Model

For Einstein and several other physicists, the solution to the *paradox* uncovered above comes from the fact that the states of quantum mechanics, in particular the singlet state above, provide and incomplete description of reality. A *complete* theory (for predicting spin measurements, in the present case) should incorporate additional variables or parameters, whose knowledge would render measurements independent for two spatially separated objects. However, present experiments cannot determine the values of these parameters, which are therefore called *hidden variables*. The experimental result should then consist of some averaging over these unknown parameters.

In the case of interest, a very simplified example of such a theory is the following. We assume that, after each dissociation, the system is in a factorized state $|e : +\phi\rangle \otimes |p : -\phi\rangle$, but that the direction ϕ varies from one event to another. In this case ϕ is the hidden variable. We assume that all directions of ϕ are equally probable, i.e., the probability density that the decay occurs with direction ϕ is uniform and equal to $1/2\pi$.

Since we are ignorant of the value of ϕ, the expectation value of an observable $\hat{A}$ is now *defined* to be

$$\left\langle \hat{A} \right\rangle = \frac{1}{2\pi} \int_0^{2\pi} \langle e : +\phi| \otimes \langle p : -\phi| \, \hat{A} \, |e : +\phi\rangle \otimes |p : -\phi\rangle \tag{11.18}$$

Let us now use this new definition of the expectation value to investigate the correlation coefficient. We have from earlier

$$E(\alpha,\beta) = \frac{\left\langle \hat{S}_{e\alpha} \otimes \hat{S}_{p\beta} \right\rangle - \left\langle \hat{S}_{e\alpha} \right\rangle \left\langle \hat{S}_{p\beta} \right\rangle}{\left(\left\langle \hat{S}^2_{e\alpha} \right\rangle \left\langle \hat{S}^2_{p\beta} \right\rangle \right)^{1/2}}$$

Using our earlier results we have

$$\left\langle \hat{S}_{e\alpha} \right\rangle = \frac{\hbar^2}{4} \int \cos(\phi - \alpha) \frac{d\phi}{2\pi} = 0 \tag{11.19}$$

and similarly $\left\langle \hat{S}_{p\beta} \right\rangle = 0$. We also have

$$\left\langle \hat{S}_{e\alpha} \otimes \hat{S}_{p\beta} \right\rangle = -\frac{\hbar^2}{4} \int \cos(\phi - \alpha) \cos(\phi - \beta) \frac{d\phi}{2\pi} = -\frac{\hbar^2}{8} \cos(\alpha - \beta) \tag{11.20}$$

Therefore, in this simple hidden variable model

$$E(\alpha,\beta) = -\frac{1}{2} \cos(\alpha - \beta) \tag{11.21}$$

In such a model, one finds a non-vanishing correlation coefficient, which is an interesting observation. Even more interesting is that the correlation is smaller than the prediction of quantum mechanics by a factor of 2.

The first precise experimental tests of hidden variable descriptions versus quantum mechanics have been performed on correlated pairs of photons emitted in an atomic cascade. Although, we are not dealing with spin−1/2 particles in this case (see discussion later in this chapter), the physical content is basically the same as in this case. As an example Figure 11.1 below presents the experimental results of Aspect, et al,

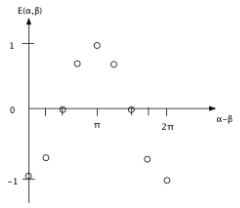


Figure 11.1: Data from Aspect, etal

It gives the variation of $E(\alpha,\beta)$ as a function of the difference $\alpha - \beta$, which is found experimentally to be the only relevant quantity, i.e., the results do not depend in any way on α or β separately! The circles indicate the size of experimental errors.

The experimental points agree with the predictions of quantum mechanics and clearly disagree, therefore, with the predictions of this particular hidden variables theory.

11.2 Bell's Theorem and Experimental Results

As proved by Bell in 1965, the disagreement between the predictions of quantum mechanics and those of hidden variable theories is actually very general when one considers correlation measurements on entangled states.

We can, however, show that the correlation results for hidden variable theories are constrained by what is known as Bell's inequality, which is violated by quantum mechanics.

Consider a hidden variable theory, whose results consists of two functions $A(\lambda, \hat{u}_\alpha)$ and $B(\lambda, \hat{u}_\beta)$ giving respectively the results of the electron and proton spin measurements. Each of these two functions takes only two values $+\hbar/2$ and $-\hbar/2$. It depends on the value of the hidden variable λ for the considered electron-proton pair. The nature of the hidden variable need not be further specified for this discussion. The result A of course depends on the axis $\hat{u}_\alpha$ chosen for the measurement of the electron spin, but it *does* not depend on the axis $\hat{u}_\beta$. Similarly B does not depend on $\hat{u}_\alpha$. This *locality* hypothesis is essential for the following discussion.

Note that we assume here that the hidden variable theory reproduces the one operator averages found for the singlet state:

$$\left\langle \hat{S}_{e\alpha} \right\rangle = \int P(\lambda) A(\lambda, \hat{u}_\alpha) d\lambda = 0 \tag{11.22}$$

$$\left\langle \hat{S}_{p\beta} \right\rangle = \int P(\lambda) B(\lambda, \hat{u}_\beta) d\lambda = 0 \tag{11.23}$$

If this was not the case, such a hidden variable theory should clearly be rejected since it would not reproduce a well-established experimental result.

Let us now consider the quantity

$$A(\lambda, \hat{u}_\alpha)B(\lambda, \hat{u}_\beta) + A(\lambda, \hat{u}_\alpha)B(\lambda, \hat{u}'_\beta) + A(\lambda, \hat{u}'_\alpha)B(\lambda, \hat{u}'_\beta) - A(\lambda, \hat{u}'_\alpha)B(\lambda, \hat{u}_\beta) \tag{11.24}$$

for any set $\hat{u}_\alpha, \hat{u}_\beta, \hat{u}'_\alpha, \hat{u}'_\beta$. We can rewrite this as

$$A(\lambda, \hat{u}_\alpha)(B(\lambda, \hat{u}_\beta) + B(\lambda, \hat{u}'_\beta)) + A(\lambda, \hat{u}'_\alpha)(B(\lambda, \hat{u}'_\beta) - B(\lambda, \hat{u}_\beta)) \tag{11.25}$$

Now the two quantities $B(\lambda, \hat{u}_\beta)$ and $B(\lambda, \hat{u}'_\beta)$ can take on only two values $\pm\hbar/2$. Therefore, one has either

$$B(\lambda, \hat{u}_\beta) + B(\lambda, \hat{u}'_\beta) = \pm\hbar \quad , \quad B(\lambda, \hat{u}'_\beta) - B(\lambda, \hat{u}_\beta) = 0 \tag{11.26}$$

or

$$B(\lambda, \hat{u}_\beta) + B(\lambda, \hat{u}'_\beta) = 0 \quad , \quad B(\lambda, \hat{u}'_\beta) - B(\lambda, \hat{u}_\beta) = \pm\hbar \tag{11.27}$$

Therefore, since $|A(\lambda, \hat{u}_\alpha)| = |A(\lambda, \hat{u}'_\alpha)| = \hbar/2$, we have the result

$$A(\lambda, \hat{u}_\alpha)(B(\lambda, \hat{u}_\beta) + B(\lambda, \hat{u}'_\beta)) + A(\lambda, \hat{u}'_\alpha)(B(\lambda, \hat{u}'_\beta) - B(\lambda, \hat{u}_\beta)) = \pm\hbar^2/2 \tag{11.28}$$

We then define the quantity S as

$$S = E(\alpha,\beta) + E(\alpha,\beta') + E(\alpha',\beta') - E(\alpha',\beta) \tag{11.29}$$

and we get

$$\int P(\lambda)d\lambda \left[\begin{array}{c} A(\lambda,\hat{u}_\alpha)B(\lambda,\hat{u}_\beta) + A(\lambda,\hat{u}_\alpha)B(\lambda,\hat{u}'_\beta) \\ +A(\lambda,\hat{u}'_\alpha)B(\lambda,\hat{u}'_\beta) - A(\lambda,\hat{u}'_\alpha)B(\lambda,\hat{u}_\beta) \end{array} \right]$$
$$= \frac{\hbar^2}{4}\left[E(\alpha,\beta) + E(\alpha,\beta') + E(\alpha',\beta') - E(\alpha',\beta)\right]$$
$$= \pm\frac{\hbar^2}{2}\int P(\lambda)d\lambda = \pm\frac{\hbar^2}{2}$$

or $|S| \leq 2$, which is Bell's inequality.

Now let us consider a special case $\alpha - \beta = \beta' - \alpha = \alpha' - \beta' = \pi/4$. The quantum mechanical result for S is

$$S_Q = -\cos(\alpha-\beta) - \cos(\alpha-\beta') - \cos(\alpha'-\beta') + \cos(\alpha'-\beta) \tag{11.30}$$

If we set $\theta_1 = \alpha - \beta$, $\theta_2 = \beta' - \alpha$, $\theta_3 = \alpha' - \beta'$, we can look for the extrema of

$$f(\theta_1,\theta_2,\theta_3) = \cos(\theta_1+\theta_2+\theta_3) - \cos\theta_1 - \cos\theta_2 - \cos\theta_3 \tag{11.31}$$

The extrema correspond to $\theta_1 = \theta_2 = \theta_3$ and $\sin\theta_1 = \sin 3\theta_1$ whose solutions between 0 and π are $\theta_1 = 0,\ \pi/4,\ 3\pi/4,\ \pi$.

Defining the function $g(\theta_1) = -3\cos\theta_1 + \cos 3\theta_1$ we have

$$g(0) = -2\ ,\ g(\pi/4) = -2\sqrt{2}\ ,\ g(3\pi/4) = 2\sqrt{2}\ ,\ g(\pi) = 2 \tag{11.32}$$

The plot in Figure 11.2 below shows $g(\theta)$:

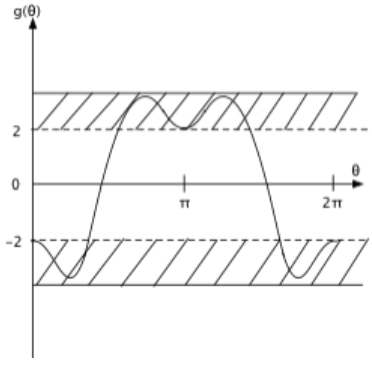


Figure 11.2: Bell Function

The shaded areas correspond to results which *cannot* be explained by hidden variable theories. This system therefore constitutes a test of the predictions of quantum mechanics versus any local hidden variable theory.

11.3 The EPR(Einstein-Podolsky-Rosen) Argument-Quick Overview

The nondeterministic character of quantum mechanics is very disturbing for the classical physicist. Hence, there were repeated attempts to replace quantum theory by a statistical theory.

According to these theories, there exist hidden variables whose values prescribe the values of all observables for any particular object, except that the hidden variables are unknown to the experimenter, thus yielding the probabilistic character of the theory. The probabilistic character of quantum mechanics would then be quite analogous to that of classical statistical mechanics, where one can imagine that the motion of all the particles is, in principle, known.

For example, let us consider a particle of spin = 1/2 in an eigenstate of S_x with eigenvalue $\hbar/2$. According to quantum mechanics, the $z-$component is not fixed. If one measures it for a very large number of such particles, one finds $\hbar/2$ 50% of the time and $-\hbar/2$ 50% of the time. According to the idea of hidden variables, for each particle, parameters unknown to us would determine whether $\hbar/2$ or $-\hbar/2$ results. These hidden variables would prescribe $\pm\hbar/2$ each 50% of the time.

By means of a number of thought experiments, Einstein attempted to demonstrate the incompleteness of the quantum mechanical description and to get around the indeterminism and the uncertainty relation. Each of these arguments was refuted, in turn, by Bohr.

An argument - sometimes referred to as a paradox - due to Einstein, Podolsky and Rosen (EPR) , played a pivotal role in the discussion of indeterminism and the existence of hidden variables; we consider the argument as reformulated by David Bohm.

Let two spin = 1/2 particles in the singlet state

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle\right) \tag{11.33}$$

be emitted from a source an then move apart in space. Even if the two particles are separated by an arbitrarily large distance and can no longer communicate with one another, one finds the following correlations in this state during a measurement of the one particle spin states:

If one measures the $z-$component of the spin and finds particle 1 spin up, particle 2 is spin down. If one finds particle 1 spin down, particle 2 has spin up.

If, instead, one measures S_x, then $+\hbar/2$ for particle 1 implies $-\hbar/2$ for particle 2, etc.

This expresses the *nonlocality* of quantum theory.

The experiment on particle 1 influences the result of the experiment on particle 2, although they are widely separated. The nonlocality is a consequence of the existence of correlated many-particle states such as the direct product

$$|\uparrow\rangle\,|\downarrow\rangle \tag{11.34}$$

and the fact that one can linear superimpose such states.

The nonlocality of quantum mechanics does not lead to contradictions with relativity theory. Although a measurement of a spin component of particle 1 immediately reveals the value of that component for particle 2, no information can be transmitted in this way. Since particle 1 takes values $\pm\hbar/2$ each 50% of the time, this remains true for particle 2, even after the measurement of particle 1. Only by a subsequent(slow) comparison of the results is it possible to verify the correlation.

Einstein, Podolsky and Rosen gave the following argument in favor of hidden parameters in conjunction with the EPR thought experiment.

By the measurement of S_z or S_x of particle 1, the values of S_z or S_x of particle 2 are known. Because of the separation of the particles, there was no influence on particle 2, and therefore the values of S_z, S_x etc, must have been fixed before the experiment. Thus, there must be a more complete theory with hidden variables.

In the EPR argument, the predictions of the quantum states

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle\,|\downarrow\rangle - |\downarrow\rangle\,|\uparrow\rangle\right)$$

are used, but the inherent nonlocality is denied.

In the remainder of our discussion, we will consider local hidden variables. These would predetermine which value each of the components of $\vec{S}$ of particle 1 has and likewise for particle 2. Each of the particles would carry this information independently of the other.

11.3.1 The Bell Inequality again

We now show again that such local hidden variables lead to predictions different from those of quantum mechanics. We then compare the predictions with experiment.

We consider a correlation experiment in which a particle of total spin = 0 decays into two particles each with spin = 1/2. At a sufficiently large distance from the source, a rotatable polarizer and a detector are set up for each particle as shown below, so that the particles can be detected and we can investigate whether any correlation in the spin orientations exists. The setup is shown in Figure 11.3 below.

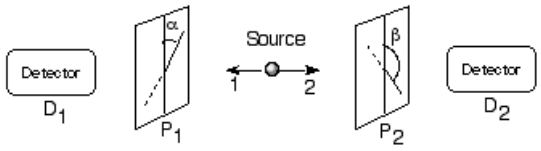


Figure 11.3: Bell-EPR Experiment

Polarizer 1 with angular setting α only lets particle 1 through if its spin in the direction $\hat{n}_\alpha$ has the value $+\hbar/2$ and polarizer 2 with angular setting β only lets particle 2 through if its spin in the direction $\hat{n}_\beta$ has the value $+\hbar/2$. The particles are counted by detectors 1 and 2. If they respond, then the spin is positive, otherwise it is negative.

We consider the correlation between various angular settings of the polarization experiment.

A measure of the correlation is $N(\alpha;\beta)$, defined as the relative number of experiments resulting in particle 1 at angle α being positive *and* particle 2 at angle β being positive.

Using the spin projection operator

$$P_\theta = \frac{1}{2}\left(1 + \vec{\sigma}\cdot\hat{n}_\theta\right) \tag{11.35}$$

quantum mechanics gives

$$\begin{aligned} N(\alpha;\beta) &= \langle 0,0|\,\frac{1}{2}\left(1+\vec{\sigma}_1\cdot\hat{n}_\alpha\right)\frac{1}{2}\left(1+\vec{\sigma}_2\cdot\hat{n}_\beta\right)|0,0\rangle \\ &= \langle 0,0|\,\frac{1}{2}\left(1+\vec{\sigma}_1\cdot\hat{n}_\alpha\right)\frac{1}{2}\left(1-\vec{\sigma}_1\cdot\hat{n}_\beta\right)|0,0\rangle \\ &= \frac{1}{4}\left(1-\hat{n}_\alpha\cdot\hat{n}_\beta\right) \end{aligned} \tag{11.36}$$

since $\langle 0,0|\,\vec{\sigma}_1\,|0,0\rangle = 0$ in the singlet state. For coplanar detectors this reduces to

$$N(\alpha;\beta) = \frac{1}{2}\sin^2\frac{\beta-\alpha}{2} \tag{11.37}$$

If hidden variables were really present, we could represent $N(\alpha;\beta)$ by the following sum

$$N(\alpha;\beta) = N(\alpha\gamma;\beta) + N(\alpha;\gamma\beta) \tag{11.38}$$

Here, $N(\alpha\gamma;\beta)$ is the relative number of particle pairs in which particle 1 has positive spin at angles α and γ and negative spin at β, while $N(\alpha;\gamma\beta)$ is the relative number of particle pairs in which particle 1 has negative spin at $|gamma$ instead. In theories with hidden variables, all of these quantities are assumed to be known.

Now one has $N(\alpha\gamma;\beta) \le N(\gamma;\beta)$ since $N(\gamma;\beta) = N(\alpha\gamma;\beta) + N(\gamma;\beta\alpha)$ and both quantities on the right-hand side of the equation are nonnegative. Similarly, $N(\alpha;\gamma\beta) \le N(\alpha;\gamma)$. Thus,

$$N(\alpha;\beta) \le N(\alpha;\gamma) + N(\gamma;\beta) \tag{11.39}$$

This is a simple version of the *Bell inequality*.

Remarks

1. In experiments one often works with the correlation defined by

   $$P(\alpha;\beta) = \langle 0,0|\,(\vec{\sigma}_1\cdot\hat{n}_\alpha)\,(\vec{\sigma}_2\cdot\hat{n}_\beta)\,|0,0\rangle = 4N(\alpha;\beta) - 1 \tag{11.40}$$

   instead of $N(\alpha;\beta)$ itself. Using

   $$N(\alpha;\beta) = \frac{1}{2}\sin^2\frac{\beta-\alpha}{2} \tag{11.41}$$

   we get

   $$P(\alpha-\beta) \equiv P(\alpha;\beta) = -\cos(\alpha-\beta) \tag{11.42}$$

   and Bell's inequality becomes

   $$P(\alpha;\beta) - 1 \le P(\alpha;\gamma) + P(\gamma;\beta) \tag{11.43}$$

2. The limit prescribed by the Bell inequality can be determined as follows. In

   $$N(\alpha;\beta) \le N(\alpha;\gamma) + N(\gamma;\beta) \tag{11.44}$$

   we substitute for α,β,γ the values $0,\pi,\pi/2$ respectively to obtain

   $$N(0;\pi) \le N(0;\pi/2) + N(\pi/2;\pi) \tag{11.45}$$

   In the singlet state,

   $$N(0;\pi) = \frac{1}{2} \quad , \quad N(0;\pi/2) = N(\pi/2;\pi) \tag{11.46}$$

so that

$$N(0;\pi/2) \geq \frac{1}{4} \tag{11.47}$$

Other values can be obtained by different combinations of angles.

Finally, we contrast the consequences of the Bell inequality with quantum mechanics and compare with experiments.

To this end, we compute $N(\alpha;\beta)$, $N(\alpha;\gamma)$ and $N(\gamma;\beta)$ for the three angles $\alpha = 0°$, $\gamma = 45°$ and $\beta = 90°$ using

$$N(\alpha;\beta) = \frac{1}{2}\sin^2\frac{\beta-\alpha}{2} \tag{11.48}$$

to get

$$\frac{1}{2}\sin^2 45° \leq 2\sin^2 22.5° \tag{11.49}$$

or

$$0.5 \leq 0.29 \tag{11.50}$$

which is clearly not true! Therefore, quantum mechanics and hidden variables are incompatible.

The comparison of quantum mechanics and the Bell inequality is shown in Figure 11.4 below, which gives the correlation $P(\theta) = P(\theta;0)$ according to quantum mechanics and the Bell inequality.

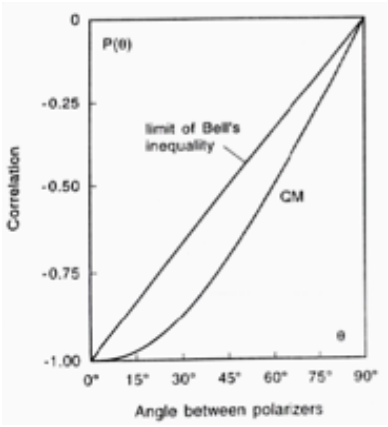


Figure 11.4: Bell-EPR Experiment Theoretical Predictions

The experimental demonstration of the violation of the Bell inequality by Lamehl-Rachti for protons and Aspect for photons is shown in Figure 11.5 below which gives the experimental results on the spin correlation of proton pairs

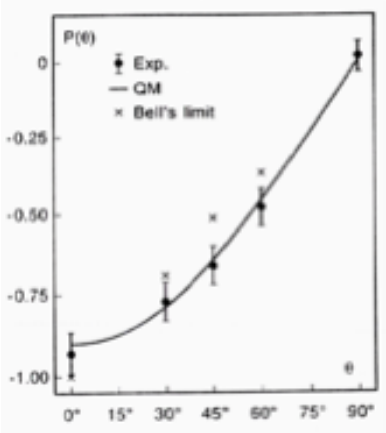


Figure 11.5: Bell-EPR Experiment - Data

Clearly, quantum mechanics is correct. This means that any theory that has the same probabilistic predictions as quantum mechanics must be *nonlocal*.

11.4 EPR and Bell - The Details

Let us first rethink some quantum mechanical ideas in a context needed for this discussion. This review will hopefully reinforce the ideas you have learned so far.

11.4.1 Single-Photon Interference

All good discussions on quantum mechanics present a long an interesting analysis of the double slit experiment. The crux of the discussion comes when *the light intensity is reduced sufficiently for photons to be considered as presenting themselves at the entry slit one by one*. For a long time this point was very contentious, because correlations between two successive photons cannot be ruled out a priori.

Since 1985, however, the situation has changed. An experiment was performed by Grangier, Roger and Aspect. It was an interference experiment with only a single photon. They used a light source devised for an EPR experiment which guarantees that photons arrive at the entry slit singly.

The experiment is difficult to do in practice, but is very simple in principle and

it provides an excellent experimental introduction to the concepts of quantum mechanics.

The light source is a beam of calcium atoms, excited by two focused laser beams having wavelengths $\lambda' = 406\,nm$ and $\lambda'' = 581\,nm$ respectively. Two-photon excitation produces a state having the quantum number $J = 0$. When it decays, this state emits two monochromatic photons having the wavelengths $\lambda_1 = 551.3\,nm$ and $\lambda_2 = 422.7\,nm$ respectively, in a cascade of two electronic transitions from the initial $J = 0$ level to the final $J = 0$ state, passing through an intermediate $J = 1$ state, as shown in Figure 11.6 below

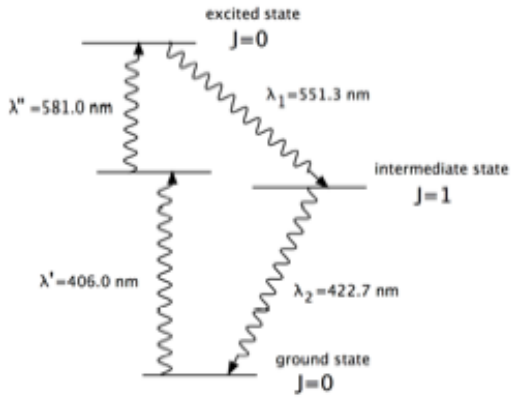


Figure 11.6: Calcium based light source - Energy Levels

The mean lifetime of the intermediate state is $4.7\,ns$. To simplify the terminology, we shall call the $\lambda_1 = 551.3\,nm$ light green, and the $\lambda_2 = 422.7\,nm$ light violet.

Next we describe the experiment, exhibiting its three stages which reveal the complications of the apparatus in progressively greater detail (Figures 11.7-11.9 below).

1. The first stage is a trivial check that the apparatus is working properly; nevertheless it is already very instructive (Figure 11.7 below).

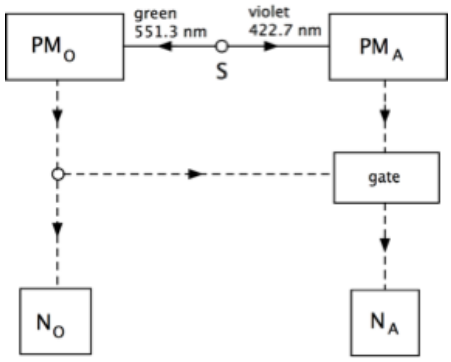


Figure 11.7: Single Photon Interference Experiment - Stage 1

Figure 11.7 shows interference with a single photon (first stage). In the sketch, solid lines are optical paths and dashed lines are electrical connections.

On either side of the source S one positions two photomultiplier tubes PM_O and PM_A. These are very sensitive, and can detect the arrival of a single photon. Detection proceeds through photoelectric absorption, followed by amplification which produces an electric signal proportional to the energy of the incident photon. The associated electronic logic circuits can identify the photons absorbed by each detector: the channel PM_O responds only to green light, and the channel PM_A responds only to violet light. The electronic gate is opened (for $9\,ns$ - this is twice the mean lifetime and corresponds to an 85% probability that the photon has been emitted) when green light is detected by PM_O. If, while the gate is open, violet light is emitted by the same atom towards(not all of the violet photons go towards the source) PM_A, then PM_A detects this photon, producing a signal that passes through the gate and is counted in N_A. The counter N_O registers the number of green photons detected by PM_O. It turns out that $N_A \ll N_O$. As the observation period becomes very long(approximately 5 hours), the ratio N_A/N_O tends to a limit that is characteristic of the apparatus. It represents the probability of detecting a violet photon in PM_A during the $9\,ns$ following the detection of a green photon by PM_O.

The purpose of this arrangement is to use a green photon in order to open a $9\,ns$ time window, in which to detect a violet photon emitted by the same atom. As we shall see, there is only an extremely small probability of detecting through the same window another violet photon emitted by a different atom.

We will assume that a second observer is in the lab. This observer always

feels compelled to present what he thinks are *simple-minded truths* using ordinary words. We will called this second observer Albert. Albert, as we shall see, has a tendency to use, one after another, the three phrases, *I observe*, *I conclude*, and *I envisage*. Consulted about the above experiment, Albert states, with much confidence,

I observe that the photomultiplier PM_A detects violet light when the source S is on, and that it ceases to detect anything when the source is off. I conclude that the violet light is emitted by S, and that it travelled from S to PM_A.

I observe that energy is transferred between the light and the photomultiplier PM_A always in the same amount, which I will call a quantum.

I envisage the quanta as particles, emitted by the source, propagating freely from S to PM_A, and absorbed by the detector. I shall call this quanta photons.

Albert stops talking at this point.

2. The second stage of the experiment introduces the concept of individual photons in Figure 11.8 below which is interference with a single photon (second stage).

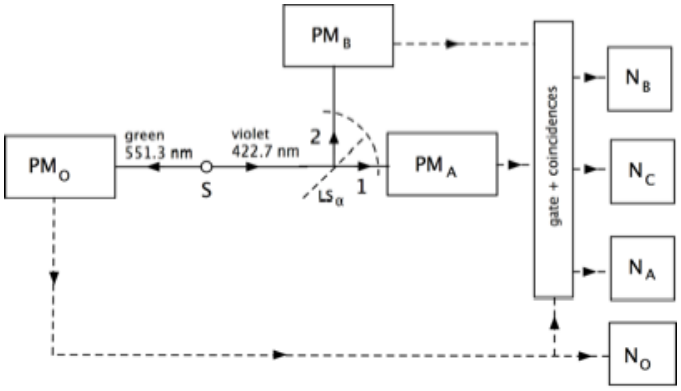


Figure 11.8: Single Photon Interference Experiment - Stage 2

Across the path of the violet light one places a half-silvered mirror LS_α, which splits the primary beam into two secondary beams(equal intensity), one transmitted and detected by PM_A, the other reflected and detected by PM_B. As in the first stage, the gate is opened for $9\,ns$, by PM_O. While it is open, one registers detection by either PM_A (counted as N_A) or by PM_B (counted as N_B) or by both, which we call a coincidence (counted as N_C). The experiment runs for 5 hours again and yields the following

results:

(a) The counts N_A and N_B are both of the order of 10^5. By contrast, N_C is much smaller, being equal to 9.

(b) The sequence of counts from PM_A is random in time, as is the sequence of counts from PM_B.

(c) The very low value of N_C shows that counts in PM_A and PM_B are mutually exclusive (do not occur at same time).

The experimenters analyze the value of N_C in depth; their reasoning can be outlined as follows:

(a) Suppose two different atoms each emit a violet photon, one being transmitted to PM_A and the other reflected to PM_B with both arriving during the $9\,ns$ opening of the gate; then the circuitry records a coincidence. In the regime under study, and for a run of 5 hours, quantum theory predicts that the number of coincidences should be $N_C = 9$. The fact that this number is so small means that, in practice, any given single photon is either transmitted or reflected.

(b) If light is considered as a wave, split into two by LS_α and condensed into quanta on reaching PM_A and PM_B, then one would expect the photon counts to be correlated in time, which would entail $N_C \gg 9$. Classically speaking this would mean that we cannot have a transmitted wave without a reflected wave.

(c) Experiment yields $N_C = 9$; this quantum result differs from the classical value by 13 standard deviations; hence the discrepancy is very firmly established, and allows us to assert that we are indeed dealing with a source of individual photons.

Albert leaves such logical thinking to professionals. Once he notes that N_C is very small, he is quite prepared to treat it as if it were zero. He therefore says *I observe that light travels from the source to PM_A or to PM_B, because detection ceases when the source is switched off.*

I observe the counts N_A and N_B correspond to a game of heads or tails, in that the two possibilities are mutually exclusive, and that the counts are random.

I observe that the optical paths 1 and 2 are distinguishable, because the experiment allows me to ascertain, for each quantum, whether it has travelled path 1 (detection by PM_A) or path 2 (detection by PM_B).

I envisage that, on arrival at the half-silvered mirror, each photon from the source is directed at random either along path 1 or along path 2; and I assert that it is the nature of photons to play heads or tails..

Digression: The Mach-Zender Interferometer and Quantum Interference

(a) The next experiment uses a Mach-Zender interferometer so let us see how it operates.

Background information: Consider a single photon incident on a 50-50 beam splitter (that is, a partially transmitting, partially reflecting mirror, with equal coefficients). Whereas classical electromagnetic energy divides equally, the photon is indivisible. That is, if a photon-counting detector is placed at each of the output ports (see Figure 11.9 below), only *one* of them clicks. Which one clicks is completely random (that is, we have no better guess for one over the other).

Figure 11.9: 50-50 Beam Splitter

The input-output transformation of the waves incident on 50-50 beam splitters and perfectly reflecting mirrors are shown in Figure 11.10 below.

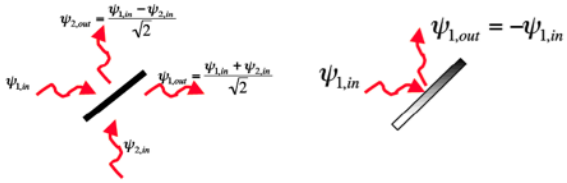


Figure 11.10: Input/output Transformations

We can easily show that with these rules, there is a 50-50 chance of either of the detectors shown in the first figure above to click. According to the rules given in the figure

$$\psi_{1,out} = \frac{1}{\sqrt{2}}\psi_{in} \quad , \quad \psi_{2,out} = \frac{1}{\sqrt{2}}\psi_{in} \tag{11.51}$$

since nothing enters port #2.

By our probability postulate the probability to find a photon at po-

sition 1 or 2 is

$$\left.\begin{aligned} P_{1,out} &= \int |\psi_{1,out}|^2 dx = \tfrac{1}{2}\int |\psi_{in}|^2 dx = \tfrac{1}{2} \\ P_{2,out} &= \int |\psi_{2,out}|^2 dx = \tfrac{1}{2}\int |\psi_{in}|^2 dx = \tfrac{1}{2} \end{aligned}\right\} \Rightarrow 50-50\%\ chance$$

Note: As we see from the experimental discussion below, the photon is found at one detector or the other, never both. The photon is indivisible. This contrasts with classical waves where half of the intensity goes one way and half the other; an antenna would also receive energy. We interpret this as the mean value of a large number of photons.

(b) Now we set up a Mach-Zender interferometer(shown in Figure 11.11 below):

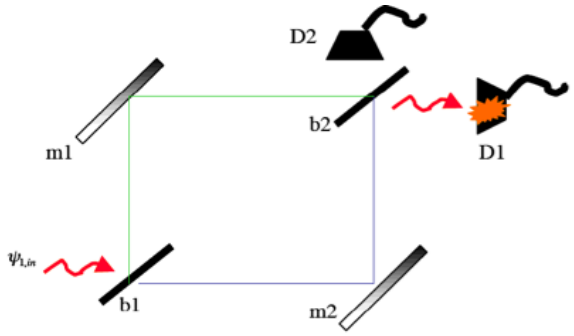


Figure 11.11: Mach-Zehnder Interferometer

?The wave is split at beam-splitter b1, where it travels either path b1-m1-b2(call it the green path) or the path b1-m2-b2 (call it the blue path). Mirrors are then used to recombine the beams on a second beam splitter, b2. Detectors D1 and D2 are placed at the two output ports of b2.

Assuming the paths are perfectly balanced (that is equal length), we can show that the probability for detector D1 to click is 100% - no randomness!

To find the wavefunctions impinging on detectors D1 and D2 let us apply the transformation rules sequentially.

(1) Beamsplitter #1

$$\psi_{1,out} = \frac{1}{\sqrt{2}}\psi_{in} \quad , \quad \psi_{2,out} = \frac{1}{\sqrt{2}}\psi_{in}$$

(2) Propagation a distance $L/2$ along each path mean that the phase of the wavefunction changes by $e^{ikL/2}$ so that the wavefunctions

are

$$\psi_{1,at-mirror} = \frac{1}{\sqrt{2}} e^{ikL/2}\psi_{in} \quad , \quad \psi_{2,at-mirror} = \frac{1}{\sqrt{2}} e^{ikL/2}\psi_{in}$$

(3) Reflection off mirrors means wavefunctions become

$$\psi_{1,after-mirror} = -\frac{1}{\sqrt{2}} e^{ikL/2}\psi_{in} \tag{11.52}$$
$$\psi_{2,after-mirror} = -\frac{1}{\sqrt{2}} e^{ikL/2}\psi_{in}$$

(4) Propagation a distance $L/2$ along each path mean that the phase of the wavefunction changes by $e^{ikL/2}$ so that the wavefunctions are

$$\psi_{1,at-b2} = \frac{1}{\sqrt{2}} e^{ikL}\psi_{in} \quad , \quad \psi_{2,at-b2} = \frac{1}{\sqrt{2}} e^{ikL}\psi_{in}$$

(5) After beamsplitter #2

$$\psi_{out,1} = \frac{\psi_{1,at-b2} + \psi_{2,at-b2}}{\sqrt{2}} = e^{ikL}\psi_{in}$$
$$\psi_{out,2} = \frac{\psi_{1,at-b2} - \psi_{2,at-b2}}{\sqrt{2}} = 0$$

Therefore,

$$P_{1,out} = \int |\psi_{out,1}|^2 dx = \int |\psi_{in}|^2 dx = 1$$
$$P_{2,out} = \int |\psi_{out,2}|^2 dx = 0$$

Thus, we have a 100% chance of detector D1 firing and a 0% chance of detector D2 firing. There is no randomness.

(c) Classical logical reasoning would predict a probability for D1 to click given by

$$P_{D1} = P(transmission\ at\ b1|green\ path)P(green\ path) + P(reflection\ at\ b2|blue\ path)P(blue\ path)$$

Now we know that there is a 50-50 probability for the photon to take the blue or green path which implies that

$$P(green) = P(blue) = 1/2$$

Also with the particle incident at b2 along the green path there is a 50% chance of transmission and similarly for reflection of the blue path.

Therefore,

$$P(transmission\ at\ b2|green) = P(reflection\ at\ b2|blue) = 1/2$$

and

$$P_{D1} = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2}$$

so that classical reasoning implies a 50-50 chance of D1 firing, that is, it is completely random!

The quantum case is different because the two paths which lead to detector D1 interfere. For the two paths leading to D1 we have

$$\psi_{total} = \frac{\frac{1}{\sqrt{2}}e^{ikL}\psi_{in} + \frac{1}{\sqrt{2}}e^{ikL}\psi_{in}}{\sqrt{2}}$$
$$P_{D1} = \int |\psi_{total}|^2 dx = P_{D1} = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = 1$$

where the last two terms are the so-called interference terms. Thus, $P_{D1} = 1$. The paths that lead to detector D2 destructively interfere so that $P_{D2} = 0$.

We now ask how would you set up the interferometer so that detector D2 clicked with 100% probability? How about making them click at random?

Leave the basic geometry the same, that is, do not change the direction of the beam splitters or the direction of the incident light.

We now want constructive interference for the paths leading to D2 and destructive interference for D1.

We can achieve this by changing the relative phase of the two paths by moving the mirror so that the path lengths are not the same.

Suppose we move the mirror on the green path (at an angle of 45°) so that the path lengths in the green path are both changed to $L + \Delta L$. We then have

$$\psi_{out,1} = \frac{\frac{1}{\sqrt{2}}e^{ik(L+\Delta L)} + \frac{1}{\sqrt{2}}e^{ikL}}{\sqrt{2}} = \psi_{in}$$
$$\psi_{out,2} = \frac{\frac{1}{\sqrt{2}}e^{ik(L+\Delta L)} - \frac{1}{\sqrt{2}}e^{ikL}}{\sqrt{2}} = \psi_{in}$$

and

$$P_{D1} = \int dx\,|\psi_{2,out}|^2 = \frac{1}{4}\int dx\,|\psi_{in}|^2\left|e^{ikL}\right|^2\left|e^{ik\Delta L}+1\right|^2$$
$$= \frac{1}{4}\left(e^{ik\Delta L}+1\right)\left(e^{ik\Delta L}+1\right)^* = \frac{1}{4}\left(e^{ik\Delta L}+1\right)\left(e^{-ik\Delta L}+1\right)$$
$$= \frac{1}{4}(2+e^{ik\Delta L}+e^{-ik\Delta L}) = \frac{1+\cos(k\Delta L)}{2} = \cos^2\left(\frac{k\Delta L}{2}\right)$$

Similarly we have

$$P_{D2} = \int dx\,|\psi_{1,out}|^2 = \frac{1}{4}\int dx\,|\psi_{in}|^2\left|e^{ikL}\right|^2\left|e^{ik\Delta L}-1\right|^2$$
$$= \frac{1}{4}\left(e^{ik\Delta L}-1\right)\left(e^{ik\Delta L}-1\right)^* = \frac{1}{4}\left(e^{ik\Delta L}-1\right)\left(e^{-ik\Delta L}-1\right)$$
$$= \frac{1}{4}(2-e^{ik\Delta L}-e^{-ik\Delta L}) = \frac{1-\cos(k\Delta L)}{2} = \sin^2\left(\frac{k\Delta L}{2}\right)$$

Therefore, to achieve $P_{D1}=0$ and $P_{D2}=1$ we choose

$$k\Delta L = m\pi \quad (m\ odd) \Rightarrow \Delta L = m\frac{\pi}{k} = m\frac{\lambda}{2}$$

We can make the both random if

$$\cos^2\left(\frac{k\Delta L}{2}\right) = \sin^2\left(\frac{k\Delta L}{2}\right) = \frac{1}{2} \Rightarrow \frac{k\Delta L}{2} = p\frac{\pi}{4} \quad (p\ odd)$$
$$\Delta L = p\frac{\pi}{2k} = p\frac{\lambda}{4}$$

Returning to our discussion.....

3. The third stage then consists of an interference experiment as shown in Figure 11.12 below, which is the interference with a single photon (third stage).

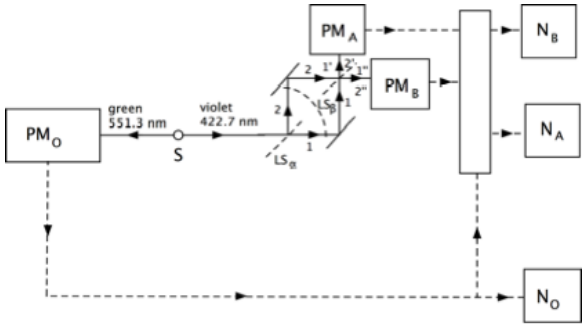


Figure 11.12: Mach-Zehnder Interferometer Inserted

A so-called Mach-Zehnder interferometer is used, allowing one to obtain two interference profiles. The beam of violet light from the source S is split into two by the mirror LS_α. After reflection from two different mirrors, these secondary beams meet on a second half-silvered mirror LS_β. Here, each secondary beam is further split into two; thus one establishes two interference regions, region (1',2') where one places PM_A, and region (1",2") where one places PM_B.

A very high precision piezoelectric system allows one of the mirrors to be displaced so as to vary the path difference between the two arms of the interferometer. In this way one can shift the pattern of interference fringes by regular steps, without moving the detectors PM_A and PM_B; the standard step corresponds to a change of $\lambda/50$ in the difference between the two optical paths.

A sweep, taking 15 sec for each standard step, yields two interference plots corresponding, respectively, to the paths (1',2') and (1",2"); the fringes have good contrast(difference in intensity between maxima and minima), and their visibility

$$(N_{A,\max} - N_{A,\min}) \,/\, (N_{A,\max} + N_{A,\min}) \tag{11.53}$$

was measured as 98% as shown in Figure 11.13 below which gives the two interference plots obtained with the Mach-Zehnder interferometer. Note that the maximum counting rates in PM_A correspond to minima in PM_B, indicating a relative displacement of $\lambda/2$ between the two interference patterns

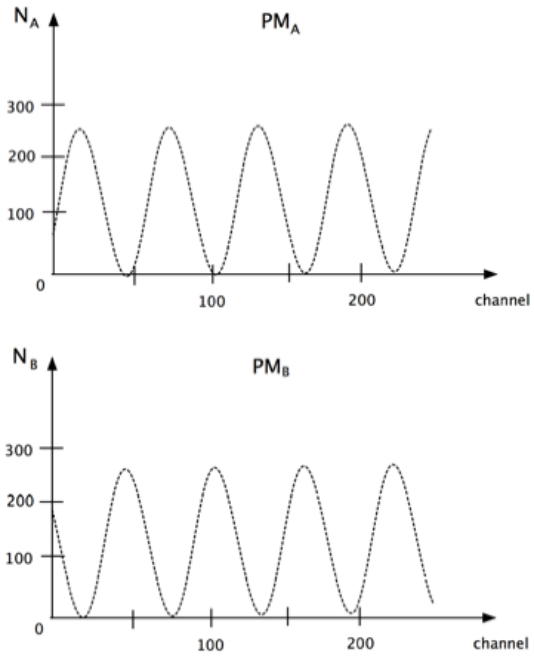


Figure 11.13: Interference Results

If we recall that we are reasoning in terms of photons, and that the photons are being processed individually, then we must admit that the interference does not stem from any interaction between successive photons, but that each photon interferes with itself.

What would Albert have to say? He seems exasperated but is still polite. His statements are brief:

I observe that the optical paths differ in length between LS_α and LS_β, and are then coincident over (1',2') and over (1",2").

In PM_A I observe a process that seems perfectly natural to me, namely

$$light + light \rightarrow light$$

In PM_B I observe a process that I find astounding, namely

$$light + light \rightarrow darkness$$

Such superposition phenomena with light I shall call interference; constructive in PM_A and destructive in PM_B.

In the situation considered before, I envisaged light as consisting of particles called photons, which travelled either along path 1 or along path 2.

In the present situation I want to know for each individual photon which path it has travelled; to this end I should like to ask you to close off path 2, since this will ensure that the photons travel by path 1..

Clearly Albert is perturbed. He awaits the new experimental results with some anxiety.

On closing either path, whether 1 or 2, one observes that all interference phenomena disappear. For instance, instead of a very high count N_A and a very low count N_B, we now obtain essentially equal counts from PM_A and PM_B.

Albert is visibly displeased and now very wary. He then continues with his analysis of the experiment:

I observe that in order to produce interference phenomena it is necessary to have two optical paths of different lengths, both open.

Whenever a photon is detected, I note my inability to ascertain whether the light has travelled by path 1 or by path 2, because I have no means for distinguishing between the two cases.

If I were to suppose that photons travel only along 1, then this would imply that path 2 is irrelevant, which is contrary to what I have observed. Similarly, if I were to suppose that photons travel only along 2, then this would imply that path 1 is irrelevant, which is also contrary to my observations.

If I envisage the source S as emitting particles, then I am forced to conclude that each individual photon travels simultaneously along both paths 1 and 2; but this result contradicts the results of the previous experiment (second stage), which compelled me to envisage that every photon chooses, at random, either path 1 or path 2.

I conclude that the notion of particles is unsuited to explaining interference phenomena.

I shall suppose instead that the source emits a wave; this wave splits into two at LS_α, and the two secondary waves travel one along path 1 and the other along path 2. They produce interference by mutual superposition on LS_β constructively in (1',2') and destructively in (1",2"). At the far end of (1',2') or of (1",2") I envisage each of the waves condensing into particles, which are then detected by the photomultipliers (essentially by PM_A since the contrast is 98% means only very few photons are detected by PM_B

It seems to me that I am beginning to understand the situation. I en-

visage light as having two complementary forms: depending on the kind of experiment that is being done, it can manifest itself either as a wave, or as a particle, but never as both simultaneously and in the same place. Thus, in the experiment where the path followed by the light cannot be ascertained (third stage), light behaves first like a wave, producing interference phenomena; but it behaves like a particle when, afterwards, it is detected through the photoelectric effect. I conclude that light behaves rather strangely, but nevertheless I have the impression that its behavior can be fully described once one has come to terms with the idea of wave-particle duality..

Albert leaves the room slowly, hesitantly, even reluctantly. He might be impressed by all the completeness of all that he has just described or maybe he is worried that more needs to be said.

In fact, something does remain to be said, since the problem of causality remains open. Let us look carefully at the experimental layouts in the second and third stages: we see that they have LS_α in common, and that they differ only beyond some boundary (indicated by the dashed circle downstream from LS_α). We have stated that light behaves like a particle or like a wave depending on whether or not one can ascertain the path it takes through the apparatus; but in the two experiments under consideration, the choice between the alternatives must be decided on LS_α, *before* the light has crossed the crucial boundary, that is, at a stage where nothing can as yet distinguish between the two kinds of apparatus, since they differ only beyond the point of decision. It is as if the light *chose* whether to behave like a wave or like a particle before *knowing* whether the apparatus it will pass through will elicit interference phenomena or the photoelectric effect. Hence the question of causality is indeed opened up with vengeance.

Albert comes back abruptly. He is disconcerted and wearily says:

Originally I supposed that light would behave like a wave or like a particle, depending on the kind of experiment to which it was being subjected.

I observe that the choice must be made on the half-silvered mirror LS_α, before the light reaches that part of the apparatus where the choice is actually implemented; this would imply that the effect precedes the cause.

I know that both waves and particles obey the principle of causality, that is, that cause precedes effect.

I conclude that light is neither wave nor particle; it behaves neither like waves on the sea, nor like projectiles fired from a gun, nor like any other kind of object that I am familiar with.

I must ask you to forget everything I have said about this experiment, which seems to me to be thoroughly mysterious.

Albert leaves, but quickly returns with a contented smile, and his final statement is not without a touch of malice. *I observe in all cases that the photomultipliers register quanta when I switch on the light source.*

I conclude that "something" has travelled from the source to the detector. This "something" is a quantum object, and I shall continue to call it a photon, even though I know that it is neither a wave nor a particle.

I observe that the photon gives rise to interference when one cannot ascertain which path it follows; and that interference disappears when it is possible to ascertain the path.

For each detector, I observe that the quanta it detects are randomly distributed in time.

If I repeat the experiment several times under identical conditions, then I observe that the photon counts registered by each photomultiplier are reproducible in a statistical sense. For example, suppose that in the first and in the second experiments PM_A registers N'_A and N''_A respectively; then one can predict that N''_A has a probability of 0.68 of being in the interval $N'_A \pm (N'_A)^{1/2}$

Thus, these counts enable me to determine experimentally, for any kind of apparatus, the probability that a given detector will detect a quantum, and it is precisely such probabilities that constitute the results of experiments.

I assert that the function of a physical theory is to predict the results of experiments.

What I expect from theoretical physicists is a theory that will enable me to predict, through calculation, the probability that a given detector will detect a photon. This theory will have to take into account the random behavior of the photon, and the absence or presence of interference phenomena depending on whether the paths followed by the light can or cannot be ascertained..

Albert leaves, wishing the physicists well in their future endeavors.

Physicist have indeed worked hard and the much desired theory has indeed come to light, namely, quantum mechanics, as we have seen in our discussions. As we have seen, it applies perfectly not only to photons, but equally well to electrons, protons, neutrons, etc; in fact, it applies to all the particles of microscopic physics. For the last 75 years it has worked

to the general satisfaction of physicists.

Meanwhile, it has produced two very interesting problems of a philosophical nature.

1. Chance as encountered in quantum mechanics lies in the very nature of the coupling between the quantum object and the experimental apparatus. No longer is it chance as a matter of ignorance or incompetence: it is *chance quintessential and unavoidable.*

2. Quantum objects behave quite differently from the familiar objects of our everyday experience: whenever, for pedagogical reasons, one allows an analogy with macroscopic models like waves or particles, one always fails sooner or later, because the analogy is never more than partially valid. Accordingly, the first duty of a physicist is to force her grey cells, that is her concepts and her language, into unreserved compliance with quantum mechanics (as we have been attempting to do); eventually this will lead her to view the actual behavior of microsystems as perfectly normal. As a teacher of physics, our duties are if anything more onerous still, because we must convince the younger generations that quantum mechanics is not a branch of mathematics, but an expression of our best present understanding of physics on the smallest scale; and that, like all physical theories, it is predictive.

In this context, let us review the basic formalism of quantum mechanics.

11.4.2 Basic Formalism

We will introduce the elements of quantum mechanics as axioms. Physicists have devised a new mathematical tool. The transition amplitude from initial to final state, and it is this amplitude that enables one to calculate the needed probabilities.

1. For the experiment where the photon travels from the source S to the detector PM_A (see Figure 11.14(a) below), we write the transition amplitude from S to PM_A as

$$\langle \text{photon arriving at } PM_A \mid \text{photon leaving S} \rangle$$

which we symbolize simply as

$$\langle f_1 \mid i \rangle \quad , \quad \langle f_2 \mid i \rangle$$

In this case there are two probabilities:

$$|\langle f_1 \mid i \rangle|^2 \quad , \quad |\langle f_2 \mid i \rangle|^2$$

The total probability is their sum:

$$|\langle f_1 \mid i \rangle|^2 + |\langle f_2 \mid i \rangle|$$

More generally, we would write

$$|\langle f \mid i\rangle|^2 = \sum_k |\langle f_k \mid i\rangle|^2$$

Figure 11.14 below shows three arrangements sufficient to determine the transition amplitude: (a) a single optical path; (b) two paths,allowing us to ascertain which path has actually been taken; (c) two paths, not allowing us to ascertain which path has actually been taken.

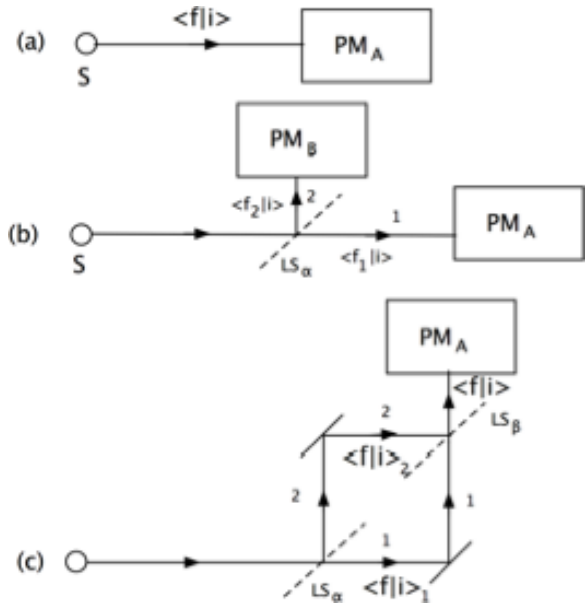


Figure 11.14: Three Arrangements

2. If a photon is emitted by the source S can take either of two paths, but it is impossible to ascertain which path it does take (Figure 11.14(c) above), then there are again two transition amplitudes:

$$\langle \text{photon arriving at } PM_A \mid \text{photon leaving S}\rangle_{\text{along path 1}}$$
$$\langle \text{photonarrivingat} PM_B \mid \text{photon leaving S}\rangle_{\text{along path 2}}$$

which we symbolize simply as

$$\langle f \mid i\rangle_1 \quad , \quad \langle f \mid i\rangle_2$$

To allow for interference, we assert that in this case it is the amplitudes that must be added; the total amplitude reads

$$\langle f \mid i\rangle = \langle f \mid i\rangle_1 + \langle f \mid i\rangle_2$$

The total probability is then:

$$|\langle f \mid i\rangle_1 + \langle f \mid i\rangle_2|^2$$

More generally, we would write

$$\text{total amplitude: } \langle f \mid i\rangle = \sum_k \langle f \mid i\rangle_k$$

$$\text{total probability: } |\langle f \mid i\rangle|^2 = \left|\sum_k \langle f \mid i\rangle_k\right|^2$$

where the sums are over all possible paths.

3. If one wants to analyze the propagation of the light more closely, one can take into account its passage through the half-silvered mirror LS_α, considering this as an intermediate state (Figure 11.14(b) above). The total amplitude for path 1 is

$$\langle \text{photon arriving at } PM_A \mid \text{photon leaving S}\rangle$$

However, it results from two successive intermediate amplitudes:

$$\langle \text{photon arriving at } LS_\alpha \mid \text{photon leaving } S\rangle$$
$$\langle \text{photon arriving at } PM_A \mid \text{photon leaving } LS_\alpha\rangle$$

Here we consider the total amplitude as the product of the successive intermediate amplitudes; symbolically, labelling the intermediate state as v, we have

$$\langle f \mid i\rangle = \langle f \mid \nu\rangle\langle \nu \mid i\rangle$$

Finally, consider a system of two mutually independent photons. If photon 1 undergoes a transition from a state i_1 to a state f_1, and photon 2 from a state i_2 to a state f_2, then

$$\langle f_1 f_2 \mid i_1 i_2\rangle = \langle f_1 \mid i_1\rangle\langle f_2 \mid i_2\rangle$$

The four rules just given suffice to calculate the detection probability in any possible experimental situation. They assume their present form as a result of a long theoretical evolution; but they are best justified *a posteriori*, because in 75 years they have never been found to be wrong. Accordingly, we may consider them as the basic principles governing the observable behavior of all microscopic objects, that is, objects whose action on each other are of order (Planck's constant). From these principles (they are equivalent to our earlier postulates - just look different because we are using the amplitude instead of the state vector as the fundamental mathematical object in the theory) one can derive all the requisite formalism, that is, all of quantum mechanics.

Quantum mechanics as we have described it earlier and also above, works splendidly, like a well-oiled machine. It, and its basic principles, might therefore be

expected to command the assent of every physicist; yet it has evoked, and on occasion continues to evoke, reservations both explicit and implicit. For this there are two reasons:

1. Quantum mechanics introduces unavoidable chance, meaning that its characteristic randomness is inherent in the microscopic phenomena themselves.

2. It attributes to microscopic objects properties so unprecedented that we cannot represent them through any macroscopic analogs or models.

Both features are revolutionary, and it is natural that they should have provoked debate. On the opposite sides of this debate we find two great physicists, Neils Bohr and Albert Einstein, and we will now discuss how the debate evolved from its beginnings in 1927 to its conclusion in 1983 (that is 56 years!).

11.5 Inseparable Photons (the EPR Paradox) including some history

Though ornithologists have known about inseparable parrots for a long time, to physicists the existence of inseparable photons has been brought home only during the last two decades, through a beautiful series of experiments by Alain Aspect and his research group at Orsay Laboratory in Paris. The experiments are exemplary, in virtue both of the difficulties they had to overcome and the results achieved, which are exceptionally clear-cut. In fact, the significance of the experiments extends beyond the strict confines of physics, because they provide the touchstone for settling a philosophical debate that has divided physicist for 75 years. The division dates back to the appearance of two mutually contradictory interpretations of quantum mechanics at the Como conference in 1927. To sketch the debate, we start with a brief summary of the philosophy of physics.

11.5.1 The Philosophical Stakes in the Debate

Our summary is best presented diagrammatically as shown in Figure 11.15 below where we present the philosophical elements in a debate between physicists.

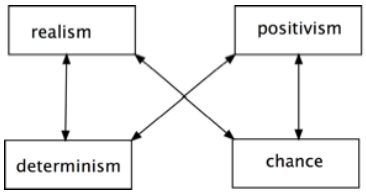


Figure 11.15: Philosophical Ideas

1. For the physicist who is a *realist*, a physical theory reflects the behavior of real objects, whose existence is not brought into question.

2. For the physicist who is a *positivist*, the purpose of a physical theory is to describe the relations between measurable quantities. The theory does not tell one whether anything characterized by these quantities really exists, nor even whether the question makes sense.

3. For the physicist who is a *determinist*, exact knowledge of the initial conditions and of the interactions allows the future to be predicted exactly. Determinism is held to be a universal characteristic of natural phenomena, even about those which we know, as yet, little or nothing. In this framework, any recourse to chance merely reflects our own ignorance.

4. For the physicist who is a *probabilist*, chance is inherent in the very nature of microscopic phenomena. To her, determinism is a consequence, on the macroscopic level, if the laws of chance operating on the microscopic level; it is appropriate to measurements of mean values of quantities whose relative fluctuations are very weak.

From these four poles, realism, positivism, determinism, and chance, the physicist chooses two, one on each axis. Though sometimes the choice is made in full awareness of what it entails, most often it is made subconsciously. In our description of quantum mechanics, we might adopt without reservations, the point of view of the elementary particle physicist. For a start, she believes firmly in the existence of particles, since she spends her time in accelerating, deflecting, focusing, and detecting them. Even though she has never seen or touched them, to her their objective existence is not in any doubt. Next she observes that they impinge on the detectors quite erratically, whence she has no doubts, either, that their behavior is random. Accordingly, the elementary particle experimentalist has chosen realism and chance, most often without realizing that she has made choices at all.

There are other philosophical options that can be adopted with eyes fully open: realism and determinism are the choices of Albert Einstein; positivism and chance are those of Neils Bohr. They are well acquainted and each thinks very highly of the other: which is no bar to their views being incompatible, nor to the two men representing opposite poles of the debate.

11.5.2 From Como to Brussels (1927-30)

On September 26,1927, in Como, Niels Bohr delivered a memorable lecture. His stance is that of an enthusiastic champion of the new quantum mechanics. He puts special weight on the inequalities proved by Heisenberg the year before:

$$\Delta x \Delta p_x \geq \frac{1}{2}\hbar \quad , \quad \Delta t \Delta E \geq \frac{1}{2}\hbar \tag{11.54}$$

They imply that it is impossible to define exact initial conditions for a microscopic object, which automatically makes it impossible to construct, on the microscopic scale, a deterministic theory patterned on classical mechanics. Only a probabilistic theory is possible, and that theory is quantum mechanics.

Einstein disagrees with this point of view, and his opposition to Bohr's theses becomes public at the Brussels conference in 1930: he adopts the role of a dissenter who knows precisely how to press home the most difficult questions. Deeply shocked by the retreat from determinism, he tries to show via his thought (gedanken) experiments he can contravene the Heisenberg inequalities.

At the cost of several sleepless nights devoted to analyzing the objections of his adversary, Bohr refutes all of Einstein's criticisms, and emerges from the conference as the undoubted winner.

11.5.3 From Brussels to the EPR Paradox (1930-35)

Having lost the argument at Brussels, Einstein tries to define his objections with ever greater precision. Believing as he does that position and momentum exist emphobjectively and simultaneously, he considers quantum mechanics to be incomplete and merely provisional. The points of view of the two antagonists at this stage of the debate can be spelled out as follows.

For Einstein, a physical theory must be a deterministic and a complete representation of the objective reality underlying the phenomena. It features known variables that are observable, and others, unknown as yet, called *hidden variables*. Because of our provisional ignorance of the hidden variables, matter at the microscopic level appears to us to behave arbitrarily, and we describe it by means of a theory that is incomplete and probabilistic, namely by quantum mechanics.

For Bohr, a physical theory makes sense only as a set of relations between observable quantities. Quantum mechanics supplies a correct and complete description of the behavior of objects at the microscopic level, which means that the theory itself is likewise complete. The observed behavior is probabilistic, implying that chance is inherent in the nature of the phenomena.

Between chance as a matter of ignorance, as advocated by Einstein, and chance unavoidable, as advocated by Bohr, the debate does not remain merely philosophical. Quite naturally it returns to the plane of physics with the thought experiment proposed by Einstein, Podolsky and Rosen in 1935, which in their view proves that quantum mechanics is indeed incomplete. Their thought experiment is published as a paper in the Physical Review, but it is so important that it reverberates as far as the New York Times. Physicists call the proposal the EPR paradox, after its proponents. It will take fifty years to untangle the question, first in theory and then by experiment. We will not, of course, follow these

fifty years blow by blow; instead, we confine attention to three decisive stages reached respectively in 1952, 1964, and 1983. But we start with an illustration that helps one see what the EPR paradox actually is.

11.5.4 Elementary Introduction to the EPR Paradox

Consider two playing cards, one red(diamond) and one black(spade) as shown in Figure 11.16 below where we use two playing cards help us understand the stakes in the EPR paradox.

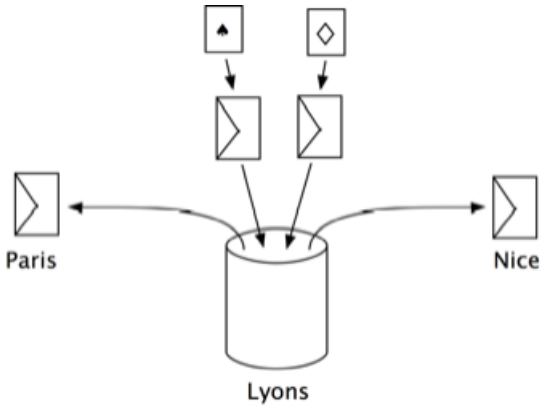


Figure 11.16: EPR Setup

An experimenter in Lyons puts them into separate envelopes which she then seals. She is thus provided with two envelopes looking exactly alike, and she puts both into a container. She shakes the container so as the *shuffle the pack*, and the system is ready for the experiment.

At 8:00 two travellers, one from Paris and one from Nice, come to the container (in Lyons), take one envelope each, and then return to Paris and Nice, respectively. At 14:00 they are back at their starting points; each opens her envelope, looks at the card, and telephones to Lyons reporting the color. The experiment is repeated every day for a year, and the observer in Lyons keeps a careful record of the results. At the end of the year the record stands as follows:

1. The reports from Paris are *red* or *black*, and the sequence of these reports is random. The situation is exactly the same as in a game of heads or tails, and probability of each outcome is 1/2.

2. The reports from Nice are *red* or *black*, and the sequence of these reports is random. Here too probability of each outcome is 1/2.

3. When Paris reports *red*, Nice reports *black*; when Paris reports *black*, Nice reports *red*. One sees that there is perfect(anti) correlation between the report from Paris and the report from Nice.

Accordingly, the experiment we have described displays two features:

1. It is *unpredictable* and thereby random at the level of individual observations in Paris and Nice.

2. It is *predictable*, by virtue of the correlation, at the level where one observes the Paris and the Nice results simultaneously.

Einstein and Bohr might have interpreted the correlation as follows.

According to Einstein

The future of the system is decided at 8:00 when the envelopes are chosen, because he believes that the contents of the two envelopes differ. Suppose, for instance, that Paris has (without knowing it) drawn a red card, and Nice the black. The colors so chosen exist in reality, even though we do not know them. The two cards are moved, separately, by the travellers between 8:00 and 14:00, during which time they do not influence each other in any way. The results on opening the envelopes read *red* in Paris and *black* in Nice. Since the choice at 8:00 was made blind, the opposite outcome is equally possible, but the results at 14:00 are always correlated (either red/black or black/red). This correlation at 14:00 is determined by the separation of the colors at 8:00, and we say the theory proposed by Einstein is *realist, deterministic, and separable(or local)*, by virtue of a hidden variable, namely, the color.

According to Bohr

There is a crucial preliminary factor, inherent in the preparation of the system. On shaking the container with the two envelopes, one loses information regarding the colors. Afterwards, one only knows that each envelope contains either a red card (probability 1/2) or a black card (probability 1/2). We will therefore say that a given envelope is in a *brown state*, which is a superposition of a red state and of a black state having equal probabilities. At 8:00 the two envelopes are identical: both are in a *brown state*, and the future of the system is still undecided. There is no solution until the envelopes are opened at 14:00, since it is only the action of opening them that makes the colors observable. The result is probabilistic. There is a probability 1/2 that in Paris the envelope will be observed to go from the *brown state* to the red, while the envelope in Nice is observed to go from the *brown state* to the black; there is the same probability 1/2 of observing the opposite. But the results of the observations on the two envelopes are always correlated, which means that there is a mutual influence between them, in particular at 14:00; in fact it is better to say that, jointly, they constitute a single and non-separable system, even though one is in Paris and the other is in Nice. Accordingly, the theory proposed by Bohr is positivist, probabilistic(non-deterministic) and non-separable(non-local), interrelating as it does the colors that are actually observed.

Einstein's view appears to be common sense, while it must be admitted that

Bohr's is very startling; however, the point of this macroscopic example is, precisely, to stress how different the quantum view is from the classical.

Proceeding with impeccable logic but from different premises, both theories predict the same experimental results. Can we decide between them? At the level considered here it seems we cannot: for even if the envelopes were opened prematurely while still in Lyons, one would merely obtain the same results at a different time, and without affecting the validity of either interpretation. The solution to the problem must be looked for at the atomic level, by studying the true EPR set-up itself.

11.5.5 The EPR Paradox (1935-52)

Albert Einstein, Boris Podolsky, and Nathan Rosen meant to look for an experiment that could measure, indirectly but simultaneously, two mutually exclusive quantities like position and momentum. Such results would contravene the predictions of quantum mechanics, which allows the measurement of only one such quantity at any one time; that is why the thought experiment is called the EPR paradox.

In 1952, David Bohm showed that the paradox could be set up not only with continuously varying quantities like position and momentum, but also with discrete quantities like spin. This was the first step towards any realistically conceivable experiment. Meanwhile, objectives have evolved, and nowadays it is more usual to talk of the EPR scenario, meaning some sensible experiment capable of discriminating between quantum theory and hidden-variable theories. Such a set-up is sketched in Figure 11.17 below where we present the simplest EPR scenario.

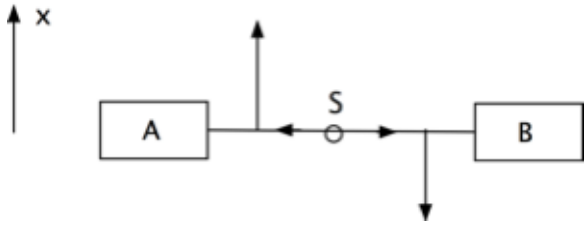


Figure 11.17: EPR Scenario

A particle with spin 0 decays, at S, into two particles of spin 1/2, which diverge from S in opposite directions. Two Stern-Gerlach type detectors A and B measure the $x-$components of the spins. Two types of response are possible:

1. *spin up* at A, *spin down* at B, a result denoted by $(+1, -1)$
2. *spin down* at A, *spin up* at B, a result denoted by $(-1, +1)$

Thus far everyone is agreed, but the interpretation is yet to come.

Einstein reasons that if pairs of particles produced at S elicit different responses $(+1,-1)$ and $(-1,+1)$ from the detector system A,B, then the pairs must have differed already at S, immediately after the decay.

It must be possible to represent this difference by a hidden variable λ, which has an objective meaning, and *which governs the future of the system.* After the decay the two particles separate without influencing each other any further, and eventually they trigger the detectors A and B.

Bohr reasons that all the pairs produced at S are identical. Each pair constitutes a non-separable system right up to the time when the photons reach the detectors A and B. At that time we observe the response of the detectors, which is probabilistic, admitting two outcomes $(+1,-1)$ and $(-1,+1)$. To sum up, Einstein restricts the operation of chance to the instant of decay (at S), whose details we ignore, but which we believe creates pairs whose hidden variables λ are different.

By contrast, Bohr believes that chance operates at the instant of detection, and that it is inherent in the very nature of the detection process: this chance is unavoidable.

We are still in the realms of thought, and stay there up to 1964.

In 1964, the landscape changes: John Bell, a theorist at CERN, shows that it is possible to distinguish between the two interpretations experimentally.

The test applies to the EPR scenario; it is refined by Clauser, Horne, Shimony, and Holt, whence it is called the BCHSH inequality after its five originators.

11.5.6 The BCHSH Inequality (1964)

To set up an EPR scenario, one first needs a source that emits particle pairs. Various experimental possibilities have been explored:

1. atoms emitting two photons in cascade
2. electron-positron annihilation emitting two high-energy photons
3. elastic proton-proton scattering

It is solution (1) that has eventually proved the most convenient; it has been exploited by Alain Aspect at the Institute for Optics in Paris, in particular.

Next one needs detectors whose response can assume one of two values, represented conventionally by +1 and −1. Such a detector might be

1. for spin 1/2 particles, a Stern-Gerlach apparatus responding to *spin up* or *spin down*
2. for photons, a polarizer responding to *parallel polarization* or *perpendicular polarization*

Our sketch of the EPR scenario can now be completed as in Figure 11.18 below where we present the most general EPR scenario.

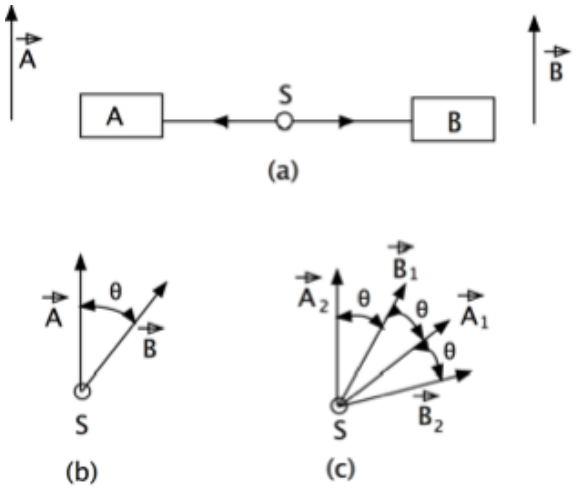


Figure 11.18: BCHSH Setup

Figure 11.18(a) views the apparatus perpendicularly to axis, showing the two detectors A and B, with their polarizing directions denoted as $\vec{A}$ and $\vec{B}$.

Figure 11.18(b) views the apparatus along its axis, and shows that the analyzing directions of the two detectors are not parallel, but inclined to each other at an angle θ.

In Figure 11.18((c) we also a view along the axis of the apparatus, and shows the actual settings chosen by Aspect: two orientations are allowed for each detector, $\vec{A}_1$ or $\vec{A}_2$ for one, and $\vec{B}_1$ or $\vec{B}_2$ for the other.

We adopt the following conventions:

(1) $\alpha = \pm 1$ is the response of detector A when oriented along $\vec{A}$

(2) $\beta = \pm 1$ is the response of detector B when oriented along $\vec{B}$

Since each detector has two possible orientations, called 1 and 2, we shall denote their responses as α_1, α_2 and β_1, β_2 respectively. Now consider the quantity $\langle\gamma\rangle$ defined by

$$\langle\gamma\rangle = \langle\alpha_1\beta_1\rangle + \langle\alpha_1\beta_2\rangle + \langle\alpha_2\beta_1\rangle - \langle\alpha_2\beta_2\rangle \tag{11.55}$$

where the symbol $\langle ... \rangle$ denotes the mean value over very many measured events. We call $\langle \gamma \rangle$ the *correlation function* of the system.

The BCHSH inequality reads $-2 \leq \langle \gamma \rangle \leq 2$. Its authors have proved that it must be satisfied if mechanics at the microscopic level constitutes a theory that is realist, deterministic, and separable: or in other words if the theory contains a hidden variable. A sketch of the a proof is shown below.

A Proof of Bell's Inequality

A theory that is deterministic and separable

Suppose that the pair a, b emerging from S can be characterized by a hidden variable λ. The responses of the detectors A, B are $\alpha(\vec{A}, \lambda)$ and $\beta(\vec{B}, \lambda)$ respectively as shown in Figure 11.19 below.

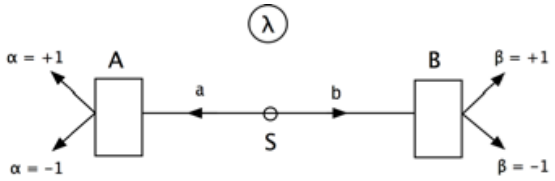


Figure 11.19: Bell Inequality Setup

The theory is deterministic and separable:

1. *deterministic*, because the results are determined by the hidden variables plus the settings $\vec{A}$ and $\vec{B}$
2. *separable*, because the response of A is independent of the response of B, and vice versa

Since the value of λ is unknown and different for each pair, the responses of A and B seem random. Lacking information about λ, we characterize it by choosing a statistical distribution $\rho(\lambda)$, which then allows us to derive the distribution of the responses $\alpha(\vec{A}, \lambda)$ and $\beta(\vec{B}, \lambda)$, which can be compared with experiment.

Bell's inequalities have the great virtue that they apply to any hidden variable theory, irrespective of the choice of $\rho(\lambda)$.

Theorem 1: Consider the four numbers α_1, α_2 and β_1, β_2, each of which can assume only the values 1 or −1. Then the combination

$$\gamma = \alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_2\beta_2 \tag{11.56}$$

can assume only the values 2 and −2.

To prove the theorem, one constructs a truth table for all 16 possibilities, which shows that 2 and -2 are indeed the only possible values of γ.

α_1	α_2	β_1	β_1	γ
1	1	1	1	2
1	1	1	-1	2
1	1	-1	1	-2
1	1	-1	-1	-2
1	-1	1	1	2
1	-1	1	-1	2
1	-1	-1	1	-2
1	-1	-1	-1	-2
-1	1	1	1	-2
-1	1	1	-1	2
-1	1	-1	1	-2
-1	1	-1	-1	2
-1	-1	1	1	-2
-1	-1	1	-1	-2
-1	-1	-1	1	2
-1	-1	-1	-1	2

Table 11.1: γ Values

Theorem 2: Consider very many sets of four numbers α_1, α_2 and β_1, β_2. The mean value of γ lies in the range $[-2, 2]$. In other words,

$$-2 \leq \langle \gamma \rangle \leq 2 \tag{11.57}$$

This is obvious, because every value of γ lies in this range, and so therefore must the mean. The endpoints are included in order to allow for limiting cases. Note that both theorems are purely mathematical, neither involves any assumptions about physics.

11.5.7 BCHSH Inequality(Bell's inequality in real world)

Within the framework of a theory that is realist, deterministic, and separable, we can describe the photon pair in detail. Realism leads us to believe that polarization is an objective property of each member of the pair, independent of any measurements that may be made later. Determinism leads us to believe that the polarizations are uniquely determined by the decay cascade, and that they are fully specified by the hidden variable λ, which governs the correlation of the polarizations in A and B. Finally, separability leads us to believe that the measurements in A and B do not influence each other, which means in particular that the response of detector A is independent of the orientation of detector B.

Now consider a pair of photons a, b, characterized by a hidden variable λ. The

response of the apparatus in its four settings would be as follows:

$$\alpha_1 \text{ and } \beta_1 \text{ in the orientation } (\vec{A}_1, \vec{B}_1)$$
$$\alpha_2 \text{ and } \beta_2 \text{ in the orientation } (\vec{A}_2, \vec{B}_2)$$
$$\alpha_1' \text{ and } \beta_2' \text{ in the orientation } (\vec{A}_1, \vec{B}_2)$$
$$\alpha_2' \text{ and } \beta_1' \text{ in the orientation } (\vec{A}_2, \vec{B}_1)$$

Recall that the variables α and β can only take on the values 1 and -1. It is impossible in practice to make four measurements on one and the same pair of photons, because each photon is absorbed in the first measurement made on it; that is why we have spoken conditionally, that is, of what results would be(a **COUNTERFACTUAL** statement). But if we believe that the photon correlations are governed by a theory that is realist, deterministic, and separable, then we are entitled to assume that the responses, of type α or type β, depend on properties that the photons possess before the measurement, so that the responses correspond to some objective reality. In such a framework we can appeal to the principle of separability, which implies, for instance, that detector A would give the same response to the orientations $(\vec{A}_1, \vec{B}_1)$ and $(\vec{A}_1, \vec{B}_2)$, because the response of A is independent of the orientation of B.

Mathematically, this is expressed by the relation $\alpha_1 = \alpha_1'$.

Similarly one finds $\alpha_2 = \alpha_2'$, $\beta_1 = \beta_1'$, $\beta_2 = \beta_2'$.

Thus, we have shown that, for a given pair of photons, all possible responses of the apparatus in its four chosen settings can be specified by means of only four two-valued variables α_1, α_2 and β_1, β_2. This reduction from eight to four variables depends on the *principle of separability.* In this way, we are led to a situation covered by Theorem 2, and therefore $-2 \le \langle\gamma\rangle \le 2$.

By making many measurements for each of the four settings we can determine the four mean values $\langle\alpha_1\beta_1\rangle$, $\langle\alpha_1\beta_2\rangle$, $\langle\alpha_2\beta_1\rangle$, $\langle\alpha_2\beta_2\rangle$ and thus the mean value of the correlation

$$\langle\gamma\rangle = \langle\alpha_1\beta_1\rangle + \langle\alpha_1\beta_2\rangle + \langle\alpha_2\beta_1\rangle - \langle\alpha_2\beta_2\rangle$$

According to quantum mechanics (which is positivist, probabilistic, and non-separable), there are cases where the BCHSH inequality is violated. In particular, one can show that for photons in the configuration chosen by Aspect quantum mechanics yields(we will derive this shortly)

$$\langle\gamma\rangle = 3\cos 2\theta - \cos 6\theta \tag{11.58}$$

This leads to values well outside the interval $[-2, 2]$; for example to $\langle\gamma\rangle = 2\sqrt{2}$ when $\theta = 22.5°$ and to $\langle\gamma\rangle = -2\sqrt{2}$ when $\theta = 67.5°$.

Proof: The laboratory reference frame $Oxyz$ serves to specify the orientations of detectors and polarizers as shown in Figure 11.20 below:

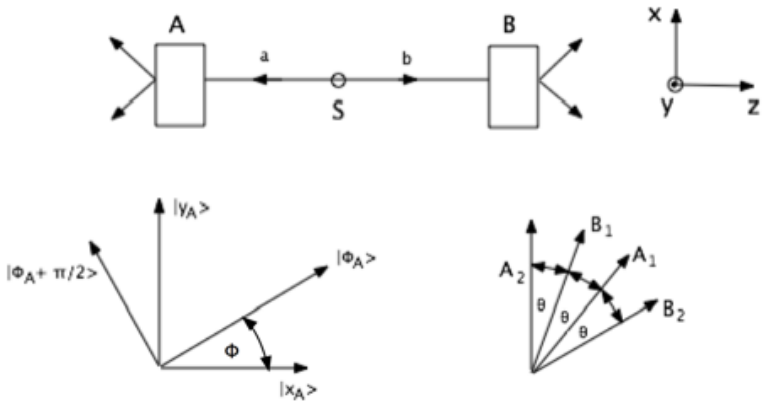


Figure 11.20: Detector Orientations

Before any measurements have been made, the photon pair a, b forms a non-separable entity, represented by the vector

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\left(|x_A, x_B\rangle + |y_A, y_B\rangle\right) \tag{11.59}$$

The act of measurement corresponds to passage to the φ-basis. Hence, we require the transition amplitudes from the two states $|x_A, x_B\rangle, |y_A, y_B\rangle$ to the four states

$$|\varphi_A, \varphi_B\rangle, |\varphi_A, \varphi_B + \pi/2\rangle, |\varphi_A + \pi/2, \varphi_B\rangle, |\varphi_A + \pi/2, \varphi_B + \pi/2\rangle \tag{11.60}$$

In the φ-basis we have

$$\begin{aligned} |\Phi\rangle = \frac{1}{\sqrt{2}}[&\cos(\varphi_B - \varphi_A)\,|\varphi_A, \varphi_B\rangle \\ &- \sin(\varphi_B - \varphi_A)\,|\varphi_A, \varphi_B + \pi/2\rangle \\ &+ \sin(\varphi_B - \varphi_A)\,|\varphi_A + \pi/2, \varphi_B\rangle \\ &+ \cos(\varphi_B - \varphi_A)\,|\varphi_A + \pi/2, \varphi_B + \pi/2\rangle] \end{aligned} \tag{11.61}$$

The square of each amplitude featured here represents the detection probability. For example, the probability of simultaneously detecting photon a polarized at the angle φ_A and the photon b polarized at the angle φ_B is

$$\left(\frac{1}{\sqrt{2}}\cos(\varphi_B - \varphi_A)\right)^2 = \frac{1}{2}\cos^2(\varphi_B - \varphi_A) \tag{11.62}$$

By convention, we write the responses of detector A to a photon in state $|\varphi_A\rangle$ (respectively $|\varphi_A + \pi/2\rangle$) as $\alpha = 1$ and similarly with β for detector B.

Let us analyze the four possible responses:

1. $|\varphi_A, \varphi_B\rangle$ gives $\alpha = 1, \beta = 1$ so $\alpha\beta = 1$; the probability is

$$P_{++} = \frac{1}{2}\cos^2(\varphi_B - \varphi_A) \tag{11.63}$$

2. $|\varphi_A, \varphi_B + \pi/2\rangle$ gives $\alpha = -1, \beta = 1$ so $\alpha\beta = -1$; the probability is

$$P_{+-} = \frac{1}{2}\sin^2(\varphi_B - \varphi_A) \tag{11.64}$$

3. $|\varphi_A + \pi/2, \varphi_B\rangle$ gives $\alpha = 1, \beta = -1$ so $\alpha\beta = -1$; the probability is

$$P_{-+} = \frac{1}{2}\sin^2(\varphi_B - \varphi_A) \tag{11.65}$$

4. $|\varphi_A + \pi/2, \varphi_B + \pi/2\rangle$ gives $\alpha = -1, \beta = -1$ so $\alpha\beta = 1$; the probability is

$$P_{--} = \frac{1}{2}\cos^2(\varphi_B - \varphi_A) \tag{11.66}$$

The mean value of $\langle\alpha\beta\rangle_{AB}$ follows immediately as

$$\langle\alpha\beta\rangle_{AB} = P_{++} - P_{+-} - P_{-+} + P_{--} = \cos 2(\varphi_B - \varphi_A) \tag{11.67}$$

The settings chosen by Aspect are as shown in Figure 16.20 above. Corresponding to it we have the four terms

$$\langle\alpha_1\beta_1\rangle = \langle\alpha\beta\rangle_{A_1B_1} = \cos 2(\varphi_{B_1} - \varphi_{A_1}) = \cos 2\theta \tag{11.68}$$
$$\langle\alpha_1\beta_2\rangle = \langle\alpha\beta\rangle_{A_1B_2} = \cos 2(\varphi_{B_1} - \varphi_{A_2}) = \cos 2\theta \tag{11.69}$$
$$\langle\alpha_2\beta_1\rangle = \langle\alpha\beta\rangle_{A_2B_1} = \cos 2(\varphi_{B_2} - \varphi_{A_1}) = \cos 2\theta \tag{11.70}$$
$$\langle\alpha_2\beta_2\rangle = \langle\alpha\beta\rangle_{A_2B_2} = \cos 2(\varphi_{B_2} - \varphi_{A_2}) = \cos 6\theta \tag{11.71}$$

For comparison with Bell's inequality, we introduce the correlation function

$$\langle\gamma\rangle = \langle\alpha_1\beta_1\rangle + \langle\alpha_1\beta_2\rangle + \langle\alpha_2\beta_1\rangle - \langle\alpha_2\beta_2\rangle = 3\cos 2\theta - \cos 6\theta \tag{11.72}$$

Thus, the BCHSH test turns the EPR scenario into an arena for rational confrontation between the two interpretations; it remains only to progress from thought experiments to experiments conducted in the laboratory.

11.5.8 The Beginnings of the Experiment at Orsay (1976)

Alain Aspect's experiment studies the correlation between the polarizations of the members of photon pairs emitted by calcium. The light source is a beam of calcium atoms, excited by two focused laser beams having wavelengths $\lambda' = 406\,nm$ and $\lambda'' = 581\,nm$ respectively. Two-photon excitation produces

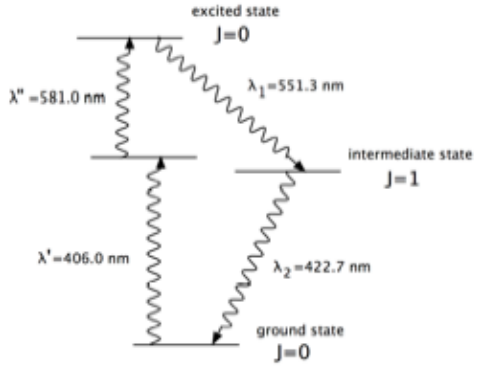


Figure 11.21: Calcium based light source - Energy Levels

a state having the quantum number $J = 0$. When it decays, this state emits two monochromatic photons having the wavelengths $\lambda_1 = 551.3\,nm$ and $\lambda_2 = 422.7\,nm$ respectively, in a cascade of two electronic transitions from the initial $J = 0$ level to the final $J = 0$ state, passing through an intermediate $J = 1$ state, as shown in Figure 11.21 below which shows the excitation and decay of the calcium atom.

The mean lifetime of the intermediate state is $4.7\,ns$. To simplify the terminology, we shall call the $\lambda_1 = 551.3\,nm$ light green, and the $\lambda_2 = 422.7\,nm$ light violet.

The polarizer, which works like a Wollaston prism is shown in Figure 11.22 below where we can see the two-valued response of a Wollaston prism.

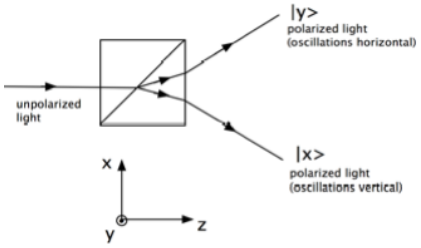


Figure 11.22: Wollason Prism - Polarizer

The Wollaston prism is made of quartz or of calcite. It splits an incident beam of natural (unpolarized) light into two beams of equal intensity, polarized at 90° to each other. If only a single unpolarized photon is incident, it emerges either in the state $|x\rangle$, with probability 1/2, or in the state $|y\rangle$, with probability 1/2. Thus, the response of the system is two-valued.

The photon is detected by the photomultiplier tubes (PM) downstream from the prism. Every electric pulse from these detectors corresponds to the passage of a photon, allowing the photons to be counted. The experimental layout is sketched in Figure 11.23 below, which shows a sketch of the first Orsay experiment.

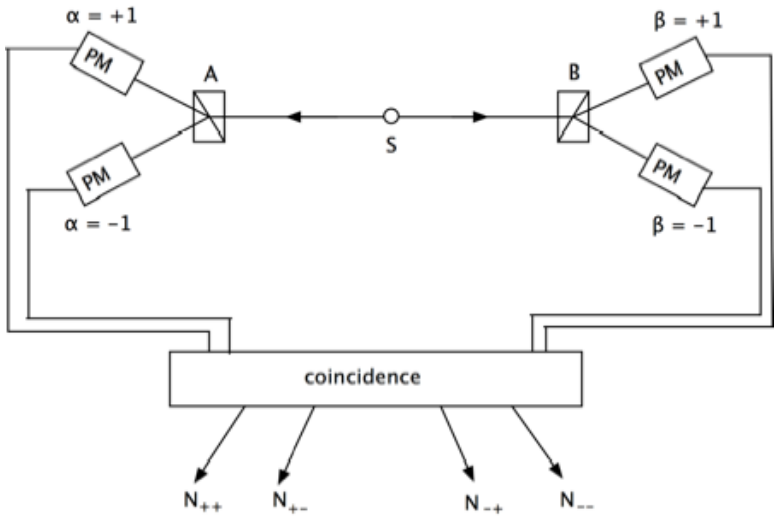


Figure 11.23: First Orsay Experiment

?It uses a coincidence circuit which registers an event whenever two photons are detected in cascade. In this way four separate counts are recorded simultaneously, over some given period of time. In the EPR scenario envisaged by Bohm, where $\theta = 0$, the only possible responses are $(+1, -1)$ or $(-1, +1)$ In the situation realized by Aspect, the angle θ is non-zero, and four different responses are possible.

1. N_{++} the number of coincidences corresponding to $\alpha = 1$ and $\beta = 1$, that is, to $\alpha\beta = 1$

2. N_{+-} the number of coincidences corresponding to $\alpha = 1$ and $\beta = -1$, that is, to $\alpha\beta = -1$

3. N_{-+} the number of coincidences corresponding to $\alpha = -1$ and $\beta = 1$, that is, to $\alpha\beta = -1$

4. N_{--} the number of coincidences corresponding to $\alpha = -1$ and $\beta = -1$, that is, to $\alpha\beta = 1$

The resolving time of the coincidence circuit is $10\,ns$, meaning that it reckons two photons as coincident if the they are separated in time by no more than $10\,ns$. The mean life of the intermediate state of the calcium atom is $4.7\,ns$.

Therefore, after a lapse of $10\,ns$, that is more than twice the mean lifetime, almost all the atoms have decayed (actually 88%). In other words, the efficiency of the coincidence counter is very high.

The experiment consists in counting, over some given time interval, the four kinds of coincidence: N_{++}, N_{+-}, N_{-+} and N_{--}. The total number of events is $N = N_{++} + N_{+-} + N_{-+} + N_{--}$.

Accordingly, the different kinds of coincidence have probabilities

$$\begin{aligned} P_{++} &= N_{++}/N \text{ corresponding to } \alpha\beta = 1 \\ P_{+-} &= N_{+-}/N \text{ corresponding to } \alpha\beta = -1 \\ P_{-+} &= N_{-+}/N \text{ corresponding to } \alpha\beta = -1 \\ P_{--} &= N_{--}/N \text{ corresponding to } \alpha\beta = 1 \end{aligned}$$

and the measured average of $\alpha\beta$ is

$$\langle \alpha\beta \rangle = \frac{N_{++} - N_{+-} - N_{-+} + N_{--}}{N} \tag{11.73}$$

Each set of four coincidence counts corresponds to one particular setting of $(\vec{A}, \vec{B})$, and yields a mean value $\langle \alpha\beta \rangle$. But in order to determine the correlation function $\langle \gamma \rangle$ used in the BCHSH inequality, we need four mean values $\langle \alpha\beta \rangle$. Therefore, we choose, in succession four different settings as shown in Figure 16.18(c); four counting runs then yield the four mean values $\langle \alpha_1\beta_1 \rangle$, $\langle \alpha_1\beta_2 \rangle$, $\langle \alpha_2\beta_1 \rangle$, $\langle \alpha_2\beta_2 \rangle$ which then determine the value of $\langle \gamma \rangle$ via

$$\langle \gamma \rangle = \langle \alpha_1\beta_1 \rangle + \langle \alpha_1\beta_2 \rangle + \langle \alpha_2\beta_1 \rangle - \langle \alpha_2\beta_2 \rangle \tag{11.74}$$

The Results of the First Experiment at Orsay

The results of the first Orsay experiment are shown in Figure 11.24 below. The angle θ which specifies the setting of the polarizers is plotted horizontally, and the mean value $\langle \gamma \rangle$ vertically.

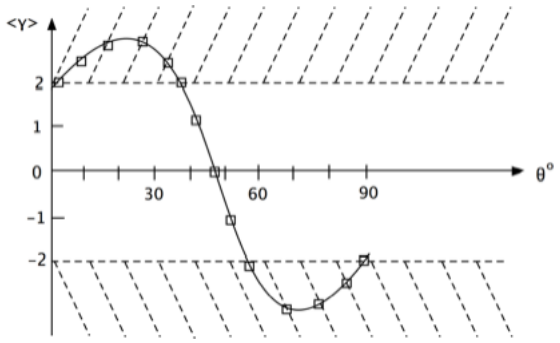


Figure 11.24: Results from First Orsay Experiment

From earlier, the correlation function predicted by quantum mechanics reads

$$\langle\gamma\rangle = 3\cos 2\theta - \cos 6\theta \tag{11.75}$$

It is drawn as the solid curve on the graph(the curve has been corrected for instrumental effects, which explains why its ends are not precisely at 2 and −2). According to the BCHSH inequality

$$-2 \le \langle\gamma\rangle \le 2 \tag{11.76}$$

so that hidden-variable theories exclude the cross-hatched regions of the plane, which correspond to $\langle\gamma\rangle > 2$ or $\langle\gamma\rangle < -2$.

The experimental results from 17 different values of θ are indicated on the figure by squares, where the vertical size of the square gives plus or minus one standard deviation (a measure of the experimental error).

Clearly, there can be no doubt that the BCHSH inequality is violated; many of the experimental points fall outside the interval $[-2, 2]$. At the point where the violation is maximal ($\theta = 22.5°$), one finds

$$\langle\gamma\rangle = 2.70 \pm 0.015 \tag{11.77}$$

which represents a departure of over 40 standard deviations from the extreme value of 2. What is even more convincing is the precision with which the experimental points lie on the curve predicted by quantum mechanics.

Quite evidently, for the EPR scenario one must conclude not only that hidden-variable theories fail, but also that quantum mechanics is positively the right theory for describing the observations.

The Relativistic Test

The EPR experiment just described shows that the measurements in A and B are correlated. What is the origin of the correlations?

According to quantum theory, before the measurement each particle pair constitutes a single system extending from A to B, whose two parts are non-separable and correlated. This interpretation corresponds to a violation of Bell's inequality and agreement with experiment.

According to hidden-variables theories, the particle pair is characterized, at the instant of decay, by its hidden variable , which determines the correlation between the polarizations measured in A and B. This interpretation satisfies Bell's inequality but disagrees with experiment.

Accordingly, the Orsay experiment supports the quantum interpretation (in terms of the correlation between two parts A and B of a single system).

However, to clinch this conclusion, one must ensure that no influence is exerted in the ordinary classical sense through some interaction propagated between the two detectors A and B, that is, no influence which might take effect after the decay at S, and which might be responsible for the correlation actually observed.

Let us therefore examine the Orsay apparatus in more detail as in Figure 11.25 below where we attempt to test Einsteinian non-separability.

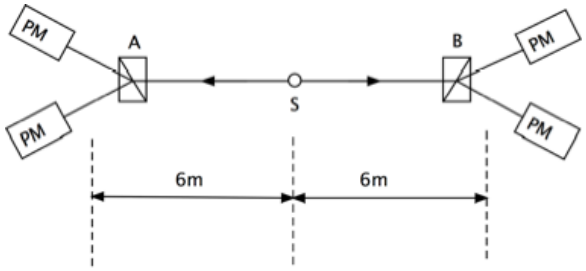


Figure 11.25: Orsay Experiment - Details

When the detectors at A and B record a coincidence, this means that both have been triggered within a time interval of at most $10\,ns$, the resolving time of the circuit. Could it happen that, within this interval, A sends to B a signal capable of influencing the response of B? In the most favorable case, such a signal would travel with the speed of light in vacuum, which according to relativity theory is the upper limit on the propagation speed of information, and thereby of energy. To cover the distance AB, which is $12\,m$ in the figure, such a signal would need $40\,ns$. This is too long by at least $30\,ns$, and rules out any causal links between A and B in the sense of classical physics. One says that *the interval between A and B is space-like.*

One of the advantages of the Orsay experiment is that it uses a very strong light-source, allowing sufficient distance between the detectors A and B while

still preserving reasonable counting rates. By increasing the distance AB step by step, Aspect could check that the correlation persists, even when the interval between A and B becomes space-like. This is the check that guarantees that the two-photon system is non-separable irrespective of the distance AB.

It has become the custom to speak of *the principle of Einsteinian separability* in order to denote the absence of correlations between two events separated by a space-like interval. This is the principle that the Orsay experiment invites us to reconsider, even though our minds, used to the world at the macroscopic level, find it difficult to conceive of two *microscopic* photons $12\,m$ apart as a single indivisible object.

The Final Stage of the Experiment at Orsay (1983)

Though the results of the first Orsay experiment are unarguable and clear-cut, the conclusion they invite is so startling that one should not be surprised at the appearance of a last-ditch objection, which as it happens gave the experimenters a great deal of trouble. In the preceding section we discussed the possible role of interactions between A and B operating after the decay at S, and duly eliminated the objection. But one can also ask whether correlations might be introduced through an interaction operating before the decay. We could imagine that the decay itself is preconditioned by the setting of detectors A and B, such influences taking effect through the exchange of signals between the detectors and the source. No such mechanism is known a priori, but we do know that, if there is one, then Einsteinian non-separability would cease to be a problem, because the mechanism could come into action long before the decay, removing any reason for expecting a minimum $30\,ns$ delay. Though such a scenario is very unlikely, the objection is a serious one and must be taken into account; to get around it, the experimenter must be able to choose the orientation of the detectors A and B at random after the decay has happened at S. In more picturesque language, we would say that the two photons must leave the source without knowing the orientations of the polarizers A and B. Briefly put, this means that it must be possible to change the detector orientations during the $20\,ns$ transits over SA and SB.

The solution adopted at Orsay employs periodic switching every $10\,ns$. These changes are governed by two independent oscillators, one for channel A and one for channel B. The oscillators are stabilized, but however good the stabilization it cannot eliminate small random drifts that are different in the two channels, seeing that the oscillators are independent. This ensures that the changes of orientation are random even though the oscillations are periodic, provided the experiment lasts long enough (1 to 3 hours).

The key element of the second Orsay experiment is the optical switch shown in Figure 11.26 below.

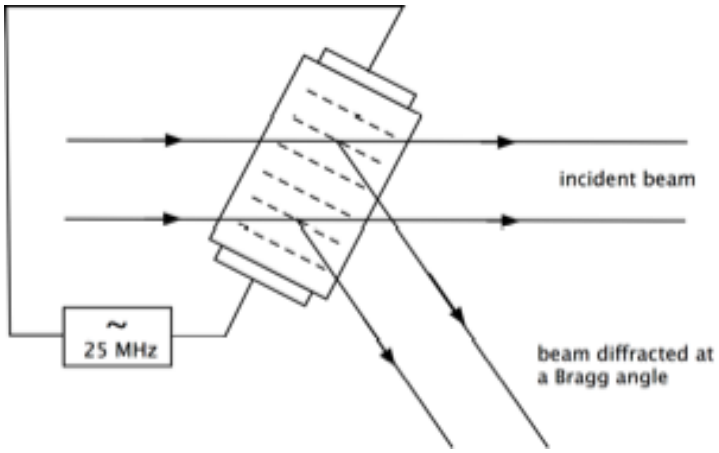


Figure 11.26: Second Orsay Experiment - Optical Switch

In a water tank, a system of standing waves is produced by electro-acoustic excitation at a frequency of $25\,MHz$(corresponds to $10\,ns$ between switchings).

The fluid keeps changing from a state of perfect rest to one of maximum agitation and back again. In the state of rest, the light beam is simply transmitted, In the state of maximum agitation, the fluid arranges itself into a structure of parallel and equidistant plane layers, alternately stationary (nodal planes) or agitated (antinodal planes). Thus, one sets up a lattice of net-like diffracting planes; the diffracted intensity is maximum at the so-called Bragg angles, just as in scattering from a crystal lattice. Here the light beam is deviated through 10^2 radians (the angles in the figure are exaggerated for effect). The two numerical values, $25\,MHz$ and 10^2 radians, suffice to show the magnitude of the technical achievement. With the acoustic power of 1 watt, the system functions as an ideally efficient switch.

The second Orsay experiment (using optical switches) is sketched in Figure 11.27 below.

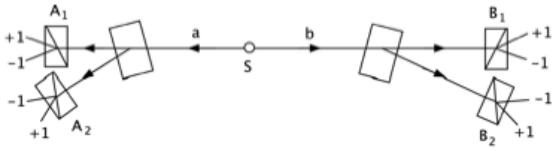


Figure 11.27: Second Orsay Experiment using Optical Switches

In this set-up, the photons a and b leave S without *knowing* whether they will go, the first to A_1 or A_2 and the second to B_1 or B_2.

The second experiment is less precise than the first, because the light beams must be very highly collimated in order to ensure efficient switching. Nevertheless, its results exhibit an unambiguous violation of Bell's inequality, reaching 5 standard deviations at the peak; moreover the results are entirely compatible with the predictions of quantum mechanics.

11.6 The Principle of Non-Separability

Experiment has spoken. Half a century after the Como conference, Bohr's interpretation once again beats Einstein's, in a debate more subtle and also more searching. There were two conflicting theories:

Einstein	Bohr
hidden variables	quantum mechanics
realist	positivist
deterministic	probabilistic
separable	non-separable

Table 11.2: Two Conflicting Theories

The violation of the BCHSH inequality argues for Bohr's interpretation, all the more so as the measured values of $\langle\gamma\rangle$ are in close agreement with the predictions of quantum mechanics.

It remains to ask oneself just why hidden-variable theories do fail. Of the three basic assumptions adopted by such theories, namely realism, determinism, and separability, at least one must be abandoned. In the last resort, it is separability that seems to be the most vulnerable assumption. Indeed, one observes experimentally that the violation of the BCHSH inequality is independent of the distance between the two detectors A and B, even when this distance is $12\,m$ or more. There are still die-hard advocates of determinism, who try to explain non-separability through non-local hidden variables. Such theories, awkward and barely predictive, are typically ad hoc, and fit only a limited number of phenomena. They are weakly placed to defend themselves against interpretations furnished by quantum mechanics, which have the virtues of simplicity, elegance, efficiency, and generality, and which are invariably confirmed by experiment.

The principle of Einsteinian separability asserts that *there are no correlations between two phenomena separated by a space-like interval.* In other words, no interaction can propagate faster than light in vacuum. In an EPR scenario this principle must be abandoned, and replaced by a principle asserting non-separability:

in a quantum system evolving free of external perturbations,and from well-defined initial conditions, all parts of thesystem remain correlated, even when the interval between them is space-like

This assertion reflects the properties of the state vector of a quantum system. For an EPR system, the state vector after the decay of the source reads

$$|\Phi\rangle = \frac{1}{\sqrt{2}} \left(|x_A, x_B\rangle + |y_A, y_B\rangle \right) \tag{11.78}$$

This expression combines the elements A and B in a non-separable manner, which is what explains the observed correlations. The truth is that all this has been well known ever since the beginnings of quantum mechanics, with the concept of the electron cloud as the most telling illustration. It is for instance hard to imagine separability between the 92 electrons of a uranium atom. What is new is that quantum mechanics, considered hitherto as a microscopic theory applicable on the atomic scale, is now seen to apply to a two-particle system macroscopically, on the scale of meters. The truly original achievement of Aspect's experiment is the demonstration of this fact.

Quantum objects have by no means exhausted their capacity to astonish us by their difference from the properties of the macroscopic objects in our everyday surroundings. In the preceding sections we saw that a photon can interfere with itself and we have shown that two photons $12\,m$ apart constitute but a single object. Thus, it becomes ever more difficult to picture a photon through analogies with rifle bullets, surface waves in water, clouds in the sky, or with any other object of our familiar universe. Such partial analogies fail under attempts to make them more complete, and through their failure we discover new properties pertaining to quantum objects. The only fruitful procedure is to follow the advice of Niels Bohr, namely, to bend one's mind to the new quantum concepts until they become habitual and thereby intuitive. Earlier generations of physicists have had to face similar problems. They had to progress from Aristotle's mechanics to Newton's, and then from Newton's to Einstein's. The same effort is now required of us, at a time favorable in that, by mastering the EPR paradox, quantum mechanics has passed a particularly severe test with flying colors.

From this point of view, the principle of non-separability seems as important as the principle of special relativity , and Aspect's experiment plays the same role now that the Michelson-Morley experiment played then.

11.7 An Example and a Solution - Bell's Theorem with Photons

Two photons fly apart from one another, and are in oppositely oriented circularly polarized states. One strikes a polaroid film with axis parallel to the unit vector $\hat{a}$, the other a polaroid with axis parallel to the unit vector $\hat{b}$. Let $P_{++}(\hat{a},\hat{b})$ be the joint probability that both photons are transmitted through their respective polaroids. Similarly, $P_{--}(\hat{a},\hat{b})$ is the probability that both photons are absorbed by their respective polaroids, $P_{+-}(\hat{a},\hat{b})$ is the probability that the photon at the $\hat{a}$ polaroid is transmitted and the other is absorbed, and finally, $P_{-+}(\hat{a},\hat{b})$ is the probability that the photon at the $\hat{a}$ polaroid is absorbed and the other is transmitted.

The classical realist assumption is that these probabilities can be separated:

$$P_{ij}(\hat{a},\hat{b}) = \int d\lambda\, \rho(\lambda) P_i(\hat{a},\lambda) P_j(\hat{b},\lambda) \tag{11.79}$$

where i and j take on the values + and −, where λ signifies the so-called hidden variables, and where $\rho(\lambda)$ is a weight function. This equation is called the separable form.

The correlation coefficient is defined by

$$C(\hat{a},\hat{b}) = P_{++}(\hat{a},\hat{b}) + P_{--}(\hat{a},\hat{b}) - P_{+-}(\hat{a},\hat{b}) - P_{-+}(\hat{a},\hat{b}) \tag{11.80}$$

and so we can write

$$C(\hat{a},\hat{b}) = \int d\lambda\, \rho(\lambda) C(\hat{a},\lambda) C(\hat{b},\lambda) \tag{11.81}$$

where

$$C(\hat{a},\lambda) = P_+(\hat{a},\lambda) - P_-(\hat{a},\lambda) \tag{11.82}$$

$$C(\hat{b},\lambda) = P_+(\hat{b},\lambda) - P_-(\hat{b},\lambda) \tag{11.83}$$

It is required that

(a) $\rho(\lambda) \geq 0$

(b) $\int d\lambda\, \rho(\lambda) = 1$

(c) $-1 \leq C(\hat{a},\lambda) \leq 1\ ,\ -1 \leq C(\hat{b},\lambda) \leq 1$

The Bell coefficient

$$B = C(\hat{a},\hat{b}) + C(\hat{a},\hat{b}') + C(\hat{a}',\hat{b}) - C(\hat{a}',\hat{b}') \tag{11.84}$$

combines four different combinations of the polaroid directions.

(1) Show that the above classical realist assumptions imply that $|B| \leq 2$.

(2) Show that quantum mechanics predicts that $C(\hat{a}, \hat{b}) = 2\left(\hat{a} \cdot \hat{b}\right)^2 - 1$.

(3) Show that the maximum value of the Bell coefficient is $2\sqrt{2}$, according to quantum mechanics.

(4) Cast the quantum mechanical expression for $C(\hat{a}, \hat{b})$ into a separable form. Which of the classical requirements, (a), (b), or (c) above is violated?

Solution

(1) With the separability assumption, we have (16.81)

$$C(\hat{a}, \hat{b}) = \int d\lambda\, \rho(\lambda) C(\hat{a}, \lambda) C(\hat{b}, \lambda)$$

It follows that the Bell coefficient can be written in the form

$$\begin{aligned} B &= C(\hat{a}, \hat{b}) + C(\hat{a}, \hat{b}') + C(\hat{a}', \hat{b}) - C(\hat{a}', \hat{b}') \\ &= \int d\lambda \rho(\lambda) [C(\hat{a}, \lambda)(C(\hat{b}, \lambda) + C(\hat{b}', \lambda)) \\ &\qquad + C(\hat{a}', \lambda)(C(\hat{b}, \lambda) - C(\hat{b}', \lambda))] \end{aligned} \tag{11.85}$$

Since $|C(\hat{a}, \lambda)| \leq 1$, $|C(\hat{a}', \lambda)| \leq 1$ and $\rho(\lambda) \geq 0$, we have

$$|B| \leq \int d\lambda \rho(\lambda) \left(\left| C(\hat{b}, \lambda) + C(\hat{b}', \lambda) \right| + \left| C(\hat{b}, \lambda) - C(\hat{b}', \lambda) \right| \right) \tag{11.86}$$

Now suppose that for a given λ, $C_M(\lambda)$ is the maximum and $C_m(\lambda)$ is the minimum of $C(\hat{b}, \lambda)$ and $C(\hat{b}', \lambda)$, so that $C_M(\lambda) \geq C_m(\lambda)$. Then

$$|B| \leq \int d\lambda \rho(\lambda) \left(|C_M(\lambda) + C_m(\lambda)| + C_M(\lambda) - C_m(\lambda) \right) \tag{11.87}$$

There are two cases to consider.

For the case $C_M(\lambda) \geq 0$, we have $|C_M(\lambda) + C_m(\lambda)| = C_M(\lambda) + C_m(\lambda)$ so that

$$\begin{aligned} |B| &\leq \int d\lambda \rho(\lambda) \left(C_M(\lambda) + C_m(\lambda) + C_M(\lambda) - C_m(\lambda) \right) \\ &= 2 \int d\lambda \rho(\lambda) C_M(\lambda) \leq 2 \int d\lambda \rho(\lambda) |C_M(\lambda)| \leq 2 \int d\lambda \rho(\lambda) = 2 \end{aligned}$$

For the case $C_M(\lambda) < 0$, we have $|C_M(\lambda) + C_m(\lambda)| = -C_M(\lambda) - C_m(\lambda)$ so that

$$\begin{aligned} |B| &\leq \int d\lambda \rho(\lambda) \left(-C_M(\lambda) - C_m(\lambda) + C_M(\lambda) - C_m(\lambda) \right) \\ &= 2 \int d\lambda \rho(\lambda) (-C_m(\lambda)) \leq 2 \int d\lambda \rho(\lambda) |C_m(\lambda)| \leq 2 \int d\lambda \rho(\lambda) = 2 \end{aligned}$$

Thus, in all cases $|B| \leq 2$.

(2) A photon, traveling in the $y-$direction, might have right- or left-handed circular polarization. The corresponding quantum states are written $|R\rangle$ and $|L\rangle$ respectively. These circular polarization states can be expressed as coherent superpositions of linearly polarized states in the z and x directions:

$$|R\rangle = \frac{1}{\sqrt{2}}\left(|z\rangle + i\,|x\rangle\right) \quad , \quad |L\rangle = \frac{1}{\sqrt{2}}\left(|z\rangle - i\,|x\rangle\right) \tag{11.88}$$

Under a rotation of the coordinate axes by an angle θ about the y direction, $|R\rangle \to e^{i\theta}\,|R\rangle$ and $|L\rangle \to e^{-i\theta}\,|L\rangle$ or equivalently

$$\begin{pmatrix} |z\rangle \\ |x\rangle \end{pmatrix} \to \begin{pmatrix} |z'\rangle \\ |x'\rangle \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} |z\rangle \\ |x\rangle \end{pmatrix} \tag{11.89}$$

If each photon is in a state of right-handed circular polarization, we write the corresponding state vector as $|R_1\rangle\,|R_2\rangle$. However, since the photons are moving in opposite directions, one along the positive, and the other along the negative y axis, it follows that the actual directions in which the electric fields rotate, in time, in the vicinity of the two photons, are opposed to one another. The same holds for the state $|L_1\rangle\,|L_2\rangle$, corresponding to each photon being in a state of left-handed circular polarization.

The linear combination of these two states,

$$|EPR\rangle = \frac{1}{\sqrt{2}}\left(|R_1\rangle\,|R_2\rangle + |L_1\rangle\,|L_2\rangle\right) \tag{11.90}$$

corresponds to the more general situation in which the photons are in oppositely oriented states of circular polarization, where the sense of this polarization is not specified. We can write this *entangled* or *Einstein-Podolsky-Rosen* state in the form

$$|EPR\rangle = \frac{1}{\sqrt{2}}\left(|z_1\rangle\,|z_2\rangle - |x_1\rangle\,|x_2\rangle\right) \tag{11.91}$$

which is a superposition of states of linear polarization.

Suppose now that a measurement of linear polarization is made on photon 1 in the z direction, and of photon 2 in the z' direction, that is, the z direction after a rotation of the axes about the y axis. The probability amplitude associated with this measurement on the EPR state is

$$\begin{aligned} \langle EPR \mid z_1 z_2'\rangle &= \frac{1}{\sqrt{2}}\left(\left(\langle z_1|\,\langle z_2| - \langle x_1|\,\langle x_2|\right)\left(|z_1\rangle\left(\cos\theta\,|z_2\rangle - \sin\theta\,|x_2\rangle\right)\right)\right) \\ &= \frac{1}{\sqrt{2}}\cos\theta \end{aligned} \tag{11.92}$$

where we have used $\langle z_1 \mid x_1\rangle = 0$. The probability that photon 1 is found to have linear polarization in the direction z, and photon 2 in the direction z' is

$$P_{++}(\hat{a},\hat{b}) = \left|\langle EPR \mid z_1 z_2'\rangle\right|^2 = \frac{1}{2}\cos^2\theta \tag{11.93}$$

where we have assumed that $\hat{a}$ is in the z direction and $\hat{b}$ is in the z' direction.

Suppose next that the linear polarization of the linear polarization of photon 1 were measured in the x direction, and that of photon 2 again in the z' direction. The probability amplitude is

$$\begin{aligned}\langle EPR \mid x_1 z_2' \rangle &= \frac{1}{\sqrt{2}} \left(\langle z_1 | \langle z_2 | - \langle x_1 | \langle x_2 | \right) \left(|x_1\rangle \left(\cos\theta \, |z_2\rangle - \sin\theta \, |x_2\rangle \right) \right) \\ &= \frac{1}{\sqrt{2}} \sin\theta \end{aligned} \tag{11.94}$$

If photon 1 has polarization in the x direction, then it will not be transmitted by a polarizer in the z direction - it will be absorbed. Hence,

$$P_{-+}(\hat{a}, \hat{b}) = |\langle EPR \mid x_1 z_2' \rangle|^2 = \frac{1}{2} \sin^2\theta \tag{11.95}$$

Similarly,

$$P_{+-}(\hat{a}, \hat{b}) = |\langle EPR \mid z_1 x_2' \rangle|^2 = \frac{1}{2} \sin^2\theta \tag{11.96}$$

$$P_{--}(\hat{a}, \hat{b}) = |\langle EPR \mid x_1 x_2' \rangle|^2 = \frac{1}{2} \cos^2\theta \tag{11.97}$$

The correlation coefficient is then

$$\begin{aligned} C(\hat{a}, \hat{b}) &= P_{++}(\hat{a}, \hat{b}) + P_{--}(\hat{a}, \hat{b}) - P_{+-}(\hat{a}, \hat{b}) - P_{-+}(\hat{a}, \hat{b}) \\ &= \cos^2\theta - \sin^2\theta = 2\cos^2\theta - 1 = \cos 2\theta \end{aligned} \tag{11.98}$$

Since the unit vectors $\hat{a}$ and $\hat{b}$ are at an angle θ with respect to one another, it follows that $\hat{a} \cdot \hat{b} = \cos\theta$ and therefore

$$C(\hat{a}, \hat{b}) = 2\cos^2\theta - 1 = 2(\hat{a} \cdot \hat{b})^2 - 1 \tag{11.99}$$

(3) Suppose that the angle between the vectors $\hat{a}'$ and $\hat{a}$ is $x/2$, between $\hat{a}$ and $\hat{b}$ is $y/2$ and between $\hat{b}$ and $\hat{b}'$ is $z/2$. Then the angle between $\hat{a}'$ and $\hat{b}'$ is $(x + y + z)/2$ and according to quantum mechanics, the Bell coefficient has the form

$$B = \cos x + \cos y + \cos z - \cos(x + y + z) \tag{11.100}$$

This function has extrema when

$$\begin{aligned} \frac{\partial B}{\partial x} &= -\sin x + \sin(x + y + z) = 0 \\ \frac{\partial B}{\partial y} &= -\sin y + \sin(x + y + z) = 0 \\ \frac{\partial B}{\partial z} &= -\sin z + \sin(x + y + z) = 0 \end{aligned}$$

or

$$\sin x = \sin y = \sin z = \sin(x + y + z) \tag{11.101}$$

This has the solution

$$x = y = z \text{ and } 3x = \pi - x \rightarrow x = \pi/4 \tag{11.102}$$

For this extremum

$$B = 3\cos\frac{\pi}{4} - \cos\frac{3\pi}{4} = \frac{3}{\sqrt{2}} + \frac{1}{\sqrt{2}} = 2\sqrt{2} \tag{11.103}$$

This is a maximum, since at this point

$$\frac{\partial^2 B}{\partial x^2} = \frac{\partial^2 B}{\partial y^2} = \frac{\partial^2 B}{\partial z^2} = -\cos\frac{\pi}{4} + \cos\frac{3\pi}{4} = -\sqrt{2} < 0$$

(4) Let the vector $\hat{a}$ be at an angle θ_a with respect to some direction in the xz plane, and let $\hat{b}$ be at an angle θ_b with respect to the same direction. Then

$$\begin{aligned} C(\hat{a}, \hat{b}) &= \cos 2(\theta_a - \theta_b) \\ &= \cos 2\theta_a \cos 2\theta_b + \sin 2\theta_a \sin 2\theta_b \\ &= \int d\lambda\, \rho(\lambda) C(\hat{a}, \lambda) C(\hat{b}, \lambda) \end{aligned} \tag{11.104}$$

and with the assignments

$$\begin{aligned} &\rho(\lambda) = \delta(\lambda + 1) + \delta(\lambda - 1) \\ &C(\hat{a}, 1) = \cos 2\theta_a \quad , \quad C(\hat{a}, -1) = \sin 2\theta_a \\ &C(\hat{b}, 1) = \cos 2\theta_b \quad , \quad C(\hat{b}, -1) = \sin 2\theta_b \end{aligned}$$

we then see that

$$\begin{aligned} &\rho(\lambda) \geq 0 \\ &-1 \leq C(\hat{a}, \lambda)\ ,\ C(\hat{b}, \lambda) \leq 1 \text{for} \lambda = \pm 1 \end{aligned}$$

but

$$\int d\lambda\, \rho(\lambda) = 1 + 1 = 2$$

so that the normalization condition (b) is violated.

11.8 Non-Locality, EPR and Bell - a last time

As we discussed earlier, the second major problem confronting hidden variables and possessed properties was first understood in the context of the EPR paradox and then reinforced by the Bell inequalities. Let us look back at these ideas in light of the above discussions.

The original EPR analysis was rather complex in a technical sense and most discussions now use a simpler version due to Bohm. He considered a particle whose decay produces two spin$-1/2$ particles whose total spin angular momentum is zero. These particles move away from each other in opposite directions, and the components of their spins along various directions are subsequently measured by two observers, N and L, say. The constraint on the total spin means that if both observers agree to measure the spin along a certain direction $\hat{n}$, and if N measures $+\hbar/2$, then L will necessarily get the result $-\hbar/2$, and if N measures $-\hbar/2$, then L will necessarily get the result $+\hbar/2$.

There are no surprises if such correlations are analyzed in the context of classical physics. If one particle emerges from the decay with its internal angular momentum vector pointing along some particular direction, then because of conservation of angular momentum, the second particle is guaranteed to emerge with its spin vector pointing in the opposite direction. Thus, the 100% anti-correlations found in the measurements made by the two observers are simply the result of the fact that both particles possess actual, and (anti-)correlated, values of internal angular momentum and this is true from the time they emerge from the decay to the time the measurements are made. There are no paradoxes here, and everything is in accord with the simple realist view of classical physics.

The situation in quantum theory is radically different. Suppose first that the measurements are made along the $z-$axes of the two observers. The spin part of the state of the two particles can be written in terms of the associated eigenvectors as

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle |\downarrow\rangle - |\downarrow\rangle |\uparrow\rangle \right) \tag{11.105}$$

where, for example, $|\uparrow\rangle |\downarrow\rangle$ is the state in which particles 1 and 2 have spin $+\hbar/2$ and $-\hbar/2$ respectively. Thus

$$\hat{S}_z |\uparrow\rangle = +\hbar/2 |\uparrow\rangle \ , \ \hat{S}_z |\downarrow\rangle = -\hbar/2 |\downarrow\rangle \tag{11.106}$$

The pragmatic or instrumentalist interpretation of the entangled state $|\psi\rangle$ is straightforward. If, in a series of repeated measurements by N, a selection is made of the pairs of particles for which the measurement of particle gave spin-up, then - with probability one - a series of measurements by L on her particle in these pairs will yield spin-down. Similarly, if N finds spin-down then, with probability one, L will find spin-up. This correlation can be explained by saying that the measurements by N(computed with the operator $\hat{S}_z \otimes \hat{I}$) **cause** a reduction of the state vector from $|\psi\rangle$ to $|\uparrow\rangle |\downarrow\rangle$ or $|\downarrow\rangle |\uparrow\rangle$ respectively according to whether the spin-up or the spin-down result is selected. This new state is an eigenstate of the operator $\hat{I} \otimes \hat{S}_z$ associated with the second particle, and with an eigenvalue that is the opposite of the result obtained by N.

This description is acceptable within the confines of the pragmatic approach to quantum theory, but difficulties arise if one tries to enforce a more realist

interpretation of the above entangled state.

The obvious question is how the information about each observer's individual results *gets to* the other particle to guarantee that the result obtained by the second observer will be the correct one.

One might be tempted to invoke the reasoning of classical physics and argue that both particles *possess* the appropriate value all the time.

However, the only way in standard quantum theory of guaranteeing that a certain result will be obtained is if the state is an eigenvector of the observable concerned. But the state $|\psi\rangle$ above is *not* of this type. In fact, it displays the typical features of quantum entanglement - it is a superposition of states. Any attempt to invoke a hidden variable resolution will have to cope with the implications of the Kochen-Specker theorem.

There is also a question of whether this picture is compatible with special relativity. If the measurements by the two observers are space-like separated (which can be easily arranged) then which of them makes the first measurement and hence, in the standard interpretation, causes the state-vector reduction is clearly reference-frame dependent.

The problem is compounded by considering what happens if the observers decide to measure, say, the x-component of the spins, rather than the z-components. The state above can now be written in terms of $\hat{S}_x$ eigenvectors as

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|\leftarrow\rangle |\rightarrow\rangle - |\rightarrow\rangle |\leftarrow\rangle \right) \tag{11.107}$$

where $|\rightarrow\rangle$ and $|\leftarrow\rangle$ correspond to eigenvalues $+\hbar/2$ and $-\hbar/2$ respectively of the operator $\hat{S}_x$.

In one sense, this new entangled state is what might have been expected, and confirms that there is the same type of 100% anti-correlation between $\hat{S}_x$ measurements as that found for the observable $\hat{S}_z$. Indeed, this argument can be generalized to show that for any unit vector $\hat{n}$, the entangled state can be rewritten as a sum of two anti-correlated terms containing eigenvectors of the projection $\hat{n} \cdot \hat{S}$ of the spin operator along $\hat{n}$. Thus, if one adopts the classical type argument, one is obliged to conclude that both particles possessed exact values of spin along any axis from the moment they left the decay. This might not be easy to reconcile with the uncertainty relations associated with the angular momentum commutators.

EPR considered these issues, and concluded that the difficulties could be resolved in one of only two ways:

1. When N makes her measurement, the result communicates itself at once

in some way to particle 2, and converts its state into the appropriate eigenvector.

or

2. Quantum theory is incomplete and provides only a partial specification of the actual state of the system.

In contemplating the first possibility it must be appreciated that the two particles may have moved a vast distance apart before the first measurement is made and, therefore, any *at once* mode of communication would be in violent contradiction with the spirit (if not the law) of special relativity. It is not surprising that Einstein was not very keen on this alternative! An additional objection involves the lack within quantum theory itself of any idea about how this non-local effect is supposed to take place, so in this sense the theory would be incomplete anyway.

EPR came to the conclusion that the theory is indeed incomplete, although they left open the correct way in to *complete* it. One natural path is to suppose that there exist *hidden variables* whose values are not accessible to measurement in the normal way but which determine the actual values of what we normally regard as observables - in the same way as do the microstates in classical statistical physics.

However, it is not a trivial matter to construct a hidden variable theory that reproduces all of the empirical results of quantum mechanics (which are experimentally correct!). In particular, such a theory would need to explain why it is that certain observables are incompatible (those with non-vanishing commutator) in the sense that one cannot prepare a state of the system that violates the predictions of the uncertainty relations. In addition, there is the need to come to terms with the implications of the Kochen-Specker theorem.

Hidden variable theories capable of reproducing the results of conventional quantum theory do in fact exist (Bohm for example) but they exhibit a non-locality which is every bit as peculiar as that discussed above. One might think that this is a deficiency of these particular theories and that others might exist without this problem.

However, as we discussed earlier and will review here again, a very famous result of John Bell shows that this is not possible, that is, any hidden-variable theory that exactly replicates the results of quantum theory will necessarily possess striking non-local features.

This result is of major importance in understanding and appreciating the conceptual challenge posed by quantum theory.

11.8.1 The Bell Inequalities

The non-locality property we are about to discuss is not just a feature of hidden variables theories. It applies to *any* realist interpretation of quantum theory in which it is deemed meaningful to say that an individual system *possesses* values for its physical quantities in a way that is analogous to that in classical physics.

We will derive an inequality that is satisfied by certain correlation functions in any such theory *which is also local*. We will then see that the predictions (which are experimentally correct) of quantum theory violate this inequality.

The considerations of EPR were concerned with two observers who make measurements along the same axis. Bell found his famous inequalities by asking what happens if the observers measure the spin of the particles along different axes. In particular, we consider a pair of unit vectors $\hat{a}$ and $\hat{a}'$ for one observer and another pair $\hat{b}$ and $\hat{b}'$ for the other observer.

Now suppose a series of repeated measurements is made on a collection of systems whose quantum state is described by the entangled state vector $|\psi\rangle$. For example, we could look at a series of decays, each of which produces a pair of particles with zero total spin angular momentum. The central realist assumption we are testing is that each particle has a definite value at all times for any direction of spin. We let a_n denote $2/\hbar$ times the value of $\hat{a}\cdot\hat{S}$ possessed by particle 1 in the n^{th} element of the collection. Thus $a_n = \pm 1$ if $\hat{a}\cdot\hat{S} = \pm\hbar/2$.

The key ingredient in the derivation of the Bell inequalities is the correlation between measurements made by the two observers along these different directions. For directions $\hat{a}$ and $\hat{b}$ this is defined by

$$C(\hat{a},\hat{b}) := \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} a_n b_n \tag{11.108}$$

and similarly for the other directions. Note that if the results are always totally correlated(spins always in the same direction) then $C(\hat{a},\hat{b}) = +1$, whereas if they are totally anti-correlated(spins always in opposite directions) we get $C(\hat{a},\hat{b}) = -1$.

Now look at the quantity

$$g_n := a_n b_n + a_n b'_n + a'_n b_n + a'_n b'_n \tag{11.109}$$

For any member n of the collection, each term in this sum will take on the value $+1$ or -1 . Furthermore, the fourth term on the right hand side is equal to the product of the first three (because $(a_n)^2 = 1 = (b_n)^2$). Then thinking about the various possibilities shows that g_n can take on only the values ± 2. Therefore, the right hand side of the expression

$$\left|\frac{1}{N}\sum_{n=1}^{N} g_n\right| = \left|\frac{1}{N}\sum_{n=1}^{N} a_n b_n + \frac{1}{N}\sum_{n=1}^{N} a_n b'_n + \frac{1}{N}\sum_{n=1}^{N} a'_n b_n - \frac{1}{N}\sum_{n=1}^{N} a'_n b'_n\right|$$

representing the average value of g_n must be less than or equal to 2. Thus in the limit as $N \to \infty$, we get

$$\left|C(\hat{a},\hat{b}) + C(\hat{a},\hat{b}') + C(\hat{a}',\hat{b}) - C(\hat{a}',\hat{b}')\right| \leq 2 \tag{11.110}$$

which is one of the famous Bell inequalities.

It is important to emphasize the only assumptions that have gone into proving this inequality are:

1. For each particle it is meaningful to talk about the actual values of the projection of the spin along any direction.

2. There is locality in the sense that the value of any physical quantity is not changed by altering the position of a remote piece of measuring equipment.

This means that both occurrences of a_n in the expression for the average value of g_n have the same value, that is, they do not depend on the direction ($\hat{b}$ or $\hat{b}'$) along which the other observer chooses to measure the spin of particle 2. In particular, we are ruling out the type of context-dependent values that arose in our discussion of the Kochen-Specker theorem.

We will now show that the predictions of quantum theory violate this inequality over a range of directions for the spin measurements. The quantum mechanical prediction for the correlation between the spin measurements along axes $\hat{a}$ or $\hat{b}$ is

$$C(\hat{a},\hat{b}) := \left(\frac{2}{\hbar}\right)^2 \langle\psi|\,\hat{a}\cdot\hat{S}_{(1)} \otimes \hat{b}\cdot\hat{S}_{(2)}\,|\psi\rangle \tag{11.111}$$

where $\hat{S}_{(1)}$ and $\hat{S}_{(2)}$ are the spin operators for particles 1 and 2 respectively, and the tensor product is as we defined earlier in this chapter. Since the total angular momentum of the entangled vector $|\psi\rangle$ is zero, it is invariant under the unitary operators which generate rotations of the coordinate systems.

This means that $C(\hat{a},\hat{b})$ is a function of $\hat{a}\cdot\hat{b} = \cos\theta_{ab}$ only and, hence, there is no loss of generality in assuming that $\hat{a}$ points along the $z-$axis and that $\hat{b}$ lies in the $x-z$ plane. Then the expression for $C(\hat{a},\hat{b})$ becomes

$$C(\hat{a},\hat{b}) = \langle\psi|\,\sigma_{1z} \otimes (\sigma_{2z}\cos\theta_{ab} + \sigma_{2x}\sin\theta_{ab})\,|\psi\rangle \tag{11.112}$$

It is then straightforward to show that

$$C(\hat{a},\hat{b}) = -\cos\theta_{ab} \tag{11.113}$$

Now we restrict our attention to the special case in which (1) the four vectors $\hat{a},\hat{a}',\hat{b},\hat{b}'$ are coplanar and (2) $\hat{a}$ or $\hat{b}$ are parallel and (3) $\theta_{ab'} = \theta_{a'b} = \varphi$ say. Then the Bell inequality will be satisfied provided that

$$|1 + 2\cos\varphi - \cos 2\varphi| \leq 2 \tag{11.114}$$

This is violated for all values of φ between $0°$ and $90°$. This means that if the predictions of quantum theory are experimentally valid in this region then any idea of systems possessing individual values for observables must necessarily involve an essential non-locality. This applies in particular to any hidden variable theory that is completely consistent with the results of quantum theory. Thus, the important questions are:

1. Are the Bell inequalities empirically violated?

2. If so, are such violations in accord with the predictions of quantum theory?

In many experiments over the last two decades, the overwhelming conclusion is that the predictions of quantum theory *are* vindicated and so we are obliged either to stick with a pragmatic approach or a strict instrumentalist interpretation or else to accept the existence of a strange non-locality that seems hard to reconcile with our normal concepts of spatial separation between independent entities.

11.9 Bayesian Probability in QM

We turn to Bayesian probability arguments to deal with a realist.

11.9.1 Using Bayesian Ideas in Analysis of Experiments

In actual experimental tests, there are no infinite ensembles for accurate measurements of mean values. Experimental physicists perform a finite number of tests and then they state their results accompanied by a confidence level.

The real problem of theory versus experimental analysis is of a different nature however.

I am a theorist and I believe that quantum mechanics gives a reliable description of nature. I have a friend, however, who is a local realist.

We only have a finite number of trials of a Bell inequality experiment at our disposal.

How many tests are needed to make my realist friend feel uncomfortable?

The problem is not whether the validity of a Bell inequality can be salvaged by invoking clever loopholes, as some realists try to trick us into, but whether there can be any local realistic theory that reproduces the experimental results.

To simplify the discussion, I will assume that there are ideal detectors and that the rate at which particles are produced by the apparatus is perfectly known.

Experimental Results Change Beliefs

First, we consider a yes-no test.

Quantum mechanics(QM) predicts that the probability of the **yes** result is q and an alternative local realistic(LR) theory predicts a probability r.

An experimental test is performed n times and yields m **yes** results.

What can we infer about the likelihood of the two theories?

The answer is given by Baye's theorem

$$P(B|A \wedge C) = P(A|B \wedge C)\frac{P(B|C)}{P(A|C)} \tag{11.115}$$

Denote by $p'_q = P(q|I)$ and $p'_r = P(r|I)$ the prior probabilities that we assign to the validity of the two theories. These are subjective probabilities expressing our personal beliefs.

For example, if my friend is willing to bet 100 to 1 (for example) that LR is correct and QM is wrong, then

$$\frac{p'_r}{p'_q} = 100 \tag{11.116}$$

The question is: how many experimental tests are needed to change my friend's opinion to

$$\frac{p''_r}{p''_q} = 0.01 \tag{11.117}$$

say, before he is driven to bankruptcy. This is a reversal (in belief) by a factor of 10^4.

In this case, $P(r|\{m,n\} \wedge I) = p''_r$ is the new *prior* probability for my friend after the experiments are finished and similarly we have $P(q|\{m,n\} \wedge I) = p''_q$.

If we define

$$E_r = P(\{m,n\}|q \wedge I) \quad , \quad E_q = P(\{m,n\}|q \wedge I) \tag{11.118}$$

which are just the probabilities of the experimentally found result (the actual data - m successes in n trials) according to the two theories.

These follow from the binomial theorem

$$E_r = \frac{n!}{m!(n-m)!}r^m(1-r)^{n-m} \quad , \quad E_q = \frac{n!}{m!(n-m)!}q^m(1-q)^{n-m} \tag{11.119}$$

It then follows from Baye's theorem that

$$P(\{m,n\}|r\wedge I)P(r|I) = P(r|\{m,n\}\wedge I)P(\{m,n\}|I) \tag{11.120}$$
$$P(\{m,n\}|q\wedge I)P(q|I) = P(q|\{m,n\}\wedge I)P(\{m,n\}|I) \tag{11.121}$$

or that

$$\frac{P(\{m,n\}|r\wedge I)P(r|I)}{P(r|\{m,n\}\wedge I)} = P(\{m,n\}|I) = \frac{P(\{m,n\}|q\wedge I)P(q|I)}{P(q|\{m,n\}\wedge I)}$$

or

$$\frac{P(\{m,n\}|r\wedge I)P(r|I)}{P(\{m,n\}|q\wedge I)P(q|I)} = \frac{p'_r E_r}{p'_q E_q} = \frac{P(r|\{m,n\}\wedge I)}{P(q|\{m,n\}\wedge I)} = \frac{p''_r}{p''_q} \tag{11.122}$$

We define the ratio

$$\frac{P(\{m,n\}|q\wedge I)}{P(\{m,n\}|r\wedge I)} = \frac{E_q}{E_r} = D = \left(\frac{q}{r}\right)^m \left(\frac{1-q}{1-r}\right)^{n-m} \tag{11.123}$$

as the *confidence depressing factor* for the hypothesis LR with respect to the hypothesis QM.

11.9.2 Simple Example

Suppose that we flip coins and the yes-no question is: Did the coin come up *heads*?

I, the theorist, will assume that the coin is unbiased and that therefore $q = 0.5$ and $m = n/2$ (assuming that I am correct). We then have

$$D = \left(\frac{1}{2r}\right)^{n/2} \left(\frac{1}{2(1-r)}\right)^{n/2} = \left(\frac{1}{2}\right)^n \left(\frac{1}{r(1-r)}\right)^{n/2} \tag{11.124}$$

Since we want 10^4, we find

$$\left(\frac{5}{3}\right)^n = 10^4 \rightarrow n\log\frac{5}{3} = 4 \rightarrow n = \frac{4}{\log\frac{5}{3}} = \frac{4}{0.22} \approx 16 \tag{11.125}$$

So that it would take only 16 coin flips to reverse my untrusting friend's belief.

Now let us return to the Bell inequality.

11.9.3 Simple Ideas

Let us consider a device that has three widely separated detectors each of which has two switch settings as shown in Figure 11.28 above.

When a detector is triggered it flashes either red or green. The detectors are far apart from the source, there are no connections between the detectors and no connections between the source and the detectors other than those mediated

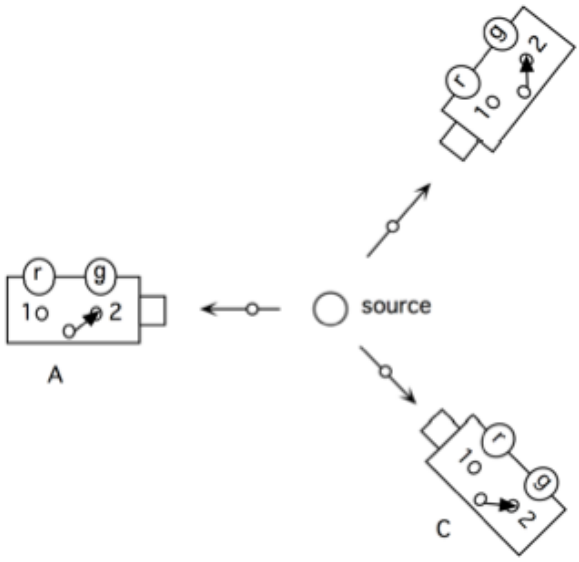


Figure 11.28: Three Detector Experiment

by the group of three particles (as shown) that originate at the source and fly away, one to each detector.

A run of the experiment consists of setting the switch on each detector to one of its two positions (labeled 1 or 2), pressing a button at the source (to release a trio of particles, one aimed at each detector), and recording the color subsequently flashed at each detector.

There are eight possible switch settings:

$$111\ 112\ 121\ 122\ 211\ 212\ 221\ 222 \tag{11.126}$$

We consider only the data acquired for four of the eight possible switch settings, namely, those in which the number of detectors set to 1 is odd.

$$111\ 122\ 212\ 221 \tag{11.127}$$

The other set

$$112\ 122\ 212\ 222 \tag{11.128}$$

will lead to similar results ($1 \leftrightarrow 2$). As shown in Figure 16.28 we call the detectors A, B, and C, and specify pertinent facts about them by listing three pieces of information (switch settings or colors flashed) in that order.

If we run the experiment many times, then the observational results are the following. If just one detector is set to 1 (and the others to 2), then an odd

number of red lights *always* flash, that is, either all three detectors flash red or there is one red flash and two green ones.

If all three detectors are set to 1, then an odd number of red lights is *never* observed to flash - either two of the three flash red or all three flash green.

This is summarized by the table:

Settings	Result 1	Result 2	Result 3	Result 4
111	GGG	GRR	RGR	RRG
122	RGG	GRG	GGR	RRR
212	RGG	GRG	GGR	RRR
221	RGG	GRG	GGR	RRR

Table 11.3: Results

All four outcomes are *equally likely* in each case(this particular detail is not important).

We will discuss a real, physical system that exhibits this behavior later.

Let us set aside, for the moment, the 111 case and consider the 122, 212, and 221 cases in which just one detector is set to 1. Because an odd number of red lights always flash in any of these three cases, whenever the switches are so set we can predict with *certainty* what one of the three detectors will do in a run, merely by noting what happens to the other two. For should the other two flash the same color (RR or GG), then the third will have to flash red, but should the other two flash different colors (RG or GR), then the third will have to flash green.

Now we follow the path set out by EPR to draw an inference that will seem inescapable. Along the way we will use the so-called *EPR reality criterion*.

Since there are no direct connections between the detectors, their behavior can *only be coordinated* due to the fact that all three are triggered by particles that came from a *common source*. This fact and this fact alone must contain the explanation for why we can learn in advance what color will flash at a given detector, say A, from measurements made far away at B and C. Information telling the detector at A what color to flash in order to maintain the observed consistency with the colors flashed at B and C must somehow be *encoded* in the particle that triggers A. Since that particle could indeed have been coordinated with the particles that triggered B and C when all three were back at their common source, this explanation seems *both inevitable and entirely reasonable*.

We can apply this reasoning to any one of the three detectors (by moving it

farther from the source so that before it flashes we have had the opportunity to observe what colors flash at the other two). We conclude that in each run of the experiment each particle *must* be carrying to its detector *instructions* on what color to flash, and that an odd number of the particles must specify red.

Thus, for a given choice of the switch settings (say 122) the particles heading for detectors A, B, and C must respectively be carrying instructions RRR, RGG, GRG, or GGR, but never GRR, RGR, RRG, or GGG.

Which of the four allowed groups of instructions they collectively carry is revealed only when the lights flash. All of the above reasoning applies equally well, of course, to 212 and 221 runs.

In the absence of connections between the detectors and the source, a particle has no information about how the switch of its detector will be set until it arrives there. Since in each run any detector might turn out to be the one set to 1 or one of the ones set to 2, to preserve the perfect record of always having an odd number of red flashes in 122, 212, and 221 runs, it would *seem to be essential* for each particle to be carrying instructions for how its detector should flash for *either* of the two possible switch settings it might find upon arrival.

The existence of *instructions* of this sort is the *EPR reality criterion*.

The instructions carried by each particle can be symbolized by a pair of letters

$$\begin{array}{llllll} 1 \rightarrow & R & R & G & or & G \\ 2 \rightarrow & R & G & R & & G \end{array}$$

would result in RRR if the switch settings were 122, GGR for 212, and GRG for 221.

Since each of the three possible switch settings result in an odd number of red flashes, this is indeed a legal set of instructions.

An example of an illegal set of instructions is

$$\begin{array}{lll} R & R & G \\ G & R & R \end{array}$$

for this gives an even number of red flashes GRR for the switch setting 212, which is never observed.

Since there are eight ways the lights can flash, namely,

$$\text{RRR RRG RGG RGR GRR GRG GGR GGG} \tag{11.129}$$

the total number of possible instruction sets is $8 \times 8 = 64$.

It is not hard to enumerate all the legal (odd number of red flashes) instruction sets.

First note that three of the six positions in a legal instruction set corresponding to any one of the three choices 122, 212, or 221 for the switch settings, must contain an odd number of R's, since that particular setting might be encountered in any run, and since only odd numbers of red flashes are ever observed. Thus, the only possible entries for the positions corresponding to the switch settings 122 are (leaving blank the entries not relevant to those three settings):

$$\begin{array}{cccccc} & ABC & ABC & ABC & ABC \\ 1\rightarrow & R\,-\,- & R\,-\,- & G\,-\,- & G\,-\,- \\ 2\rightarrow & -\,R\,R & -\,GG & -\,RG & -\,GR \end{array}$$

so that 122 gives RRR, RGG, GRG, or GGR independent of the other entries.

We can next count the way to fill in the blanks in these four forms so as to produce the correct data for switch settings 221. Since each of the four already specifies the color flashed at detector B for setting 2, namely, R G R G, to ensure that any 221 run produces an odd number of red flashes there are only two choices available for the other two (A and C) unspecified 221 entries for each of the four forms: RR or GG if the specified entry is R and RG or GR if the specified entry is G so that we have

$$\begin{array}{cccccc} & \text{ABC} & \text{ABC} & \text{ABC} & \text{ABC} \\ 1\rightarrow & \text{R–R} & \text{R–G} & \text{G–G} & \text{G–R} \\ 2\rightarrow & \text{RRR} & \text{RGG} & \text{GRG} & \text{GGR} \\ & \text{ABC} & \text{ABC} & \text{ABC} & \text{ABC} \\ 1\rightarrow & \text{R–R} & \text{R–R} & \text{G–R} & \text{G–G} \\ 2\rightarrow & \text{GRR} & \text{GGG} & \text{RRG} & \text{RGR} \end{array}$$

This raises the number of possible forms to eight, each of which leaves only the entry for setting 1 at the detector B unspecified. But that entry is now entirely determined by the entries at settings 2 for detectors A and C (having to be R, if the latter two entries are the same color and G, if they are different).

$$\begin{array}{cccccc} & \text{ABC} & \text{ABC} & \text{ABC} & \text{ABC} \\ 1\rightarrow & \text{RRR} & \text{RGG} & \text{GRG} & \text{GGR} \\ 2\rightarrow & \text{RRR} & \text{RGG} & \text{GRG} & \text{GGR} \\ & \text{ABC} & \text{ABC} & \text{ABC} & \text{ABC} \\ 1\rightarrow & \text{RGR} & \text{RRR} & \text{GGR} & \text{GRG} \\ 2\rightarrow & \text{RRR} & \text{GGG} & \text{RRG} & \text{RGR} \end{array}$$

They are arranged in the same horizontal order as the forms in (1), with the two possibilities for each form placed directly above one another. It is easy to

check explicitly that each instruction set (2) does indeed give an odd number of red flashes when a single detector is set to 1.

122	gives	RRR	RGG	GRG	GGR
		RRR	RGG	GRG	GGR
221	gives	RRR	RGG	GRG	GGR
		GRG	GGR	RRR	RGG
212	gives	RRR	RGG	GRG	GGR
		GGR	GRG	RGG	RRR

Clearly, (2) represents the eight legal sets.

Now, finally, we consider the fourth type of run, in which all three detectors are set to 1, and an odd number of red flashes is never observed.

The above instruction sets must determine the outcomes of these runs as well. For who is to prevent somebody from flipping the two switches set to 2 over to 1, just before the particles arrive?

An inspection of the upper rows in (2) reveals that *every one of the eight allowed instruction sets results in an odd number of red flashes when all three switches are set to 1.*

If the instruction sets existed, then 111 runs would *always* have to produce an odd number of red flashes. But they *never* do.

Thus, a *single* 111 run suffices all by itself to give data inconsistent with the otherwise compelling inference of instruction sets.

Here the instruction sets(realistic theory) *require* an odd number of red flashes in *every* 111 run, but quantum mechanics(experiment) *prohibits* an odd number of red flashes in *every* 111 run.

Something is wrong with the EPR idea of *instruction* sets or EPR reality.

The Quantum Mechanical Explanation

Here is how the device works. What emerges from the source are three spin−1/2 particles (a, b, and c) in a spin state whose structure is given below. The particles fly apart to the detectors in the horizontal plane. We define the z−direction for each particle to be along the line of flight. The detectors contain Stern-Gerlach magnets which measure the vertical (x) component of the spin when the switch is set to 1 and the horizontal component (y) perpendicular to the line

of flight when their switch is set to 2. They are set so that we get red flashes for spin-up, and green flashes for spin-down.

Let us describe a spin state that produces the remarkable correlations (GHZ-state) described earlier.

We measure angular momentum for each particle in units of ? so that the spin operators for each particle can be taken to be the Pauli matrices. Now consider the three commuting Hermitian operators

$$\sigma_x^a \sigma_y^b \sigma_y^c \quad , \quad \sigma_y^a \sigma_x^b \sigma_y^c \quad , \quad \sigma_y^a \sigma_y^b \sigma_x^c \tag{11.130}$$

They commute because all pairs of the six spin operators out of which they are constructed commute, except for those associated with the x and y components of the spin of a single particle, which anticommute. This does not cause any trouble, however, because converting the product in one order to the product in then other order always involves and even number of such anticommuting exchanges.

Being commuting and Hermitian, the three operators above can be provided with simultaneous eigenvectors. Since the square of each operator is the identity, the eigenvalues of each can only be ±1.

The actual spin state that produces the remarkable correlations (the *Greenberger-Horne-Zeilinger or GHZ-state*) is described by

$$|GHZ\rangle = \frac{1}{\sqrt{2}} \left(|1,1,1\rangle - |-1,-1,-1\rangle \right) \tag{11.131}$$

where ±1 specifies spin-up or spin-down along the appropriate $z-$axis.

For simplicity in the following argument, here we pick the state with all three eigenvalues equal to +1, which preserves the symmetry among the particles. The argument works for any such symmetric state and for any linear combination of such states as in the above state.

Since the components of the spin vectors of different particles commute, we can simultaneously measure the x component for one particle and the y components for the other two. Because the spin state is an eigenvector of all three of the operators

$$\sigma_x^a \sigma_y^b \sigma_y^c \quad , \quad \sigma_y^a \sigma_x^b \sigma_y^c \quad , \quad \sigma_y^a \sigma_y^b \sigma_x^c$$

with eigenvalue +1, the product of the results of each of the three single spin measurements has to be +1, regardless of which particle we pick for the x spin measurement. Since +1 flashes red and −1 flashes green, there must indeed be an even number of green flashes and thus an odd number of red flashes.

What about the result of three $x-$spin measurements, declared earlier never to

result in an odd number of red flashes? Translating this into spin language tells us that the product of the three results must always be −1. The Hermitian operator corresponding to that product is

$$\sigma_x^a \sigma_x^b \sigma_x^c \tag{11.132}$$

so for the declaration to be correct, it must be that the eigenvector of the first three operators with eigenvalue +1 is also an eigenvector of the last operator(above) with eigenvalue −1.

This is easily confirmed. Indeed, the last operator is just minus the product of the other three operators

$$\sigma_x^a \sigma_x^b \sigma_x^c = -(\sigma_x^a \sigma_y^b \sigma_y^c)(\sigma_y^a \sigma_x^b \sigma_y^c)(\sigma_y^a \sigma_y^b \sigma_x^c) \tag{11.133}$$

Since we are in an eigenvector with eigenvalue +1 of each of the three operators appearing on the right, we are indeed also in an eigenvector of $\sigma_x^a \sigma_x^b \sigma_x^c$ with eigenvalue −1.

The consequence of the EPR reality criterion specified earlier, if translated into quantum theoretic terminology, would also assert that the state was an eigenvector of the operator $\sigma_x^a \sigma_x^b \sigma_x^c$, *but with the wrong eigenvalue.* In this sense, the GHZ experiment provides the strongest possible contradiction between quantum mechanics and the EPR reality criterion.

Alternatively, we can say it this way. We may measure, on each particle, either σ_x or σ_y, without disturbing the other particles. The results of these measurements will be called m_x or m_y, respectively. From

$$\sigma_x^a \sigma_y^b \sigma_y^c |111\rangle = |111\rangle$$
$$\sigma_y^a \sigma_x^b \sigma_y^c |111\rangle = |111\rangle$$
$$\sigma_y^a \sigma_y^b \sigma_x^c |111\rangle = |111\rangle$$

and

$$\sigma_x^a \sigma_x^b \sigma_x^c |111\rangle = -|111\rangle$$

we can predict with certainty that, if the three σ_x are measured, the results satisfy

$$m_{ax} m_{bx} m_{cx} = -1 \tag{11.134}$$

Therefore, each of the operators σ_x^a, σ_x^b and σ_x^c corresponds to an EPR element of reality, because its value can be predicted with certainty by performing measurements on the two *other*, distant particles.

However, it follows from

$$\sigma_x^a \sigma_y^b \sigma_y^c |111\rangle = |111\rangle$$
$$\sigma_y^a \sigma_x^b \sigma_y^c |111\rangle = |111\rangle$$
$$\sigma_y^a \sigma_y^b \sigma_x^c |111\rangle = |111\rangle$$

that we can predict with certainty the value of σ_x^a by measuring σ_y^b and σ_y^c rather than σ_x^b and σ_x^b. We then have

$$m_{ax}m_{by}m_{cy} = +1 \tag{11.135}$$

and likewise, by cyclic permutation,

$$m_{ay}m_{bx}m_{cy} = +1 \tag{11.136}$$

and

$$m_{ay}m_{by}m_{cx} = +1 \tag{11.137}$$

The product of the last four results gives

$$m_{ax}m_{bx}m_{cx}m_{ax}m_{by}m_{cy}m_{ay}m_{bx}m_{cy}m_{ay}m_{by}m_{cx} = -1 \tag{11.138}$$

$$(m_{ax})^2(m_{bx})^2(m_{cx})^2(m_{by})^2(m_{cy})^2(m_{ay})^2 = -1 \tag{11.139}$$

But,

$$(m_{jx})^2 = 1 \tag{11.140}$$

so that we get a contradiction.

There is a tacit assumption in the above argument, that m_{ax} in $m_{ax}m_{bx}m_{cx} = -1$ is the same as m_{ax} in $m_{ax}m_{by}m_{cy} = +1$, in spite of the fact that these two ways of obtaining m_{ax} involve mutually exclusive experiments - measuring σ_x^b and σ_x^c or measuring σ_y^b and σ_y^c.

This tacit assumption is of counterfactual nature, and cannot be experimentally verified. It obviously adheres to the EPR reality criterion - but is simply wrong!

Saying it another way, the crucial minus sign in

$$\sigma_x^a\sigma_x^b\sigma_x^c = -(\sigma_x^a\sigma_y^b\sigma_y^c)(\sigma_y^a\sigma_x^b\sigma_y^c)(\sigma_y^a\sigma_y^b\sigma_x^c)$$

which is totally destructive of the possibility of these instruction sets, comes from the fact that in working out the identity it is necessary to interchange the anticommuting operators σ_x^b and σ_y^b in order to get rid of all the y components (using $(\sigma_y^i)^2 = 1$) and be left with a product of three x components. It is only that one instance of uncompensated anticommutation that produces the conclusion so devastating to the hypothesis of instruction sets.

This is extremely pleasing, for it is just the fact the x and y components of the spin of a single particle do not commute, which leads the well-educated quantum mechanician to reject from the start the inference instruction sets (which have to specify the value of both of these non-commuting observables), making it necessary for me to disguise what was going on earlier so that you would not have dismissed this discussion as rubbish before reaching the interesting part.

There is no other Bellian refutation of EPR in which the mathematical details of the refutation so closely reflect the broad interpretive doctrines of quantum theory that EPR tried to challenge. The entries in the instruction sets are precisely the conjectured c–number values for all the σ_x^i and σ_y^i–values that appear to be the only explanation for the remarkable correlations. In addition, the logic of the red and green lights in the simple model precisely parallels the algebraic behavior of the four operators used here except for that one devastating anticommutation.

Let us return now to using Bayesian ideas to convince our realist friends about the validity of quantum mechanics within the context of the Bell inequalities.

11.9.4 More about the Greenberger-Horne-Zeilinger(GHZ) State

We consider the GHZ state for a three-particle system

$$|GHZ\rangle = \frac{1}{\sqrt{2}}\left(|1,1,1\rangle - |-1,-1,-1\rangle\right) \tag{11.141}$$

where -1 and $+1$ denote any two orthogonal states of each of the three particle subsystems.

We have three distant observers examine the three subsystems. The first observer has the choice of two tests. The first test can give two different results that we label $a = \pm 1$, and likewise the other test yields $a' = \pm 1$. Symbols b, b', c and c' are similarly defined for the other two observers. Any possible values of their results satisfy

$$a'bc = ab'c = abc' = -a'b'c' = +1 \tag{11.142}$$

Mermin has then shown that we have the inequality

$$-2 \le \langle a'bc + ab'c + abc' - a'b'c' \rangle \le +2 \tag{11.143}$$

As we saw above quantum mechanics makes a very simple prediction for the GHZ state: there are well chosen tests that give with certainty

$$a'bc = ab'c = abc' = -a'b'c' = +1 \tag{11.144}$$

It is important to remember that performing any such test can verify the value 1 for only one of these products (at a time) since each product corresponds to a different experimental setup.

If, however, we take all these results together, they manifestly conflict with $a'bc + ab'c + abc' - a'b'c' = \pm 2$.

Many physicists have erroneously, at this point, stated that a single experiment is sufficient to invalidate local realism. This is sheer nonsense: a single

experiment can only verify one occurrence of one of the terms in

$$a'bc = ab'c = abc' = -a'b'c' = +1 \tag{11.145}$$

What does our realist friend think?

He believes that, in each experimental run, each term in the above result has a definite value even if that term is not actually measured in that run.

We ask him to propose a rule giving the *average* values of the products in

$$a'bc + ab'c + abc' - a'b'c' = \pm 2 \tag{11.146}$$

Suppose that he assumes

$$\langle a'bc \rangle = \langle ab'c \rangle = \langle abc' \rangle = \langle -a'b'c' \rangle = 0.5 \tag{11.147}$$

This clearly attains the right hand side of (Mermin's) inequality. This assumption then leads to the prediction that if we measure $a'bc$ we shall find the result 1 (that is, *yes*) in 75% of the cases and the opposite result in 25% and like wise for the other tests. This simply corresponds to the averages proposed above.

In our earlier discussion about confidence depressing factors, this corresponds to

$$q = 1 \text{ and } r = 0.75 \tag{11.148}$$

If we assume that quantum mechanics is correct, then $m = n$ (that is what $q = 1$ means). Therefore, we have

$$D = \left(\frac{q}{r}\right)^m \left(\frac{1-q}{1-r}\right)^{n-m} = \left(\frac{1}{r}\right)^n = \left(\frac{4}{3}\right)^n \tag{11.149}$$

Therefore it would take

$$D = \left(\frac{4}{3}\right)^n = 10^4 \rightarrow n = \frac{4}{\log(1.33)} \approx 32 \tag{11.150}$$

tests to undo the realist's beliefs.

11.10 Problems

11.10.1 Bell Inequality with Stern-Gerlach

A pair of spin−1/2 particles is produced by a source. The spin state of each particle can be measured using a Stern-Gerlach apparatus (see diagram below).

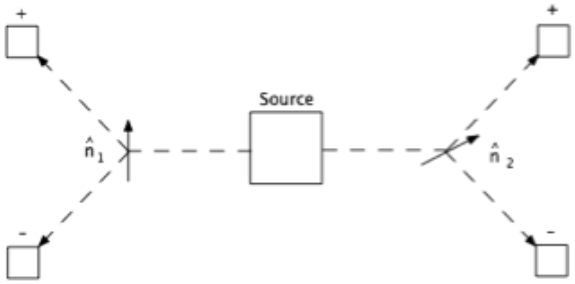


Figure 11.29: EPR Stern-Gerlach Setup

(a) Let $\hat{n}_1$ and $\hat{n}_2$ be the field directions(arrows in diagram) of the Stern-Gerlach magnets. Consider the commuting observables

$$\sigma^{(1)} = \frac{2}{\hbar}\hat{n}_1 \cdot \vec{S}_1 \quad , \quad \sigma^{(2)} = \frac{2}{\hbar}\hat{n}_2 \cdot \vec{S}_2$$

corresponding to the spin component of each particle along the direction of the Stern-Gerlach apparatus associated with it. What are the possible values resulting from measurement of these observables and what are the corresponding eigenstates?

(b) Consider the observable $\sigma^{(12)} = \sigma^{(1)} \otimes \sigma^{(2)}$ and write down its eigenvectors and eigenvalues. Assume that the pair of particles is produced in the singlet state

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left(|S_z+\rangle^{(1)}|S_z-\rangle^{(2)} - |S_z-\rangle^{(1)}|S_z+\rangle^{(2)}\right)$$

What is the expectation value of $\sigma^{(12)}$?

(c) Make the assumption that the spin of a particle has a meaningful value even when it is not being measured. Assume also that the only possible results of the measurement of a spin component are $\pm\hbar/2$. Then show that the probability of finding the spins pointing in two given directions will be proportional to the overlap of the hemispheres that these two directions define. Quantify this criterion and calculate the expectation value of $\sigma^{(12)}$.

(d) Assume the spin variables depend on a *hidden variable* λ. The expectation value of the spin observable $\sigma^{(12)}$ is determined in terms of the normalized distribution function $f(\lambda)$:

$$\left\langle \sigma^{(12)} \right\rangle = \frac{4}{\hbar^2}\int d\lambda f(\lambda) S_z^{(1)}(\lambda) S_\varphi^{(2)}(\lambda)$$

Prove *Bell's inequality*

$$\left|\left\langle \sigma^{(12)}(\varphi)\right\rangle - \left\langle \sigma^{(12)}(\varphi')\right\rangle\right| \leq 1 + \left|\left\langle \sigma^{(12)}(\varphi - \varphi')\right\rangle\right|$$

(e) Consider Bell's inequality for $\varphi' = 2\varphi$ and show that it is not true when applied in the context of quantum mechanics.

11.10.2 Bell's Theorem with Photons

Two photons fly apart from one another, and are in oppositely oriented circularly polarized states. One strikes a polaroid film with axis parallel to the unit vector $\hat{a}$, the other a polaroid with axis parallel to the unit vector $\hat{b}$. Let $P_{++}(\hat{a},\hat{b})$ be the joint probability that both photons are transmitted through their respective polaroids. Similarly, $P_{--}(\hat{a},\hat{b})$ is the probability that both photons are absorbed by their respective polaroids, $P_{+-}(\hat{a},\hat{b})$ is the probability that the photon at the $\hat{a}$ polaroid is transmitted and the other is absorbed, and finally, $P_{-+}(\hat{a},\hat{b})$ is the probability that the photon at the $\hat{a}$ polaroid is absorbed and the other is transmitted.

The classical realist assumption is that these probabilities can be separated:

$$P_{ij}(\hat{a},\hat{b}) = \int d\lambda \rho(\lambda) P_i(\hat{a},\lambda) P_j(\hat{b},\lambda)$$

where i and j take on the values + and −, where λ signifies the so-called hidden variables, and where $\rho(\lambda)$ is a weight function. This equation is called the separable form.

The correlation coefficient is defined by

$$C(\hat{a},\hat{b}) = P_{++}(\hat{a},\hat{b}) + P_{--}(\hat{a},\hat{b}) - P_{+-}(\hat{a},\hat{b}) - P_{-+}(\hat{a},\hat{b})$$

and so we can write

$$C(\hat{a},\hat{b}) = \int d\lambda \rho(\lambda) C(\hat{a},\lambda) C(\hat{b},\lambda)$$

where

$$C(\hat{a},\lambda) = P_+(\hat{a},\lambda) - P_-(\hat{a},\lambda) \quad , \quad C(\hat{b},\lambda) = P_+(\hat{b},\lambda) - P_-(\hat{b},\lambda)$$

It is required that

(a) $\rho(\lambda) \geq 0$

(b) $\int d\lambda \rho(\lambda) = 1$

(c) $-1 \leq C(\hat{a},\lambda) \leq 1$, $-1 \leq C(\hat{b},\lambda) \leq 1$

The Bell coefficient

$$B = C(\hat{a},\hat{b}) + C(\hat{a},\hat{b}') + C(\hat{a}',\hat{b}) - C(\hat{a}',\hat{b}')$$

combines four different combinations of the polaroid directions.

(1) Show that the above classical realist assumptions imply that $|B| \leq 2$

(2) Show that quantum mechanics predicts that $C(\hat{a},\hat{b}) = 2\left(\hat{a}\cdot\hat{b}\right)^2 - 1$

(3) Show that the maximum value of the Bell coefficient is $2\sqrt{2}$ according to quantum mechanics

(4) Cast the quantum mechanical expression for $C(\hat{a},\hat{b})$ into a separable form. Which of the classical requirements, (a), (b), or (c) above is violated?

11.10.3 Bell's Theorem with Neutrons

Suppose that two neutrons are created in a singlet state. They fly apart; the spin of one particle is measured in the direction a, the other in the direction b.

(a) Calculate the relative frequencies of the coincidences $R(up, up)$, $R(up, down)$, $R(down, up)$ and $R(down, down)$, as a function of θ, the angle between a and b.

(b) Calculate the correlation coefficient

$$C(a, b) = R(up, up) - R(up, down) - R(down, up) + R(down, down)$$

(c) Given two possible directions, a and a', for one measurement, and two possible directions, b and b', for the other, deduce the maximum possible value of the Bell coefficient, defined by

$$B = C(a, b) + C(a', b) + C(a', b') - C(a, b')$$

(d) Show that this prediction of quantum mechanics is inconsistent with classical local realism.

11.10.4 Greenberger-Horne-Zeilinger State

The Greenberger-Horne-Zeilinger (GHZ) state of three identical spin−1/2 particles is defined by

$$|GHZ\rangle = \frac{1}{\sqrt{2}} \left(|z_a+\rangle \, |z_b+\rangle \, |z_c+\rangle - |z_a-\rangle \, |z_b-\rangle \, |z_c-\rangle \right)$$

where z_a+ is the eigenvector of the $z-$component of the spin operator of particle a belonging to eigenvalue $+\hbar/2$ ($z-$spin up), z_a- is the eigenvector of the $z-$component of the spin operator of particle a belonging to eigenvalue $-\hbar/2$ ($z-$spin down), and similarly for b and c. Show that, if spin measurements are made on the three particle in the $x-$ or $y-$directions,

(a) the product of three spins in the $x-$direction is always $-\hbar^3/8$

(b) the product of two spins in the $y-$direction and one spin in the $x-$direction is always $+\hbar^3/8$

(c) Consider a prize game for a team of three players, A, B, and C. The players are told that they will be separated from one another and that each will be asked one of two questions, say X or Y, to which each must give one of two allowed answers, namely, +1 or −1. Moreover, *either*

   (a) all players will be asked the same question X

   or

   (b) one of the three players will be asked X and the other two Y

After having been asked X or Y, no player may communicate with the others until after all three players have given their answers, +1 or −1. To win the game, the players must give answers such that, in case (a) the product of the three answers is −1, whereas in case (b) the product of the three answers is +1.

(a) Show that no classical strategy gives certainty of a win for the team

(b) Show that a quantum strategy, in which each player may take one of the GHZ particles with her, exists for which a win is certain

Chapter 12

Identical Particles

12.1 Theoretical ideas

We now apply these quantum mechanical methods we have developed to multi-electron atoms.

We will create a model to handle these atoms that follows from the one-electron case considered earlier. These systems are very complex and all the results that we derive will be approximations.

If we consider an atom or system of N particles, the wave function

$$\begin{aligned}\psi(1,2,3,4,.....,N,t) &= \psi(\vec{r}_1 s_1, \vec{r}_2 s_2, \vec{r}_3 s_3, \vec{r}_4 s_4, ..., \vec{r}_N s_N, t) \\ &= \langle 1,2,3,4,.....,N,t \mid \psi \rangle \qquad (12.1)\end{aligned}$$

where $\vec{r}_j s_j =$ (radius vector,spin) of the j^{th} particle

describing the system will be a function of $3N$ spatial coordinates, time and all of the particle spin variables. The $3N$ spatial coordinates form a multi-dimensional configuration space.

The Hamiltonian of the system is given by

$$\hat{H} = \hat{T} + \hat{V} \qquad (12.2)$$

where

$$\hat{T} = \sum_{j=1}^{N} \hat{T}_j = -\sum_{j=1}^{N} \left(\frac{\hbar^2}{2m_j} \nabla_j^2 \right) \text{ and } \hat{V} = \hat{V}(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N, t) \qquad (12.3)$$

The time-dependent Schrodinger equation is

$$\hat{H}\psi(1,2,3,4,.....,N,t) = i\hbar \frac{\partial}{\partial t} \psi(1,2,3,4,.....,N,t) \qquad (12.4)$$

The probability density is defined as

$$\rho(1,2,3,4,.....,N,t) = \psi^*(1,2,3,4,.....,N,t)\psi(1,2,3,4,.....,N,t) \qquad (12.5)$$

so that

$$\rho(1,2,3,4,.....,N,t)d^3\vec{r}_1 d^3\vec{r}_2.....d^3\vec{r}_N = \text{ probability of finding}$$
$$\text{particle 1 at } \vec{r}_1 \text{ in } d^3\vec{r}_1 \text{ , and particle 2 at } \vec{r}_2 \text{ in } d^3\vec{r}_2 \text{ ,}$$
$$\text{....., and particle } N \text{ at } \vec{r}_N \text{ in } d^3\vec{r}_N \text{ all at time time } t$$

We assume that the N-particle wave function is normalized.

The energy eigenstates or stationary states are solutions of

$$\hat{H}\psi_E(1,2,3,4,.....,N) = E\psi_E(1,2,3,4,.....,N) \tag{12.6}$$

which implies the time dependence

$$\psi(1,2,3,4,.....,N,t) = \psi_E(1,2,3,4,.....,N)e^{-i\frac{E}{\hbar}t} \tag{12.7}$$

where E is the energy of the system.

When we consider an $N-$electron atom, the system really has $N+1$ particles (we must include the nucleus). However, the nucleus is so much more massive than the electrons that we can make the approximation that it has infinite mass and is fixed.

We put the nucleus of charge Ze at the origin and define

$$\vec{r}_j = \text{ position vector of the } j^{th} \text{ electron}$$
$$r_{jk} = |\vec{r}_j - \vec{r}_k| = \text{ separation between the } j^{th} \text{ and } k^{th} \text{ electrons}$$

The potential energy is

$$V(1,2,3,.....,N) = -\sum_{j=1}^{N}\frac{Ze^2}{r_j} + \sum_{j=1}^{N}\sum_{i>j}^{N}\frac{e^2}{r_{ij}} \tag{12.8}$$
$$= \text{ Coulomb energy between nucleus and electrons}$$
$$+ \text{ Coulomb energy between electrons}$$

We will assume no spin-dependence or time dependence in the potential energy.

All electrons in the atom are considered to be *identical* or *indistinguishable*. This means that there are no interactions that can, in any way, distinguish them from each other.

Alternatively, we can say that, if we interchange the coordinates and spins of two particles, then it is not possible to determine via any physical measurement that any change was made in the system.

This says that all measurable quantities or the operators representing them must

remain unchanged by the interchange of indistinguishable particles.

In particular, the Hamiltonian must remain unchanged, i.e., we must have

$$\hat{H}(1,2,3,4,...,j,k,...,N) = \hat{H}(1,2,3,4,...,k,j,...,N) \tag{12.9}$$

This property of $\hat{H}$ is called *exchange symmetry*. Operators that have this property are *symmetric* functions of their indices $1,2,3,4,.....,N$ and they are called *symmetric* operators.

Now, every symmetry of a physical system must be represented by an operator that commutes with $\hat{H}$. In this case, we introduce the *particle interchange* or *permutation operator* $\hat{P}_{ij}$ such that

$$\begin{aligned}\hat{P}_{ij}\psi(1,2,3,.....,i,j,.....,N) &= \langle 1,2,3,.....,i,j,.....,N|\,\hat{P}_{ij}\,|\psi\rangle \\ &= \langle 1,2,3,.....,j,i,.....,N\mid\psi\rangle = \psi(1,2,3,.....,j,i,.....,N)\end{aligned}$$

In words, we say

$\hat{P}_{ij}\psi(1,2,3,.....,i,j,.....,N)$ gives the amplitude for
finding the j^{th} particle at $\vec{r}_i$ with spin s_i
and i^{th} particle at $\vec{r}_j$ with spin s_j

Now, the transformed Hamiltonian operator is given by

$$\hat{H}' = \hat{P}_{ij}\hat{H}\hat{P}_{ij}^{-1} = \hat{H} \text{ (by assumption)} \tag{12.10}$$

This implies that

$$\hat{P}_{ij}\hat{H} = \hat{H}\hat{P}_{ij} \rightarrow \left[\hat{H},\hat{P}_{ij}\right] = 0 \tag{12.11}$$

as we expected. The same result holds for all symmetric operators.

Now, suppose that the state vector $|\psi\rangle$ is an eigenvector of the symmetric, $N-$particle $\hat{H}$ with energy E. We then have

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle \tag{12.12}$$

$$\hat{H}\hat{P}_{ij}\,|\psi\rangle = \hat{P}_{ij}\hat{H}\,|\psi\rangle = E\hat{P}_{ij}\,|\psi\rangle \tag{12.13}$$

which says that

$$\hat{P}_{ij}\,|\psi\rangle \text{ is also an eigenvector of } \hat{H} \text{ with the same energy} \tag{12.14}$$

This holds for any pair (i,j). So $\hat{H}$ and $\hat{P}_{ij}$ share a common eigenbasis as expected. This phenomenon is called *exchange degeneracy*.

For simplicity, we assume that $N = 2$. We then have

$$\hat{H}(1,2) \text{ and } \psi(1,2) \tag{12.15}$$

and

$$\left[\hat{H}, \hat{P}_{12}\right] = 0 \tag{12.16}$$

What are the simultaneous eigenfunctions? We have

$$\hat{P}_{12}\psi(1,2) = \psi(2,1)$$
$$\hat{P}_{12}^2\psi(1,2) = \hat{P}_{12}\psi(2,1) = \psi(1,2)$$

which says that

$$\hat{P}_{12}^2 = \hat{I} \tag{12.17}$$

and that $\hat{P}_{12}$ has eigenvalues ± 1. Now if

$$\hat{H}\psi(1,2) = E\psi(1,2) \tag{12.18}$$

then

$$\hat{H}\psi(2,1) = E\psi(2,1) \tag{12.19}$$

and these two state functions are degenerate. Then, we can write

$$\psi_S(1,2) = \psi(1,2) + \psi(2,1) \rightarrow \text{ symmetric wave function}$$
$$\psi_A(1,2) = \psi(1,2) - \psi(2,1) \rightarrow \text{ antisymmetric wave function}$$

which are the simultaneous eigenfunctions with

$$\hat{H}\psi_S = E\psi_S \qquad \hat{H}\psi_A = E\psi_A \tag{12.20}$$
$$\hat{P}_{12}\psi_S = +\psi_S \qquad \hat{P}_{12}\psi_A = -\psi_A \tag{12.21}$$

It is an *experimental fact* that the behavior of wave functions under pairwise particle interchange depends only on the kind of particles involved, in particular on their *spin*.

All known particles divide themselves in to *two classes*:

1. Bosons $\rightarrow$ particles with integer spin, $s = 0, 1, 2, 3, 4, ...$
2. Fermions $\rightarrow$ particles with half-integer spin, $s = 1/2, 3/2, 5/2, ...$

and

1. Fermions have antisymmetric wave functions under particle interchange
2. Bosons have symmetric wave functions under particle interchange

This relationship between spin and wave function symmetry cannot be proved in non-relativistic quantum mechanics. It can, however, be proved if we add relativity and construct the relativistic waves equations for bosons and fermions.

As we shall see, this symmetry/antisymmetry connection of spin and wave functions will generalize to more complex systems with more particles.

Before proceeding to study real atoms with N electrons, let us see what we can learn from a one-dimensional systems containing either two identical bosons or two identical fermions.

The general Hamiltonian for a one-dimensional two-particle system is

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \hat{U}(x_1 - x_2) \tag{12.22}$$

$$\hat{H}_1 = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_1^2} + \hat{V}(x_1) \tag{12.23}$$

$$\hat{H}_2 = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_2^2} + \hat{V}(x_2) \tag{12.24}$$

where

$$\hat{U}(x_1 - x_2) = \text{ the particle - particle interaction} \tag{12.25}$$

We will assume that $\hat{U}(x_1 - x_2)$ is small enough that we can apply perturbation theory. We then use direct product states and write

$$\hat{H} = \hat{H}_0 + \hat{U} \tag{12.26}$$

$$\hat{H}_1\psi_{n_1}^{(0)}(x_1) = E_{n_1}^{(0)}\psi_{n_1}^{(0)}(x_1) \tag{12.27}$$

$$\hat{H}_2\psi_{n_2}^{(0)}(x_2) = E_{n_2}^{(0)}\psi_{n_2}^{(0)}(x_2) \tag{12.28}$$

$$\begin{aligned}\hat{H}\psi_{n_1n_2}^{(0)}(x_1,x_2) &= \hat{H}_0\psi_{n_1}^{(0)}(x_1)\psi_{n_2}^{(0)}(x_2)\\ &= (\hat{H}_1 + \hat{H}_2)\psi_{n_1}^{(0)}(x_1)\psi_{n_2}^{(0)}(x_2)\\ &= (E_{n_1}^{(0)} + E_{n_2}^{(0)})\psi_{n_1}^{(0)}(x_1)\psi_{n_2}^{(0)}(x_2)\\ &= E_{n_1n_2}^{(0)}\psi_{n_1n_2}^{(0)}(x_1,x_2)\end{aligned} \tag{12.29}$$

We will construct the unperturbed(zero order) eigenfunctions and energies from these direct product states.

For the moment, we will also ignore spin.

The simple direct product states will not work for a description of the two particle system since the eigenfunctions of $\hat{H}_0$ must be either symmetric or antisymmetric under particle interchange.

The correct choice is ψ_S or ψ_A where

$$\psi_{n_1n_2}^{(0)S} = \frac{1}{\sqrt{2}}\left[\psi_{n_1}^{(0)}(x_1)\psi_{n_2}^{(0)}(x_2) + \psi_{n_1}^{(0)}(x_2)\psi_{n_2}^{(0)}(x_1)\right] \tag{12.30}$$

$$\psi_{n_1n_2}^{(0)A} = \frac{1}{\sqrt{2}}\left[\psi_{n_1}^{(0)}(x_1)\psi_{n_2}^{(0)}(x_2) - \psi_{n_1}^{(0)}(x_2)\psi_{n_2}^{(0)}(x_1)\right] \tag{12.31}$$

Both of these states have energy $E_{n_1n_2}^{(0)} = E_{n_1}^{(0)} + E_{n_2}^{(0)}$.

12.2 Bosons with Spin = 0

We assume that $s_1 = s_2 = 0$. This says that there are no new degrees of freedom and hence no reason to change the wave functions.

Indistinguishable bosons of spin = 0 require a symmetric wave function and thus we choose as the properly *symmetrized* zero-order wave functions

$$\psi^{(0)S}_{n_1 n_2} = \frac{1}{\sqrt{2}}\left[\psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2) + \psi^{(0)}_{n_1}(x_2)\psi^{(0)}_{n_2}(x_1)\right] \tag{12.32}$$

The ground state corresponds to $n_1 = n_2 = 1$ or

$$\psi^{(0)S}_{11} = \psi^{(0)}_1(x_1)\psi^{(0)}_1(x_2) \tag{12.33}$$

In perturbation theory, the first order energy is then

$$\begin{aligned} E_{11} &= 2E^{(0)}_1 + \left\langle \psi^{(0)S}_{11}\middle|\hat{U}(x_1 - x_2)\middle|\psi^{(0)S}_{11}\right\rangle \\ &= 2E^{(0)}_1 \\ &\quad + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx_1 dx_2 dx'_1 dx'_2 \\ &\quad \times \left\langle \psi^{(0)S}_{11}\middle|\left(\left|\psi^{(0)}_1(x_2)\right\rangle\middle|\psi^{(0)}_1(x_1)\right\rangle\left\langle\psi^{(0)}_1(x_1)\right|\left\langle\psi^{(0)}_1(x_2)\right|\right) \\ &\quad \times \hat{U}(x_1 - x_2)\left(\left|\psi^{(0)}_1(x'_1)\right\rangle\middle|\psi^{(0)}_1(x'_2)\right\rangle\left\langle\psi^{(0)}_1(x'_1)\right|\left\langle\psi^{(0)}_1(x'_2)\right|\right)\middle|\psi^{(0)S}_{11}\right\rangle \end{aligned} \tag{12.34}$$

Now

$$\begin{aligned} &\left\langle\psi^{(0)}_1(x_1)\right|\left\langle\psi^{(0)}_1(x_2)\right|\hat{U}(x_1 - x_2)\left|\psi^{(0)}_1(x'_1)\right\rangle\left|\psi^{(0)}_1(x'_2)\right\rangle \\ &\qquad = U(x_1 - x_2)\delta(x_1 - x'_1)\delta(x_2 - x'_2) \end{aligned} \tag{12.35}$$

which implies that

$$\begin{aligned} E_{11} &= 2E^{(0)}_1 \\ &\quad + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx_1 dx_2 \left\langle\psi^{(0)S}_{11}\middle|\left(\left|\psi^{(0)}_1(x_2)\right\rangle\middle|\psi^{(0)}_1(x_1)\right\rangle\right) \\ &\quad \times U(x_1 - x_2)\left(\left\langle\psi^{(0)}_1(x_1)\right|\left\langle\psi^{(0)}_1(x_2)\right|\right)\left|\psi^{(0)S}_{11}\right\rangle \end{aligned} \tag{12.36}$$

or

$$E_{11} = 2E^{(0)}_1 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx_1 dx_2 \left|\psi^{(0)}_1(x_1)\right|^2 U(x_1 - x_2)\left|\psi^{(0)}_1(x_2)\right|^2 \tag{12.37}$$

For later use we define the general direct integral

$$J_{n_1 n_2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx_1 dx_2 \left|\psi^{(0)}_{n_1}(x_1)\right|^2 U(x_1 - x_2)\left|\psi^{(0)}_{n_2}(x_2)\right|^2 \tag{12.38}$$

In this case, we have

$$E_{11} = 2E_1^{(0)} + J_{11} \tag{12.39}$$

Now we look at the first excited state of this system. We assume that for the zero-order states, the first excited state corresponds to $n_1 = 1$ and $n_2 = 2$. Therefore, the zero-order symmetric wave function for the first excited state is

$$\psi_{12}^{(0)S} = \frac{1}{\sqrt{2}}\left[\psi_1^{(0)}(x_1)\psi_2^{(0)}(x_2) + \psi_1^{(0)}(x_2)\psi_2^{(0)}(x_1)\right] \tag{12.40}$$

and the first order energy is

$$E_{12} = E_1^{(0)} + E_2^{(0)} + \left\langle \psi_{12}^{(0)S} \middle| \hat{U}(x_1 - x_2) \middle| \psi_{12}^{(0)S} \right\rangle \tag{12.41}$$

Using the same procedure as before we get

$$E_{12} = E_1^{(0)} + E_2^{(0)} + J_{12} + K_{12} \tag{12.42}$$

where

$$K_{n_1 n_2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx_1 dx_2 \psi_{n_1}^{(0)*}(x_1)\psi_{n_2}^{(0)}(x_1)U(x_1 - x_2)\psi_{n_2}^{(0)*}(x_2)\psi_{n_1}^{(0)}(x_2) \tag{12.43}$$

is called the *exchange integral*.

Now let us look at a possible physical meaning of these direct and exchange integrals.

We define

$$\left|\psi_{n_1}^{(0)}(x_1)\right|^2 = \rho_1 = \text{ probability density for particle 1 in state } n_1 \tag{12.44}$$

and

$$\left|\psi_{n_2}^{(0)}(x_2)\right|^2 = \rho_2 = \text{ probability density for particle 2 in state } n_2 \tag{12.45}$$

Therefore, the direct integrand takes the form

$$\rho_1 \rho_2 U(r_{12}) \tag{12.46}$$

To see what this means let

$$U(r_{12}) = \frac{e^2}{r_{12}} \tag{12.47}$$

which corresponds to a repulsive Coulomb potential. The direct integral is then

$$\int\int \frac{(e\rho_1)(e\rho_2)}{r_{12}}\, dx_1\, dx_2 \tag{12.48}$$

This represents the total energy of two classical charge distributions interacting with the potential energy $U(r_{12})$.

The exchange integral, however, has no such classical counterpart. It is the result of symmetrizing the wave function and therefore arises because of the invariance of $\hat{H}$ with respect to particle interchange.

The energy level diagram to first order might look like Figure 12.1 below.

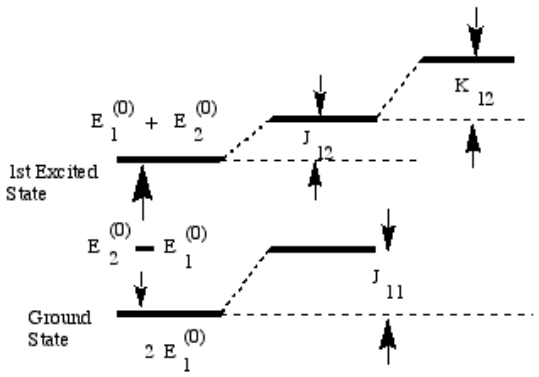


Figure 12.1: Typical Boson Energy Level Diagram

The more interesting case is a two spin = 1/2 fermion system (since electrons are spin = 1/2 fermions).

12.3 Spin = 1/2 Fermions

The particles now have internal degrees of freedom. The single particle state vectors must now have both spatial and spin parts

$$|space\rangle\,|spin\rangle \tag{12.49}$$

For example,

$$\left|\psi_{n_1}^{(0)}\right\rangle|+\rangle_1 \tag{12.50}$$

presents a fermion in the $\psi_{n_1}^{(0)}$ spatial state with spin *up*.

We write the corresponding wave function as

$$\left\langle x_1 \,\middle|\, \psi_{n_1}^{(0)}\right\rangle|+\rangle_1 = \psi_{n_1}^{(0)}(x_1)\alpha(1) \tag{12.51}$$

and so on, where we define the labels $\alpha(j) = |+\rangle_j$ and $\beta(j) = |-\rangle_j$.

We must choose the antisymmetric combination for the zero-order wave func-

tions. We have 4 possible direct product states given n_1 and n_2, i.e.,

$$\psi_1(1,2) = \psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2)\alpha(1)\alpha(2) \tag{12.52}$$

$$\psi_2(1,2) = \psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2)\alpha(1)\beta(2) \tag{12.53}$$

$$\psi_3(1,2) = \psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2)\beta(1)\alpha(2) \tag{12.54}$$

$$\psi_4(1,2) = \psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2)\beta(1)\beta(2) \tag{12.55}$$

These are not antisymmetric, however. A useful operator allows us to construct antisymmetric states. Consider the operator

$$\hat{R} = \frac{1}{\sqrt{2}}(1 - \hat{P}_{12}) \tag{12.56}$$

Now for any function $A(1,2)$ we have

$$\hat{R}A(1,2) = \frac{1}{\sqrt{2}}(1 - \hat{P}_{12})A(1,2) = \frac{1}{\sqrt{2}}\left[A(1,2) - A(2,1)\right] \tag{12.57}$$

which is antisymmetric. The factor $1/\sqrt{(}2)$ keeps the state normalized. We now use $\hat{R}$ to construct four antisymmetric states from the four direct product states (12.52).

$$\begin{aligned}\psi^{(0)}_{n_1 n_2 ++}(x_1,x_2) &= \frac{1}{\sqrt{2}}(1 - \hat{P}_{12})\psi_1(1,2) \\ &= \frac{1}{\sqrt{2}}\left[\psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2)\alpha(1)\alpha(2) - \psi^{(0)}_{n_2}(x_1)\psi^{(0)}_{n_1}(x_2)\alpha(1)\alpha(2)\right]\end{aligned} \tag{12.58}$$

$$\begin{aligned}\psi^{(0)}_{n_1 n_2 +-}(x_1,x_2) &= \frac{1}{\sqrt{2}}(1 - \hat{P}_{12})\psi_2(1,2) \\ &= \frac{1}{\sqrt{2}}\left[\psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2)\alpha(1)\beta(2) - \psi^{(0)}_{n_2}(x_1)\psi^{(0)}_{n_1}(x_2)\alpha(2)\beta(1)\right]\end{aligned} \tag{12.59}$$

$$\begin{aligned}\psi^{(0)}_{n_1 n_2 -+}(x_1,x_2) &= \frac{1}{\sqrt{2}}(1 - \hat{P}_{12})\psi_3(1,2) \\ &= \frac{1}{\sqrt{2}}\left[\psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2)\beta(1)\alpha(2) - \psi^{(0)}_{n_2}(x_1)\psi^{(0)}_{n_1}(x_2)\alpha(1)\beta(2)\right]\end{aligned} \tag{12.60}$$

$$\begin{aligned}\psi^{(0)}_{n_1 n_2 --}(x_1,x_2) &= \frac{1}{\sqrt{2}}(1 - \hat{P}_{12})\psi_4(1,2) \\ &= \frac{1}{\sqrt{2}}\left[\psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2)\beta(1)\beta(2) - \psi^{(0)}_{n_2}(x_1)\psi^{(0)}_{n_1}(x_2)\beta(1)\beta(2)\right]\end{aligned} \tag{12.61}$$

where the subscripts imply

+ + means both spins *up*

+ - or - + means one spin *up* and one spin *down*

- - means both spins *down*

Each of these wave functions is antisymmetric and each is an eigenfunction of $\hat{H}_0$ (since $\hat{H}_0$ does not contain any spin dependent terms) with the same energy. This implies that, at this point, we have a 4-fold degenerate zero-order system with energy $E_{n_1}^{(0)} + E_{n_2}^{(0)}$.

We could use these states as the zero-order wave function to start perturbation theory. It would be like doing the spin-orbit calculation using the $|\ell s m_\ell m_s\rangle$ basis, rather than the $|\ell s j m_j\rangle$ basis where $\hat{H}_{so}$ is diagonal. It is always important to choose zero-order wave functions, if it is not to difficult to do, that incorporate as much of the symmetry of the system as possible. In other words, choose zero-order wave functions that are simultaneous eigenstates of the maximal set of commuting observables. This will hopefully produce a diagonal perturbation matrix or at least so many zeros that it is easy to diagonalize the rest of the matrix.

In this case, we not only have $[\hat{H}, \hat{P}_{12}] = 0$ which told us to choose antisymmetric zero-order states, but we also have $[\hat{H}, \vec{S}_{op}^2] = 0$ and $[\hat{H}, \hat{S}_z] = 0$ where

$\vec{S}_{op} = \vec{S}_{1,op} + \vec{S}_{2,op} =$ the total spin angular momentum

$\hat{S}_z = \hat{S}_{1z} + \hat{S}_{2z} =$ the z - component of the total spin angular momentum

Therefore we should choose antisymmetric state functions which are also eigenfunctions of $\hat{H}_0$, $\vec{S}_{op}^2$ and $\hat{S}_z$ as our zero-order states.

From our earlier work we know that the possible values of the total spin are $S = 0, 1$ and the state vectors that are eigenstates of $\vec{S}_{op}^2$ and $\hat{S}_z$ are

$$|1,1\rangle = \alpha(1)\alpha(2) = \chi_{11} \tag{12.62}$$

$$|1,0\rangle = \frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) + \alpha(2)\beta(1)\right) = \chi_{10} \tag{12.63}$$

$$|1,-1\rangle = \beta(1)\beta(2) = \chi_{1,-1} \tag{12.64}$$

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) - \alpha(2)\beta(1)\right) = \chi_{00} \tag{12.65}$$

Notice that the $\chi_{1,m=\pm1,0}$ are symmetric under $\hat{P}_{12}$ and χ_{00} is antisymmetric.

Therefore, we will maintain overall antisymmetry by writing the wave functions as products of spatial wave function and spin functions such that the spatial function is symmetric when combined with χ_{00} and the spatial function is antisymmetric when combined with $\chi_{1,m=\pm1,0}$.

The symmetric spatial function is

$$\psi^{(0)S}_{n_1 n_2} = \frac{1}{\sqrt{2}}\left[\psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2) + \psi^{(0)}_{n_1}(x_2)\psi^{(0)}_{n_2}(x_1)\right] \tag{12.66}$$

and the antisymmetric spatial function is

$$\psi^{(0)A}_{n_1 n_2} = \frac{1}{\sqrt{2}}\left[\psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2) - \psi^{(0)}_{n_1}(x_2)\psi^{(0)}_{n_2}(x_1)\right] \tag{12.67}$$

The four zero-order wave functions, which are now eigenfunctions of $\hat{P}_{12}$, $\hat{H}_0$, $\vec{S}^2_{op}$ and $\hat{S}_z$ are then

$$\psi^{(0)}_{n_1 n_2 00} = \frac{1}{\sqrt{2}}\left[\psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2) + \psi^{(0)}_{n_1}(x_2)\psi^{(0)}_{n_2}(x_1)\right]\chi_{00} \tag{12.68}$$

$$\psi^{(0)}_{n_1 n_2 1 m_s} = \frac{1}{\sqrt{2}}\left[\psi^{(0)}_{n_1}(x_1)\psi^{(0)}_{n_2}(x_2) - \psi^{(0)}_{n_1}(x_2)\psi^{(0)}_{n_2}(x_1)\right]\chi_{1 m_s} \quad m_s = \pm 1, 0 \tag{12.69}$$

Notice that if we have identical spatial states, i.e., $n_1 = n_2$, the $S = 1$ states vanish identically. This says that two fermions in an $S = 1$ spin state cannot be in the same spatial state(the wavefunction vanishes). This is the first example of a general principle we will discuss later called the *Pauli Exclusion Principle*.

An alternative way to find these zero-order wave functions is to go back to first principles and use CG coefficients. For example

$$\psi^{(0)}_{n_1 n_2 s m_s} = \sum_{\substack{m_{s_1}, m_{s_2} \\ m_{s_1} + m_{s_2} = m_s}} a_{m_{s_1}, m_{s_2}} \psi^{(0)}_{n_1 n_2 m_{s_1} m_{s_2}} \tag{12.70}$$

where

$$a_{m_{s_1} m_{s_2}} = \langle s_1 s_2 m_{s_1} m_{s_2} \mid s_1 s_2 s m_s \rangle \tag{12.71}$$

Now

$$\left\langle \frac{1}{2}\frac{1}{2} m_{s_1} m_{s_2} \;\middle|\; \frac{1}{2}\frac{1}{2} 1 1 \right\rangle = \delta_{m_{s_1}, \frac{1}{2}} \delta_{m_{s_2}, \frac{1}{2}} \tag{12.72}$$

which implies that

$$\psi^{(0)}_{n_1 n_2 1 1} = \psi^{(0)}_{n_1 n_2 + +} \tag{12.73}$$

as written above.

Similarly, for $s = 1, m_s = 0$, the only nonzero CG coefficients are

$$\left\langle \frac{1}{2}\frac{1}{2}\frac{1}{2} -\frac{1}{2} \;\middle|\; \frac{1}{2}\frac{1}{2} 1 0 \right\rangle = \frac{1}{\sqrt{2}} = \left\langle \frac{1}{2}\frac{1}{2} -\frac{1}{2}\frac{1}{2} \;\middle|\; \frac{1}{2}\frac{1}{2} 1 0 \right\rangle \tag{12.74}$$

which implies that

$$\begin{aligned} \psi^{(0)}_{n_1 n_2 1 0} &= \frac{1}{\sqrt{2}}\psi^{(0)}_{n_1 n_2 + -} + \frac{1}{\sqrt{2}}\psi^{(0)}_{n_1 n_2 - +} \\ &= \frac{1}{\sqrt{2}}\left[\psi^{(0)}_{n_1}(1)\psi^{(0)}_{n_2}(2) - \psi^{(0)}_{n_1}(2)\psi^{(0)}_{n_2}(1)\right]\chi_{10} \end{aligned} \tag{12.75}$$

as written above. We now have the appropriate zero-order wave functions and can apply perturbation theory to the two fermion system.

As with the two boson case, the zero-order ground state for two fermions corresponds to $n_1 = n_2 = 1$ with zero-order energy $2E_1^{(0)}$. Since the $S = 1$ or triplet states have identically zero state functions in the case (since the spatial function are antisymmetric), we have $\psi^{(0)}_{11,1m_s} = 0$. The unperturbed ground state must then have $S = 0, S_z = 0$ or it is $\psi^{(0)}_{11,00}(1,2)$. This involves a singlet state with $m_s = 0$ only

$$\chi_{00} = \frac{1}{\sqrt{2}}\left(\alpha(1)\beta(2) - \beta(1)\alpha(2)\right) \tag{12.76}$$

In this state, the particle spins are always opposite or antiparallel.

The ground state energy to first order is

$$\begin{aligned} E_{11} &= 2E_1^{(0)} + \left\langle \psi^{(0)}_{11,00} \middle| \hat{U}(x_1 - x_2) \middle| \psi^{(0)}_{11,00} \right\rangle \\ &= 2E_1^{(0)} + J_{11} \end{aligned} \tag{12.77}$$

which is the *same* energy as in the two boson system(we are assuming the *same Hamiltonian* applies).

The spatial part of the wave function is the same also, namely,

$$\psi_1^{(0)}(1)\psi_2^{(0)}(2) \tag{12.78}$$

We must use a symmetric spatial wave function here because the spin vector is antisymmetric in the ground state of two fermions. The presence of the spin internal degrees of freedom(and the Pauli principle) has a more dramatic effect on the first excited state for two fermions.

We again assume that the first excited state corresponds to $n_1 = 1, n_2 = 2$. This gives the energy to first order as

$$E_{12} = E_1^{(0)} + E_2^{(0)} + \left\langle \psi^{(0)}_{12,sm_s} \middle| \hat{U}(x_1 - x_2) \middle| \psi^{(0)}_{12,sm_s} \right\rangle \tag{12.79}$$

We can write the energy this way, i.e., we do not need to write a 4×4 matrix $\langle \hat{U} \rangle$ and diagonalize it because the $\langle \hat{U} \rangle$ matrix is already diagonal in this basis due the orthogonality of the spin functions and the fact that the perturbing potential does not depend on spin. This first order energy is different for the triplet and singlet states. If we do the integrals (they are the same as the boson case) we get

$$E_{12} = E_1^{(0)} + E_2^{(0)} + J_{12} \pm K_{12} \tag{12.80}$$

where

$$\begin{aligned} + &\rightarrow \text{ singlet } s = 0, m_s = 0 \\ - &\rightarrow \text{ triplet } s = 1, m_s = \pm 1, 0 \end{aligned}$$

All the triplet states have the same energy because they have the same spatial wave function and the perturbing potential does not depend on spin.

We thus get the energy level structure shown in Figure 12.2 below.

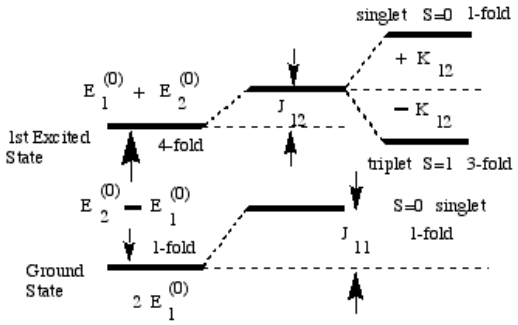


Figure 12.2: Typical Fermion Energy Level Diagram

The energies now depend on the total spin S *even though* the Hamiltonian $\hat{H}$ *does not* explicitly depend on spin. A very dramatic effect!!

This level splitting is not due to any additional terms added to the Hamiltonian such as $\hat{H}_{so}$ or $\hat{H}_{Zeeman}$. This effect is strictly due to *symmetry* requirements. The requirement of symmetry or antisymmetry forced on the spatial wave functions by the symmetry or antisymmetry of the spin vectors causes this level splitting. The entire effect is due to the invariance of the Hamiltonian under pairwise particle interchange.

Physically, we can argue as follows:

1. symmetric spatial functions are large for $x_1 \approx x_2$, while antisymmetric spatial functions are ≈ 0 for $x_1 \approx x_2$

2. $U(x_1 - x_2)$ is expected to be largest for $x_1 \approx x_2$

3. this implies that for $S = 1$ fermions $\langle \hat{U} \rangle$ is relatively small while for $S = 0$ fermions $\langle \hat{U} \rangle$ is relatively large

4. two identical fermions with antiparallel spins have a large probability of being close together – they *attract* each other

5. two identical fermions with parallel spins have zero probability of being close together – they *repel* each other

This *repulsion* is spin dependent and not due to the Coulomb repulsion between the electrons.

The first order energy for the singlet state is larger than for the triplet states because the repulsive interaction is enhanced in the singlet state. This overall effect is called *spin pairing* and it is a purely quantum mechanical effect.

12.4 The N-Electron Atom

We now extend our discussion to a system with N electrons (fermions). We write

$$\hat{H} = \hat{H}_0 + \hat{H}' \tag{12.81}$$

where

$$\hat{H}_0 = \sum_{i=1}^{N} \left[-\frac{\hbar^2}{2m_e} \nabla_i^2 + \hat{V}(\vec{r}_i) \right] \tag{12.82}$$

$$\hat{H}' = \sum_{i=1}^{N} \sum_{j>i}^{N} \hat{U}(\vec{r}_i - \vec{r}_j) \tag{12.83}$$

The energy eigenstates for the $N-$electron atom are solutions of the time independent Schrodinger equation

$$\hat{H}\psi_E = E\psi_E \tag{12.84}$$

where $\psi_E = \psi_E(1, 2, 3, 4, ..., N)$ and $1 = (\vec{r}_1, s_1)$ and so on.

The indistinguishability of the N electrons implies that

$$\left[\hat{H}, \hat{P}_{ij}\right] = 0 \qquad i, j = 1, 2, 3, 4, ..., N \; ; \; i \neq j \tag{12.85}$$

where the $\hat{P}_{ij}$ interchange *all attributes* of the electrons, i.e., both the spatial and spin degrees of freedom.

The wave function must be antisymmetric under pairwise electron interchange

$$\hat{P}_{ij}\psi_E = -\psi_E \qquad i, j = 1, 2, 3, 4, ..., N \; ; \; i \neq j \tag{12.86}$$

The general problem of N interacting electrons is very complex. At this stage we only want to extract general properties that will also hold in real 3−dimensional atomic systems. It turns out to be instructive to consider the case of non-interacting electrons - the so-called independent particle model. In this model we neglect the electron-electron interactions and look only at the zeroth order.

In particular, we consider N identical non-interacting particles in a potential well $V(\vec{r})$. The Hamiltonian for any particle in the well is

$$\hat{H}_0(k) = \frac{\vec{p}_{k,op}^{\,2}}{2m} + V(\vec{r}_k) \tag{12.87}$$

where

$$\hat{H}_0(k)\phi_n(\vec{r}_k) = \varepsilon_n\phi_n(\vec{r}_k) \qquad n = 0, 1, 2, 3, 4, ... \tag{12.88}$$

Thus, any single particle sees the energy level structure as shown in Figure 12.3 below.

Figure 12.3: Single Particle Energy Level Structure

The N-particle Hamiltonian is then

$$\hat{H} = \hat{H}_0(1) + \hat{H}_0(2) + \hat{H}_0(3)..... + \hat{H}_0(N) \tag{12.89}$$

with solutions given by

$$\hat{H}\psi(1, 2, 3, ..., N) = E\psi(1, 2, 3, ..., N) \tag{12.90}$$

where

$$\psi(1, 2, 3, ..., N) = \phi_a(1)\phi_b(2).....\phi_n(N) \tag{12.91}$$

and

$$E = \varepsilon_a + \varepsilon_b + ... + \varepsilon_n \tag{12.92}$$

This solution implies that

particle 1 is in state a with energy ε_a
particle 2 is in state b with energy ε_b
.....
.....
.....
particle N is in state n with energy ε_n

Electrons have spin = 1/2. Thus, corresponding to any single particle energy level, say a, there are two possible single particle states, namely,

$$\phi_a(1)\alpha(1) \text{ and } \phi_a(1)\beta(1) \tag{12.93}$$

From now on when we write $\phi_a(1)$, where the subscript a will be understood to also include the spin information.

The simple product state solutions are not physically admissible solutions since

they are not antisymmetric under particle interchange for any two particles.

All such states with particles interchanged pairwise have the same energy. In fact, any permutation of the indices produces a state with the same energy. We need to construct a completely antisymmetric linear combination of all of these solutions.

If these were bosons we would have to construct a completely symmetric linear combination of all these solutions.

If we define

$$\wp\psi(1,2,3,4,.....,N) = \text{ a permutation of the particles} \tag{12.94}$$

then, the completely symmetric state is easy to construct. It is

$$\psi_S(1,2,3,.....,N) = \sum_{\wp} \wp\psi(1,2,3,4,.....,N) \tag{12.95}$$

where the sum means a sum over all possible permutations or arrangements. There are $N!$ such permutations.

Examples

$$N = 2 \to N! = 2$$
$$\psi_S(1,2) = \phi_1(1)\phi_2(2) + \phi_1(2)\phi_2(1)$$
$$N = 3 \to N! = 6$$
$$\psi_S(1,2,3) = \phi_1(1)\phi_2(2)\phi_3(3) + \phi_1(2)\phi_2(1)\phi_3(3) + \phi_1(3)\phi_2(2)\phi_3(1)$$
$$+ \phi_1(1)\phi_2(3)\phi_3(2) + \phi_1(3)\phi_2(1)\phi_3(2) + \phi_1(2)\phi_2(3)\phi_3(1)$$

How do we construct a completely antisymmetric state? Let us define a general permutation operator by (illustrate for $N = 5$)

$$\hat{\wp}_{13452}\psi(1,2,3,4,5) = \psi(3,1,4,5,2)$$
$$\hat{\wp}_{23451}\psi(1,2,3,4,5) = \psi(2,3,4,5,1)$$

Any such permutation operator can be written as the product of the 2−particle interchange operators $\hat{P}_{ij}$, i.e.,

$$\psi(2,3,1) = \hat{\wp}_{231}\psi(1,2,3)$$
$$= \hat{P}_{12}\psi(3,2,1) = \hat{P}_{12}\hat{P}_{13}\psi(1,2,3)$$

Thus, any permutation $\hat{\wp}$ can be written in terms of an odd or even number of pair interchanges or pair permutations and we call it an odd or even permutation accordingly. All pair permutations are odd.

Therefore, for a completely antisymmetric state we must have

$$\hat{\wp}\psi_A = \begin{cases} +\psi_A & \text{if } \hat{\wp} \text{ is an even permutation} \\ -\psi_A & \text{if } \hat{\wp} \text{ is an odd permutation} \end{cases} \tag{12.96}$$

We therefore form a completely antisymmetric state as follows. We let

$$(-1)^{\hat{\wp}} = \begin{cases} +1 & \text{if } \hat{\wp} \text{ is an even permutation} \\ -1 & \text{if } \hat{\wp} \text{ is an odd permutation} \end{cases} \tag{12.97}$$

and then

Examples

$$N = 2 \to N! = 2$$
$$\psi_A(1,2) = \phi_1(1)\phi_2(2) - \phi_1(2)\phi_2(1)$$
$$N = 3 \to N! = 6$$
$$\psi_A(1,2,3) = \phi_1(1)\phi_2(2)\phi_3(3) - \phi_1(2)\phi_2(1)\phi_3(3) - \phi_1(3)\phi_2(2)\phi_3(1)$$
$$- \phi_1(1)\phi_2(3)\phi_3(2) + \phi_1(3)\phi_2(1)\phi_3(2) + \phi_1(2)\phi_2(3)\phi_3(1)$$

It is clear that if any two states are identical (put 2 = 3 above), then ψ_A is identically = 0 as it should be for fermions.

This implies that we can put at most 2 electrons in each energy level of the potential well. The two electrons in the k^{th} level would then have wave functions

$$\phi_k\alpha \text{ and } \phi_k\beta \tag{12.98}$$

i.e., they must have opposite spins. This says that N spin = 1/2 fermions must occupy at least $N/2$ different states in the well.

This is very different than for bosons where all the N bosons can be in any energy level.

Another way to write the completely antisymmetric wave function for fermions is the so-called Slater determinant

$$\psi_A(1,2,3,.....,N) = \begin{vmatrix} \phi_a(1) & \phi_a(2) & . & \phi_a(N) \\ \phi_b(1) & \phi_b(2) & . & \phi_b(N) \\ . & . & . & . \\ \phi_n(1) & \phi_n(2) & . & \phi_n(N) \end{vmatrix} \tag{12.99}$$

The last thing we must do is to normalize these state vectors.

$$\langle\psi_A \mid \psi_A\rangle = \sum_{s_1,s_2,....,s_N} \int d^3\vec{r}_1.....d^3\vec{r}_N \times \sum_{\wp\wp'} (-1)^{\wp}(-1)^{\wp'} [\wp\phi_a^*(1)....\phi_n^*(N)][\wp'\phi_a(1)....\phi_n(N)]$$

Now if $\wp \neq \wp'$, then $[\wp\phi_a(1)....\phi_n(N)]$ and $[\wp'\phi_a(1)....\phi_n(N)]$ are orthogonal and the integration for that term is zero.

Therefore, we get

$$\langle\psi_A \mid \psi_A\rangle = \sum_{s_1,s_2,....,s_N} \int d^3\vec{r}_1.....d^3\vec{r}_N \sum_{\wp} \wp\,|\phi_a(1)|^2....|\phi_n(N)|^2 \tag{12.100}$$

But

$$\sum_{s_k} \int d^3\vec{r}_k\,|\phi_k(j)|^2 = 1 \tag{12.101}$$

so we finally get

$$\langle\psi_A \mid \psi_A\rangle = \sum_{\wp} 1 = \text{ number of possible permutations } = N! \tag{12.102}$$

and therefore, the properly normalized completely antisymmetric wave function is

$$\psi_A(1,2,3,.....,N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_a(1) & \phi_a(2) & . & \phi_a(N) \\ \phi_b(1) & \phi_b(2) & . & \phi_b(N) \\ . & . & . & . \\ \phi_n(1) & \phi_n(2) & . & \phi_n(N) \end{vmatrix} \tag{12.103}$$

In a similar manner

$$\langle \psi_S \mid \psi_S \rangle = \frac{N!}{N_a!.....N_n!} \tag{12.104}$$

where

$N_k =$ the number of times the single particle state ϕ_k occurs

What is the difference between the ground state of N fermions and N bosons?

For N bosons, all N particles occupy the lowest level ϕ_0 and the wavefunction is

$$\psi_S(1,2,3,....,N) = \phi_0(1)\phi_0(2)\phi_0(3).....\phi_0(N) \tag{12.105}$$

with energy

$$E_0 = N\varepsilon_0 \tag{12.106}$$

This is true no matter how large N might be, even for macroscopic systems where $N \approx 10^{23}$. As we shall see in later discussions, this is one of the physical requirements for phenomena like superconductivity, superfluidity and Bose-Einstein condensation.

Such a state is not allowed for fermions however. We must have

N even	N odd
2 in ϕ_0	2 in ϕ_0
2 in ϕ_1	2 in ϕ_1
.....	.....
2 in $\phi_{\frac{N}{2}-1}$	2 in $\phi_{\frac{N-1}{2}}$
2 in $\phi_{\frac{N}{2}}$	1 in $\phi_{\frac{N+1}{2}}$

This difference for systems with even or odd numbers of fermions will lead to dramatic physical consequences later for some atomic systems.

The ground state energy for N fermions is

$$2(\varepsilon_0 + \varepsilon_1 + ... + \varepsilon_{\frac{N}{2}}) \qquad \text{for } N \text{ even}$$

$$2(\varepsilon_0 + \varepsilon_1 + ... + \varepsilon_{\frac{N-1}{2}}) + \varepsilon_{\frac{N+1}{2}} \quad \text{for } N \text{ odd}$$

Either of these two energies is always greater than the N boson ground state energy.

The extra energy is called the *zero point energy* and it arises from particle interchange invariance or it arises from the *Pauli Exclusion Principle* which states

> No two identical fermions in a physical system
> can have the same set of quantum numbers

It is equivalent to the antisymmetry of the wave function requirement for fermions.

12.5 The Helium Atom

We now consider the simplest multielectron atom, namely, helium, which has two electrons. The Hamiltonian is

$$\hat{H} = \hat{H}(1) + \hat{H}(2) + \hat{V} = \frac{\vec{p}_{1,op}^{\,2}}{2m} - \frac{Ze^2}{r_1} + \frac{\vec{p}_{2,op}^{\,2}}{2m} - \frac{Ze^2}{r_2} + \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \tag{12.107}$$

where

$\hat{H}(i) =$ hydrogen atom Hamiltonian with nuclear charge Ze (instead of e)

$\hat{V} =$ electrostatic repulsion between the electrons

We start by neglecting the electrostatic repulsion between the electrons. This gives us a zero-order solution that we can use in perturbation theory. This is equivalent to the independent particle model we just discussed.

Since $\hat{H} = \hat{H}(1) + \hat{H}(2)$ in this model, we can write

$$|\psi\rangle = |n_1\ell_1 m_1\rangle\,|n_2\ell_2 m_2\rangle \tag{12.108}$$

where

$$\hat{H}(1)\,|n_1\ell_1 m_1\rangle = E_{n_1}^{(0)}\,|n_1\ell_1 m_1\rangle \text{ and } \hat{H}(2)\,|n_2\ell_2 m_2\rangle = E_{n_2}^{(0)}\,|n_2\ell_2 m_2\rangle$$

$$\begin{aligned}\hat{H}\,|\psi\rangle &= (\hat{H}(1) + \hat{H}(2))\,|\psi\rangle = (\hat{H}(1) + \hat{H}(2))\,|n_1\ell_1 m_1\rangle\,|n_2\ell_2 m_2\rangle \\ &= E_{n_1 n_2}^{(0)}\,|\psi\rangle = (E_{n_1}^{(0)} + E_{n_2}^{(0)})\,|n_1\ell_1 m_1\rangle\,|n_2\ell_2 m_2\rangle = (E_{n_1}^{(0)} + E_{n_2}^{(0)})\,|\psi\rangle\end{aligned}$$

and

$$E_n^{(0)} = -\frac{Z^2 e^2}{2a_0 n^2} \quad (\text{Z} = 2 \text{ for helium}) \tag{12.109}$$

We will be working out the numbers in this problem so that we can compare our results to experiment. The zero order energies are shown in Table 12.1 below:

n_1	n_2	$E_{n_1 n_2}^{(0)}(Ry)$	$E_{n_1 n_2}^{(0)}(eV)$
1	1	-8	-108.8
1	2	-5	-68.0
1	3	-40/9	-64.4
1	..	..	..
1	..	-4	-54.5
2	2	-2	-27.2

Table 12.1: Zero Order Energies

where

$$1 \text{ Ry(Rydberg)} = \frac{e^2}{2a_0} = 13.6\,eV \tag{12.110}$$

The ground state energy is

$$E_{gs} = E_{11}^{(0)} = 2E_1^{(0)} = -8\,Ry \tag{12.111}$$

and the energy of the system when one electron has been ionized (no longer bound) is

$$E_{ion} = E_1^{(0)} + E_\infty^{(0)} = -4\,Ry \tag{12.112}$$

Therefore, it requires the addition of $4\,Ry$ to create singly ionized helium. Notice that the $(2,2)$ state has an energy greater than E_{ion}, which implies that it is not a bound state of the helium atom. All the states $(1,n)$ are bound states. The energy level spectrum looks as shown in Figure 12.4 below.

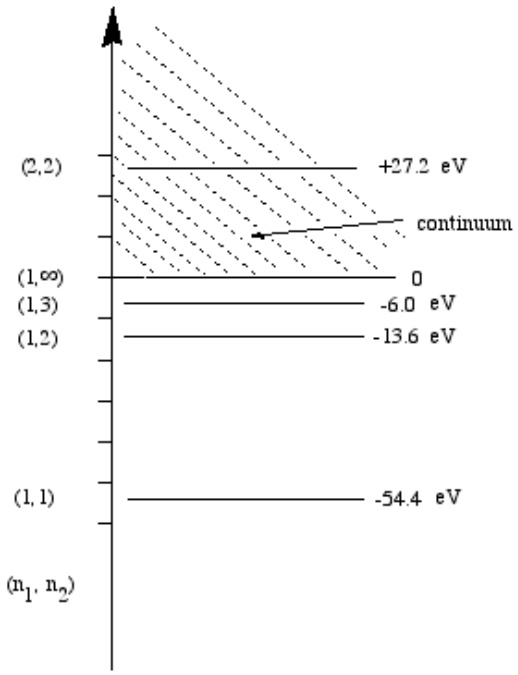


Figure 12.4: Helium Energy level Spectrum

Since the particles are electrons we must antisymmetrize the wave functions. We have two spin = 1/2 fermions. The spin functions are

$$|s, m_s\rangle = \begin{cases} |1,(\pm 1,0)\rangle & \rightarrow \text{symmetric} \\ |0,0\rangle & \rightarrow \text{antisymmetric} \end{cases} \tag{12.113}$$

The spatial part of the wave function must be of opposite symmetry to the spin functions so that the product is antisymmetric. By convention we label the states as follows:

Parahelium

$$\begin{aligned}&(\text{symmetric space part})\chi_{00}\\&(|100\rangle|100\rangle)|00\rangle\\&\frac{1}{\sqrt{2}}\left[|100\rangle|2\ell m\rangle+|2\ell m\rangle|100\rangle\right]\chi_{00}\end{aligned}$$

and so on.

Orthohelium

$$\begin{aligned}&(\text{antisymmetric space part})\chi_{1m_s}\\&\frac{1}{\sqrt{2}}\left[|100\rangle|2\ell m\rangle-|2\ell m\rangle|100\rangle\right]\chi_{1m}\end{aligned}$$

and so on.

These are the zero-order wave functions. We now handle the

$$\frac{e^2}{|\vec{r}_1-\vec{r}_2|}=\frac{e^2}{r_{12}} \tag{12.114}$$

term by perturbation theory.

The first order ground state energy correction is

$$\begin{aligned}\Delta E&=\langle 100|\langle 100|\frac{e^2}{r_{12}}|100\rangle|100\rangle\langle 00\,|\,00\rangle\\&=e^2\int\int d^3\vec{r}_1 d^3\vec{r}_2\frac{|\psi_{100}(\vec{r}_1)|^2|\psi_{100}(\vec{r}_2)|^2}{r_{12}}\end{aligned} \tag{12.115}$$

where

$$\psi_{100}(\vec{r})=\frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{3/2}e^{-\frac{Zr}{a_0}} \tag{12.116}$$

Therefore,

$$\Delta E=\frac{1}{\pi^2}\left(\frac{Z}{a_0}\right)^3 e^2\int_0^\infty dr_1 r_1^2 e^{-\frac{2Zr_1}{a_0}}\int_0^\infty dr_2 r_2^2 e^{-\frac{2Zr_2}{a_0}}\int\int d\Omega_1 d\Omega_2\frac{1}{r_{12}} \tag{12.117}$$

Even though this calculation does not give a very accurate result, it is still very instructive to learn the tricks necessary to evaluate the integrals.

We first need to find a useful expression for $1/r_{12}$. We have

$$r_{12}=|\vec{r}_1-\vec{r}_2|=\sqrt{(\vec{r}_1-\vec{r}_2)\cdot(\vec{r}_1-\vec{r}_2)} \tag{12.118}$$

$$\begin{aligned}r_{12}^2&=(\vec{r}_1-\vec{r}_2)\cdot(\vec{r}_1-\vec{r}_2)=r_1^2+r_2^2-2\vec{r}_1\cdot\vec{r}_2\\&=r_1^2+r_2^2-2r_1r_2\cos\beta\end{aligned} \tag{12.119}$$

where

$$\beta = \text{ angle between } \vec{r}_1 \text{ and } \vec{r}_2 \tag{12.120}$$

Therefore,

$$\frac{1}{r_{12}} = \frac{1}{\left(r_1^2 + r_2^2 - 2r_1 r_2 \cos\beta\right)^{1/2}} \tag{12.121}$$

In the subsequent development, we let the larger of r_1, r_2 be called $r_>$ and the smaller be called $r_<$. We then have

$$\begin{aligned}\frac{1}{r_{12}} &= \frac{1}{r_> \left(1 - 2\frac{r_<}{r_>}\cos\beta + \left(\frac{r_<}{r_>}\right)^2\right)^{1/2}} \\ &= \frac{1}{r_>} + \frac{1}{2r_>}\left(2\frac{r_<}{r_>}\cos\beta - \left(\frac{r_<}{r_>}\right)^2\right) - \frac{3}{8r_>}\left(2\frac{r_<}{r_>}\cos\beta - \left(\frac{r_<}{r_>}\right)^2\right)^2 \\ &\quad + \frac{15}{48r_>}\left(2\frac{r_<}{r_>}\cos\beta - \left(\frac{r_<}{r_>}\right)^2\right)^3 - \ldots \end{aligned} \tag{12.122}$$

or

$$\frac{1}{r_{12}} = \frac{1}{r_>}\left[1 + \frac{r_<}{r_>}\cos\beta + \left(\frac{r_<}{r_>}\right)^2\left(\frac{3}{2}\cos^2\beta - \frac{1}{2}\right) + \ldots\right] \tag{12.123}$$

$$= \frac{1}{r_>}\left[P_0(\cos\beta) + \frac{r_<}{r_>}P_1(\cos\beta) + \left(\frac{r_<}{r_>}\right)^2 P_2(\cos\beta) + \ldots\right] \tag{12.124}$$

Therefore,

$$\frac{1}{r_{12}} = \frac{1}{r_>}\sum_{\lambda=0}^{\infty}\left(\frac{r_<}{r_>}\right)^\lambda P_\lambda(\cos\beta) \tag{12.125}$$

Now, the addition theorem for spherical harmonics, which is proved at the end of this chapter, gives

$$P_\lambda(\cos\beta) = \frac{4\pi}{2\lambda+1}\sum_{m=-\lambda}^{\lambda} Y_{\lambda m}(\Omega_1)Y^*_{\lambda m}(\Omega_2) \tag{12.126}$$

Therefore, we finally have

$$\frac{1}{r_{12}} = \frac{1}{r_>}\sum_{\lambda=0}^{\infty}\left(\frac{r_<}{r_>}\right)^\lambda \frac{4\pi}{2\lambda+1}\sum_{m=-\lambda}^{\lambda} Y_{\lambda m}(\Omega_1)Y^*_{\lambda m}(\Omega_2) \tag{12.127}$$

Now, the factor

$$\int\int d\Omega_1 d\Omega_2 \frac{1}{r_{12}} \tag{12.128}$$

contains terms like

$$\int\int d\Omega_1 d\Omega_2 Y_{\lambda m}(\Omega_1)Y^*_{\lambda m}(\Omega_2) \tag{12.129}$$

and

$$\int d\Omega Y_{\lambda m}(\Omega) \propto \int d\Omega Y_{\lambda m}(\Omega)Y_{00}(\Omega) = \delta_{\lambda 0}\delta_{m0} \tag{12.130}$$

Therefore, the only term that contributes from the sum is $\lambda = m = 0$ and we get

$$\int\int d\Omega_1 d\Omega_2 \frac{1}{r_{12}} = \frac{1}{r_>} \tag{12.131}$$

and therefore we have

$$\Delta E = \frac{1}{\pi^2}\left(\frac{Z}{a_0}\right)^3 e^2 \int_0^\infty dr_1 r_1^2 e^{-\frac{2Zr_1}{a_0}} \int_0^\infty dr_2 r_2^2 e^{-\frac{2Zr_2}{a_0}} \frac{1}{r_>} \tag{12.132}$$

or

$$\begin{aligned}\Delta E = &\frac{1}{\pi^2}\left(\frac{Z}{a_0}\right)^3 e^2 \int_0^\infty dr_1 r_1^2 e^{-\frac{2Zr_1}{a_0}} \int_0^{r_1} dr_2 r_2^2 e^{-\frac{2Zr_2}{a_0}} \frac{1}{r_1} \\ &+ \frac{1}{\pi^2}\left(\frac{Z}{a_0}\right)^3 e^2 \int_0^\infty dr_1 r_1^2 e^{-\frac{2Zr_1}{a_0}} \int_{r_1}^\infty dr_2 r_2^2 e^{-\frac{2Zr_2}{a_0}} \frac{1}{r_2}\end{aligned} \tag{12.133}$$

which gives

$$\Delta E = \frac{5}{8}\frac{Ze^2}{a_0} = J_{1s,1s} = J_{10,10} = 2.5\ Ry = 34\ eV \tag{12.134}$$

for $Z = 2$.

The ground state energy corrected to first order is then

$$E_{11} = E_{11}^{(0)} + \Delta E = -74.8\ eV = -5.5\ Ry \tag{12.135}$$

The experimental value is

$$(E_{11})_{exp\,t} = -78.975\ eV = -5.807\ Ry \tag{12.136}$$

This first order result is amazingly good for this complex system!

Now we deal with the first excited state.

The first order energy shifts are once again given by standard perturbation theory since the $\langle \hat{V} \rangle$ matrix is diagonal in this basis due to the orthonormality of the spin vectors and the fact that $\hat{V}$ is independent of spin.

We thus have

$$\Delta E_{n\ell}^{s,t} = \frac{1}{2}\int\int d^3\vec{r}_1 d^3\vec{r}_2 \left|\psi_{100}(1)\psi_{n\ell 0}(2) \pm \psi_{100}(2)\psi_{n\ell 0}(1)\right|^2 \frac{e^2}{r_{12}} \tag{12.137}$$

where $s,t \rightarrow$ singlet, triplet $\rightarrow S = 0,1 \rightarrow -,+$. As shown before, we need only calculate the $m = 0$ case because $[\vec{L}_{op}, \hat{V}] = 0$ where

$$\vec{L}_{op} = \vec{L}_{1,op} + \vec{L}_{2,op} = \text{ total orbital angular momentum} \tag{12.138}$$

which implies that the result is independent of m.

Therefore

$$\Delta E^{s,t}_{n\ell} = e^2 \int\int d^3\vec{r}_1 d^3\vec{r}_2 \left|\psi_{100}(1)\right|^2 \left|\psi_{n\ell 0}(2)\right|^2 \frac{1}{r_{12}} \tag{12.139}$$
$$+ e^2 \int\int d^3\vec{r}_1 d^3\vec{r}_2 \psi^*_{100}(1)\psi^*_{n\ell 0}(2)\psi_{100}(2)\psi_{n\ell 0}(1)\frac{1}{r_{12}} = J_{n\ell} \pm K_{n\ell}$$

where

$$J_{n\ell} = \text{ electrostatic repulsion between two charge distributions } \left|\psi_{100}(1)\right|^2 \text{and} \left|\psi_{n\ell 0}(2)\right|^2 = \text{the direct integral}$$

and

$$K_{n\ell} = \text{the exchange integral which arises from antisymmetrization of the wave function}$$

with

$$+ = \text{singlet and} - = \text{triplet}$$

A convenient way of representing this result is as follows.

$$\vec{S}_{op} = \vec{S}_{1,op} + \vec{S}_{2,op}$$
$$\rightarrow 2\vec{S}_{1,op}\cdot\vec{S}_{2,op} = \vec{S}^2_{op} - \vec{S}^2_{1,op} - \vec{S}^2_{2,op} = \hbar^2\left(S(S+1) - \frac{3}{2}\right) \tag{12.140}$$

$$2\vec{S}_{1,op}\cdot\vec{S}_{2,op} = \hbar^2 \begin{cases} +\frac{1}{2} & \text{triplet} \\ -\frac{3}{2} & \text{singlet} \end{cases} \tag{12.141}$$

and therefore

$$\Delta E^{s,t}_{n\ell} = J_{n\ell} - \frac{1}{2\hbar^2}\left(1 + 4\vec{S}_{1,op}\cdot\vec{S}_{2,op}\right)K_{n\ell} \tag{12.142}$$

The calculation results (in eV) are shown in Table 12.2 below and the energy levels are shown in Figure 12.5 below. Not bad!!

State	1s2s		1s2p	
$n\ell$	10,20	10,20	10,21	10,21
	singlet	triplet	singlet	triplet
0^{th} order	-68.0	-68.0	-68.0	-68.0
J	11.4	11.4	13.2	13.2
K	1.2	1.2	0.9	0.9
1^{st} order	-55.4	-57.8	-53.9	-55.7
E_{expt}	-58.4	-59.2	-57.8	-58.0
error	3.0=5.1%	1.4=2.4%	3.9=6.7%	2.3=4.0%

Table 12.2: Calculation Results

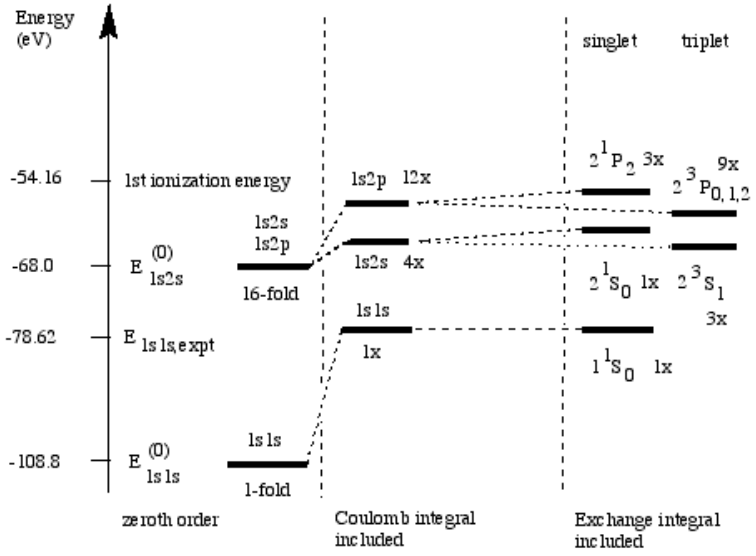


Figure 12.5: Helium Energy Levels

For comparison, we will also calculate the ground state energy using the variational method. We neglect spin in this case. The simplest choice of a trial function is the product of two hydrogen atom wave functions as in (12.143) below, which would be an exact solution if the electron-electron repulsion was neglected.

$$\psi(\vec{r}_1,\vec{r}_2) = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{3/2} e^{-\frac{Zr_1}{a_0}} \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{3/2} e^{-\frac{Zr_2}{a_0}} \tag{12.143}$$

Since we expect the true wave function to be approximately represented by the above function, we change Z to α and thus obtain the best possible value for E_0 for this type of trial function.

We do the calculation as follows. We write

$$\begin{aligned}\hat{H} &= \hat{H}_Z(1) + \hat{H}_Z(2) + \hat{V} = \frac{\vec{p}_{1,op}^2}{2m} - \frac{Ze^2}{r_1} + \frac{\vec{p}_{2,op}^2}{2m} - \frac{Ze^2}{r_2} + \frac{e^2}{r_{12}} \\ &= \frac{\vec{p}_{1,op}^2}{2m} - \frac{\alpha e^2}{r_1} + \frac{\vec{p}_{2,op}^2}{2m} - \frac{\alpha e^2}{r_2} + \frac{(\alpha - Z)e^2}{r_1} + \frac{(\alpha - Z)e^2}{r_2} + \frac{e^2}{r_{12}} \\ &= \hat{H}_\alpha(1) + \hat{H}_\alpha(2) + \frac{(\alpha - Z)e^2}{r_1} + \frac{(\alpha - Z)e^2}{r_2} + \frac{e^2}{r_{12}} \end{aligned} \tag{12.144}$$

Now

$$\hat{H}_\alpha(1)\psi_{100}^\alpha(1) = E_{100}^{(0)}(\alpha)\psi_{100}^\alpha(1) = -\alpha^2\psi_{100}^\alpha(1) \tag{12.145}$$

$$\hat{H}_\alpha(2)\psi_{100}^\alpha(2) = E_{100}^{(0)}(\alpha)\psi_{100}^\alpha(2) = -\alpha^2\psi_{100}^\alpha(2) \tag{12.146}$$

in Rydbergs. Therefore

$$\begin{aligned} f(\alpha) = -2\alpha^2 &+ \langle\psi_{100}^\alpha(1)| \frac{(\alpha - Z)e^2}{r_1} |\psi_{100}^\alpha(1)\rangle \\ &+ \langle\psi_{100}^\alpha(2)| \frac{(\alpha - Z)e^2}{r_2} |\psi_{100}^\alpha(2)\rangle + \langle\psi(\alpha)| \frac{e^2}{r_{12}} |\psi(\alpha)\rangle \end{aligned} \tag{12.147}$$

But

$$\begin{aligned} \langle\psi_{100}^\alpha(1)| \frac{(\alpha - Z)e^2}{r_1} |\psi_{100}^\alpha(1)\rangle &= \langle\psi_{100}^\alpha(2)| \frac{(\alpha - Z)e^2}{r_2} |\psi_{100}^\alpha(2)\rangle \\ &= \langle 100| \frac{(\alpha - Z)e^2}{r} |100\rangle \end{aligned} \tag{12.148}$$

Therefore,

$$f(\alpha) = -2\alpha^2 + 2e^2(\alpha - Z)\langle 100| \frac{1}{r} |100\rangle + \langle 100|\langle 100| \frac{e^2}{r_{12}} |100\rangle |100\rangle \tag{12.149}$$

Using some earlier calculations we get

$$f(\alpha) = -2\alpha^2 + 4\alpha(\alpha - Z) + \frac{5}{4}\alpha \tag{12.150}$$

Minimizing

$$\frac{df}{d\alpha} = 0 = -2\alpha + 2Z - \frac{5}{8} \tag{12.151}$$

or

$$\alpha = Z - \frac{5}{16} \tag{12.152}$$

and

$$E_0^{\text{variational}} = f(Z - \frac{5}{16}) = -2(Z - \frac{5}{16})^2 = -2Z^2 + \frac{5}{4}Z - 2(\frac{5}{16})^2 \tag{12.153}$$

The first two terms are just the first order perturbation theory result. The third term lowers the energy relative to perturbation theory.

For $Z = 2$, we get

$$E_0^{variational} = -5.7\ Ry = -77.48\ eV$$
$$E_0^{experimental} = -78.975\ eV$$
$$E_0^{perturbation\ theory} = -74.8\ eV$$

Even with the simple trial function, we get a significantly better result using the variational method. The reduction in the value of Z represents the effect of the inner electron screening the outer electron so it see a smaller nuclear charge.

12.6 Multielectron Atoms

We now return to the case of N electrons ($N > 2$). We have

$$\hat{H}\psi_\alpha = E_\alpha \psi_\alpha \tag{12.154}$$

where α = all quantum numbers needed to specify the N-electron state and the Hamiltonian $\hat{H}$ is

$$\hat{H} = \sum_{i=1}^{N}\left[-\frac{\hbar^2}{2m_e}\nabla_i^2 - \frac{Ze^2}{r_i}\right] + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{e^2}{r_{ij}} \tag{12.155}$$

For the moment we are neglecting many small(weak) interactions (spin-orbit, etc). We are also not including the electromagnetic field at this stage. We will consider it later when we talk about time-dependent perturbation theory and we will see that its presence leads to instability of atoms with respect to photon absorption/emission.

The electrons are all indistinguishable, which says that

$$\left[\hat{H}, \hat{P}_{ij}\right] = 0 \quad i, j = 1, 2, 3, ..., N\ ;\ i \neq j \tag{12.156}$$

This implies, since electrons are fermions, that the wave functions must be completely antisymmetric, i.e.,

$$\hat{P}_{ij}\psi_\alpha = -\psi_\alpha \quad i, j = 1, 2, 3, ..., N\ ;\ i \neq j \tag{12.157}$$

The full Hamiltonian is much too complex to solve exactly. We will approach the solution as a series of increasingly better approximations and obtain a qualitative picture of the energy level structure of these complex atoms.

Since the difficulties arise from the e^2/r_{ij} terms that represent the electron-electron repulsion, we start with a model where each electron moves independently of all the other electrons (an independent particle model). In this model

each electron will be described by a single-particle wavefunction called an orbital.

This leads us to write an approximate Hamiltonian in terms of the single particle Hamiltonians

$$\hat{H}_{0i} = -\frac{\hbar^2}{2m_e}\nabla_i^2 - \hat{V}_i(\vec{r}_i) \tag{12.158}$$

where we assume that the potential energy of the i^{th} electron $\hat{V}_i(\vec{r}_i)$ does not depend on the coordinates of the other $N-1$ electrons. We then have the approximate Hamiltonian for the entire system

$$\hat{H}_0^A = \sum_{i=1}^{N}\hat{H}_{0i} = \sum_{i=1}^{N}\left[-\frac{\hbar^2}{2m_e}\nabla_i^2 - \hat{V}_i(\vec{r}_i)\right] \tag{12.159}$$

This Hamiltonian is separable, i.e., we can assume that the system wavefunction is a product of single-particle wavefunctions or orbitals.

$$\psi_\alpha = \psi_{\varepsilon_1}(\vec{r}_1)\psi_{\varepsilon_2}(\vec{r}_2)\psi_{\varepsilon_3}(\vec{r}_3).....\psi_{\varepsilon_N}(\vec{r}_N) \tag{12.160}$$

where the subscript ε_k represents all applicable single particle quantum numbers for the k^{th} electron, that is,

$$\varepsilon_k = (n_i\ell_i m_{\ell_i} m_{s_i}) \tag{12.161}$$

Each single particle wave function is a product of the form

$$\psi_\varepsilon = (\text{spatial wave function})(\text{spin vector}) \tag{12.162}$$

We will assume that the system wavefunction is a completely antisymmetric combination of the product states ψ_α.

Each term in $\hat{H}_0^A$ has an eigenvalue equation of the form

$$\left[-\frac{\hbar^2}{2m_e}\nabla_i^2 - \hat{V}_i(\vec{r}_i)\right]\psi_{n_i\ell_i m_{\ell_i} m_{s_i}} = E_{n_i\ell_i m_{\ell_i} m_{s_i}}\psi_{n_i\ell_i m_{\ell_i} m_{s_i}} \tag{12.163}$$

and there are N such equations.

To solve these equations, we must know the potential energy functions $\hat{V}_i(\vec{r}_i)$. As a first approximation within the orbital approximations, we ignore the electron-electron repulsion so that the electrons only interact with the nucleus and we have

$$\hat{V}_i(\vec{r}_i) = \hat{V}(r_i) = -\frac{Ze^2}{r_i} \tag{12.164}$$

In this approximation, all the other electrons do not matter at all and each electron satisfies

$$\left[-\frac{\hbar^2}{2m_e}\nabla_i^2 - \frac{Ze^2}{r_i}\right]\psi_{n_i\ell_i m_{\ell_i} m_{s_i}}(\vec{r}_i) = E_{n_i\ell_i m_{\ell_i} m_{s_i}}\psi_{n_i\ell_i m_{\ell_i} m_{s_i}}(\vec{r}_i) \tag{12.165}$$

This is a hydrogen atom with charge Ze. The single-particle wavefunctions are given by

$$\psi_{n\ell m_\ell m_s}(\vec{r}) = R_{n\ell}(r)Y_{\ell m_\ell}(\theta,\varphi)\chi_{sm_s} = \psi_\varepsilon(\vec{r}) \tag{12.166}$$

where

$$E_{n_k} \approx -\frac{m_e Z^2 e^4}{2\hbar^2 n_k} \qquad n_k = 1,2,3,..... \tag{12.167}$$

The wave function corresponding to a set of orbitals $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_N)$ is then properly antisymmetrized by writing it as

$$\psi_\alpha = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{\varepsilon_1}(1) & \psi_{\varepsilon_2}(1) & .. & \psi_{\varepsilon_N}(1) \\ \psi_{\varepsilon_1}(2) & .. & .. & .. \\ .. & .. & .. & .. \\ \psi_{\varepsilon_1}(N) & .. & .. & \psi_{\varepsilon_N}(N) \end{vmatrix} \tag{12.168}$$

where

$$E_\alpha = E_{\varepsilon_1} + E_{\varepsilon_2} + ... + E_{\varepsilon_N} \tag{12.169}$$

We certainly can write down an answer in this approximation but the result, not surprisingly, is terrible. Any real electron is dramatically affected by the others, even when there is only one other electron as we saw in helium.

We move on by making the next incrementally better approximation. This involves the concept of screening.

12.6.1 Screening

Any electron, on the average, if it is far from the nucleus, does not feel all of the nuclear charge Ze and hence has a weaker Coulomb attraction then we have assumed. This is clear in the helium variational calculation where we found that the best value of the charge variational parameter Z' was

$$Z' = Z - \frac{5}{16} \tag{12.170}$$

This implies that, on the average, the distant electrons are *shielded* or *screened* from the nucleus by the other electrons.

What is the simplest correction that we can make to take this effect into account for multi-electron atoms and still leave us with solvable equations?

Suppose we write

$$V_i(r_i) = -\frac{Ze^2}{r_i} + V_i^{eff}(r_i) \tag{12.171}$$

where $V_i^{eff}(r_i)$ includes the screening effects of the other $N-1$ electrons. An important feature of this assumption is that $V_i^{eff}(r_i)$ is independent of (θ,φ). This says that the angular part of the wave function is still

$$Y_{\ell_i m_{\ell_i}}(\theta,\varphi) \tag{12.172}$$

The radial function, however, now satisfies a modified equation

$$\frac{1}{r_i^2}\frac{d}{dr_i}\left[r_i^2\frac{d}{dr_i} - \frac{\ell(\ell+1)}{r_i^2} + \frac{2m_e}{\hbar^2}\left(E_{n_i} - V_i(r_i)\right)\right]R_{n_i\ell_i}(r_i) = 0 \qquad (12.173)$$

This is called the *central field approximation*. We still have N difficult equations to solve.

Hartree proposed the following solution using a successive approximations technique.

1. an initial potential function is guessed (a very educated guess)
2. this potential function is used to derive new wave functions
3. the new wave functions generate a new potential energy function
4. the procedure is continued until the final wave functions determine a *self-consistent* potential, i.e., it stops changing as we iterate

The Hartree method is equivalent to a variational calculation, where the trial function is taken to be a simple product of single-particle orbitals and the variation is performed by varying each orbital in an arbitrary way.

Using single-particle wave functions, however, we are still neglecting the correlations between the electrons. Although the simple single-particle orbital product functions ignore antisymmetry, some effect of the Pauli exclusion principle (PEP) can be included in the calculations by choosing the single-particle quantum numbers so they do not violate the PEP.

If we make the calculation more complicated, we can include antisymmetry by using Slater determinant wave functions. This is called the Hartree-Fock theory. Correlation effects arising from the $1/r_{ij}$ terms can then be added using perturbation theory. At the level of this text, we assume that this can be done (see Bethe/Jackiw for details).

12.6.2 Shell Structure

The hydrogen atom solution exhibited a kind of shell structure. We found that the energy levels were given by

$$E_n = -\frac{Z^2e^2}{2a_0n^2} \qquad (12.174)$$

and each level had a degeneracy equal to n^2 arising from the allowed ranges

$$\ell = 0, 1, 2, ..., n-1$$
$$m_\ell = -\ell, ..., \ell$$

We say that each n value defines a shell with energy E_n and within each shell we have subshells defined by ℓ. Thus,

$$n = 1 \to \ell = 0 \to 1s\text{subshell}$$
$$n = 2 \to \ell = 0, 1 \to 2s, 2p\text{subshells}$$
$$n = 3 \to \ell = 0, 1, 2 \to 3s, 3p, 3d\text{subshells}$$

We can generalize this idea to N electron atoms.

We assume that the atom consists of shells (n) and subshells ($n\ell$). Electrons are placed into these shells so that we do not violate the PEP, that is, since the electrons are fermions only two electrons can be in each energy level. We define in this model

$$\langle r \rangle_{n0} = \textit{radius of a shell}$$
$$\langle r \rangle_{n\ell} = \textit{radius of a subshell}$$

and we have

$$E_{n\ell} = E_{n\ell'} \text{ (degenerate)} \tag{12.175}$$

(this is not true in the central field approximation).

For $n > 1$, the $s-$orbital has a nonzero probability near the origin $r = 0$ (the nucleus). This implies that it *penetrates* the $n = 1$ shell more than the corresponding $p-$orbitals do. This implies that the s subshell electrons feel a stronger nuclear charge then the p subshell electrons.

Therefore, we expect in this model that the energy levels will look like Figure 12.6 below.

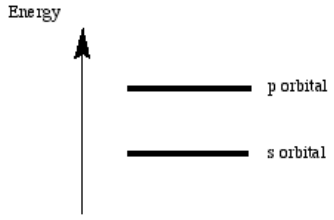


Figure 12.6: Expected Level Structure

which is the *screening* effect.

Complex screening arguments of this type lead to the *Aufbau* principle, which tells us how electrons fill shells.

We see how it works by figuring out the ground state of an N electron atom. The ground state corresponds to that state where all the *lowest energy levels* are filled with a *maximum* of two electrons per level. It is clear that this is the state of lowest energy.

The single-particle Hamiltonian in the central field approximation commutes with ℓ_{iz} and s_{iz} which implies that the energy levels are independent of m_{ℓ_i} and m_{s_i}. Each energy level is therefore characterized by the $2N$ quantum numbers

$$n_i, \ell_i \quad i = 1, 2, 3, ..., N \tag{12.176}$$

We define the electronic configuration of an atomic state as the set of quantum numbers (n_i, ℓ_i) for all the electrons in the atom. We use the symbolic representation

$$shell - label = (n - label)(\ell - label)^{number\ of\ electrons} \tag{12.177}$$

i.e.,

$$\begin{aligned} &ground\ state\ of\ hydrogen = 1s^1 = 1s \\ &\qquad (a\ superscript\ 1\ is\ always\ understood) \\ &ground\ state\ of\ helium = 1s^2 \\ &ground\ state\ of\ lithium = 1s^2 2s \end{aligned}$$

The electronic configuration of helium is an example of a *closed* or *full* shell. The $1s$ subshell has the maximum number electrons as allowed by the PEP, i.e.,

$$n = 1, \ell = 0, m_\ell = 0, s = \frac{1}{2}, m_s = \pm\frac{1}{2} \tag{12.178}$$

The Aufbau principle says that we fill the shells so that we obey the PEP or in the order

$$1s, 2s, 2p, 3s, 3p, 3d, 4s, ...$$

with energy increasing from left to right.

The screening arguments of the type we just discussed imply that for a given n (a given shell) the energy order is $s, p, d, ...$ and generally the energy of a shell increases with n = the principal quantum number. The closed shells correspond to

$$\begin{aligned} s - shell \rightarrow &\text{ maximum number of electrons} = 2 \\ &= 2(2\ell + 1) \\ p - shell \rightarrow &\text{ maximum number of electrons} = 6 \\ d - shell \rightarrow &\text{ maximum number of electrons} = 10 \end{aligned}$$

and so on.

Electrons in the shell beyond the last closed shell are called *valence* electrons.

Much of the form and shape of the periodic table is determined by the Aufbau principle. For instance

number of valence electrons → chemical properties

The valence electrons are the ones that participate in bonding and chemical reactions.

This implies that

$$\begin{aligned} &carbon \rightarrow 1s^2 2s^2 2p^2 \\ &\text{and} \\ &silicon \rightarrow 1s^2 2s^2 2p^6 3s^2 3p^2 \end{aligned}$$

which each have two p valence electrons should have similar chemical properties, which is the case.

As with all simple principles of this type, anomalies and breakdowns soon appear. For the Aufbau principle this occurs at the $n = 3$ shell.

In real atoms, when the $3d$ and $4s$ subshells are partially full, the $4s$ level fills ups before the $3d$ level. This means that

$$\begin{aligned} &potassium \rightarrow 1s^2 2s^2 2p^2 3s^2 3p^6 4s \\ &\text{and not} \qquad 1s^2 2s^2 2p^6 3s^2 3p^6 3d \end{aligned}$$

The $4s$ state has a larger probability of being near $r = 0$ then the $3d$ state and hence its energy is lower.

For neutral atoms, an experimental ordering scheme is

shell	increasing energy →																
$n = 1$	$1s$																
$n = 2$		$2s$	$2p$														
$n = 3$				$3s$	$3p$		$3d$										
$n = 4$						$4s$		$4p$		$4d$			$4f$				
$n = 5$									$5s$		$5p$			$5d$			
$n = 6$												$6s$			$6p$		

At the level of this text we will not be doing any actual energy calculations.

12.7 Angular Momentum Coupling

The N electrons each have spin and orbital angular momentum and thus have associated magnetic moments.

For the full Hamiltonian $\hat{H}$ including the electron-electron repulsion terms we have

$$\left[\hat{H}, \vec{\ell}_{i,op}\right] \neq 0 \quad i = 1, 2, 3, ..., N \tag{12.179}$$

However,

$$\left[\hat{H}, \vec{L}_{op}\right] = 0 \tag{12.180}$$

where

$$\vec{L}_{op} = \sum_{i=1}^{N} \vec{\ell}_{i,op} = \text{ the total orbital angular momentum} \tag{12.181}$$

Therefore, the individual ℓ_i are not conserved, but $\vec{L}$ is conserved.

Thus, the electron-electron repulsion or electrostatic terms in the Hamiltonian couple the electron orbital angular momenta.

In addition, we must add in spin-orbit and other magnetic interactions (spin-spin, etc).

The spin-orbit interaction leads to terms of the form $\vec{\ell}_i \cdot \vec{s}_i$ and thus couple a particle's orbital and spin angular momentum leading to (j_i, m_{j_i}) values, where $\vec{j}_i = \vec{\ell}_i + \vec{s}_i$ as we saw earlier in hydrogen.

In most light atoms, the magnetic interactions are usually weaker than the electrostatic interactions, i.e., electrostatic $\approx 1\ eV$ and spin-orbit $\approx 10^{-4} - 10^{-5}\ eV$.

The angular momentum coupling in a light atom goes like:

1. the orbital angular momenta $\vec{\ell}_i$ couple to form a total orbital angular momentum

$$\vec{L}_{op} = \sum_{i=1}^{N} \vec{\ell}_{i,op} \tag{12.182}$$

2. the spin angular momenta $\vec{s}_i$ couple to form a total spin angular momentum

$$\vec{S}_{op} = \sum_{i=1}^{N} \vec{s}_{i,op} \tag{12.183}$$

These two couplings occur when we include the electrostatic interactions in the Hamiltonian $\hat{H}$.

3. the weaker magnetic interactions then couple $\vec{L}$ and $\vec{S}$ to form the total angular momentum of the atom

$$\vec{J} = \vec{L} + \vec{S} \tag{12.184}$$

This coupling scheme or order where the electrostatic interactions dominate the magnetic interactions is called LS or *Russell-Saunders* coupling.

In heavy atoms, the spin-orbit magnetic interactions dominate the electrostatic interactions and we get an alternative coupling scheme:

1. each electron's $\vec{\ell}_i$ and $\vec{s}_i$ couple via the magnetic interactions to form

$$\begin{aligned}\vec{j}_i &= \vec{\ell}_i + \vec{s}_i \\ &= \text{the total angular momentum for the } i^{th} \text{ electron}\end{aligned}$$

2. the electrostatic interactions then couple the $\vec{j}_i$ to form

$$\vec{J}_{op} = \sum_{i=1}^{N} \vec{j}_{i,op} \tag{12.185}$$

This scheme is called *jj-coupling*.

We will now investigate the energy level structure in detail for these two different schemes.

Our discussion of helium has shown that exchange symmetry, which requires that the wave functions are completely antisymmetric, has dramatic observable consequences. We saw a spin-spin correlation energy that is characterized by the rule:

> There is a tendency for electrons with parallel spins to repel (avoid) each other. This fact, together with the electrostatic repulsion between electrons implies a strong exchange correlation that cause the spins to tend to align with each other

12.7.1 LS Coupling

In this regime we have the observables and quantum numbers as shown in Table 12.3 below:

Operator	Quantum Number
$\vec{L}^2_{op}$	L
$\vec{J}^2_{op}$	J
$\vec{S}^2_{op}$	S
$\hat{L}_z$	M_L
$\hat{S}_z$	M_S

Table 12.3: LS Coupling Quantum Numbers

When we discussed the spin-orbit interaction in hydrogen we found

1. when we neglect $\hat{H}_{so}$ we can use either $|n, L, S, M_L, M_S\rangle$ or $|n, L, S, J, M_J\rangle$ as basis states
2. when we add in $\hat{H}_{so}$, M_L and M_S are no longer conserved (not good quantum numbers) and therefore we must use $|n, L, S, J, M_J\rangle$ as basis states

Now in the orbital approximation we have

$$M_L = \sum_{i=1}^{N} m_{\ell_i} \quad , \quad M_S = \sum_{i=1}^{N} m_{s_i} \tag{12.186}$$

and

$$\hat{L}_z\psi_\alpha = \hbar M_L\psi_\alpha \qquad \hat{S}_z\psi_\alpha = \hbar M_S\psi_\alpha$$

What are the possible L, S values? We can use our addition of angular momentum rules to find out.

Consider two p–electrons , i.e., an vp^2 configuration. We have

$$\ell_1 = \ell_2 = 1 \rightarrow L = 0, 1, 2$$
$$s_1 = s_2 = \frac{1}{2} \rightarrow S = 0, 1$$

and

$$\text{for a given } L \quad M_L = -L, ..., L$$
$$\text{for a given } S \quad M_S = -S, ..., S$$

and

$$M_J = M_L + M_S$$
$$J = |L - S|, ..., L + S$$

Therefore we get the possibilities shown in Table 12.4 below:

L	S	J	State(s)
0	0	0	$^1S_0$
0	1	1	$^3S_1$
1	0	1	$^1P_1$
1	1	0,1,2	$^3P_{0,1,2}$
2	0	2	$^1D_1$
2	1	1,2,3	$^3D_{1,2,3}$

Table 12.4: Possible States

We will discuss the state notation shortly.

For a closed shell we must have $L = S = 0$ or we would violate the PEP. For example,

$$s^2 \rightarrow \ell_1 = \ell_2 = 0 \rightarrow L = 0$$
$$s_1 = s_2 = \frac{1}{2} \rightarrow S = 0 \; or \; 1$$

However, there is only one way to choose the m quantum numbers without violating the PEP which is shown in Table 12.5 below.

m_{ℓ_1}	m_{ℓ_2}	m_{s_1}	m_{s_2}
0	0	1/2	-1/2

Table 12.5: s^2 m-values

We therefore have $M_L = M_S = 0$ (only possibility) which implies that

$$L = S = J = 0 \rightarrow {}^1S_0 \text{ state} \tag{12.187}$$

For

$$p^6 \rightarrow \ell_1 = \ell_2 = \ell_3 = \ell_4 = \ell_5 = \ell_6 = 1$$
$$s_1 = s_2 = s_3 = s_4 = s_5 = s_6 = \frac{1}{2}$$

Once again it turns out there is only one way to choose the values without violating the PEP. This is shown in Table 12.6 below.

m_{ℓ_1}	m_{ℓ_2}	m_{ℓ_3}	m_{ℓ_4}	m_{ℓ_5}	m_{ℓ_6}	m_{s_1}	m_{s_2}	m_{s_3}	m_{s_4}	m_{s_5}	m_{s_6}
1	1	0	0	-1	-1	1/2	-1/2	1/2	-1/2	1/2	-1/2

Table 12.6: p^6 m-values

We therefore have $M_L = M_S = 0$ (only possibility) which implies that

$$L = S = J = 0 \rightarrow {}^1S_0 \text{ state} \tag{12.188}$$

This result is true for *all* closed shells.

In the presence of $\hat{H}_{so}$ the energy levels will depend on L, S and J but not on M_L, M_S or M_J, which is why we label their *atomic terms* by

$${}^{2S+1}L_J \tag{12.189}$$

where

$$S, P, D, F, \ldots\ldots\ldots \text{ means } L = 0, 1, 2, 3, \ldots\ldots \tag{12.190}$$

The superscript $2S + 1$ s the *multiplicity* of the level (singlet, doublet, triplet, etc).

If we ignore $\hat{H}_{so}$ then we have $(2S + 1)(2L + 1)$ degeneracy for a given level.

Adding $\hat{H}_{so}$ splits the J states. Each term ${}^{2S+1}L_J$ remains $2J + 1$ degenerate (the M_J values).

This degeneracy is removed by an external magnetic field which splits the $2J+1$ M_J levels(Zeeman effect).

How do we determine the ground state for a particular atom in this scheme?

First, we fill up as many closed shells as possible. The remaining (valence) electrons determine the ground state configuration.

Let us consider carbon which has two equivalent (same subshell) $2p$-electrons in the unfilled shell. We have

$$2p^2 \rightarrow \ell_1 = \ell_2 = 1 \rightarrow L = 0, 1 \; or \; 2$$
$$s_1 = s_2 = \frac{1}{2} \rightarrow S = 0 \; or \; 1$$

We get Table 12.7 below by applying these rules:

L	S	J	Term	Sublevels
0	0	0	^1S	$^1S_0$
1	0	1	^1P	$^1P_1$
2	0	2	^1D	$^1D_2$
0	1	1	^3S	$^3S_1$
1	1	0,1,2	^3P	$^3P_{0,1,2}$
2	1	1,2,3	^3D	$^3D_{1,2,3}$

Table 12.7: Possible States from LS Rules

Not all of these sublevels are allowed by the PEP however. To see this we must look at the individual electron quantum numbers. Table 12.8 below shows those $m_{\ell_1}, m_{\ell_2}, m_{s_1}, m_{s_2}$ values allowed by the PEP (i.e., no two electrons have the same set of quantum numbers).

Before proceeding to the table, in this case, we can use symmetry arguments to determine the allowed levels. In the special case of only two electrons in an unfilled shell, we can easily determine the symmetry of the spin vectors

$$S = 0 \rightarrow \text{ antisymmetric spin function}$$
$$S = 1 \rightarrow \text{ symmetric spin function}$$

We also know the symmetry of the spatial state in general. The symmetry follows from the symmetry of the angular part of the 2–electron wave function. Since we have a central field approximation, the angular part of the wave function is given by the Y_{LM_L} spherical harmonics. The radial function is always symmetric. The symmetry of the spherical harmonics is $(-1)^L$. Therefore,

$$L = odd \rightarrow \text{ antisymmetric space function}$$
$$L = even \rightarrow \text{ symmetric space function}$$

The product of the spin vector and the spatial function must always be antisymmetric. Therefore we have

$$S = 0 \text{ always combines with even } L$$
$$S = 1 \text{ always combines with odd } L$$

This method is only simple to carry out for 2–electron unfilled shells. In the case of carbon we get the allowed states

$$L = 2, S = 0 \rightarrow 5 \text{ states } = (2L+1)(2S+1)$$
$$L = 1, S = 1 \rightarrow 9 \text{ states}$$
$$L = 0, S = 0 \rightarrow 1 \text{ states}$$

for a total of 15 allowed states. The individual quantum numbers table corresponding to these 15 states is

Entry	$m_\ell = -1$	$m_\ell = 0$	$m_\ell = +1$	M_L	M_S
1	↑↓			-2	0
2		↑↓		0	0
3			↑↓	2	0
4	↑	↑		-1	1
5	↑	↓		-1	0
6	↓	↑		-1	0
7	↓	↓		-1	-1
8	↑		↑	0	1
9	↑		↓	0	0
10	↓		↑	0	0
11	↓		↓	0	-1
12		↑	↑	1	1
13		↑	↓	1	0
14		↓	↑	1	0
15		↓	↓	1	-1

Table 12.8: Individual Quantum Numbers

Any other combinations will violate the PEP. This table can be constructed just using the PEP.

We now construct an implied terms table which tells us how many states exist with a particular pair of (M_L, M_S) values. It is shown as Table 12.9 below.

M_L/M_S	1	0	-1
2	0	1	0
1	1	2	1
0	1	3	1
-1	1	2	1
-2	0	1	0

Table 12.9: Implied terms

We use this table to determine which atomic terms are allowed for carbon. The steps are as follows:

1. Consider the largest possible values of L and S, L = 2,S = 1 which correspond to the ${}^3D$ terms.

   Now if a ${}^3D$ atomic term existed , then we would necessarily have M_L = 2,M_S = 1 terms. However, there are no such terms, which implies that the ${}^3D$ term is not allowed and thus the sublevels ${}^3D_1, {}^3D_2, {}^3D_3$ are ruled out

by the PEP.

2. We now look at the next largest values, namely, L = 2,S = 0 or the ^1D term. A L = 2,S = 0 term requires M_L = 2,M_S = 0 terms which do exist. Therefore the ^1D term and the sublevel $^1D_2$ exist. This has J = 2 and thus $2J + 1$ = 5 M_J levels. This accounts for 5 of the 15 entries in the table.

3. We subtract these 5 states to get a *second-implied terms* table

M_L/M_S	1	0	-1
2	0	0	0
1	1	1	1
0	1	2	1
-1	1	1	1
-2	0	0	0

Table 12.10: Implied terms

4. We now look at the next largest values, namely, L = 1,S = 1 or the ^3P term. Since entries with $M_L = -1, 0, +1$ and $M_S = -1, 0, +1$ still exist in the new table, the ^3P term and the sublevels $^3P_2, ^3P_1, ^3P_0$ are allowed. These correspond to a total of $(2L + 1)(2S + 1) = 5 + 3 + 1 = 9$ states.

5. We subtract these 9 states to get the *third-implied terms* table

M_L/M_S	1	0	-1
2	0	0	0
1	0	0	0
0	0	1	0
-1	0	0	0
-2	0	0	0

Table 12.11: Implied terms

   Only one state is left with M_L = M_S = 0, which is a ^1S atomic term. Therefore the $^1S_0$ sublevel is allowed.

This accounts for all the 15 entries in the table. No more states are allowed, which means that the ^3S and ^1P atomic terms and their associated sublevels are forbidden by the PEP. This result agrees with the allowed states we obtained from symmetry arguments.

We always need to use the implied-terms tables in the general cases (more than

2 electrons in an unfilled shell) because the corresponding symmetry arguments are very complex to apply.

Finally, for carbon we have the ${}^1S$, ${}^3P$ and ${}^1D$ allowed by the PEP.

This result is true for all atoms with 2 equivalent p–electrons outside closed subshells.

To complete the picture, we must now determine how the allowed terms are ordered in energy.

12.7.2 Hund's Rules

Each sublevel in carbon ${}^1S_0$, ${}^3P_2$, ${}^3P_1$, ${}^3P_0$ and ${}^1D_2$ has a different energy when $\hat{H}_{so}$ is included in $\hat{H}$.

A set of rules exists for qualitatively ordering the levels. They are called *Hund's rules*.

Hund's rules apply when we are ordering the energy levels and sublevels for equivalent electrons in the ground state.

Rule 1 Terms in the ground state configuration with maximum multiplicity $2S + 1$ lie lowest in energy.

This follows from the fact that same spins (unpaired spins) *repel* and different spins (paired spins) *attract*.

High multiplicity implies a greater number of electrons with parallel spin than low multiplicity in multielectron atoms. Since parallel spin electrons avoid each other, the e^2/r_{ij} effect decreases and the energy of high multiplicity states lies below that of low multiplicity states.

Rule 2 Of several levels with the same multiplicity S, the one with maximum L lies lowest in energy

In some sense, the maximum L state implies that all electrons are *orbiting in the same direction*. These electrons tend to remain separated from each other and so have a lower energy than those *orbiting in the opposite direction*, which get close to each other some of the time.

Rule 2 Of several sublevels with the same multiplicity S and same L

1. the sublevel with the minimum value of J lies lowest in energy if the shell is less than half-filled. These are called *regular multiplets*
2. the sublevel with the maximum value of J lies lowest in energy if the shell

is more than half-filled. These are called *inverted multiplets*

This results follows from $\hat{H}_{so}$ and the fact that $-e^2/r$ increases as $r \to \infty$. Applying Hund's rules to an np^2 configuration we get the energy level scheme in Figure 12.7 below.

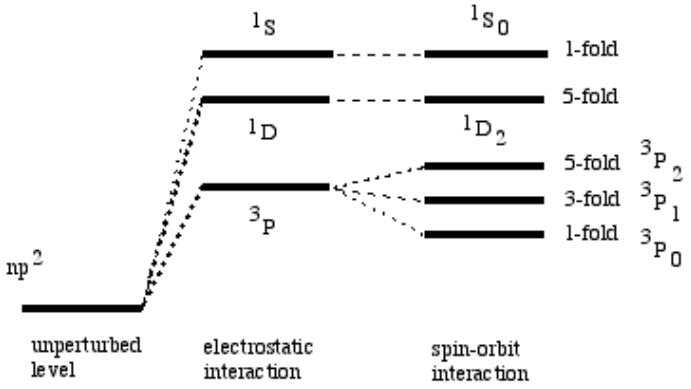


Figure 12.7: np^2 Level Scheme

Hund's rules are not perfect since they are based on the orbital approximation.

To determine an electronic configuration, we must specify how the electrons are placed into subshells. It turns out there is a phenomenon called configuration interaction or configuration mixing which forces the quantum mechanical state to sometimes be a mixture of more than one configuration.

12.7.3 JJ-Coupling

In heavy atoms, the magnetic interactions which couple the $\vec{\ell}_i$ and $\vec{s}_i$ together into the $\vec{j}_i$, dominate over the electrostatic interactions which led to LS coupling. The configurations are then better described by the so-called *jj-coupling* scheme.

Since $s_i = 1/2$ for all electrons, we have for $\ell_i \geq 1$

$$j_i = \ell_i \pm \frac{1}{2}$$
$$m_{j_i} = -j_i, ..., j_i$$

The individual $\vec{j}_i$ then couple together to give the total $\vec{J}$.

In a two-electron configuration the levels are labelled by J, j_1, j_2, M_J where

$$J = |j_1 - j_2|, ..., j_1 + j_2$$
$$M_J = -J, ..., J$$

Let us consider the Pb (lead) atom, which has np^2 valence electrons (built on many closed shells). We have

$$\begin{aligned}&\ell_1 = \ell_2 = 1\\ &\rightarrow j_1 = \frac{1}{2}, \frac{3}{2} \text{ and } j_2 = \frac{1}{2}, \frac{3}{2}\end{aligned}$$

The possible total J values are then

$$\begin{aligned}&\frac{3}{2} \otimes \frac{3}{2} = 3, 2, 1, 0\\ &\frac{1}{2} \otimes \frac{3}{2} = 2, 1\\ &\frac{1}{2} \otimes \frac{1}{2} = 1, 0\end{aligned}$$

Not all of these states are allowed by the PEP. For example,

$$\begin{aligned}J = 3, M_J = 3 \rightarrow &j_1 = j_2 = \frac{3}{2}, m_{j1} = m_{j2} = \frac{3}{2}\\ &\ell_1 = \ell_2 = 1, m_{\ell 1} = m_{\ell 2} = 1\\ &s_1 = s_2 = \frac{1}{2}, m_{s1} = m_{s2} = \frac{1}{2}\end{aligned}$$

Both electrons need to have identical quantum numbers for this state to exist. Thus, this state is not allowed. In a similar manner,

$$\begin{aligned}&J = 3 \quad , \quad j_1 = j_2 = \frac{3}{2}\\ &J = 1 \quad , \quad j_1 = j_2 = \frac{1}{2}\end{aligned}$$

can be shown to be forbidden by the PEP. Therefore we have

$$\begin{aligned}&\frac{3}{2} \otimes \frac{3}{2} = 2, 0\\ &\frac{1}{2} \otimes \frac{3}{2} = 2, 1\\ &\frac{1}{2} \otimes \frac{1}{2} = 0\end{aligned}$$

Usually, the level with the lowest J for a given pair (j_1, j_2) has the lowest energy (this is not a strict rule).

For medium weight atoms, neither LS nor jj coupling is valid.

There is a connection between the levels in the two schemes as illustrated by the energy level diagram in Figure 12.8 below.

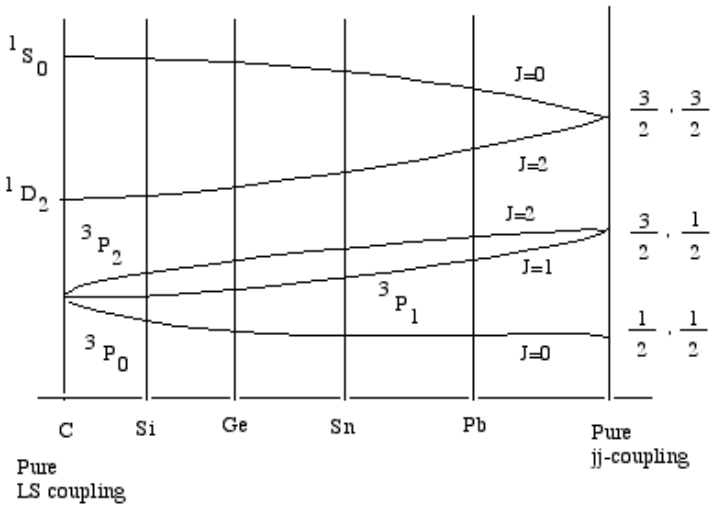


Figure 12.8: LS - jj Energy Level Connection

The connection between the two schemes is clear.

The spacing between J-levels in LS coupling is given as follows.

$$\left\langle \hat{H}_{so} \right\rangle = \frac{1}{2} C \left[(J(J+1) - L(L+1) - S(S+1) \right] \tag{12.191}$$

For the same L, S we have

$$\begin{aligned} E_{J+1} - E_J &= \frac{1}{2} C \left[(J+1)(J+2) - J(J+1) \right] \\ &= C(J+1) \end{aligned} \tag{12.192}$$

This says that the spacing between consecutive levels of a fine structure multiplet is proportional to the larger J value involved. This is the *Lande interval rule.*

We end this discussion with an example of two electrons that are not equivalent(in different shells). The discussion is more straightforward since we do not have to worry about the PEP (all possibilities are allowed).

We consider two electrons in a $4p4d$ configuration. In the LS coupling scheme

we have:

$$\ell_1 = 1\ ,\ \ell_2 = 2 \rightarrow L = 1, 2, 3$$
$$s_1 = s_2 = \frac{1}{2} \rightarrow S = 0, 1$$
$$3 \otimes 1 \rightarrow J = 4, 3, 2$$
$$3 \otimes 0 \rightarrow J = 3 \rightarrow 7 \text{ states}$$
$$2 \otimes 1 \rightarrow J = 3, 2, 1 \rightarrow 16 \text{ states}$$
$$2 \otimes 0 \rightarrow J = 2 \rightarrow 5 \text{ states}$$
$$1 \otimes 1 \rightarrow J = 2, 1, 0 \rightarrow 9 \text{ states}$$
$$1 \otimes 0 \rightarrow J = 1 \rightarrow 3 \text{ states}$$

or

$$J = 4 \quad in \quad 1 \quad level$$
$$J = 3 \quad in \quad 3 \quad levels$$
$$J = 2 \quad in \quad 4 \quad levels$$
$$J = 1 \quad in \quad 3 \quad levels$$
$$J = 0 \quad in \quad 1 \quad level$$

Thus, we have 12 total levels. The LS coupling energy level diagram is shown in Figure 12.9

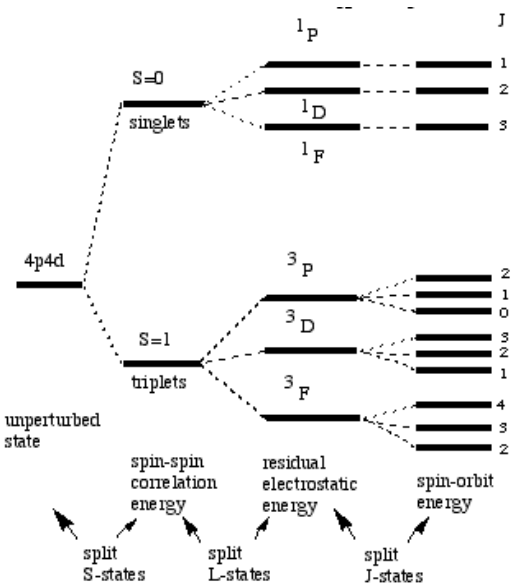


Figure 12.9: LS 4p4d Energy Levels

In the jj-coupling scheme the energy leveldiagram is shown in Figure 12.10

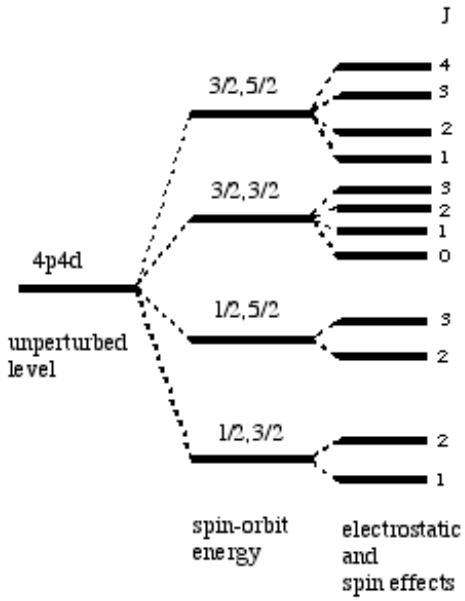


Figure 12.10: jj 4p4d Energy Levels

Notice that the final J values are identical, but their arrangement in energy is very different.

12.8 Spherical Harmonics Addition Theorem

In Chapter 9 we defined the properties of the spherical harmonics. We found the following results.

12.8.1 Orbital Angular Momentum

Abstractly,

$$\left[\hat{L}_i, \hat{L}_j\right] = i\hbar\varepsilon_{ijk}\hat{L}_k \text{ and } \left[\hat{L}^2_{op}, \hat{L}_j\right] = 0 \tag{12.193}$$

$$\hat{L}^2_{op}|\ell m\rangle = \hbar^2\ell(\ell+1)|\ell m\rangle \text{ and } \hat{L}_3|\ell m\rangle = \hbar m|\ell m\rangle \tag{12.194}$$

$$\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y \tag{12.195}$$

$$\ell = \frac{\text{integer}}{2} \geq 0 \tag{12.196}$$

For a given value of ℓ, m takes on the $2\ell+1$ values

$$m = -\ell, -\ell+1, -\ell+2, ..., \ell-2, \ell-1, \ell$$

In ordinary 3–dimensional space, if we define

$$Y_{\ell m}(\theta,\varphi) = \langle \theta\varphi \mid \ell m\rangle = \textit{spherical harmonic} \tag{12.197}$$

then we have the defining equations for the $Y_{\ell m}(\theta,\varphi)$ given by

$$\begin{aligned}\langle\theta\varphi|\,\vec{L}^2_{op}\,|\ell m\rangle &= \vec{L}^2_{op}\,\langle\theta\varphi\mid\ell m\rangle = \vec{L}^2_{op}Y_{\ell m}(\theta,\varphi)\\ &= \hbar^2\ell(\ell+1)\,\langle\theta\varphi\mid\ell m\rangle = \hbar^2\ell(\ell+1)Y_{\ell m}(\theta,\varphi) \qquad (12.198)\\ \langle\theta\varphi|\,\hat{L}_3\,|\ell m\rangle &= \hat{L}_3\,\langle\theta\varphi\mid\ell m\rangle = \hat{L}_3Y_{\ell m}(\theta,\varphi)\\ &= \hbar m\,\langle\theta\varphi\mid\ell m\rangle = \hbar m Y_{\ell m}(\theta,\varphi) \qquad (12.199)\end{aligned}$$

The general result is

$$Y_{\ell m}(\theta,\varphi) - \frac{(-1)\ell}{2^\ell\ell!}\sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell+m)!}{(\ell-m)!}}\frac{e^{im\varphi}}{(\sin\theta)^m}\left(\frac{d}{d\cos\theta}\right)^{\ell-m}(\sin\theta)^{2\ell} \tag{12.200}$$

Some examples are:

$$Y_{00} = \frac{1}{\sqrt{4\pi}} \tag{12.201}$$

$$Y_{10} = \sqrt{\frac{3}{4\pi}}\cos\theta\ ,\ Y_{1,\pm1} = \mp\sqrt{\frac{3}{8\pi}}e^{\pm i\varphi}\sin\theta \tag{12.202}$$

$$Y_{20} = \sqrt{\frac{5}{16\pi}}(3\cos^2\theta-1)\ ,\ Y_{2,\pm1} = \mp\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{\pm i\varphi} \tag{12.203}$$

$$Y_{2,\pm2} = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{\pm 2i\varphi} \tag{12.204}$$

Some Properties

Complex Conjugate

$$Y_{\ell,-m}(\theta,\varphi) = (-1)^m Y^*_{\ell,m}(\theta,\varphi) \tag{12.205}$$

Under the parity operation

$$\vec{r}\to-\vec{r}\quad or\quad r\to r,\theta\to\pi-\theta,\varphi\to\varphi+\pi$$

which says that

$$\begin{aligned}e^{im\varphi}\to e^{im\varphi}e^{im\pi} &= (-1)^m e^{im\varphi}\\ \sin\theta\to\sin(\pi-\theta)&\to\sin\theta\\ \cos\theta\to\cos(\pi-\theta)&\to-\cos\theta\end{aligned}$$

which imply that

$$Y_{\ell,m}(\theta,\varphi)\to(-1)^\ell Y_{\ell,m}(\theta,\varphi) \tag{12.206}$$

Therefore,

if ℓ is even, then we have an even parity state
if ℓ is odd, then we have an odd parity state

Since they form a complete set, any function of (θ, φ) can be expanded in terms of the $Y_{\ell,m}(\theta, \varphi)$ (the $Y_{\ell,m}(\theta, \varphi)$ are a basis), i.e., we can write

$$f(\theta, \varphi) = \sum_{\ell,m} f_{\ell m} Y_{\ell,m}(\theta, \varphi) \tag{12.207}$$

where

$$f_{\ell m} = \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta Y^*_{\ell' m'}(\theta, \varphi) f(\theta, \varphi) \tag{12.208}$$

and we have used the orthonormality relation

$$\int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta Y^*_{\ell' m'}(\theta, \varphi) Y_{\ell m}(\theta, \varphi) = \delta_{\ell' \ell} \delta_{m' m} \tag{12.209}$$

The spherical harmonics also satisfy these relations:

Closure:

$$\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} Y^*_{\ell m}(\theta, \phi) Y_{\ell m}(\theta', \phi') = \frac{\delta(\theta - \theta')\delta(\phi - \phi')}{\sin\theta} \equiv \delta(\hat{r}, \hat{r}') \tag{12.210}$$

i.e., the solid angle delta function is equal to zero unless the two vectors $\hat{r}(\theta, \phi)$, $\hat{r}'(\theta', \phi')$ coincide. It has the property

$$\int f(\hat{r}') \delta(\hat{r}, \hat{r}') d\Omega' = f(\hat{r}) \tag{12.211}$$

for any function $f(\vec{r})$ of the spatial direction specified by θ, φ.

Recursion:

$$\begin{aligned} \hat{L}_{\pm} Y_{\ell m} &= \left[\ell(\ell+1) - m(m \pm 1)\right]^{1/2} Y_{\ell, m\pm 1} \\ &= \left[(\ell \mp m)(\ell + 1 \pm m)\right]^{1/2} Y_{\ell, m\pm 1} \end{aligned} \tag{12.212}$$

$$\begin{aligned} \cos\theta Y_{\ell m} = &\left[\frac{(\ell + 1 + m)(\ell + 1 - m)}{(2\ell + 1)(2\ell + 3)}\right]^{1/2} Y_{\ell+1, m} \\ &+ \left[\frac{(\ell + m)(\ell - m)}{(2\ell + 1)(2\ell - 1)}\right]^{1/2} Y_{\ell-1, m} \end{aligned} \tag{12.213}$$

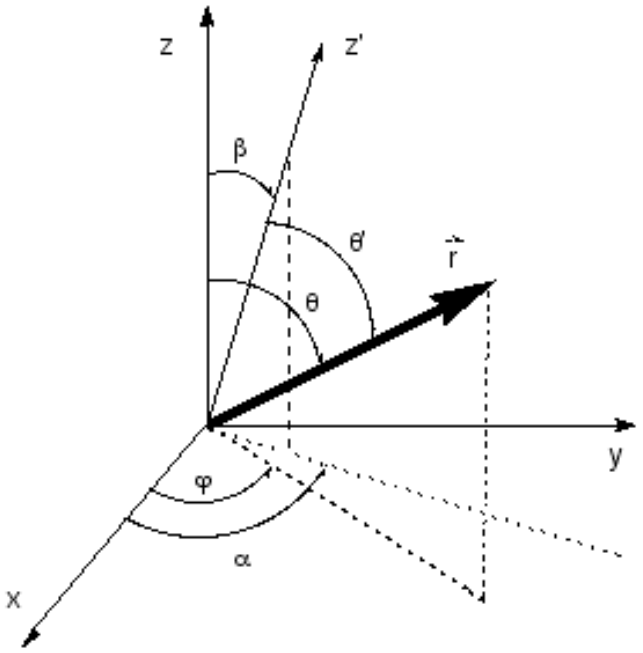


Figure 12.11: Angles Used in the Addition Theorem

12.8.2 The Addition Theorem

Consider two coordinate systems xyz and $x'y'z'$. The addition theorem is the formula expressing the eigenfunction $P_\ell(\cos\theta')$ of the angular momentum $\hat{L}_{z'}$ about the $z'-$axis in terms of the eigenfunctions $Y_{\ell,m}(\theta,\varphi)$ of $\hat{L}_z$. . See Figure 12.11 below for orientations.

The angles α and β are the azimuth and the polar angles of the z' axis in the Cartesian xyz coordinate frame. They are also the first two Euler angles specifying the orientation of the Cartesian coordinate system $x'y'z'$ with respect to xyz. The third Euler angle γ is left unspecified here and the x' and y' axes are not shown. The projections of the z' axis and the radius vector on the xy plane are dashed lines.

As we can see the position vector $\vec{r}$ has angular coordinates θ,φ and θ',φ' in the two coordinate systems.

The direction of the z' axis in space is specified by its polar angle β and its azimuth angle α with respect to the xyz system.

Since P_ℓ is an eigenfunction of $\vec{L}_{op}^2$, only spherical harmonics with the same subscript ℓ can appear in the expansion.

An interchange of θ, φ and β, α is equivalent to the transformation $\theta \to -\theta'$ and must leave the expansion unchanged because $P_\ell(\cos\theta')$ is an even function of θ'. This means that $P_\ell(\cos\theta')$ must be a function of $\varphi - \alpha$.

All of these requirements are satisfied only if we write

$$P_\ell(\cos\theta') = \sum_{m=-\ell}^{\ell} c_m Y_{\ell,-m}(\beta,\alpha) Y_{\ell,m}(\theta,\phi) \tag{12.214}$$

We determine the coefficients c_m using the conditions

$$\hat{L}_{z'} P_\ell(\cos\theta') = 0 \tag{12.215}$$

We also use the identity (from rotation of a vector component)

$$\begin{aligned} \hat{L}_{z'} &= \sin\beta\cos\alpha \hat{L}_x + \sin\beta\sin\alpha \hat{L}_y + \cos\beta \hat{L}_z \\ &= \frac{1}{2}\sin\beta e^{-i\alpha}\hat{L}_+ + \frac{1}{2}\sin\beta e^{i\alpha}\hat{L}_- + \cos\beta\hat{L}_z \end{aligned} \tag{12.216}$$

and invoke the linear independence of the spherical harmonics to obtain (after some algebra)

$$c_{m\pm1} = -c_m \to c_m = (-1)^m c_0 \tag{12.217}$$

so that we only need to determine c_0. We specialize to $\beta = 0$ or $\theta' = \theta$ so that we have the relations

$$Y_{\ell m}(0,\phi) = \sqrt{\frac{2\ell+1}{4\pi}}\delta_{m0} \tag{12.218}$$

$$Y_{\ell 0}(\theta,\phi) = \sqrt{\frac{2\ell+1}{4\pi}} P_\ell(\cos\theta) \tag{12.219}$$

We then have

$$\begin{aligned} P_\ell(\cos\theta) &= \sum_{m=-\ell}^{\ell} c_m Y_{\ell,-m}(0,\alpha) Y_{\ell,m}(\theta,\phi) = \sum_{m=-\ell}^{\ell} c_m \sqrt{\frac{2\ell+1}{4\pi}}\delta_{m0} Y_{\ell,m}(\theta,\phi) \\ &= c_0 \sqrt{\frac{2\ell+1}{4\pi}} Y_{\ell,0}(\theta,\phi) = c_0 \sqrt{\frac{2\ell+1}{4\pi}}\sqrt{\frac{2\ell+1}{4\pi}} P_\ell(\cos\theta) \end{aligned} \tag{12.220}$$

or

$$c_0 = \frac{4\pi}{2\ell+1} \tag{12.221}$$

and we end up with the addition theorem

$$P_\ell(\cos\theta') = \frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{\ell} c_m Y^*_{\ell m}(\beta,\alpha) Y_{\ell,m}(\theta,\phi) \tag{12.222}$$

where θ' = angle between the directions (β, α) and (θ, φ).

If we combine the closure relation with the addition theorem we get the identity

$$\sum_{\ell=0}^{\infty} (2\ell+1) P_\ell(\hat{r}\cdot\hat{r}') = 4\pi\delta(\hat{r},\hat{r}') \tag{12.223}$$

Since we can write

$$\delta(\vec{r}-\vec{r}') = \frac{\delta(r-r')}{r^2}\delta(\hat{r},\hat{r}') \tag{12.224}$$

we then have the identity

$$\delta(\vec{r}-\vec{r}') = \frac{\delta(r-r')}{r^2}\sum_{\ell=0}^{\infty}\frac{2\ell+1}{4\pi}P_\ell(\hat{r}\cdot\hat{r}') \tag{12.225}$$

Another useful relation is

$$e^{ikz} = \sum_{\ell=0}^{\infty}(2\ell+1)i^\ell j_\ell(kr)P_\ell(\cos\theta) \tag{12.226}$$

or, in general

$$e^{i\vec{k}\cdot\vec{r}} = 4\pi\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} i^\ell j_\ell(kr)Y^*_{\ell m}(\theta_{\vec{k}},\phi_{\vec{k}})Y_{\ell m}(\theta_{\vec{r}},\phi_{\vec{r}}) \tag{12.227}$$

12.9 Problems

12.9.1 Two Bosons in a Well

Two identical spin-zero bosons are placed in a 1−dimensional square potential well with infinitely high walls, i.e., $V = 0$ for $0 < x < L$, otherwise $V = \infty$. The normalized single particle energy eigenstates are

$$u_n(x) = \sqrt{\frac{2}{L}}\sin(n\pi x/L)$$

(a) Find the wavefunctions and energies for the ground state and the first two excited states of the system.

(b) Suppose that the two bosons interact with each other through the perturbing potential

$$H'(x_1,x_2) = -LV_0\delta(x_1-x_2)$$

Compute the first-order correction to the ground state energy of the system.

12.9.2 Two Fermions in a Well

Two identical spin−1/2 bosons are placed in a 1−dimensional square potential well with infinitely high walls, i.e., $V = 0$ for $0 < x < L$, otherwise $V = \infty$. The normalized single particle energy eigenstates are

$$u_n(x) = \sqrt{\frac{2}{L}} \sin(n\pi x/L)$$

(a) What are the allowed values of the total spin angular momentum quantum number, J ? How many possible values are there fore the z−component of the total angular momentum?

(b) If single-particle spin eigenstates are denoted by $|\uparrow\rangle = u$ and $|\downarrow\rangle = d$, construct the two-particle spin states that are either symmetric or antisymmetric. How many states of each type are there?

(c) Show that the $j = 1$, $m = 1$ state must be symmetric. What is the symmetry of the $J = 0$ state?

(d) What is the ground-state energy of the two-particle system, and how does it depend on the overall spin state?

12.9.3 Two spin−1/2 particles

The Hamiltonian for two spin−1/2 particles, one with mass m_1 and the other with m_2, is given by

$$\hat{H} = \frac{\vec{p}_1^{\,2}}{2m_1} + \frac{\vec{p}_2^{\,2}}{2m_2} + V_a(r) + \left(\frac{1}{4} - \frac{\vec{S}_1 \cdot \vec{S}_2}{\hbar^2}\right) V_b(r)$$

where $|\vec{r}| = \vec{r}_1 - \vec{r}_2$, $|\vec{r}| = r$ and

$$V_a(r) = \begin{cases} 0 & \text{for } r < a \\ V_0 & \text{for } r > a \end{cases} \quad , \quad V_b(r) = \begin{cases} 0 & \text{for } r < b \\ V_0 & \text{for } r > b \end{cases}$$

with $b < a$ and V_0 very large (assume V_0 is infinite where appropriate) and positive.

(a) Determine the normalized position-space energy eigenfunction for the ground state. What is the spin state of the ground state? What is the degeneracy?

(b) What can you say about the energy and spin state of the first excited state? Does your result depend on how much larger a is than b? Explain.

12.9.4 Hydrogen Atom Calculations

We discuss here some useful tricks for evaluating the expectation values of certain operators in the eigenstates of the hydrogen atom.

(a) Suppose we want to determine $\langle 1/r \rangle_{n\ell m}$. We can interpret $\langle \lambda/r \rangle_{n\ell m}$ as the 1^{st}–order correction due to the perturbation λ/r (same dependence on r as the potential energy). Show that this problem can be solved exactly by just replacing e^2 by $e^2 - \lambda$ everywhere in the original solution. So, the exact energy is

$$E(\lambda) = -\frac{m\left(e^2-\lambda\right)^2}{2n^2\hbar^2}$$

the 1^{st}–order correction is the term linear in λ, that is,

$$E^{(1)} = \frac{me^2\lambda}{n^2\hbar^2} = \langle \lambda/r \rangle_{n\ell m}$$

Therefore we get

$$\langle 1/r \rangle_{n\ell m} = \frac{me^2}{n^2\hbar^2} = \frac{1}{n^2 a_0}$$

We note (for later use) that

$$E(\lambda) = E^{(0)} + E^{(1)} + ... = E(\lambda = 0) + \lambda\left(\frac{dE}{d\lambda}\right)_{\lambda=0} + ...$$

so that one way to extract $E^{(1)}$ from the exact answer is to calculate

$$\lambda\left(\frac{dE}{d\lambda}\right)_{\lambda=0}$$

(b) Evaluate, in a manner similar to part (a), $\left\langle \vec{p}^2/2\mu \right\rangle_{n\ell m}$ by considering the Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2\mu} - \frac{Ze^2}{r} + \lambda\frac{\hat{p}^2}{2\mu}$$

(c) Consider now $\left\langle \lambda/r^2 \right\rangle_{n\ell m}$. In this case, an exact solution is possible since the perturbation just modifies the centrifugal term as follows:

$$\frac{\hbar^2\ell(\ell+1)}{2mr^2} + \frac{\lambda}{r^2} = \frac{\hbar^2\ell'(\ell'+1)}{2mr^2}$$

where ℓ' is a function of λ. Now go back to the original hydrogen atom solution and show that the dependence of E on $\ell'(\lambda)$ is

$$E(\ell') = -\frac{mZ^2e^4}{2\hbar^2(k+\ell'+1)^2} = E(\lambda) = E^{(0)} + E^{(1)} + ...$$

Then show that

$$\left\langle \lambda/r^2 \right\rangle_{n\ell m} = E^{(1)}\lambda\left(\frac{dE}{d\lambda}\right)_{\lambda=0} = \lambda\left(\frac{dE}{d\ell'}\right)_{\ell'=\ell}\left(\frac{d\ell'}{d\lambda}\right)_{\ell'=\ell}$$
$$= \frac{\lambda}{n^3 a_0^2(\ell+1/2)}$$

or

$$\left\langle 1/r^2 \right\rangle_{n\ell m} = \frac{1}{n^3 a_0^2(\ell+1/2)}$$

(d) Finally consider $\left\langle \lambda/r^3 \right\rangle_{n\ell m}$. Since there is no such term in the hydrogen Hamiltonian, we resort to different trick. Consider the radial momentum operator

$$p_r = -i\hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)$$

Show that in terms of this operator we may write the radial part of the Hamiltonian

$$-\frac{\hbar^2}{2m}\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right)$$

as

$$\frac{p_r^2}{2m}$$

Now show that

$$\langle [H, p_r] \rangle = 0$$

in the energy eigenstates. Using this fact, and by explicitly evaluating the commutator, show that

$$\left\langle 1/r^3 \right\rangle_{n\ell m} = \frac{Z}{a_0\ell(\ell+1)}\left\langle 1/r^2 \right\rangle_{n\ell m}$$

and hence

$$\left\langle 1/r^3 \right\rangle_{n\ell m} = \frac{Z^3}{n^3 a_0^3 \ell(\ell+1)(\ell+1/2)}$$

12.9.5 Hund's rule

Explain on the basis of Hund's rules why the ground state of carbon is $^3P_0$ and that of oxygen is $^3P_2$.

12.9.6 Russell-Saunders Coupling in Multielectron Atoms

Consider a configuration of k equivalent p electrons outside a closed shell, which we denote simply by p^k, i.e., carbon= p^2, nitrogen= p^3 and oxygen= p^4.

(a) Use the implied-terms method to determine all the terms that can arise from p^3. Which of them will have the lowest energy?

(b) Repeat this calculation for p^4 and show that we get the same result as for p^2

12.9.7 Magnetic moments of proton and neutron

The magnetic dipole moment of the proton is

$$\hat{\mu}_p = g_p \frac{e}{2m_p} \hat{S}_p$$

with a measured magnitude corresponding to a value for the gyromagnetic ratio of

$$g_p = 2 \times (2.792847337 \pm 0.000000029)$$

We have not studied the Dirac equation yet, but the prediction of the Dirac equation for a point spin−1/2 particle is $g_p = 2$. We can understand the fact that the proton gyromagnetic ratio is not two as being due its compositeness, i.e., in a simple quark model, the proton is made up of three quarks, two *ups* (u), and a *down* (d). The quarks are supposed to be point spin−1/2, hence, their gyromagnetic ratios should be $g_u = g_d = 2$ (up to higher order corrections, as in the case of the electron). Let us see if we can make sense out of the proton magnetic moment.

The proton magnetic moment should be the sum of the magnetic moments of its constituents, and any moments due to their orbital motion in the proton. The proton is the ground state baryon, so we assume that the three quarks are bound together (by the strong interaction) in a state with no orbital angular momentum. The Pauli principle says that the two identical up quarks must have an overall odd wave function under interchange of all quantum numbers. We must apply this rule with some care since we will be including *color* as one of these quantum numbers.

Let us look at some properties of *color*. It is the strong interaction analog of electric charge in the electromagnetic interaction. However, instead of one fundamental dimension in charge, there are three color directions, labeled as *red* (r), *blue* (b), and *green* (g). Unitary transformations in this color space(up to overall phases) are described by elements of the group $SU(3)$, the group of unimodular 3×3 matrices (electromagnetic charge corresponds to the group $U(1)$ whose elements are local phase changes). Just like combining spins, we can combine these three colors according to a Clebsch-Gordon series, with the result

$$3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1$$

These are different rules than for the addition of spin case because that case uses the rotation group instead. We do not need to understand all aspects of the $SU(3)$ group for this problem. The essential aspect here is that there is a singlet in the decomposition, i.e., it is possible to combine three colors in a way as to get a color singlet state or a state with no net color charge. These turn out to be the states of physical interest for the observed baryons according to a postulate of the quark model.

(a) The singlet state in the decomposition above must be antisymmetric under the interchange of any two colors. Assuming this is the case, write down the color portion of the proton wave function.

(b) Now that you know the color wave function of the quarks in the proton, write down the spin wave function. You must construct a total spin state $|1/2, 1/2\rangle$ total spin angular momentum state from three spin$-1/2$ states where the two up quarks must be in a symmetric state.

(c) Since the proton is uud and its partner the neutron (the are just two states of the same particle) is ddu and $m_p \simeq m_n$, we can make the simplifying assumption that $m_u \simeq m_d$. Given the measured value of g_p, what does you model give for m_u? Remember that the up quark has electric charge 2/3 and the down quark has electric charge $-1/3$, in units of positron charge.

(d) Finally, use your results to predict the gyromagnetic moment of the neutron(neutron results follows from proton results by interchanging u and d labels) and compare with observation.

12.9.8 Particles in a 3-D harmonic potential

A particle of mass m moves in a 3–dimensional harmonic oscillator well. The Hamiltonian is

$$\hat{H} = \frac{\vec{p}^{\,2}}{2m} + \frac{1}{2}kr^2$$

(a) Find the energy and orbital angular momentum of the ground state and the first three excited states.

(b) If eight identical non-interacting (spin $1/2$) particles are placed in such a harmonic potential, find the ground state energy for the eight-particle system.

(c) Assume that these particles have a magnetic moment of magnitude μ. If a magnetic field B is applied, what is the approximate ground state energy of the eight-particle system as a function of B (what is the effect of a closed shell?). Determine the magnetization $-\partial E/\partial B$ for the ground state as a function of B. What is the susceptibility? Don't do any integrals.

12.9.9 2 interacting particles

Consider two particles of masses $m_1 \neq m_2$ interacting via the Hamiltonian

$$\hat{H} = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \frac{1}{2}m_1\omega^2 x_1^2 + \frac{1}{2}m_2\omega^2 x_2^2 + \frac{1}{2}K\left(x_1 - x_2\right)^2$$

(a) Find the exact solutions.

(b) Sketch the spectrum in weak coupling limit $K << \mu\omega^2$ where μ = reduced mass.

12.9.10 LS versus JJ coupling

Consider a multielectron atom whose electron configuration is

$$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p 4d$$

(a) To what element does this configuration belong? Is it the ground state or an excited state? Explain.

(b) Suppose that we apply the Russell-Saunders coupling scheme to this atom. Draw and energy level diagram roughly to scale for the atom, beginning with the single unperturbed configuration energy and taking into account the various interactions one at a time in the correct order. Be sure to label each level at each stage of your diagram with the appropriate term designation, quantum numbers and so on.

(c) Suppose instead we apply pure jj–coupling to the atom. Starting again from the unperturbed $n = 4$ level, draw a second energy level diagram. [HINT: Assume that for a given level (j_1, j_2), the state with the lowest J lies lowest in energy]

12.9.11 In a harmonic potential

Two identical, noninteracting spin= 1/2 particles of mass m are in a one dimensional harmonic oscillator potential for which the Hamiltonian is

$$H = \frac{p_{1x}^2}{2m} + \frac{1}{2}m\omega^2 x_1^2 + \frac{p_{2x}^2}{2m} + \frac{1}{2}m\omega^2 x_2^2$$

(a) Determine the ground-state and first-excited state kets and the corresponding energies when the two particles are in a total spin= 0 state. What are the lowest energy states and the corresponding kets for the particles if they are in a total spin= 1 state?

(b) Suppose that the two particles interact with a potential energy of interaction

$$V(|x_1 - x_2|) = \begin{cases} -V_0 & |x_1 - x_2| < a \\ 0 & \text{elsewhere} \end{cases}$$

Argue what the effect will be on the energies that you determined in (a), that is, whether the energy of each state moves up, moves down, or remains unchanged.

12.9.12 2 particles interacting via delta function

Two particles of mass m are placed in a rectangular box of sides $a > b > c$ in the lowest energy state of the system compatible with the conditions below. The particles interact with each other according to the potential $V = A\delta(\vec{r}_1 - \vec{r}_2)$. Using first order perturbation theory calculate the energy of the system under the following conditions:

(a) particles are not identical

(b) identical particles of spin= 0

(c) identical particles of spin= 1/2 with spins parallel

12.9.13 2 particles in a square well

Two identical nonrelativistic fermions of mass m, spin= 1/2 are in a 1−dimensional square well of length L with V infinitely large outside the well. The fermions are subject to a repulsive potential $V(x_1 - x_2)$, which may be treated as a perturbation.

(a) Classify the three lowest-energy states in terms of the states of the individual particles and state the spin of each.

(b) Calculate to first-order the energies of the second- and third- lowest states; leave your result in the form of an integral. Neglect spin-dependent forces throughout.

12.9.14 2 particles interacting via a harmonic potential

Two particles, each of mass M are bound in a 1−dimensional harmonic oscillator potential

$$V = \frac{1}{2}kx^2$$

and interact with each other through an attractive harmonic force $F_{12} = -K(x_1 - x_2)$. Assume that K is very small.

(a) What are the energies of the three lowest states of this system?

(b) If the particles are identical and spinless, which of the states of (a) are allowed?

(c) If the particles are identical and have spin= 1/2, which of the states of (a) are allowed?

12.9.15 The Structure of helium

Consider a Helium atom in the $1s2p$ configuration. The total angular momentum is $L = 1$ (a P−state). Due to the Fermi-Pauli symmetry this state splits into singlet and triplet multiplets as shown below.
where the superscripts 1 and 3 represent the spin degeneracy for the singlet/triplet respectively.

(a) Explain qualitatively why the triplet state has lower energy.

Now include spin-orbit coupling described by the Hamiltonian $\hat{H}_{SO} = f(r)\hat{L}\cdot\hat{S}$, where $\hat{L}$ and $\hat{S}$ are the total orbital and spin angular momentum respectively.

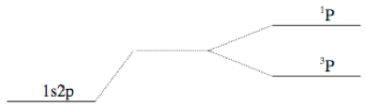


Figure 12.12: Fermi-Pauli Splittings

(b) Without the spin-orbit interaction, good quantum numbers for the angular momentum degrees of freedom are $|LM_LSM_S\rangle$. What are the good quantum numbers with spin-orbit present?

(c) The energy level diagram including spin-orbit corrections is sketched below.

Figure 12.13: Including Spin-Orbit

Label the states with appropriate quantum numbers. NOTE: Some of the levels are degenerate; the sublevels are not shown.

Chapter 13

Some Examples of Quantum Systems

13.1 Coherent and Squeezed States

We have derived relationships between the non-Hermitian operators $\hat{a}$ and $\hat{a}^\dagger$ and the position and momentum operators

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a}^+ + \hat{a}) \quad , \quad \hat{p} = i\sqrt{\frac{m\hbar\omega}{2}}(\hat{a}^+ - \hat{a}) \tag{13.1}$$

Working with the coherent states defined by

$$\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle \tag{13.2}$$

we found

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{m=0}^{\infty} \frac{\alpha^m}{\sqrt{m!}}\,|m\rangle \tag{13.3}$$

where

$$\left|\alpha^2\right| = N = \langle\alpha|\,\hat{N}_{op}\,|\alpha\rangle \tag{13.4}$$

Let us now derive some important relations.

$$\begin{aligned}\langle\alpha|\,\hat{x}\,|\alpha\rangle &= \sqrt{\frac{\hbar}{2m\omega}}\,\langle\alpha|\,(\hat{a}^+ + \hat{a})\,|\alpha\rangle \\ &= \sqrt{\frac{\hbar}{2m\omega}}\,(\alpha + \alpha^*) = \sqrt{\frac{2\hbar}{m\omega}}\mathrm{Real}(\alpha) \qquad (13.5)\\ \langle\alpha|\,\hat{p}\,|\alpha\rangle &= i\sqrt{\frac{\hbar m\omega}{2}}\,\langle\alpha|\,(\hat{a}^+ - \hat{a})\,|\alpha\rangle \\ &= i\sqrt{\frac{\hbar m\omega}{2}}\,(\alpha^* - \alpha) = \sqrt{2m\hbar\omega}\mathrm{Imag}(\alpha) \qquad (13.6)\end{aligned}$$

$$\begin{aligned}\langle\alpha|\,\hat{x}^2\,|\alpha\rangle &= \frac{\hbar}{2m\omega}\,\langle\alpha|\,(\hat{a}^+\hat{a}^+ + \hat{a}\hat{a}^+ + \hat{a}^+\hat{a} + \hat{a}\hat{a})\,|\alpha\rangle \\ &= \frac{\hbar}{2m\omega}\,\langle\alpha|\,(\alpha^{*2} + \alpha\alpha^* + \alpha^*\alpha + \alpha^2 + 1)\,|\alpha\rangle \\ &= \langle\alpha|\,\hat{x}\,|\alpha\rangle^2 + \frac{\hbar}{2m\omega} \end{aligned} \tag{13.7}$$

$$\begin{aligned}\langle\alpha|\,\hat{p}^2\,|\alpha\rangle &= -\frac{\hbar m\omega}{2}\,\langle\alpha|\,(\hat{a}^+\hat{a}^+ - \hat{a}\hat{a}^+ - \hat{a}^+\hat{a} + \hat{a}\hat{a})\,|\alpha\rangle \\ &= -\frac{\hbar m\omega}{2}\,\langle\alpha|\,(\alpha^{*2} - \alpha\alpha^* - \alpha^*\alpha + \alpha^2 - 1)\,|\alpha\rangle \\ &= \langle\alpha|\,\hat{p}\,|\alpha\rangle^2 + \frac{\hbar m\omega}{2} \end{aligned} \tag{13.8}$$

Using these relations we have

$$(\Delta x)^2 = \langle\alpha|\,\hat{x}^2\,|\alpha\rangle - \langle\alpha|\,\hat{x}\,|\alpha\rangle^2 = \frac{\hbar}{2m\omega} \tag{13.9}$$

$$(\Delta p)^2 = \langle\alpha|\,\hat{p}^2\,|\alpha\rangle - \langle\alpha|\,\hat{p}\,|\alpha\rangle = \frac{\hbar m\omega}{2} \tag{13.10}$$

Hence

$$\Delta x\Delta p = \frac{\hbar}{2} \tag{13.11}$$

which says that coherent states are minimum uncertainty states.

Now let us find the differential equation satisfied by $\langle x\,|\,\alpha\rangle$ and determine its solution. We have

$$\begin{aligned}\langle x|\,\hat{a}\,|\alpha\rangle = \alpha\,\langle x\mid\alpha\rangle &= \frac{1}{2}\,\langle x|\left(\sqrt{\frac{2m\omega}{\hbar}}\hat{x} - \frac{\hat{p}}{i}\sqrt{\frac{2}{m\omega\hbar}}\right)|\alpha\rangle \\ &= \frac{1}{2}\left(\sqrt{\frac{2m\omega}{\hbar}}\hat{x} - \frac{1}{i}\sqrt{\frac{2}{m\omega\hbar}}\left(-i\hbar\frac{d}{dx}\right)\right)\langle x\mid\alpha\rangle \end{aligned}$$

or

$$\left(\hat{x} + \frac{\hbar}{m\omega}\frac{d}{dx}\right)\langle x\mid\alpha\rangle = \sqrt{\frac{2\hbar}{m\omega}}\,\langle x\mid\alpha\rangle \tag{13.12}$$

which has the solution (check by substitution)

$$\langle x\mid\alpha\rangle = C\exp\left[-\frac{(x-\langle x\rangle)^2}{4(\Delta x)^2} + \frac{i}{\hbar}\langle p\rangle\,x\right] \tag{13.13}$$

$$= C'\exp\left[-\frac{m\omega}{2\hbar}\left(x - \sqrt{\frac{2\hbar}{m\omega}}\alpha\right)^2\right] \tag{13.14}$$

For a fixed oscillator mode, specified by a given value of $m\omega$, the coherent states are the manifold of those minimum uncertainty states that have definite values of Δx and Δp. If $m\omega = 1$, then the uncertainties in x and p are both equal to $\sqrt{\hbar/2}$.

We can construct other minimum uncertainty states with a narrower Δx, the so-called *squeezed states*, for the same oscillator by defining a new set of raising and lowering operators.

$$\hat{b} = \sqrt{\frac{m\omega'}{2\hbar}}\left(\hat{x} + i\frac{\hat{p}}{m\omega'}\right) \quad , \quad \hat{b}^+ = \sqrt{\frac{m\omega'}{2\hbar}}\left(\hat{x} - i\frac{\hat{p}}{m\omega'}\right) \tag{13.15}$$

where we introduce an arbitrarily chosen positive parameter ω'. Now we have

$$\begin{aligned}\left[\hat{b}, \hat{b}^+\right] &= \frac{m\omega'}{2\hbar}\left(-\frac{i}{m\omega'}\left[\hat{x}, \hat{p}\right] + \frac{i}{m\omega'}\left[\hat{p}, \hat{x}\right]\right) \\ &= -\frac{i}{\hbar}\left[\hat{x}, \hat{p}\right] = -\frac{i}{\hbar} i\hbar = 1\end{aligned} \tag{13.16}$$

We also have

$$\begin{aligned}\hat{b} &= \sqrt{\frac{m\omega'}{2\hbar}}\left(\hat{x} + i\frac{\hat{p}}{m\omega'}\right) \\ &= \sqrt{\frac{m\omega'}{2\hbar}}\left(\sqrt{\frac{\hbar}{2m\omega}}(\hat{a} + \hat{a}^+) + \frac{i}{m\omega'} i\sqrt{\frac{\hbar m\omega}{2}}(\hat{a}^+ - \hat{a})\right) \\ &= \left(\sqrt{\frac{m\omega'}{2\hbar}}\sqrt{\frac{\hbar}{2m\omega}} + \sqrt{\frac{m\omega'}{2\hbar}}\frac{i}{m\omega'} i\sqrt{\frac{\hbar m\omega}{2}}\right)\hat{a}^+ \\ &\quad + \left(\sqrt{\frac{m\omega'}{2\hbar}}\sqrt{\frac{\hbar}{2m\omega}} - \sqrt{\frac{m\omega'}{2\hbar}}\frac{i}{m\omega'} i\sqrt{\frac{\hbar m\omega}{2}}\right)\hat{a} \\ &= \frac{1}{2}\left(\sqrt{\frac{\omega'}{\omega}} - \sqrt{\frac{\omega}{\omega'}}\right)\hat{a}^+ + \frac{1}{2}\left(\sqrt{\frac{\omega'}{\omega}} + \sqrt{\frac{\omega}{\omega'}}\right)\hat{a} \\ &= \lambda\hat{a} + \nu\hat{a}^+\end{aligned} \tag{13.17}$$

where

$$\lambda = \frac{1}{2}\left(\sqrt{\frac{\omega'}{\omega}} - \sqrt{\frac{\omega}{\omega'}}\right) \quad , \quad \nu = \frac{1}{2}\left(\sqrt{\frac{\omega'}{\omega}} + \sqrt{\frac{\omega}{\omega'}}\right) \tag{13.18}$$

and therefore

$$\hat{b}^+ = \lambda\hat{a}^+ + \nu\hat{a} \tag{13.19}$$

since λ and ν are real. Algebra also shows that

$$\lambda^2 - \nu^2 = \frac{1}{4}\left[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \frac{\omega'}{\omega} - \frac{\omega}{\omega'} + 2 + 2\right] = 1 \tag{13.20}$$

We now invert the transformation. We have

$$\lambda\hat{b} - \nu\hat{b}^+ = \lambda^2\hat{a} + \lambda\nu\hat{a}^+ - \lambda\nu\hat{a}^+ - \nu^2\hat{a} = (\lambda^2 - \nu^2)\hat{a} = \hat{a} \tag{13.21}$$

which then implies that

$$\lambda\hat{b}^+ - \nu\hat{b} = \hat{a}^+ \tag{13.22}$$

We define the eigenstates of the lowering operator $\hat{b}$ by

$$\hat{b}\,|\beta\rangle = \beta\,|\beta\rangle \tag{13.23}$$

These new states are also minimum uncertainty states for x and p. We want to calculate

$$(\Delta x)^2 = \left\langle x^2 \right\rangle - \langle x\rangle^2 \quad , \quad (\Delta p)^2 = \left\langle p^2 \right\rangle - \langle p\rangle^2 \tag{13.24}$$

Now we have

$$\begin{aligned}
\hat{x} &= \sqrt{\frac{\hbar}{2m\omega}}\,(\hat{a} + \hat{a}^+) \\
&= \sqrt{\frac{\hbar}{2m\omega}}\left((\lambda - \nu)\,\hat{b} + (\lambda - \nu)\,\hat{b}^+\right) \\
&= \sqrt{\frac{\hbar}{2m\omega}}\,(\lambda - \nu)\left(\hat{b} + \hat{b}^+\right)
\end{aligned} \tag{13.25}$$

and

$$\begin{aligned}
\hat{p} &= i\sqrt{\frac{\hbar m\omega}{2}}\,(\hat{a}^+ - \hat{a}) \\
&= i\sqrt{\frac{\hbar m\omega}{2}}\left((\lambda + \nu)\,\hat{b}^+ - (\lambda + \nu)\,\hat{b}\right) \\
&= i\sqrt{\frac{\hbar m\omega}{2}}\,(\lambda + \nu)\left(\hat{b}^+ - \hat{b}\right)
\end{aligned} \tag{13.26}$$

These equations imply that earlier derivation we did is the same with different multiplicative factors so that

$$(\Delta x)^2 = \frac{\hbar}{2m\omega}\,(\lambda - \nu)^2 \tag{13.27}$$

$$(\Delta p)^2 = \frac{\hbar m\omega}{2}\,(\lambda + \nu)^2 \tag{13.28}$$

and therefore

$$(\Delta x)^2\,(\Delta p)^2 = \frac{\hbar^2}{4}\,(\lambda - \nu)^2\,(\lambda + \nu)^2 = \frac{\hbar^2}{4}\left(\lambda^2 - \nu^2\right)^2 = \frac{\hbar^2}{4}$$

It turns out that the operators $\hat{a}$ and $\hat{b}$ are related by a unitary transformation, i.e.,

$$\hat{b} = \hat{U}\hat{a}\hat{U}^+ \tag{13.29}$$

where

$$\hat{U} = \exp\left[\frac{\xi}{2}\left(\hat{a}^2 - \hat{a}^{+2}\right)\right] \tag{13.30}$$

and

$$e^{\xi} = \lambda + \nu \tag{13.31}$$

Proof:

$$\hat{U}\hat{a}\hat{U}^{+} = \exp\left[\frac{\xi}{2}\left(\hat{a}^2 - \hat{a}^{+2}\right)\right]\hat{a}\exp\left[-\frac{\xi}{2}\left(\hat{a}^2 - \hat{a}^{+2}\right)\right] = e^{\hat{B}}\hat{a}e^{-\hat{B}} \tag{13.32}$$

Using the identity derived earlier we have

$$e^{\hat{B}}\hat{a}e^{-\hat{B}} = \hat{a} + \left[\hat{B},\hat{a}\right] + \frac{1}{2}\left[\hat{B},\left[\hat{B},\hat{a}\right]\right] + \frac{1}{6}\left[\hat{B},\left[\hat{B},\left[\hat{B},\hat{a}\right]\right]\right] + ... \tag{13.33}$$

Now with

$$\hat{B} = \frac{\xi}{2}\left(\hat{a}^2 - \hat{a}^{+2}\right) \tag{13.34}$$

we have

$$\begin{aligned}
\left[\hat{B},\hat{a}\right] &= \frac{\xi}{2}\left(\left[\hat{a}^2,\hat{a}\right] - \left[\hat{a}^{+2},\hat{a}\right]\right) = -\frac{\xi}{2}\left[\hat{a}^{+2},\hat{a}\right] = \xi\hat{a}^{+} \\
\left[\hat{B},\left[\hat{B},\hat{a}\right]\right] &= \frac{\xi}{2}\xi\left[\hat{a}^2,\hat{a}^{+}\right] = \xi^2\hat{a} \\
\left[\hat{B},\left[\hat{B},\left[\hat{B},\hat{a}\right]\right]\right] &= \xi^3\hat{a}^{+}
\end{aligned}$$

and so on.

Therefore

$$\begin{aligned}
\hat{U}\hat{a}\hat{U}^{+} &= \hat{a} + \xi\hat{a}^{+} + \frac{\xi^2}{2}\hat{a} + \frac{\xi^3}{6}\hat{a}^{+} + ... \\
&= \hat{a}\left(1 + \frac{\xi^2}{2!} + \frac{\xi^4}{4!} + ...\right) + \hat{a}^{+}\left(\xi + \frac{\xi^3}{3!} + \frac{\xi^5}{5!} + ...\right)
\end{aligned} \tag{13.35}$$

If we define

$$\lambda = \left(1 + \frac{\xi^2}{2!} + \frac{\xi^4}{4!} + ...\right) \quad , \quad \nu = \left(\xi + \frac{\xi^3}{3!} + \frac{\xi^5}{5!} + ...\right) \tag{13.36}$$

so that

$$e^{\xi} = \lambda + \nu \tag{13.37}$$

we have

$$\hat{U}\hat{a}\hat{U}^{+} = \lambda\hat{a} + \nu\hat{a}^{+} = \hat{b} \tag{13.38}$$

Finally, we have

$$\hat{a}\,|coherent\rangle = \alpha\,|coherent\rangle$$
$$\hat{b}\,|squeezed\rangle = \beta\,|squeezed\rangle$$
$$\hat{U}\hat{a}\,|coherent\rangle = \alpha\hat{U}\,|coherent\rangle$$
$$\hat{U}\hat{a}\hat{U}^{+}\hat{U}\,|coherent\rangle = \alpha\hat{U}\hat{U}^{+}\hat{U}\,|coherent\rangle = \alpha\hat{U}\,|coherent\rangle$$
$$\hat{b}\hat{U}\,|coherent\rangle = \alpha\hat{U}\,|coherent\rangle = \alpha\hat{U}\,|coherent\rangle$$
$$\hat{b}\hat{U}\,|coherent\rangle = \alpha\hat{U}\,|coherent\rangle$$

which says that $\hat{U}\,|coherent\rangle = |squeezed\rangle$. Thus, $\hat{U}$ transforms a coherent state into a squeezed state.

13.2 Electron in a circular wire

We now consider a loop of thin wire in the shape of a circle of radius R as in Figure 13.1 below. A constant magnetic field perpendicular to the plane of the loop produces a magnetic flux passing through the loop. Imagine that the wire contains only one electron which is free to move. This electron has a wavefunction $\psi(\theta)$ which depends only on the angular coordinate θ. We neglect all interactions between the electron spin and the magnetic field as well as all magnetic fields produced by the electron itself.

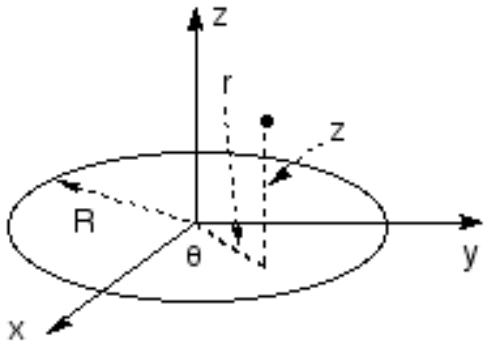


Figure 13.1: Circular Wire Configuration

We first determine the energies and energy eigenfunctions for a nonrelativistic electron of mass m moving on this ring. In particular we want to find out how ground state energy of the electron depends on the value of the applied magnetic field in this approximation.

We have

$$\nabla \times \vec{A} = B\hat{e}_z \quad , \quad B = \text{constant} \tag{13.39}$$

In cylindrical coordinates (r, θ, z), we can choose

$$A_r = A_z = 0 \ , \ A_\theta = \frac{rB}{2} \rightarrow \vec{A} = \frac{rB}{2}\hat{e}_\theta \tag{13.40}$$

The Schrodinger equation for the electron is

$$\frac{1}{2m}\left(\vec{p} - \frac{e}{c}\vec{A}\right)^2 \psi = E\psi \quad , \quad e < 0 \tag{13.41}$$

We then let

$$\psi = \psi' e^{\frac{ie}{\hbar c}\int_{\vec{r}} \vec{A}\cdot d\vec{r}} \tag{13.42}$$

which should get rid of the effect of $\vec{A}$ (equivalent to a gauge transformation). We have

$$\begin{aligned}
\left(\vec{p} - \frac{e}{c}\vec{A}\right)\psi &= \left(\vec{p} - \frac{e}{c}\vec{A}\right)\psi' e^{\frac{ie}{\hbar c}\int_{\vec{r}} \vec{A}\cdot d\vec{r}} \\
&= \frac{\hbar}{i}\nabla\left(\psi' e^{\frac{ie}{\hbar c}\int_{\vec{r}} \vec{A}\cdot d\vec{r}}\right) - \frac{e}{c}\vec{A}\psi \\
&= e^{\frac{ie}{\hbar c}\int_{\vec{r}} \vec{A}\cdot d\vec{r}}\frac{\hbar}{i}\nabla\psi' + \frac{\hbar}{i}\nabla\left(e^{\frac{ie}{\hbar c}\int_{\vec{r}} \vec{A}\cdot d\vec{r}}\right)\psi' - \frac{e}{c}\vec{A}\psi \\
&= e^{\frac{ie}{\hbar c}\int_{\vec{r}} \vec{A}\cdot d\vec{r}}\frac{\hbar}{i}\nabla\psi' + \frac{e}{c}\vec{A}\psi - \frac{e}{c}\vec{A}\psi \\
&= e^{\frac{ie}{\hbar c}\int_{\vec{r}} \vec{A}\cdot d\vec{r}}\frac{\hbar}{i}\nabla\psi' = e^{\frac{ie}{\hbar c}\int_{\vec{r}} \vec{A}\cdot d\vec{r}}\vec{p}\psi'
\end{aligned}$$

Similarly,

$$\left(\vec{p} - \frac{e}{c}\vec{A}\right)^2 \psi = e^{\frac{ie}{\hbar c}\int_{\vec{r}} \vec{A}\cdot d\vec{r}}\vec{p}^2\psi' \tag{13.43}$$

so that the Schrodinger equation becomes

$$\frac{1}{2m}\vec{p}^2\psi' = E\psi' \tag{13.44}$$

Since the electron is confined to a loop of radius R, we have

$$\psi = \psi(\theta) = \psi'(\theta)e^{\frac{ie}{\hbar c}\int_{\vec{r}} \vec{A}\cdot d\vec{r}} = \psi'(\theta)e^{\frac{ie}{\hbar c}\int_{\vec{r}} A_\theta R d\theta} = \psi'(\theta)e^{\frac{ie}{\hbar c}A_\theta R\theta} \tag{13.45}$$

Therefore we have

$$\frac{1}{2m}\vec{p}^2\psi'(\theta) = -\frac{\hbar^2}{2mR^2}\frac{d^2\psi'(\theta)}{d\theta^2} = E\psi'(\theta) = \frac{\hbar^2}{2mR^2}C_1^2\psi'(\theta) \tag{13.46}$$

which has the solution

$$\psi'(\theta) = e^{iC_1\theta} \rightarrow \psi(\theta) = e^{iC_1\theta}e^{\frac{ie}{\hbar c}A_\theta R\theta} = e^{i\left(C_1 + \frac{eBR^2}{2\hbar c}\right)\theta} \tag{13.47}$$

Now imposing single-valuedness, we have

$$\psi(\theta) = \psi(\theta + 2\pi) \tag{13.48}$$

$$e^{i\left(C_1 + \frac{eBR^2}{2\hbar c}\right)\theta} = e^{i\left(C_1 + \frac{eBR^2}{2\hbar c}\right)\theta} e^{2\pi i\left(C_1 + \frac{eBR^2}{2\hbar c}\right)} \tag{13.49}$$

which says that

$$C_1 + \frac{eBR^2}{2\hbar c} = n = 0, \pm 1, \pm 2, ... \tag{13.50}$$

or

$$C_1 = n - \frac{eBR^2}{2\hbar c} \rightarrow E_n = \frac{\hbar^2}{2mR^2}\left(n - \frac{eBR^2}{2\hbar c}\right)^2 \tag{13.51}$$

If we define $\phi_0 = -\hbar c/e$ = unit of flux, and remembering that the flux through the loop is $\phi = \pi R^2 B$ we have

$$E_n = \frac{\hbar^2}{2mR^2}\left(n + \frac{\phi}{\phi_0}\right)^2 \tag{13.52}$$

which says that the dependence of E_n on the external magnetic field B or flux ϕ is parabolic. A plot is shown in Figure 13.2 below.

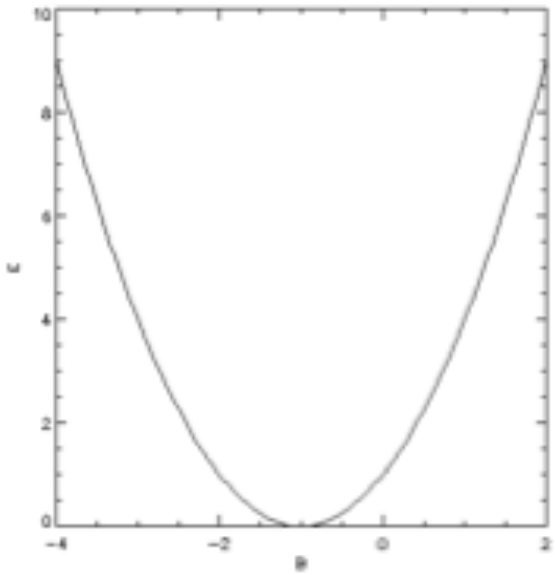


Figure 13.2: E_n versus B

Since n is an integer, the ground state energy E_g is given by

$$E_g = \frac{\hbar^2}{2mR^2}\left(n^* - \frac{eBR^2}{2\hbar c}\right)^2 \tag{13.53}$$

where n^* is the integer nearest to $eBR^2/2\hbar c$ or near

$$\frac{\phi}{\phi_0} = \frac{e\phi}{c\hbar} < 0 \tag{13.54}$$

Note that $n^* < 0$ since $e < 0$.

Now imagine that we start with the wire in its ground state in the presence of a magnetic flux ϕ. If the magnetic field is turned off determine the current in the loop. Assume that $R = 2\ cm$ and $\phi = 0.6\ gauss - cm^2$.

Suppose that we start with a state E_n which is the ground state. n will remain the same when the magnetic field is turned off. Therefore, $\psi(\theta) = Ce^{in\theta}$ and the electric current is

$$\vec{J} = \frac{e\hbar}{2mi}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right) \tag{13.55}$$

which follows from

$$\nabla\cdot\vec{J} + \frac{\partial}{\partial t}\left|\psi\right|^2 = 0 \tag{13.56}$$

We then get

$$\vec{J} = \frac{e\hbar}{2mi}(in)\frac{2}{R}\psi^*\psi\hat{e}_\theta = \frac{ne\hbar}{mR}\psi^*\psi\hat{e}_\theta \tag{13.57}$$

Now, if S = the cross-section area of the thin wire, the normalization constant is

$$\int \psi^*\psi d\ell dS = 2\pi R\left|C\right|^2 S = 1 \rightarrow \left|C\right|^2 = \frac{1}{2\pi RS} \tag{13.58}$$

Then,

$$I = \text{ current } = \int \vec{J}\cdot d\vec{S} = \frac{ne\hbar}{mR}\left|C\right|^2 S = \frac{ne\hbar}{2\pi mR^2} \tag{13.59}$$

where we have assumed that $\vec{J}$ is constant throughout the thin wire.

Since the electron is initially in the ground state, this implies that E_n is the minimum energy and we have

$$n = \text{ greatest integer not greater than } \frac{\varphi}{\varphi_0} = \frac{e\varphi}{c\hbar} \tag{13.60}$$

or

$$\text{greatest integer not greater than } \frac{\varphi}{\varphi_0} - 1 \tag{13.61}$$

For a macroscopic system we have $n \gg 1$ and we can use

$$n \approx \frac{e\varphi}{c\hbar} \rightarrow I = \frac{ne\hbar}{2\pi mR^2} \approx \frac{e^2\varphi}{4\pi mcR^2} \tag{13.62}$$

For $R = 2\ cm$ and $\phi = 0.6\ gauss - cm^2$ we have (using SI units)

$$I = \frac{(1.6\times10^{-19})^2(0.6\times10^{-4})\times10^{-4}}{4\pi^2(2\times10^{-2})^2(0.9\times10^{-30})} = 1.1\times10^{-14}\,amp \tag{13.63}$$

13.3 Spin-Orbit Coupling in Complex Atoms

Spin-orbit coupling is strictly an internal effect arising from the interaction between the electron spin and the effective magnetic field due to the apparent nuclear motion. In analogy with the one-electron atom, we can write for the $N-$electron atom

$$\hat{H}_{so} = \sum_{i=1}^{N} \xi_i(\vec{r}_i)\vec{L}_i \cdot \vec{S}_i \tag{13.64}$$

where $\xi_i(\vec{r}_i)$ is defined in a manner similar to the one-electron case, assuming an effective potential field can be defined for each electron.

For the case of weak spin-orbit, let us use both classical and quantum mechanical arguments to determine the first-order correction to the energy.

Classical Argument:

Vector Model for Combining Angular Momentum: The $\vec{L} \cdot \vec{S}$ interaction causes $\vec{L}$ and $\vec{S}$ to exert torques on each other via their magnetic moments. This means that neither of these quantities are independently constants of the motion.

However, if $\vec{\tau}_{external} = 0$, then J^2 is a constant of the motion, where $\vec{J} = \vec{L} + \vec{S}$.

In this model, there exist two extreme orientations of these vectors for a single electron as shown in Figure 13.3 below.

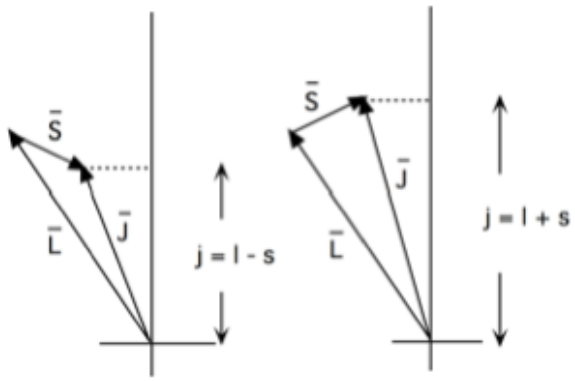


Figure 13.3: Extreme Orientations

and as shown in Figure 13.4 below the $\vec{L}$ and $\vec{S}$ vectors precess about the vector $\vec{J} = \vec{L} + \vec{S}$ at the same angular velocity.

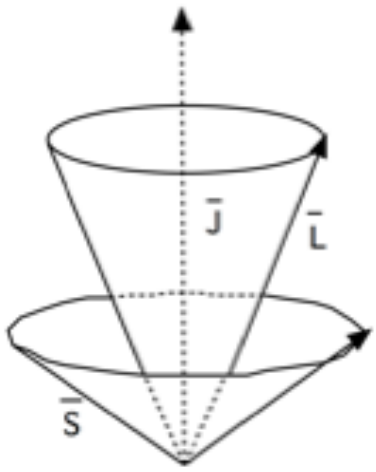


Figure 13.4: Precessing Vectors

The quantum conditions now apply to J^2 and J_z instead of L^2, L_z and S^2 and S_z separately. This means that we define the state $|J, M_J\rangle$ such that

$$\hat{J}^2 |J, M_J\rangle = J(J+1)\hbar^2 |J, M_J\rangle \tag{13.65}$$

$$\hat{J}_z |J, M_J\rangle = M_J \hbar |J, M_J\rangle \tag{13.66}$$

where $J = |L+S|, |L+S|-1, ..., |L-S|$ and each J value has $2J+1$ M_J values.

Now consider the interaction term of the form $\sum_i \vec{L}_i \cdot \vec{S}_i$. In this model

$$\text{each } \vec{L}_i \text{ precesses rapidly about } \vec{L}$$
$$\text{each } \vec{S}_i \text{ precesses rapidly about } \vec{S}$$

This means that on the average

$$\vec{L}_i = \alpha_i \vec{L} \quad , \quad \vec{S}_i = \beta_i \vec{S} \tag{13.67}$$

so that

$$\sum_i \vec{L}_i \cdot \vec{S}_i = \gamma \vec{L} \cdot \vec{S} \quad , \quad \gamma = \sum_i \alpha_i \beta_i \tag{13.68}$$

This means that (effectively)

$$\hat{H}_{so} = \gamma \vec{L} \cdot \vec{S} = \frac{\gamma}{2} \left(\vec{J}^2 - \vec{L}^2 - \vec{S}^2 \right) \tag{13.69}$$

as we assumed.

More formally (using the Wigner-Eckart theorem) we have

$$\begin{aligned}
\langle LSJM|\,\hat{H}_{so}\,|LSJM\rangle &= \sum_{\substack{M_L,M_S,M'_L,M'_S\\ M_L+M_S=M'_L+M'_S}} C^*(LSM'_LM'_S)C(LSM_LM_S)\\
&\qquad\times\langle LSM'_LM'_S|\,\hat{H}_{so}\,|LSM_LM_S\rangle\\
&= \sum_{\substack{M_L,M_S,M'_L,M'_S\\ M_L+M_S=M'_L+M'_S}} C^*(LSM'_LM'_S)C(LSM_LM_S)\\
&\qquad\times\langle LSM'_LM'_S|\sum_i \xi_i(r)\vec{L}_i\cdot\vec{S}_i\,|LSM_LM_S\rangle\\
&= \sum_{\substack{M_L,M_S,M'_L,M'_S\\ M_L+M_S=M'_L+M'_S}} C^*(LSM'_LM'_S)C(LSM_LM_S)\\
&\qquad\times\sum_i \xi_i(r)\,\langle LM'_L|\,\vec{L}_i\,|LM_L\rangle\cdot\langle SM'_S|\,\vec{S}_i\,|SM_S\rangle\\
&= \sum_{\substack{M_L,M_S,M'_L,M'_S\\ M_L+M_S=M'_L+M'_S}} C^*(LSM'_LM'_S)C(LSM_LM_S)\\
&\qquad\times\sum_i \xi_i(r)\,\langle L\,\|\vec{L}_i\|\,L\rangle\,\langle S\,\|\vec{S}_i\|\,S\rangle\\
&\qquad\times\langle LSM'_LM'_S|\,\vec{L}\cdot\vec{S}\,|LSM_LM_S\rangle
\end{aligned}\tag{13.70}$$

where the last step involves two uses of the Wigner-Eckart theorem. We then have

$$\begin{aligned}
\langle LSJM|\,\hat{H}_{so}\,|LSJM\rangle &= \sum_{\substack{M_L,M_S,M'_L,M'_S\\ M_L+M_S=M'_L+M'_S}} C^*(LSM'_LM'_S)C(LSM_LM_S)\\
&\qquad\times\sum_i \xi_i(r)\alpha_i\beta_i\,\langle LSM'_LM'_S|\,\vec{L}\cdot\vec{S}\,|LSM_LM_S\rangle\\
&= \sum_{\substack{M_L,M_S,M'_L,M'_S\\ M_L+M_S=M'_L+M'_S}} C^*(LSM'_LM'_S)C(LSM_LM_S)\\
&\qquad\times\langle LSM'_LM'_S|\gamma\vec{L}\cdot\vec{S}\,|LSM_LM_S\rangle\\
&= \langle LSJM|\gamma\vec{L}\cdot\vec{S}\,|LSJM\rangle
\end{aligned}\tag{13.71}$$

so that effectively, $\hat{H}_{so} = \gamma\vec{L}\cdot\vec{S}$.

We also note that since

$$\hat{H}_{so} = \gamma\vec{L}\cdot\vec{S} \rightarrow E(LSJ) = A(LS)\left(\frac{J(J+1)-L(L+1)-S(S+1)}{2}\right)\tag{13.72}$$

so that

$$E(L,S,J) - E(L,S,J-1) = A(LS)\left(\frac{J(J+1) - L(L+1) - S(S+1)}{2}\right) - A(LS)\left(\frac{(J-1)(J) - L(L+1) - S(S+1)}{2}\right)$$
$$= \frac{A(LS)}{2}\left(J^2 + J - J^2 + J\right) = A(LS)J \qquad (13.73)$$

which is called the Lande interval rule.

13.4 Zeeman Effect in Complex Atoms

The electronic spin and orbital angular momenta in a complex atom give rise to a magnetic moment that we can write, by analogy with the one-electron atom, as

$$\vec{M} = \sum_{i=1}^{N} \vec{M}_i = -\frac{\beta}{\hbar}\sum_{i=1}^{N}\left(\vec{L}_i + 2\vec{S}_i\right) \qquad (13.74)$$

In an external magnetic field $\vec{B} = B_z\hat{z}$, the total Hamiltonian becomes $\hat{H} = \hat{H}_0 + \hat{H}_{so} + \hat{H}_B$, where the term containing B is

$$\hat{H}_B = -\vec{M}\cdot\vec{B} = -M_z B_z \qquad (13.75)$$

Let us now use classical precession arguments to derive an approximate operator expression for $\hat{M}_z$. We will assume both a weak spin-orbit interaction and a weak magnetic field interaction, but take the spin-orbit interaction to be dominant. Remember, in the case of weak spin-orbit interaction and a weak magnetic field, we can use the precession picture of the section 14.3. The same type of vector diagrams can be used to deal with summations like

$$\sum_i \vec{L}_i\cdot\vec{L} \;\; and \;\; \sum_i \vec{S}_i\cdot\vec{S} \qquad (13.76)$$

By vector addition we have $\vec{J} = \vec{L} + \vec{S}$ with both $\vec{L}$ and $\vec{S}$ precessing rapidly about $\vec{J}$. Also by vector addition we have

$$\vec{\mu} = \vec{\mu}_L + \vec{\mu}_S = g_L\vec{L} + g_S\vec{S} = \frac{e}{2mc}(\vec{L} + 2\vec{S}) \qquad (13.77)$$

Clearly, $\vec{\mu}$ is not parallel to $\vec{J}$ since

$$g_S = \frac{e}{mc} = 2g_L \qquad (13.78)$$

In addition, $\vec{\mu}$ is precessing rapidly about $\vec{J}$. Therefore,

$$\begin{aligned}\mu_{effective} &= \vec{\mu}\cdot\frac{\vec{J}}{J} = -\frac{\mu_B}{\hbar}\frac{(\vec{L}+2\vec{S})\cdot\vec{J}}{J} = -\frac{\mu_B}{\hbar}\frac{(\vec{L}+2\vec{S})\cdot(\vec{L}+\vec{S})}{J} \\ &= -\frac{\mu_B}{\hbar}\frac{\left(L^2+2S^2+3\vec{L}\cdot\vec{S}\right)}{J} = -\frac{\mu_B}{\hbar}\frac{\left(L^2+2S^2+\frac{3}{2}(J^2-L^2-S^2)\right)}{J} \\ &= -\frac{\mu_B}{\hbar}J\frac{\left(3J^2-L^2+S^2\right)}{2J^2} = -\frac{\mu_B}{\hbar}J\left(1+\frac{J^2-L^2+S^2}{2J^2}\right) \end{aligned} \tag{13.79}$$

or

$$\vec{\mu}_{effective} = -\frac{\mu_B}{\hbar}\vec{J}\left(1+\frac{J^2-L^2+S^2}{2J^2}\right) \tag{13.80}$$

Now we use the Wigner-Eckert theorem.

13.4.1 Method #1: Plausibility Derivation

We have

$$H = -\vec{\mu}\cdot\vec{B} \tag{13.81}$$

and

$$\vec{\mu} = \frac{e}{2mc}(\vec{L}+2\vec{S}) = \frac{e}{2mc}(\vec{J}+\vec{S}) = G\vec{J} \tag{13.82}$$

by the Wigner-Eckart theorem. Therefore,

$$\begin{aligned} G\vec{J}\cdot\vec{J} &= \frac{e}{2mc}(\vec{J}\cdot\vec{J}+\vec{S}\cdot\vec{J}) \rightarrow G = \frac{e}{2mc}\frac{\vec{J}\cdot\vec{J}+\vec{S}\cdot\vec{J}}{\vec{J}\cdot\vec{J}} \\ &= \frac{e}{2mc}\frac{J(J+1)+\frac{\vec{J}\cdot\vec{J}+\vec{S}\cdot\vec{S}-\vec{L}\cdot\vec{L}}{2}}{J(J+1)} \\ &= \frac{e}{2mc}\left[1+\frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}\right] \end{aligned} \tag{13.83}$$

13.4.2 Method #2: Full Formal Derivation

The Zeeman effect Hamiltonian is given by

$$H_{zeeman} = \frac{eB}{2mc}\left(\sum_i L_{iz}+2\sum_i S_{iz}\right) = \frac{eB}{2mc}\left(L_z+2S_z\right) \tag{13.84}$$

Therefore we need to evaluate the matrix element

$$\begin{aligned} &\langle LSJM|\left(L_z+2S_z\right)|LSJM\rangle \\ &\qquad = \langle LSJM|J_z|LSJM\rangle + \langle LSJM|S_z|LSJM\rangle \\ &\qquad = M\hbar + \langle LSJ\|S\|LSJ\rangle\langle JM|J_z|JM\rangle \\ &\qquad = M\hbar\left(1+\langle LSJ\|S\|LSJ\rangle\right) \end{aligned} \tag{13.85}$$

Now

$$\begin{aligned}\langle LSJM|\,\vec{S}\cdot\vec{J}\,|LSJM\rangle &= \langle LSJ\,\|S\|\,LSJ\rangle\,\langle LSJM|\,\vec{J}\cdot\vec{J}\,|LSJM\rangle \\ &= J(J+1)\hbar^2\,\langle LSJ\,\|S\|\,LSJ\rangle \end{aligned} \tag{13.86}$$

or

$$\langle LSJ\,\|S\|\,LSJ\rangle = \frac{\langle LSJM|\,\vec{S}\cdot\vec{J}\,|LSJM\rangle}{J(J+1)\hbar^2} \tag{13.87}$$

But we have

$$\begin{aligned}\langle LSJM|\,\vec{S}\cdot\vec{J}\,|LSJM\rangle &= \frac{1}{2}\,\langle LSJM|\left(J^2+S^2-L^2\right)|LSJM\rangle \\ &= \frac{\hbar^2}{2}\left(J(J+1)+S(S+1)-L(L+1)\right) \end{aligned} \tag{13.88}$$

Putting it all together

$$\langle LSJ\,\|S\|\,LSJ\rangle = \frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)} \tag{13.89}$$

so that

$$\begin{aligned}\langle LSJM|\,(L_z+2S_z)\,|LSJM\rangle &= M\hbar(1+\langle LSJ\,\|S\|\,LSJ\rangle) \\ &= M\hbar\left(1+\frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}\right) \end{aligned} \tag{13.90}$$

Therefore, the first-order correction to the energy level $E(L,S,J)$ due to the perturbation $\hat{H}_B$ is

$$\begin{aligned}\langle H_{zeeman}\rangle &= \frac{eB}{2mc}M\hbar\left(1+\frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}\right) \\ &= \mu_0 M g(LSJ) \end{aligned} \tag{13.91}$$

where

$$g(LSJ) = 1+\frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)} \tag{13.92}$$

is the so-called Lande g-factor.

13.5 Neutron Interferometry

In the late 1970s several neutron interference experiments which are of fundamental importance in quantum mechanics and which settled debates started in 1930s, were carried out by Overhauser and collaborators.

In this section we investigate the effects on a neutron interference pattern of the

gravitational field.

We mentioned some aspects of this type of experiment in Chapter 16. Here we go into more detail.

We consider an interferometer made of three parallel equally spaced crystalline silicon strips as shown in Figure 13.5 below.

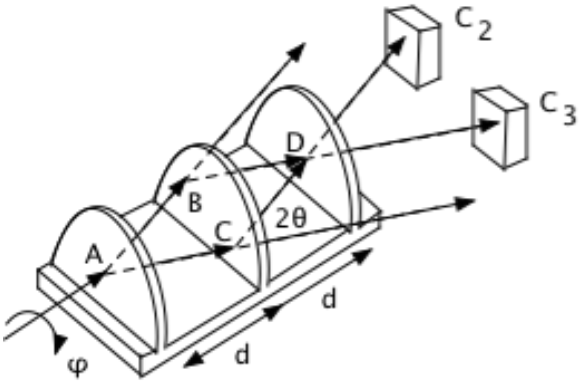


Figure 13.5: Experimental Setup

The incident neutron beam is assumed to be monochromatic. C_2 and C_3 are neutron counters.

For a particular value of the angle of incidence θ, called the Bragg angle, a plane wave

$$\psi_{inc} = e^{i(\vec{p}\cdot\vec{r}-Et)/\hbar} \tag{13.93}$$

where E is the energy of the neutrons and $\vec{P}$ their momentum, is split by the crystal into two outgoing waves which are symmetric with respect to the direction perpendicular to the crystal, as shown in Figure 13.6 below.

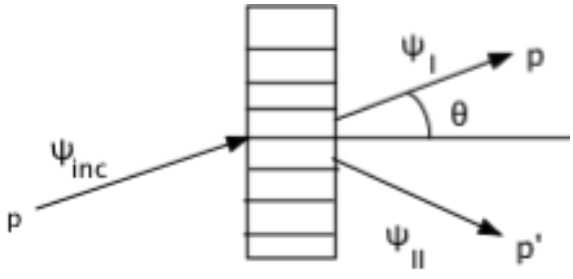


Figure 13.6: Splitting Waves

The transmitted wave and the reflected wave have complex amplitudes which can be written respectively as

$$\alpha = \cos\chi \quad , \quad \beta = i\sin\chi \qquad \chi \text{ real} \tag{13.94}$$

so that

$$\psi_I = \alpha e^{i(\vec{p}\cdot\vec{r}-Et)/\hbar} \quad , \quad \psi_{II} = \beta e^{i(\vec{p}'\cdot\vec{r}-Et)/\hbar} \tag{13.95}$$

where $|\vec{p}| = |\vec{p}\,'|$ since the neutrons scatter elastically on the nuclei of the crystal. The transmission and reflection coefficients are

$$T = |\alpha|^2 \quad , \quad R = |\beta|^2 \qquad \text{with} \quad T + R = 1 \tag{13.96}$$

In this interferometer setup the incident neutron beam is horizontal. It is split by the interferometer into a set of beams, two of which recombine and interfere at point D. The detectors C_2 and C_3 count the outgoing neutron fluxes. The neutron beam velocity corresponds to a de Broglie wavelength $\lambda = 1.445$ and the neutron mass is $M = 1.675 \times 10^{-27}\, kg$.

For calculational simplicity we are using monochromatic plane waves to represent the neutron beams; they are, however, quasi-monochromatic with finite extension in directions transverse to the beams.

13.5.1 Neutron Interferences

The measured neutron fluxes are proportional to the intensities of the waves that reach the counters. We define the intensity of the incoming wave to be 1 (units are arbitrary). For C_2 the beams $ABDC_2$ and $ACDC_2$ interfere. Omitting the propagation factors, at C_2 we have the amplitude

$$A_2 = \alpha^2\beta + \beta^3 = \beta(\alpha^2 + \beta^2) \tag{13.97}$$

Similarly, for $ABDC_2$ and $ACDC_3$,

$$A_3 = 2\alpha\beta^2 \tag{13.98}$$

The intensities at the two counters are then

$$I_2 = R - 4R^2T \quad , \quad I_3 = 4R^2T \tag{13.99}$$

Suppose that we create a phase shift δ of the wave propagating along AC, i.e., in C the wave function is multiplied by $e^{i\delta}$.

The new amplitudes at the detectors are

$$A_2 = \alpha^2\beta e^{i\delta} + \beta^3 = \beta(\alpha^2 e^{i\delta} + \beta^2) \quad , \quad A_3 = \alpha\beta^2(1 + e^{i\delta}) \tag{13.100}$$

and the new intensities become

$$I_2 = R - 2R^2T(1 + \cos\delta) \quad , \quad I_3 = 2R^2T(1 + \cos\delta) \tag{13.101}$$

Note that $I_2 + I_3$ does not depend on the phase shift δ. This is because of the conservation of the total number of particles arriving at D.

13.5.2 The Gravitational Effect

The phase difference δ between the beams ACD and ABD is created by rotating the interferometer by an angle φ around the direction of the incident beam as shown in Figure 13.7 (on the left)below.

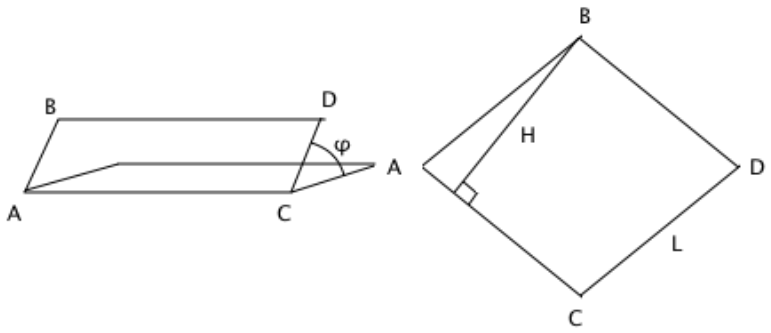


Figure 13.7: Geometrical Considerations

Now let d be the distance between the silicon strips(we neglect their thickness in this discussion). We also define L as the side of $ABCD$ and H as its height as shown in Figure 14.7 (on the right) above. We then have(simple geometry) that

$$L = \frac{d}{\cos\theta} \quad , \quad H = 2d\sin\theta \qquad \theta = \text{Bragg angle} \tag{13.102}$$

Experimentally, the values of d and θ are $d = 3.6\,cm$ and $\theta = 22.1\,^\circ$.

For an angle φ we define the gravitational potential V to be $V = 0$ along AC and $V = V_0$ along BD.

Since there is no recoil energy of the silicon atoms to be taken into account, the neutron total energy (kinetic + potential) is a constant of the motion in all of the process. The energies are given by

$$E_{AC} = \frac{p^2}{2M} = E_{BD} = \frac{(p-\Delta p)^2}{2M} + MgH\sin\phi \tag{13.103}$$

$$\Delta p \approx \frac{M^2 gH\sin\phi}{p} \tag{13.104}$$

where Δp is the difference in the neutron momentum.

The velocity $\sqrt{2gH}$ is of order $0.5\,m/s$ and the neutron velocity is

$$v = \frac{h}{M\lambda} \approx 2700\ m/s \tag{13.105}$$

The change in the velocity Δv is therefore very small, i.e.,

$$\Delta v = \frac{gH}{v} \approx 2 \times 10^{-4}\, m/s \quad \text{for} \quad \phi = \frac{\pi}{2} \tag{13.106}$$

Now the gravitational potential varies in exactly the same way along AB and CD. The neutron state in both cases is a plane wave with momentum $p = h/\lambda$ just before A or C. The same Schrodinger equation is used to determine the wave function at the end of the segments. This implies that the phases accumulated along the two segments AB and CD are equal.

When comparing segments AC and BD, the previous reasoning does not apply, since the initial state of the neutron is not the same for the two segments. The initial state is $e^{ipz/\hbar}$ for AC and $e^{i(p-\Delta p)z/\hbar}$ for BD. After traveling over a distance $L = \bar{AC} - \bar{BD}$, the phase difference between the two paths is

$$\delta = \frac{L\Delta p}{\hbar} = \frac{M^2 g \lambda d^2}{\pi \hbar^2} \tan\theta \sin\phi \tag{13.107}$$

The variation with φ of the experimentally measured intensity I_2 in the counter C_2 is shown schematically in Figure 13.8 below (the data does not display a minimum exactly at $\varphi = 0$ because of calibration difficulties).

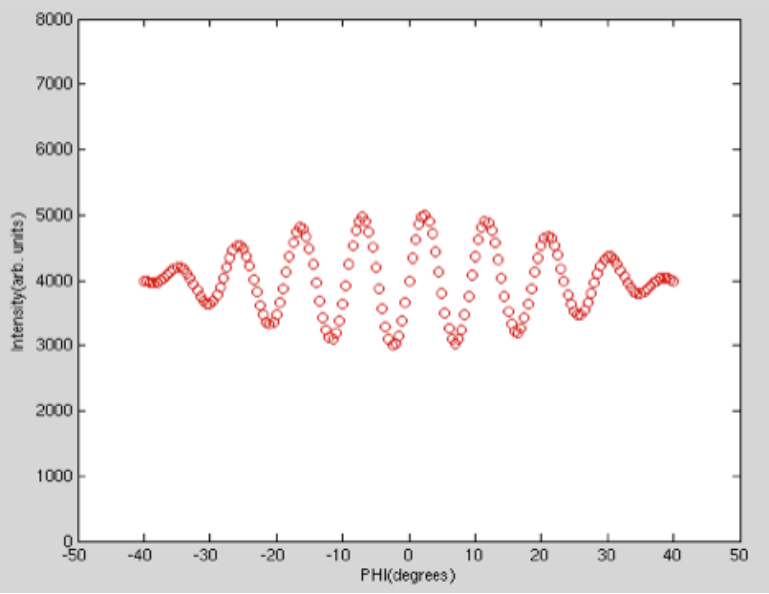


Figure 13.8: Variation of Intensity with φ

From the previous result, we have

$$\delta_2 - \delta_1 = Ag(\sin\phi_2 - \sin\phi_1) \tag{13.108}$$

where

$$A = \frac{M^2 \lambda d^2}{\pi \hbar^2} \tan\theta \tag{13.109}$$

Therefore,

$$g = \frac{\delta_2 - \delta_1}{A(\sin\phi_2 - \sin\phi_1)} \tag{13.110}$$

In the actual data there are 9 oscillations, i.e., $\delta_2 - \delta_1 = 18\pi$ between $\varphi_1 = -32^\circ$ and $\varphi_2 = +24^\circ$, which gives $g = 9.8\, m/s^2$. This clearly shows that the neutron interference effects are directly the result of the difference in the gravitational potential along two arms of the interferometer.

13.6 The Penning Trap

A Penning trap allows one to confine electrons in a finite spatial region and then allow the accurate measurement of various properties. It involves the superposition of a uniform magnetic field $\vec{B}$ directed along the $z-$axis and a quadrupole electric field which derives from an electrostatic potential of the form

$$\Phi = \mathrm{K}(2z^2 - x^2 - y^2) \tag{13.111}$$

where K is a positive constant.

An electron of charge $-q(q > 0)$ and mass m is placed in such a device. We denote its spin operator by $\vec{S}$ and its momentum operator by $\vec{p}$. The Hamiltonian of the electron in the above superposition of fields is

$$\hat{H} = \frac{1}{2m}\left(\vec{p} + q\vec{A}(\vec{r})\right)^2 + V(\vec{r}) + (1+a)\frac{q}{m}\vec{S}\cdot\vec{B} \tag{13.112}$$

$$V(\vec{r}) = \Phi = m\omega_0^2(2\hat{z}^2 - \hat{x}^2 - \hat{y}^2)/4 \tag{13.113}$$

$$= \text{ electrostatic potential energy} \tag{13.114}$$

$$\vec{A}(\vec{r}) = \vec{B}\times\vec{r}/2 \tag{13.115}$$

$$= \text{ vector potential} \tag{13.116}$$

The constant $a \approx 1.16 \times 10^{-3}$ is the gyromagnetic anomaly of the electron magnetic moment.

13.6.1 Motion of an Electron in a Penning Trap

We set $\omega_c = qB/m$, where B is the magnitude of the magnetic field, and we assume that this cyclotron frequency ω_c is much larger than ω_0.

We note that $\vec{p}\cdot\vec{A}(\vec{r}) = \vec{A}(\vec{r})\cdot\vec{p} = \vec{L}\cdot\vec{B}/2 = \hat{L}_z B/2$ and $\vec{A}^2 = B^2(x^2+y^2)/4$. We

then get

$$\hat{H} = \hat{H}_z + \hat{H}_t + \hat{H}_s \tag{13.117}$$

$$\hat{H}_z = \frac{\hat{p}_z^2}{2m} + \frac{1}{2}m\omega_0^2\hat{z}^2 \tag{13.118}$$

$$\hat{H}_t = \frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} + \frac{1}{2}m\Omega^2(\hat{x}^2 + \hat{y}^2) + \frac{1}{2}\omega_c\hat{L}_z \tag{13.119}$$

$$\hat{H}_s = (1+a)\omega_c\hat{S}_z \tag{13.120}$$

with $\Omega^2 = (\omega_c^2 - 2\omega_0^2)/4 \Rightarrow \Omega \simeq \omega_c/2 - \omega_0^2/2\omega_c$ and $\vec{L}$ is the orbital angular momentum of the electron.

The eigenstates of $\hat{H}_s$ are the eigenstates of $\hat{S}_z$, $|\pm\rangle$ with energy eigenvalues

$$\pm(1+a)\hbar\omega_c/2 = \pm\hbar\omega_s \quad , \quad \omega_s = (1+a)\omega_c/2 \tag{13.121}$$

where we have used

$$\hat{S}_z|\pm\rangle = \pm\hbar/2\,|\pm\rangle \tag{13.122}$$

Since $\hat{H}_z$, $\hat{H}_t$ and $\hat{H}_s$ act on different variables, they must commute. An eigenbasis of $\hat{H}$ can be constructed using the eigenstates $\varphi(z)$, $\psi(x,y)$ and $|\sigma\rangle$ of $\hat{H}_z$, $\hat{H}_t$ and $\hat{H}_s$, respectively. The corresponding eigenvalues are the sum of the individual eigenvalues.

In order to calculate the motion along the $z-$axis, we introduce the creation and annihilation operators

$$\hat{a}_z = \frac{1}{\sqrt{2}}\left(\alpha\hat{z} + \frac{i}{\alpha\hbar}\hat{p}_z\right) \quad , \quad \hat{a}_z^+ = \frac{1}{\sqrt{2}}\left(\alpha\hat{z} - \frac{i}{\alpha\hbar}\hat{p}_z\right) \quad , \quad \alpha = \sqrt{m\hbar\omega_0}$$

We then have

$$\begin{aligned}[\hat{a}_z, \hat{a}_z^+] &= \frac{1}{2}\left(\alpha^2[\hat{z},\hat{z}] + \frac{i}{\hbar}[\hat{p}_z,\hat{z}] - \frac{i}{\hbar}[\hat{z},\hat{p}_z] - \frac{1}{\alpha^2\hbar^2}[\hat{p}_z,\hat{p}_z]\right) \\ &= \frac{1}{2}\left(+\frac{i}{\hbar}[-i\hbar] - \frac{i}{\hbar}[i\hbar]\right) = 1\end{aligned} \tag{13.123}$$

Thus, we have the same mathematical system as the harmonic oscillator so that

$$\hat{H}_z = \hbar\omega_0(\hat{N}_z + 1/2) \quad , \quad \hat{N}_z = \hat{a}_z^+\hat{a}_z \tag{13.124}$$

$$\hat{N}_z|N_z\rangle = N_z|N_z\rangle \quad , \quad N_z = 0,1,2,3,..... \tag{13.125}$$

$$\hat{H}_z|N_z\rangle = E_{N_z}|N_z\rangle \quad , \quad E_{N_z} = \hbar\omega_0(N_z + 1/2) \tag{13.126}$$

13.6.2 The Transverse Motion

We now investigate the $x-y$ motion governed by the Hamiltonian $\hat{H}_t$. If we define the right- and left-circular creation and annihilation operators

$$\hat{a}_r = \frac{1}{2}\left(\beta(\hat{x} - i\hat{y}) + \frac{i}{\beta\hbar}(\hat{p}_x - i\hat{p}_y)\right) \tag{13.127}$$

$$\hat{a}_l = \frac{1}{2}\left(\beta(\hat{x} + i\hat{y}) + \frac{i}{\beta\hbar}(\hat{p}_x + i\hat{p}_y)\right) \tag{13.128}$$

where β is a real constant, then we can show (in same way as above) that

$$[\hat{a}_r, \hat{a}_r^+] = 1 = [\hat{a}_l, \hat{a}_l^+] \quad , \quad [\hat{a}_r, \hat{a}_l] = 0 = [\hat{a}_r, \hat{a}_l^+] \tag{13.129}$$

Defining

$$\hat{N}_r = \hat{a}_r^+ \hat{a}_r \quad , \quad \hat{N}_l = \hat{a}_l^+ \hat{a}_l \tag{13.130}$$

we have

$$\hat{N}_r = \hat{a}_r^+ \hat{a}_r = \frac{1}{4}\left(\beta^2(x^2 + y^2) + \frac{1}{\beta^2\hbar^2}(p_x^2 + p_y^2) - 2 + \frac{2L_z}{\hbar}\right) \tag{13.131}$$

$$\hat{N}_l = \hat{a}_l^+ \hat{a}_l = \frac{1}{4}\left(\beta^2(x^2 + y^2) + \frac{1}{\beta^2\hbar^2}(p_x^2 + p_y^2) - 2 - \frac{2L_z}{\hbar}\right) \tag{13.132}$$

and thus

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = \hbar(\hat{N}_r - \hat{N}_l) \tag{13.133}$$

and

$$\hat{N}_r + \hat{N}_l = \frac{1}{2}\left(\beta^2(x^2 + y^2) + \frac{1}{\beta^2\hbar^2}(p_x^2 + p_y^2)\right) - 1 \tag{13.134}$$

If we define $\beta^2 = m\Omega/\hbar$ we then have

$$\hat{H}_t = \hbar\Omega(\hat{N}_r + \hat{N}_l + 1) + \frac{\hbar\omega_c}{2}(\hat{N}_r - \hat{N}_l) \tag{13.135}$$

or

$$\hat{H}_t = \hbar\omega_c'(\hat{N}_r + 1/2) - \hbar\omega_m(\hat{N}_l + 1/2) \tag{13.136}$$

where

$$\omega_c' = \frac{\omega_c}{2} + \Omega = \frac{1}{2}\left(\omega_c + \sqrt{\omega_c^2 - 2\omega_0^2}\right) \simeq \omega_c \tag{13.137}$$

$$\omega_m = \frac{\omega_c}{2} - \Omega = \frac{1}{2}\left(\omega_c - \sqrt{\omega_c^2 - 2\omega_0^2}\right) \simeq \frac{\omega_0^2}{2\omega_c} << \omega_c \tag{13.138}$$

Since this is just the difference of two oscillators we have the energy eigenvalues

$$E_{cm} = \hbar\omega_c'(N_c + 1/2) - \hbar\omega_m(N_m + 1/2) \tag{13.139}$$

We also have from earlier

$$\pm(1 + a)\hbar\omega_c/2 = \pm\hbar\omega_s \quad , \quad \omega_s = (1 + a)\omega_c/2 \tag{13.140}$$

Thus, the energy eigenvalues of $\hat{H}$ are

$$E = \hbar\omega_0(N_z + 1/2) + \hbar\omega_c'(N_c + 1/2) - \hbar\omega_m(N_m + 1/2) + \sigma\hbar\omega_s \qquad (13.141)$$

where N_z, N_c and N_m are integers ≥ 0 and $\sigma = \pm 1$.

Note that the *magnetron* motion E_m corresponds to an *inverted* harmonic oscillator, so that its spectrum has no mower bound in the harmonic approximation used in the example. Consequently, when the system is coupled to a heat bath and relaxes towards thermal equilibrium, it should cascade down the ladder of levels of the magnetron motion, thus increasing the size of the orbit of the trapped particle in the xy–plane. Fortunately, the characteristic time corresponding to the decay of the system in this way is very long, and the electron can be confined around the center of the trap for a long time.

13.6.3 Measurement of Electron Anomalous Magnetic Moment

The electric quadrupole field is such that $\hbar\omega_0 = 2.58 \times 10^{-7}\, eV$. The magnetic field is $B = 5.87\, T$. The system is placed in liquid helium at $4.2\, K$. We then have (using $q\hbar/2m = 5.79 \times 10^{-5}\, eV$)

$$\hbar\omega_c = 6.8 \times 10^{-4} eV \sim \hbar\omega_c'$$
$$\hbar\omega_m = 4.9 \times 10^{-11} eV$$

In liquid helium, $kT = 3.5 \times 10^{-4}\, eV$ and the longitudinal and magnetron level spacings are much smaller than the thermal fluctuations. Thus, a classical description of these two motions is appropriate. In contrast, a few quanta of oscillation are thermal excited for the cyclotron motion since $kT \leq \hbar\omega\_c$. Now, the electron anomaly is $a \approx 0.00116$. Therefore we can draw the relative position of the four energy levels

$$N_z = 0;\ N_m = 0;\ N_c = 0, 1\ and\ \sigma = \pm 1 \qquad (13.142)$$

The level configuration is shown in Figure 13.9 below.

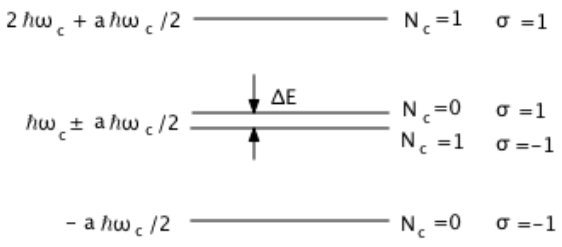


Figure 13.9: Energy Levels

The splitting ΔE between the level $N_c = 0, \sigma = +1$ and the level $N_c = 1, \sigma = -1$ is proportional to the anomaly a. We have $\Delta E = a\hbar\omega_c = 5 \times 10^{-7}\, eV$, where we have neglected the difference between ω_c and ω'_c which is $\approx 7.9 \times 10^{-11}\, eV$. The splitting corresponds to a frequency $\nu = \Delta E/\hbar = 191\, MHz$.

13.7 Schrodinger's Cat

We first just describe a simple version.

Suppose that a cat within a closed box would be killed by a $|\uparrow\rangle$ particle but not by a $|\downarrow\rangle$ particle. Now consider the effect of the state $|\uparrow\rangle + Ket\downarrow$, which can easily be produced by a properly oriented Stern-Gerlach device.

Suppose that a particle in the state $|\uparrow\rangle + Ket\downarrow$ hits the cat and that the state of the (spin + cat) makes a transition to

$$|\uparrow\rangle\,|dead\ cat\rangle + |\downarrow\rangle\,|living\ cat\rangle \tag{13.143}$$

which is a *pure state*.

When is it decided whether the cat is alive or dead?

Just when the observer opens the cat's box?

> An objective statement independent of the conscious
> mind of the observer would be impossible.

What is the consequence of including the observer herself in the quantum mechanical description?

According to the point of view presented, the cat(together with the mechanism for killing the cat, which was not mentioned above) is linked to other macroscopic objects. These are influenced differently in the two final states so that their respective wave functions do not overlap. For everything that follows, the macroscopic consequences are not recorded; the trace is taken over them. The final state of the cat is described by a mixture of states corresponding to a dead cat and a living cat: the cat is *either* dead *or* living and not in a pure state

$$|dead\ cat\rangle + |living\ cat\rangle \tag{13.144}$$

which would include both possibilities.

13.7.1 Schrodinger's Cat - a more detailed presentation

The superposition principle states that if $|\varphi_a\rangle$ and $|\varphi_b\rangle$ are two possible states of a quantum system, the quantum superposition

$$\frac{1}{\sqrt{2}}\left(|\varphi_a\rangle + |\varphi_b\rangle\right) \tag{13.145}$$

is also an allowed state for this system. This principle is essential in explaining quantum interference phenomena. However, when it is applied to *large* or *macroscopic* objects, it leads to paradoxical situations where a system can be in a superposition of states which is classical self-contradictory.

The most famous example is Schrodinger's *cat paradox* where the cat is in a superposition of the *dead* and *alive* states. The purpose of this discussion is to show that such superposition of macroscopic states are not detectable in practice. They are extremely fragile, and very weak coupling to the environment suffices to destroy the quantum superposition of the two states $|\varphi_a\rangle$ and $|\varphi_b\rangle$.

The Quasi-Classical States of a Harmonic Oscillator

We consider the high energy excitations of a one-dimensional harmonic oscillator or mass m and frequency ω. The Hamiltonian is written

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \tag{13.146}$$

We denote the eigenstates of $\hat{H}$ by $\{|n\rangle\}$ where the energy eigenvalues are given by

$$\hat{H}\,|n\rangle = E_n\,|n\rangle = \hbar\omega(n+1/2)\,|n\rangle \tag{13.147}$$

Preliminaries

We introduce the operators

$$\hat{X} = \sqrt{m\omega/\hbar}\hat{x} \quad , \quad \hat{P} = \hat{p}/\sqrt{m\hbar\omega} \tag{13.148}$$

and the annihilation and creation operators

$$\hat{a} = \frac{1}{\sqrt{2}}\left(\hat{X} + i\hat{P}\right) \quad , \quad \hat{a}^+ = \frac{1}{\sqrt{2}}\left(\hat{X} - i\hat{P}\right) \quad , \quad \hat{N} = \hat{a}^+\hat{a} \tag{13.149}$$

The commutator $[\hat{x},\hat{p}] = i\hbar$ leads to the commutators $[\hat{X},\hat{P}] = i$ and $[\hat{a},\hat{a}^+] = 1$ and the relations

$$\hat{H} = \hbar\omega(\hat{N} + 1/2) \quad , \quad \hat{N}\,|n\rangle = n\,|n\rangle \tag{13.150}$$

We also have the relations

$$\hat{P} = -i\frac{\partial}{\partial X} \quad , \quad \hat{X} = i\frac{\partial}{\partial P} \tag{13.151}$$

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle \quad , \quad \hat{a}^+\,|n\rangle = \sqrt{n+1}\,|n+1\rangle \tag{13.152}$$

We can use these relations to derive the ground state wave function in the position representation as follows:

$$\begin{aligned}
0 = \langle X|\,\hat{a}\,|0\rangle &= \frac{1}{\sqrt{2}}\,\langle X|\left(\hat{X} + i\hat{P}\right)|0\rangle \\
&= \frac{1}{\sqrt{2}} X\,\langle X \mid 0\rangle + \frac{i}{\sqrt{2}}\left(-i\frac{\partial}{\partial X}\right)\langle X \mid 0\rangle \\
&\left(X + \frac{\partial}{\partial X}\right)\langle X \mid 0\rangle = 0 \rightarrow \langle X \mid 0\rangle = Ae^{-X^2/2} = \psi_0(X) \\
&\psi_0(x) = Ae^{-m\omega x^2/2\hbar}
\end{aligned} \tag{13.153}$$

Similarly, we can derive its the ground state wave function in the momentum representation as follows:

$$\begin{aligned}
0 = \langle P|\,\hat{a}\,|0\rangle &= \frac{1}{\sqrt{2}}\,\langle P|\left(\hat{X} + i\hat{P}\right)|0\rangle \\
&= \frac{1}{\sqrt{2}} i\frac{\partial}{\partial P}\,\langle P \mid 0\rangle + \frac{i}{\sqrt{2}} P\,\langle P \mid 0\rangle \\
&\left(P + \frac{\partial}{\partial P}\right)\langle P \mid 0\rangle = 0 \rightarrow \langle P \mid 0\rangle = Ae^{-P^2/2} = \phi_0(P) \\
&\phi_0(p) = Ae^{-p^2/2m\omega\hbar}
\end{aligned} \tag{13.154}$$

These two wave functions are related by the Fourier transform, that is,

$$\begin{aligned}
\phi_0(p) = e^{-p^2/2m\omega\hbar} &\propto \int_{-\infty}^{\infty} e^{-m\omega x^2/2\hbar} e^{-ipx/\hbar} dx \\
&\propto \int_{-\infty}^{\infty} \psi_0(x) e^{-ipx/\hbar} dx
\end{aligned}$$

The Quasi-Classical States

The eigenstates of the operator $\hat{a}$ are called *quasi-classical* states, for reasons we will now discuss.

Since we are considering the question: what are the eigenstates of the lowering operator $\hat{a}$? We can write

$$\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle \text{ where } \alpha = |\alpha|\,e^{i\phi} \tag{13.155}$$

where $|\alpha\rangle$ is the eigenvector of $\hat{a}$ and α is the eigenvalue, which is not necessarily real since $\hat{a}$ is not Hermitian.

Since the vectors $|n\rangle$ are eigenvectors of a Hermitian operator, they form a

orthonormal complete set and can be used as an orthonormal basis for the vector space. We can then write

$$|\alpha\rangle = \sum_{m=0}^{\infty} b_m |m\rangle \tag{13.156}$$

where

$$\langle k \mid \alpha\rangle = \sum_{m=0}^{\infty} b_m \langle k \mid m\rangle = \sum_{m=0}^{\infty} b_m \delta_{km} = b_k \tag{13.157}$$

Now

$$\langle n-1|\, \hat{a}\, |\alpha\rangle = \alpha \langle n-1 \mid \alpha\rangle = \alpha b_{n-1} \tag{13.158}$$

and using

$$\hat{a}^+ |n-1\rangle = \sqrt{n}\,|n\rangle \rightarrow \langle n-1|\, \hat{a} = \sqrt{n}\,\langle n| \tag{13.159}$$

we have

$$\langle n-1|\, \hat{a}\, |\alpha\rangle = \sqrt{n}\,\langle n \mid \alpha\rangle = \sqrt{n} b_n \tag{13.160}$$

or

$$b_n = \frac{\alpha}{\sqrt{n}} b_{n-1} \tag{13.161}$$

This says that

$$b_1 = \frac{\alpha}{\sqrt{1}} b_0 \quad , \quad b_2 = \frac{\alpha}{\sqrt{2}} b_1 = \frac{\alpha^2}{\sqrt{2!}} b_0 \tag{13.162}$$

or

$$b_n = \frac{\alpha^n}{\sqrt{n!}} b_0 \tag{13.163}$$

We thus get the final result

$$|\alpha\rangle = b_0 \sum_{m=0}^{\infty} \frac{\alpha^m}{\sqrt{m!}} |m\rangle \tag{13.164}$$

Let us now normalize this state (choose b_0). We have

$$\begin{aligned} \langle \alpha \mid \alpha\rangle = 1 &= |b_0|^2 \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{\alpha^{*m} \alpha^k}{\sqrt{m!}\sqrt{k!}} \langle k \mid m\rangle \\ &= |b_0|^2 \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{\alpha^{*m} \alpha^k}{\sqrt{m!}\sqrt{k!}} \delta_{km} = |b_0|^2 \sum_{m=0}^{\infty} \frac{|\alpha|^2}{m!} \\ &= |b_0|^2 e^{|\alpha|^2} \end{aligned} \tag{13.165}$$

which says that

$$b_0 = e^{-\frac{1}{2}|\alpha|^2} \tag{13.166}$$

and thus

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{m=0}^{\infty} \frac{\alpha^m}{\sqrt{m!}} |m\rangle \tag{13.167}$$

Now

$$\langle n \mid \alpha \rangle = \text{ probability amplitude that the system in the state } |\alpha\rangle \text{ will be found in the state } |n\rangle$$

We have

$$\langle n \mid \alpha \rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{m=0}^{\infty} \frac{\alpha^m}{\sqrt{m!}} \langle n \mid m \rangle = e^{-\frac{1}{2}|\alpha|^2} \frac{\alpha^n}{\sqrt{n!}} \tag{13.168}$$

which then says that

$$P_n = |\langle n \mid \alpha \rangle|^2 = \frac{e^{-|\alpha|^2} |\alpha|^{2n}}{n!} = \frac{e^{-N} N^n}{n!}$$
$$= \text{ probability amplitude that the system in the state } |\alpha\rangle \text{ will be found in the state } |n\rangle$$

where we have defined $N = |\alpha|^2$. We note that

$$\langle \alpha | \hat{a}^+ \hat{a} | \alpha \rangle = \left|\alpha^2\right| \langle \alpha \mid \alpha \rangle = \left|\alpha^2\right| = N = \langle \alpha | \hat{N}_{op} | \alpha \rangle \tag{13.169}$$

or N = the average value or expectation value of the N_{op} operator in the state $|\alpha\rangle$. This type of probability distribution is called a Poisson distribution, i.e., the state $|\alpha\rangle$ has the number states or energy eigenstates distributed in a *Poisson* manner.

Since the states $|n\rangle$ are energy eigenstates, we know their time dependence, i.e.,

$$|n, t\rangle = e^{-i\frac{E_n}{\hbar}t} |n\rangle \tag{13.170}$$

Therefore, we have for the time dependence of the state $|\alpha\rangle$

$$|\alpha, t\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{m=0}^{\infty} \frac{\alpha^m}{\sqrt{m!}} |m, t\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{m=0}^{\infty} \frac{\alpha^m}{\sqrt{m!}} e^{-i\frac{E_m}{\hbar}t} |m\rangle \tag{13.171}$$

This simple operation clearly indicates the fundamental importance of the energy eigenstates when used as a basis set.

> If we are able to expand an arbitrary vector representing some physical system in the energy basis, then we immediately know the time dependence of that state vector and hence we know the time dependence of all the probabilities associated with the state vector and the system.

Now let us try to understand the physics contained in the $|\alpha\rangle$ state vector. In a

given energy eigenstate the expectation value of the position operator is given by

$$\begin{aligned}\langle n,t|\,\hat{x}\,|n,t\rangle &= \sqrt{\frac{\hbar}{2m\omega_0}}\,\langle n,t|\,(\hat{a}+\hat{a}^+)\,|n,t\rangle \\ &= \sqrt{\frac{\hbar}{2m\omega_0}}\,\langle n|\,e^{i\frac{E_n}{\hbar}t}(\hat{a}+\hat{a}^+)e^{-i\frac{E_n}{\hbar}t}\,|n\rangle \\ &= \sqrt{\frac{\hbar}{2m\omega_0}}\,\langle n|\,(\hat{a}+\hat{a}^+)\,|n\rangle \\ &= \sqrt{\frac{\hbar}{2m\omega_0}}\,\langle n|\left(\sqrt{n}\,|n-1\rangle+\sqrt{n+1}\,|n+1\rangle\right)=0\end{aligned}$$

i.e., it is equal to zero and is a constant.

On the other hand, in the state $|\alpha\rangle$ we find

$$\langle\alpha,t|\,\hat{x}\,|\alpha,t\rangle = \sqrt{\frac{\hbar}{2m\omega_0}}\sum_m\sum_k b_m^* b_k e^{i\frac{(E_m-E_k)}{\hbar}t}\,\langle m|\,(\hat{a}+\hat{a}^+)\,|k\rangle \tag{13.172}$$

Now

$$\begin{aligned}\langle m|\,(\hat{a}+\hat{a}^+)\,|k\rangle &= \langle m|\left(\sqrt{k}\,|k-1\rangle+\sqrt{k+1}\,|k+1\rangle\right) \\ &= \sqrt{k}\delta_{m,k-1}+\sqrt{k+1}\delta_{m,k+1}\end{aligned} \tag{13.173}$$

Using this result we have

$$\begin{aligned}\langle\alpha,t|\,\hat{x}\,|\alpha,t\rangle &= \sqrt{\frac{\hbar}{2m\omega_0}}\left(\sum_{k=1}^{\infty}b_{k-1}^*b_k\sqrt{k}e^{i\frac{(E_{k-1}-E_k)}{\hbar}t}+\sum_{k=0}^{\infty}b_{k+1}^*b_k\sqrt{k+1}e^{i\frac{(E_{k+1}-E_k)}{\hbar}t}\right) \\ &= \sqrt{\frac{\hbar}{2m\omega_0}}\left(\sum_{k=1}^{\infty}b_{k-1}^*b_k\sqrt{k}e^{-i\omega_0 t}+\sum_{k=0}^{\infty}b_{k+1}^*b_k\sqrt{k+1}e^{i\omega_0 t}\right) \\ &= \sqrt{\frac{\hbar}{2m\omega_0}}\left(\sum_{k=0}^{\infty}b_k^*b_{k+1}\sqrt{k}e^{-i\omega_0 t}+\sum_{k=0}^{\infty}b_{k+1}^*b_k\sqrt{k+1}e^{i\omega_0 t}\right) \\ &= \sqrt{\frac{\hbar}{2m\omega_0}}b_0^2\left(\sum_{k=0}^{\infty}\frac{\alpha^{*k}\alpha^{k+1}}{\sqrt{(k+1)!k!}}\sqrt{k}e^{-i\omega_0 t}+\sum_{k=0}^{\infty}\frac{\alpha^{*k+1}\alpha^{k}}{\sqrt{(k+1)!k!}}\sqrt{k+1}e^{i\omega_0 t}\right) \\ &= \sqrt{\frac{\hbar}{2m\omega_0}}b_0^2\sum_k\frac{1}{k!}\,|\alpha|^{2k}\left(\alpha e^{-i\omega_0 t}+\alpha^* e^{i\omega_0 t}\right)\end{aligned} \tag{13.174}$$

Now using $\alpha = |\alpha| e^{i\varphi}$ we get

$$\begin{aligned}\langle \alpha, t | \hat{x} | \alpha, t \rangle &= \sqrt{\frac{\hbar}{2m\omega_0}} b_0^2 2 |\alpha| \sum_k \frac{|\alpha|^{2k}}{k!} \text{Real}(e^{i\phi} e^{-i\omega_0 t}) \\ &= 2x_0 |\alpha| \cos(\omega_0 t - \phi)(b_0^2 \sum_k \frac{|\alpha|^{2k}}{k!}) \\ &= 2x_0 |\alpha| \cos(\omega_0 t - \phi) \quad , \quad x_0 = \sqrt{\frac{\hbar}{2m\omega_0}} \end{aligned} \tag{13.175}$$

The expectation value in the state $|\alpha\rangle$ behaves like that of a *classical oscillator*.

Before proceeding with the discussion, we will repeat the derivation using an alternate but very powerful technique.

Using the Translation Operator

In general, a displaced state $|\lambda\rangle$ is given in terms of the displacement operator (in one dimension) by

$$|\lambda\rangle = e^{-\frac{i}{\hbar}\hat{p}\lambda} |0\rangle \tag{13.176}$$

For the harmonic oscillator system

$$\hat{p} = \frac{1}{i} \sqrt{\frac{m\hbar\omega}{2}} (\hat{a} - \hat{a}^+) \tag{13.177}$$

If we choose $|0\rangle$ to be the ground state of the oscillator, then we have for the corresponding displaced ground-state

$$|\lambda\rangle = e^{\sqrt{\frac{m\omega}{2\hbar}}(\hat{a}^+ - \hat{a})\lambda} |0\rangle \tag{13.178}$$

By Glauber's theorem

$$e^{(\hat{A}+\hat{B})} = e^{\hat{A}} e^{\hat{B}} e^{-\frac{1}{2}[\hat{A},\hat{B}]} \tag{13.179}$$

we have

$$\begin{aligned} e^{\sqrt{\frac{m\omega}{2\hbar}}(\hat{a}^+ - \hat{a})\lambda} &= e^{\sqrt{\frac{m\omega}{2\hbar}}\hat{a}^+\lambda} e^{-\sqrt{\frac{m\omega}{2\hbar}}\hat{a}\lambda} e^{\frac{1}{2}\frac{m\omega}{2\hbar}[\hat{a}^+,\hat{a}]\lambda^2} \\ &= e^{\sqrt{\frac{m\omega}{2\hbar}}\hat{a}^+\lambda} e^{-\sqrt{\frac{m\omega}{2\hbar}}\hat{a}\lambda} e^{-\frac{1}{4}\frac{m\omega}{\hbar}\lambda^2} \end{aligned} \tag{13.180}$$

and thus

$$|\lambda\rangle = e^{\sqrt{\frac{m\omega}{2\hbar}}\hat{a}^+\lambda} e^{-\sqrt{\frac{m\omega}{2\hbar}}\hat{a}\lambda} e^{-\frac{1}{4}\frac{m\omega}{\hbar}\lambda^2} |0\rangle \tag{13.181}$$

Now

$$e^{-\sqrt{\frac{m\omega}{2\hbar}}\hat{a}\lambda} |0\rangle = \left(\hat{I} + \left(-\sqrt{\frac{m\omega}{2\hbar}}\lambda\hat{a} \right) + \frac{1}{2} \left(-\sqrt{\frac{m\omega}{2\hbar}}\lambda\hat{a} \right)^2 + ... \right) |0\rangle = |0\rangle$$

using $\hat{a}\,|0\rangle = 0$. Similarly, using $(\hat{a}^+)^n\,|0\rangle = \sqrt{n!}\,|n\rangle$ we have

$$\begin{aligned} e^{\sqrt{\frac{m\omega}{2\hbar}}\hat{a}^+\lambda}\,|0\rangle &= \left(\hat{I} + \left(\sqrt{\frac{m\omega}{2\hbar}}\lambda\hat{a}^+\right) + \frac{1}{2}\left(\sqrt{\frac{m\omega}{2\hbar}}\lambda\hat{a}^+\right)^2 + ...\right)|0\rangle \\ &= |0\rangle + \sqrt{\frac{m\omega}{2\hbar}}\lambda\,|1\rangle + \frac{1}{2}\left(\sqrt{\frac{m\omega}{2\hbar}}\lambda\right)^2|2\rangle + ... \\ &= \sum_{n=0}^{\infty}\frac{\left(\sqrt{\frac{m\omega}{2\hbar}}\lambda\right)^n}{\sqrt{n!}}\;|n\rangle \end{aligned} \tag{13.182}$$

or

$$|\lambda\rangle = e^{-\frac{1}{4}\frac{m\omega}{\hbar}\lambda^2}\sum_{n=0}^{\infty}\frac{\left(\sqrt{\frac{m\omega}{2\hbar}}\lambda\right)^n}{\sqrt{n!}}\;|n\rangle \tag{13.183}$$

Thus,

$$|\lambda\rangle = \sum_{n=0}^{\infty} b_n\,|n\rangle \tag{13.184}$$

where

$$b_n = \frac{e^{-\frac{N}{2}}N^{\frac{n}{2}}}{\sqrt{n!}} \quad , \quad \frac{N}{2} = \frac{m\omega}{4\hbar}\lambda^2 \tag{13.185}$$

or

$$\begin{aligned} P_n &= \text{ probability of finding the system in the state } |n\rangle \\ &= |b_n|^2 = \frac{e^{-N}N^n}{n!} \end{aligned} \tag{13.186}$$

which is a Poisson distribution. Thus, we obtain the coherent states once again.

Let us now return to the original discussion. In the state $|\alpha\rangle$ we have

$$\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle \rightarrow \langle\alpha|\,\hat{a}^+ = \alpha^*\,\langle\alpha| \tag{13.187}$$

so that

$$\langle E\rangle = \langle\alpha|\,\hat{H}\,|\alpha\rangle = \hbar\omega\,\langle\alpha|\left(\hat{N}+1/2\right)\rangle|\alpha\rangle = \hbar\omega\left(|\alpha|^2+1/2\right)\rangle \tag{13.188}$$

$$\langle x\rangle = \sqrt{\frac{\hbar}{2m\omega}}\,\langle\alpha|\,(\hat{a}+\hat{a}^+)\,|\alpha\rangle = \sqrt{\frac{\hbar}{2m\omega}}(\alpha+\alpha^*) \tag{13.189}$$

$$\langle p\rangle = -i\sqrt{\frac{m\hbar\omega}{2}}\,\langle\alpha|\,(\hat{a}-\hat{a}^+)\,|\alpha\rangle = i\sqrt{\frac{m\hbar\omega}{2}}(\alpha^*-\alpha) \tag{13.190}$$

$$\begin{aligned}(\Delta x)^2 &= \frac{\hbar}{2m\omega}\,\langle\alpha|\,(\hat{a}+\hat{a}^+)^2\,|\alpha\rangle - \langle x\rangle^2\\ &= \frac{\hbar}{2m\omega}\left[(\alpha+\alpha^*)^2+1\right] - \frac{\hbar}{2m\omega}\,(\alpha+\alpha^*)^2\\ &= \frac{\hbar}{2m\omega} \rightarrow \Delta x = \sqrt{\frac{\hbar}{2m\omega}}\end{aligned} \tag{13.191}$$

$$\begin{aligned}(\Delta p)^2 &= -\frac{m\hbar\omega}{2}\,\langle\alpha|\,(\hat{a}-\hat{a}^+)^2\,|\alpha\rangle - \langle p\rangle^2\\ &= -\frac{m\hbar\omega}{2}\left[(\alpha-\alpha^*)^2-1\right] + \frac{m\hbar\omega}{2}\,(\alpha^*-\alpha)^2\\ &= \frac{m\hbar\omega}{2} \rightarrow \Delta p = \sqrt{\frac{m\hbar\omega}{2}}\end{aligned} \tag{13.192}$$

Therefore, the Heisenberg inequality becomes an equality in this case

$$\Delta x\Delta p = \frac{\hbar}{2} \tag{13.193}$$

independent of the value of α.

We can find the wave functions corresponding to $|\alpha\rangle$ using the earlier method. We have in the position representation:

$$\begin{aligned}\langle X|\,\hat{a}\,|\alpha\rangle &= \alpha\,\langle X\mid\alpha\rangle = \frac{1}{\sqrt{2}}\,\langle X|\left(\hat{X}+i\hat{P}\right)|\alpha\rangle\\ &= \frac{1}{\sqrt{2}}X\,\langle X\mid\alpha\rangle + \frac{i}{\sqrt{2}}\left(-i\frac{\partial}{\partial X}\right)\langle X\mid\alpha\rangle\end{aligned} \tag{13.194}$$

$$\begin{aligned}\frac{1}{\sqrt{2}}\left(X+\frac{\partial}{\partial X}\right)\langle X\mid\alpha\rangle &= \alpha\,\langle X\mid\alpha\rangle\\ \rightarrow \langle X\mid\alpha\rangle &= Ae^{-(X-\alpha\sqrt{2})^2/2} = \psi_\alpha(X)\end{aligned} \tag{13.195}$$

and in the momentum representation:

$$\begin{aligned}\langle P|\,\hat{a}\,|\alpha\rangle &= \alpha\,\langle P\mid\alpha\rangle = \frac{1}{\sqrt{2}}\,\langle P|\left(\hat{X}+i\hat{P}\right)|\alpha\rangle \\ &= \frac{1}{\sqrt{2}}i\frac{\partial}{\partial P}\,\langle P\mid\alpha\rangle + \frac{i}{\sqrt{2}}P\,\langle p\mid\alpha\rangle \qquad (13.196)\end{aligned}$$

$$\begin{aligned}\frac{1}{\sqrt{2}}\left(P+\frac{\partial}{\partial P}\right)\langle P\mid\alpha\rangle &= \alpha\,\langle P\mid\alpha\rangle \\ \rightarrow \langle P\mid\alpha\rangle &= A'e^{-(P+i\alpha\sqrt{2})^2/2} = \phi_\alpha(P) \qquad (13.197)\end{aligned}$$

Suppose that at time $t=0$, the oscillator is in a quasi-classical state $|\psi(0)\rangle = |\alpha_0\rangle$ with $\alpha_0 = \rho e^{i\varphi}$ where ρ is a real positive number. Then at a later time t

$$\begin{aligned}|\psi(t)\rangle = |\alpha_0,t\rangle &= e^{-\frac{1}{2}|\alpha_0|^2}\sum_{n=0}^{\infty}\frac{\alpha_0^n}{\sqrt{n!}}\,|n,t\rangle \\ &= e^{-\frac{1}{2}|\alpha_0|^2}\sum_{n=0}^{\infty}\frac{\alpha_0^n}{\sqrt{n!}}e^{-i\frac{E_n}{\hbar}t}\,|n\rangle \\ &= e^{-\frac{1}{2}|\alpha_0|^2}e^{-i\omega t/2}\sum_{n=0}^{\infty}\frac{\alpha_0^n}{\sqrt{n!}}e^{-in\omega t}\,|n\rangle \\ &= e^{-i\omega t/2}\,|\alpha(t)\rangle \qquad (13.198)\end{aligned}$$

where $\alpha(t) = \alpha_0 e^{-i\omega t} = \rho e^{-i(\omega t-\varphi)}$.

Finally, we have

$$\begin{aligned}\langle\alpha,t|\,\hat{x}\,|\alpha,t\rangle &= \sqrt{\frac{\hbar}{2m\omega_0}}b_0^2\sum_k\frac{1}{k!}\,|\alpha|^{2k}\left(\alpha e^{-i\omega_0 t}+\alpha^* e^{i\omega_0 t}\right) \\ &= 2\sqrt{\frac{\hbar}{2m\omega_0}}\,|\alpha|\cos(\omega_0 t-\phi)(b_0^2\sum_k\frac{|\alpha|^{2k}}{k!}) \quad , \quad x_0 = \sqrt{\frac{\hbar}{2m\omega_0}} \\ &= x_0\cos(\omega_0 t-\phi) \quad , \quad x_0 = \rho\sqrt{\frac{2\hbar}{m\omega_0}} \qquad (13.199)\end{aligned}$$

and

$$\langle\alpha,t|\,\hat{p}\,|\alpha,t\rangle = -p_0\sin(\omega_0 t-\phi) \quad , \quad p_0 = \rho\sqrt{2m\hbar\omega} \qquad (13.200)$$

In addition, we have (for $\rho \gg 1$)

$$\frac{\Delta x}{x_0} = \frac{1}{2\rho} << 1 \quad , \quad \frac{\Delta p}{p_0} = \frac{1}{2\rho} << 1 \qquad (13.201)$$

This says that the relative uncertainties in the position and momentum of the oscillator are quite accurately defined at any time. Hence the name *quasi-classical state.*

Let us look at some numbers. We consider a pendulum of length 1 meter and of

mass 1 gram and assume that the state of this pendulum can be described by a quasi-classical state. At time $t = 0$ we assume that the pendulum is at $\langle x(0)\rangle = 1$ micron from its classical equilibrium position, with zero mean velocity.

An appropriate choice is $\langle x(0)\rangle = x_0$, $\langle p(0)\rangle = 0 \rightarrow \phi = 0$. We also have

$$\omega = 2\pi\nu = \sqrt{\frac{g}{\ell}} = 3.13\, s^{-1} \rightarrow \alpha(0) = 3.9 \times 10^9 \tag{13.202}$$

The relative uncertainty in the position is

$$\frac{\Delta x}{x_0} = \frac{1}{2\rho} = \frac{1}{2\alpha(0)} = 1.3 \times 10^{-10} \tag{13.203}$$

We note that after 1/4 period of oscillation,

$$T = \text{ period } = \frac{2\pi}{\omega} \tag{13.204}$$

$$\rightarrow \alpha\left(T/4\right) = \alpha(0)e^{i\omega T/4} = \alpha(0)e^{i\pi/2} = i\alpha(0) = -3.9i \times 10^9 \tag{13.205}$$

13.7.2 Construction of a Schrodinger-Cat State

Suppose that during the interval $[0, T]$ we add to the harmonic potential, the coupling (interaction)

$$\hat{W} = \hbar g \left(\hat{a}^+\hat{a}\right)^2 = \hbar g \hat{N}^2 \tag{13.206}$$

We will assume that $g \gg \omega, \omega T \ll 1$. Under these conditions, we can make the approximation that, during the interval $[0, T]$, the Hamiltonian of the system is simply $\hat{W}$. Assume that at time $t = 0$, the system is in a quasi-classical state $|\psi(0)\rangle = |\alpha\rangle$.

The eigenvectors of $\hat{W}$ are $\{|n\rangle\}$ with $\hat{W}|n\rangle = \hbar g n^2 |n\rangle$. This implies that for $|\psi(0)\rangle = |\alpha\rangle$

$$|\psi(T)\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} e^{-ign^2T} |n\rangle \tag{13.207}$$

Some special cases will be of interest later.

$$|\psi(T = 2\pi/g)\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} e^{-i2\pi n^2} |n\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle = |\alpha\rangle$$

$$|\psi(T = \pi/g)\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} e^{-i\pi n^2} |n\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} (-1)^n |n\rangle = |-\alpha\rangle$$

$$
\begin{aligned}
|\psi(T=\pi/2g)\rangle &= e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}e^{-i\pi n^2/2}|n\rangle \\
&= e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\frac{1}{2}\left[1-i+(1+i)(-1)^n\right]|n\rangle \\
&= e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\frac{1}{\sqrt{2}}\left[e^{-i\pi/4}+e^{i\pi/4}(-1)^n\right]|n\rangle \\
&= \frac{1}{\sqrt{2}}\left[e^{-i\pi/4}|\alpha\rangle + e^{i\pi/4}|-\alpha\rangle\right] \qquad (13.208)
\end{aligned}
$$

Now, suppose that α is pure imaginary, that is, $\alpha = i\rho$. In this case, in the state $|\alpha\rangle$, the oscillator has a zero mean position and a positive velocity.

$$
\begin{aligned}
\langle x\rangle &= \sqrt{\frac{\hbar}{2m\omega}}(\alpha+\alpha^*) = 0 \\
\langle p\rangle &= i\sqrt{\frac{m\hbar\omega}{2}}(\alpha^*-\alpha) = \sqrt{2m\hbar\omega}\rho
\end{aligned}
$$

Similarly, in the state $|-\alpha\rangle$, the oscillator also has a zero mean position, but a negative velocity.

If $|\alpha| \gg 1$, the states $|\alpha\rangle$ and $|-\alpha\rangle$ are macroscopically different. The state (14.208) is a quantum superposition of such states. It therefore constitutes a (harmless) version of Schrodinger's cat, where we represent *dead* and *alive* cats by simple vectors in Hilbert space.

13.7.3 Quantum Superposition Versus Statistical Mixture

We now consider the properties of the state (14.208) in a *macroscopic* situation $|\alpha| \gg 1$. We will choose $\alpha = i\rho$ pure imaginary and we set $p_0 = \sqrt{2m\hbar\omega\rho}$.

The probability distributions for position and momentum are given by

$$
\begin{aligned}
Pr(X) &\propto \left|e^{-i\pi/4}\langle X\mid\alpha\rangle + e^{i\pi/4}\langle X\mid-\alpha\rangle\right|^2 \\
&\propto \left|e^{-i\pi/4}e^{-(X-i\rho\sqrt{2})^2/2} + e^{i\pi/4}e^{-(X+i\rho\sqrt{2})^2/2}\right|^2 \\
&\propto e^{-X^2}\cos^2\left(\sqrt{2}X\rho - \frac{\pi}{4}\right) \qquad (13.209)
\end{aligned}
$$

$$
\begin{aligned}
Pr(P) &\propto \left|e^{-i\pi/4}\langle P\mid\alpha\rangle + e^{i\pi/4}\langle P\mid-\alpha\rangle\right|^2 \\
&\propto \left|e^{-i\pi/4}e^{-(P-\rho\sqrt{2})^2/2} + e^{i\pi/4}e^{-(P+\rho\sqrt{2})^2/2}\right|^2 \\
&\approx e^{-(P-\rho\sqrt{2})^2} + e^{-(P+\rho\sqrt{2})^2} \qquad (13.210)
\end{aligned}
$$

where in the last expression we have used the fact that, for $\rho \gg 1$, the two Gaussians centered at $\rho\sqrt{2}$ and $-\rho\sqrt{2}$ have a negligible overlap.

These probability distributions are plotted in Figure 13.10 below for $\alpha = 5i$

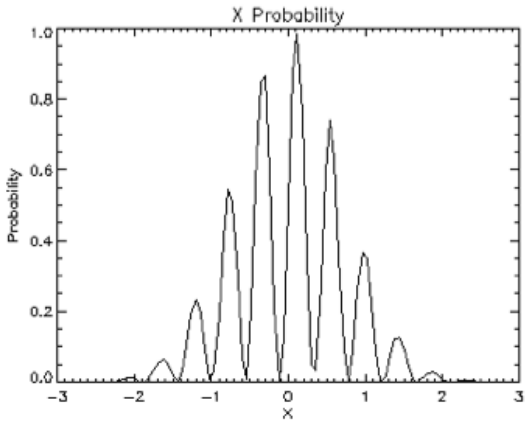


Figure 13.10: X Probability Distribution

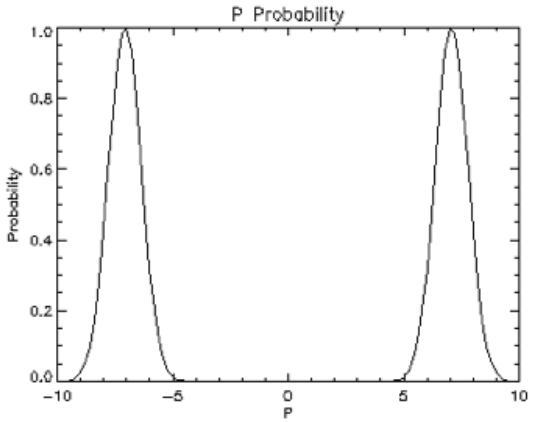


Figure 13.11: P Probability Distribution

Suppose that a physicist (Alice) prepares N independent systems all in the state (13.208) and measures the momentum of each of these systems. Suppose the measuring apparatus has a resolution δp such that:

$$\sqrt{m\hbar\omega} << \delta p << p_0 \tag{13.211}$$

For $N \gg 1$, the results of these measurements is that Alice (plotting a histogram) will find two peaks, each of which contains roughly half of the events,

centered respectively at p_0 and $-p_0$ (resembling Figure 13.11).

The state (13.208) represents the quantum superposition of two states which are macroscopically different, and therefore leads to the paradoxical situations mentioned earlier.

Another physicist (Bob) claims that the measurements done by Alice have not been performed on N quantum systems in the state (14.208), but that Alice is actually dealing with a nonparadoxical *statistical mixture*, that is, half of the N systems are in the state $|\alpha\rangle$ and the other half in the state $|-\alpha\rangle$.

Assuming that this is true, the statistical mixture of Bob leads (after N momentum measurements) to the same momentum distribution as that measured by Alice: the $N/2$ oscillators in the state $|\alpha\rangle$ all lead to a mean momentum p_0 and the $N/2$ oscillators in the state $|-\alpha\rangle$ all lead to a mean momentum $-p_0$. Up to this point, there is therefore no difference and no paradoxical behavior related to the quantum superposition (13.208).

In order to settle the matter, Alice now measures the position of each of the N independent systems, all prepared in the state (14.208). Assuming that the resolution δx of the measuring apparatus is such that

$$\delta x << \frac{1}{|\alpha|}\sqrt{\frac{\hbar}{m\omega}} \rightarrow \delta X << \frac{1}{|\alpha|} = \frac{1}{\rho} \tag{13.212}$$

Alice has sufficient resolution to observe the oscillations of the function

$$\cos^2\left(\sqrt{2}X\rho - \frac{\pi}{4}\right) \tag{13.213}$$

in the distribution $Pr(X)$. The shape of the distribution for x will therefore reproduce the probability law for X as drawn in Figure 14.10 above, that is, a modulation of period

$$\left[\hbar\pi^2/(2m\alpha^2\omega)\right]^{1/2} \tag{13.214}$$

with a Gaussian envelope.

We continue with the assumption that Bob is dealing with a statistical mixture. If Bob performs a position measurement on the $N/2$ systems in the state $|\alpha\rangle$, he will find a Gaussian distribution corresponding to the probability law

$$Pr(X) \propto |\langle X \mid \alpha\rangle|^2 \propto e^{-X^2} \tag{13.215}$$

He will find the same distribution for $N/2$ systems in the state $|-\alpha\rangle$. The sum of his results will be a Gaussian distribution, which is quite different(see Figure 14.11) from the result expected by Alice.

The position measurement should, in principle, allow one to discriminate between the quantum superposition and the statistical mixture.

In our earlier discussion of numbers for a pendulum we found that $\alpha = 3.9 \times 10^9$. Therefore, the resolution δx which is necessary in order to tell the difference between a set of N systems in a quantum superposition (14.208), and a statistical mixture consisting of $N/2$ systems in the state $|\alpha\rangle$ and $N/2$ systems in the state $|-\alpha\rangle$ is given by

$$\delta x << \frac{1}{|\alpha|}\sqrt{\frac{\hbar}{m\omega}} \approx 5 \times 10^{-26} m \tag{13.216}$$

Clearly, it is impossible to attain such a resolution in practice!

13.7.4 The Fragility of a Quantum Superposition

In a realistic physical situation, one must take into account the coupling of the oscillator with its environment, in order to estimate how long one can discriminate between the quantum superposition (13.208), that is, the *Schrodinger cat* which is *alive and dead*, and a simple statistical mixture, that is, a set of cats (systems), half of which are alive, the other half beginning dead; each cat being *either* alive *or* dead.

If the oscillator is initially in the quasi-classical state $|\alpha_0\rangle$ and if the environment is in a state $|\xi_e(0)\rangle$, the wave function of the total system is the product of the individual wave functions, and the state vector of the total system can be written as the (tensor) product of the state vectors of the two subsystems:

$$|\Phi(0)\rangle = |\alpha_0\rangle\,|\chi_e(0)\rangle \tag{13.217}$$

The coupling is responsible for the damping of the oscillator's amplitude.

At a later time t, the state vector of the total system becomes

$$|\Phi(t)\rangle = |\alpha_1\rangle\,|\chi_e(t)\rangle \tag{13.218}$$

where $\alpha_1 = \alpha(t)e^{-\gamma t}$. The number $\alpha(t)$ corresponds to the quasi-classical state one would find in the absence of damping (evaluated earlier as $\alpha(t) = \alpha_0 e^{-i\omega t}$) and γ is a real positive number.

From earlier

$$E(t) = \hbar\omega\left(|\alpha(t)|^2 + 1/2\right)\Big) = \hbar\omega\left(|\alpha_0|^2\, e^{-2\gamma t} + 1/2\right)\Big) \tag{13.219}$$

The energy decreases with time. After a time much longer than $1/\gamma$, the oscillator is in its ground state. This dissipation model corresponds to a zero temperature environment. The mean energy acquired by the environment is

$$E(0) - E(t) = \hbar\omega\,|\alpha_0|^2\,(1 - e^{-2\gamma t}) \approx 2\hbar\omega\,|\alpha_0|^2\,\gamma t \quad , \quad 2\gamma t << 1 \tag{13.220}$$

For initial states of the *Schrodinger cat* type for the oscillator, the state vector of the total system, at $t = 0$,

$$|\Phi(0)\rangle = \frac{1}{\sqrt{2}}\left(e^{-i\pi/4}\,|\alpha_0\rangle + e^{i\pi/4}\,|-\alpha_0\rangle\right)|\chi_e(0)\rangle \tag{13.221}$$

and, at a later time t,

$$|\Phi(t)\rangle = \frac{1}{\sqrt{2}}\left(e^{-i\pi/4}\,|\alpha_1\rangle\left|\chi_e^{(+)}(t)\right\rangle + e^{i\pi/4}\,|-\alpha_1\rangle\left|\chi_e^{(-)}(t)\right\rangle\right) \tag{13.222}$$

still with $\alpha_1 = \alpha(t)e^{-\gamma t}$. We assume that t is chosen such that α_1 is pure imaginary, $|\alpha_1| \gg 1$, and $\left|\chi_e^{(+)}(t)\right\rangle$ and $\left|\chi_e^{(-)}(t)\right\rangle$ are two normalized states of the environment that are a priori different (but not orthogonal).

The probability distribution of the oscillator's position, measured independently of the state of the environment, is then

$$\begin{aligned} Pr(x) = {} & \frac{1}{2}\,|\langle x \mid \alpha_1\rangle|^2 + \frac{1}{2}\,|\langle x \mid -\alpha_1\rangle|^2 \\ & + Real\left(i\,\langle x \mid \alpha_1\rangle^*\,\langle x \mid -\alpha_1\rangle\left\langle \chi_e^{(+)}(t) \mid \chi_e^{(-)}(t)\right\rangle\right) \end{aligned} \tag{13.223}$$

Let $\eta = \left\langle \chi_e^{(+)}(t) \,\middle|\, \chi_e^{(-)}(t)\right\rangle$. We then have $0 \le \eta \le 1$, η real.

This says that the probability distribution of the position keeps its Gaussian envelope, but the contrast of the oscillations (cross term) is reduced by a factor η.

The probability distribution for the momentum is given by

$$Pr(p) = \frac{1}{2}\,|\langle p \mid \alpha_1\rangle|^2 + \frac{1}{2}\,|\langle p \mid -\alpha_1\rangle|^2 + \eta \mathrm{Real}\left(i\,\langle p \mid -\alpha_1\rangle^*\,\langle p \mid \alpha_1\rangle\right) \tag{13.224}$$

Since the overlap of the two Gaussians $\langle p\,|\,\alpha_1\rangle$ and $\langle p\,|\,-\alpha_1\rangle$ is negligible for $|\alpha_1| \gg 1$, the crossed term, which is proportional to η does not contribute significantly. We recover two peaks centered at $\pm|\alpha_1|\sqrt{2m\hbar\omega}$. The distinction between a quantum superposition and a statistical mixture can be made by position measurements. The quantum superposition leads to a modulation of spatial period

$$\left[\hbar\pi^2/(2m\alpha^2\omega)\right]^{1/2} \tag{13.225}$$

with a Gaussian envelope, whereas only the Gaussian is observed for a statistical mixture.

In order to see this modulation, the parameter η must not be too small, say $\eta \ge 1/10$.

In a very simple model, the environment is represented by a second oscillator, of the same mass and frequency as the first one. We will assume that this second oscillator is initially in its ground state $|\xi_e(0)\rangle = |0\rangle$. If the coupling between the two oscillators is quadratic, we can take for granted that

1. the states $\left|\chi_e^{(\pm)}(t)\right\rangle$ are quasi-classical: $\left|\chi_e^{(\pm)}(t)\right\rangle = |\pm\beta\rangle$

2. and that, for short times ($\gamma t \ll 1$): $|\beta|^2 = 2\gamma t|\alpha_0|^2$

A simple calculation then gives

$$\langle \beta \mid -\beta \rangle = e^{-|\beta|^2} \sum_n \frac{\beta^{*n}(-\beta)^n}{n!} = e^{-|\beta|^2} e^{-|\beta|^2} = e^{-2|\beta|^2} \tag{13.226}$$

From earlier considerations we must have

$$\eta = \langle \beta \mid -\beta \rangle = e^{-2|\beta|^2} \geq 1/10 \rightarrow |\beta| \leq 1 \tag{13.227}$$

For times shorter than $1/\gamma$, the energy of the first oscillator is

$$E(t) = E(0) - 2\gamma t \left|\alpha_0\right|^2 \hbar\omega \tag{13.228}$$

The energy of the second oscillator is

$$E'(t) = \hbar\omega \left|(\beta(t)\right|^2 + 1/2) = \hbar\omega/2 + 2\gamma t \left|\alpha_0\right|^2 \hbar\omega \tag{13.229}$$

The total energy is conserved: the energy transferred during the time t is

$$\Delta E(t) = 2\gamma t \left|\alpha_0\right|^2 \hbar\omega = \hbar\omega \left|(\beta(t)\right|^2 \tag{13.230}$$

In order to distinguish between a quantum superposition and a statistical mixture, we must have $\Delta E \leq \hbar\omega$. In other words, if a single energy quantum $\hbar\omega$ is transferred, it becomes problematic to tell the difference.

If we return to the numerical example of the pendulum we have the following results: with $1/2\gamma = 1\ year = 3 \times 10^7 s$, the time it takes to reach $|\beta| = 1$ is $(2\gamma \left|\alpha_0\right|^2)^{-1} \approx 2 \times 10^{-12} s$

Conclusion

Even for a system as well protected from the environment as we have assumed for the pendulum, the quantum superpositions of macroscopic states are unobservable. After a very short time, all measurements one can make on a system initially prepared in such a state coincide with those made on a statistical mixture. It is therefore not possible, at present, to observe the effects related to the paradoxical character of a macroscopic quantum superposition. However, it is quite possible to observe *mesoscopic* kittens, for systems which have a limited number of degrees of freedom and are well isolated.

13.8 The Quantum Eraser

We now investigate a quantum process where the superposition of two probability amplitudes leads to an interference phenomenon. The two amplitudes will be associated with two quantum paths as in the double slit experiment. In the investigation we will show that the interference *disappears* if an intermediate

measurement gives information about which path has actually been followed. Then we will show how the interference can actually *reappear* if the path information is *erased* by a quantum device.

We consider a beam of neutrons, which are particles with charge zero and spin 1/2, propagating along the $x-$axis with velocity v. We will treat the motion of the neutrons classically as uniform linear motion. Only the evolution of their spin states will be treated quantum mechanically.

13.8.1 Magnetic Resonance

The eigenstates of the z component of the neutron spin are denoted $|n:\pm\rangle$. A constant uniform magnetic field $\vec{B}_0 = B_0\hat{u}_z$ is applied along the $z-$axis. The magnetic moment of the neutron is denoted by $\hat{\mu} = \gamma_n \hat{S}_n$ where γ_n is the gyromagnetic ratio and $\hat{S}_n$ is the spin operator of the neutron.

The magnetic energy levels of the neutron in the presence of the field $\vec{B}_0$ are $E_\pm = \mp\gamma_n \hbar B_0/2 = \pm\hbar\omega_0/2$ where $\omega_0 = -\gamma_n B_0$.

The neutrons cross a cavity of length L between times t_0 and $t_1 = t_0 + L/v$. Inside the cavity, in addition to the constant field $\vec{B}_0$, a rotating field $\vec{B}_1(t)$ is applied. The field $\vec{B}_1(t)$ lies in the (x, y) plane and has a constant angular frequency ω:

$$\vec{B}_1(t) = B_1\left(\cos\omega t\hat{u}_x + \sin\omega t\hat{u}_y\right) \tag{13.231}$$

Let $|\psi_n(t)\rangle = \alpha_+ |n:+\rangle + \alpha_- |n:-\rangle$ be the neutron spin state at time t and consider a neutron entering the cavity at time t_0.

The Hamiltonian for the system is

$$\begin{aligned} H &= \hat{\mu}_n \cdot \left(\vec{B}_0 + \vec{B}_1(t)\right) = \gamma_n \hat{S}_n \cdot \left(\vec{B}_0 + \vec{B}_1(t)\right) \\ &= \gamma_n \left(B_0 \hat{S}_z + B_1\left(\cos\omega t\hat{S}_x + \sin\omega t\hat{S}_y\right)\right) \end{aligned} \tag{13.232}$$

which gives in the $|n:\pm\rangle$ basis

$$H = \frac{\hbar}{2}\begin{pmatrix} \omega_0 & \omega_1 e^{-i\omega t} \\ \omega_1 e^{i\omega t} & -\omega_0 \end{pmatrix} \tag{13.233}$$

Therefore, the evolution equations are

$$\begin{aligned} H|\psi_n(t)\rangle &= \frac{\hbar}{2}\begin{pmatrix} \omega_0 & \omega_1 e^{-i\omega t} \\ \omega_1 e^{i\omega t} & -\omega_0 \end{pmatrix}\alpha_+\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha_-\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= i\hbar\frac{d}{dt}|\psi_n(t)\rangle = i\hbar\frac{d\alpha_+}{dt}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + i\hbar\frac{d\alpha_-}{dt}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned} \tag{13.234}$$

or

$$i\hbar\frac{d\alpha_+}{dt} = \frac{\hbar}{2}\omega_0\alpha_+ + \frac{\hbar}{2}\omega_1 e^{-i\omega t}\alpha_- \tag{13.235}$$

$$i\hbar\frac{d\alpha_-}{dt} = \frac{\hbar}{2}\omega_1 e^{i\omega t}\alpha_+ - \frac{\hbar}{2}\omega_0\alpha_- \tag{13.236}$$

or

$$i\frac{d\alpha_+}{dt} = \frac{1}{2}\omega_0\alpha_+ + \frac{1}{2}\omega_1 e^{-i\omega t}\alpha_- \tag{13.237}$$

$$i\frac{d\alpha_-}{dt} = \frac{1}{2}\omega_1 e^{i\omega t}\alpha_+ - \frac{1}{2}\omega_0\alpha_- \tag{13.238}$$

Defining

$$\alpha_\pm(t) = \beta_\pm(t)e^{\mp i\omega(t-t_0)/2} \tag{13.239}$$

we get

$$i\frac{d\beta_+}{dt} = \frac{\omega_0-\omega}{2}\beta_+ + \frac{\omega_1}{2}e^{-i\omega t_0}\beta_- \tag{13.240}$$

$$i\frac{d\beta_-}{dt} = \frac{\omega-\omega_0}{2}\beta_- + \frac{\omega_1}{2}e^{i\omega t_0}\beta_+ \tag{13.241}$$

which has constant coefficients.

We assume that near resonance, $|\omega-\omega_0| \ll \omega_1$, and that terms proportional to $\omega-\omega_0$ may be neglected in these equations. The equations become

$$i\frac{d\beta_+}{dt} = \frac{\omega_1}{2}e^{-i\omega t_0}\beta_- \tag{13.242}$$

$$i\frac{d\beta_-}{dt} = \frac{\omega_1}{2}e^{i\omega t_0}\beta_+ \tag{13.243}$$

whose solution is

$$\beta_\pm(t) = \beta_\pm(t_0)\cos\frac{\omega_1(t-t_0)}{2} - i\beta_\mp(t_0)e^{\mp i\omega t_0}\sin\frac{\omega_1(t-t_0)}{2} \tag{13.244}$$

Defining

$$\varphi = \frac{\omega_1(t_1-t_0)}{2} \quad , \quad \chi = \frac{\omega(t_1-t_0)}{2} \quad , \quad \delta = \frac{\omega(t_1+t_0)}{2} \tag{13.245}$$

we get

$$\alpha_+(t_1) = e^{-i\chi}\beta_+(t_1) = e^{-i\chi}\left[\alpha_+(t_0)\cos\varphi - i\alpha_-(t_0)e^{-i\omega t_0}\sin\varphi\right] \tag{13.246}$$

$$\alpha_-(t_1) = e^{i\chi}\beta_-(t_1) = e^{i\chi}\left[\alpha_-(t_0)\cos\varphi - i\alpha_+(t_0)e^{+i\omega t_0}\sin\varphi\right] \tag{13.247}$$

$$\alpha_+(t_1) = \left[\alpha_+(t_0)e^{-i\chi}\cos\varphi - i\alpha_-(t_0)e^{-i\delta}\sin\varphi\right] \tag{13.248}$$

$$\alpha_-(t_1) = \left[\alpha_-(t_0)e^{i\chi}\cos\varphi - i\alpha_+(t_0)e^{i\delta}\sin\varphi\right] \tag{13.249}$$

$$\begin{pmatrix}\alpha_+(t_1)\\ \alpha_-(t_1)\end{pmatrix} = U(t_0,t_1)\begin{pmatrix}\alpha_+(t_0)\\ \alpha_-(t_0)\end{pmatrix} \tag{13.250}$$

where

$$U(t_0,t_1) = \begin{pmatrix} e^{-i\chi}\cos\varphi & -ie^{-i\delta}\sin\varphi \\ -ie^{i\delta}\sin\varphi & e^{i\chi}\cos\varphi \end{pmatrix} \tag{13.251}$$

is the time evolution matrix.

13.8.2 Ramsey Fringes

The neutrons are initially in the spin state $|n:-\rangle$. They successively cross two identical cavities of the type described above. This is called a Ramsey configuration and it is shown in Figure 13.12 below.

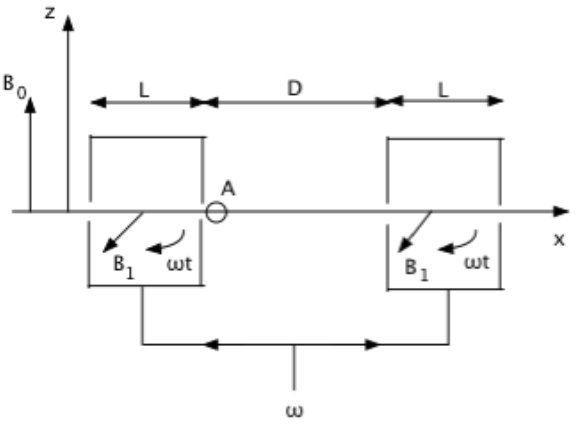


Figure 13.12: Ramsey Configuration

The object A is a detecting atom described later. The same oscillating field $\vec{B}_1(t)$, is applied in both cavities. The magnitude B_1 of this field is applied so as to satisfy the condition $\varphi = \pi/4$. The constant field $\vec{B}_0$ is applied throughout the entire experimental setup. At the end of the setup, one measures the number of outgoing neutrons which have flipped their spin and are in the final state $|n:+\rangle$. This is done for several values of ω near $\omega = \omega_0$.

The initial state condition corresponds to

$$\alpha_+(t_0) = 0 \quad , \quad \alpha_-(t_0) = 1 \tag{13.252}$$

At time t_1 the state is

$$\begin{pmatrix} \alpha_+(t_1) \\ \alpha_-(t_1) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\chi} & -ie^{-i\delta} \\ -ie^{i\delta} & e^{i\chi} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -ie^{-i\delta} \\ e^{i\chi} \end{pmatrix} \tag{13.253}$$

or

$$|\psi(t_1)\rangle = \frac{1}{\sqrt{2}} \left(-ie^{-i\delta}\,|n:+\rangle + e^{i\chi}\,|n:-\rangle \right) \tag{13.254}$$

This says that

$$\alpha_+(t_1) = -\frac{i}{\sqrt{2}}e^{-i\delta} \quad , \quad \alpha_-(t_1) = \frac{1}{\sqrt{2}}e^{i\chi} \tag{13.255}$$

and the probability of finding it in the state $|n:\pm\rangle$ is

$$P_\pm = \frac{1}{2} \tag{13.256}$$

The same neutron enters the second cavity at time $t_0' = t_1 + T$, with $T = D/v$, where D is the distance between the two cavities. Between the two cavities the spin precesses freely about $\vec{B}_0$.

We then have

$$\begin{pmatrix} \alpha_+(t_0') \\ \alpha_-(t_0') \end{pmatrix} = \begin{pmatrix} e^{-i\tilde{\chi}} & 0 \\ 0 & e^{i\tilde{\chi}} \end{pmatrix} \begin{pmatrix} \alpha_+(t_1) \\ \alpha_-(t_1) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -ie^{-i\delta}e^{-i\tilde{\chi}} \\ e^{i\chi}e^{i\tilde{\chi}} \end{pmatrix} \tag{13.257}$$

where

$$\tilde{\chi} = \frac{\omega_0 T}{2} \tag{13.258}$$

so that the spin state at t_0' is

$$\begin{pmatrix} \alpha_+(t_0') \\ \alpha_-(t_0') \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -ie^{-i\delta}e^{-i\omega_0 T/2} \\ e^{i\chi}e^{i\omega_0 T} \end{pmatrix} \tag{13.259}$$

Now let t_1' be the time when the neutron leaves the second cavity with $t_1' - t_0' = t_1 - t_0 = D/v$. Now $\delta' = \omega(t_1' + t_0')/2$ is given by

$$\begin{aligned} &t_0' = t_1 + T \quad , \quad t_1' = 2t_1 - t_0 + T \\ &\delta' = \omega(2t_1 - t_0 + T + t_1 + T)/2 = \omega(3t_1 + 2T - t_0)/2 \end{aligned}$$

so that (for the second cavity)

$$U' = U(t_0', t_1') = \begin{pmatrix} e^{-i\chi'}\cos\varphi' & -ie^{-i\delta'}\sin\varphi' \\ -ie^{i\delta'}\sin\varphi' & e^{i\chi'}\cos\varphi' \end{pmatrix} \tag{13.260}$$

where

$$\varphi' = \varphi = \omega_1(t_1 - t_0)/2 \quad , \quad \chi' = \chi = \omega_1(t_1 - t_0)/2 \tag{13.261}$$

so only the parameter δ changes into δ'.

Thus the probability amplitude for detecting the neutron in state + after the second cavity is obtained by

1. Applying U' to the vector

$$\begin{pmatrix} \alpha_+(t_0') \\ \alpha_-(t_0') \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -ie^{-i\delta}e^{-i\omega_0 T/2} \\ e^{i\chi}e^{i\omega_0 T} \end{pmatrix} \tag{13.262}$$

2. calculating the scalar product of the result with the + state.

We have

$$\alpha_+(t_1') = \frac{1}{2}\left(-ie^{-i\delta}e^{-i\chi/2}e^{-i\omega_0 T/2} - ie^{-i\delta'}e^{i\chi/2}e^{i\omega_0 T/2}\right) \tag{13.263}$$

Since

$$\delta + \chi = \omega t_1 \quad , \quad \delta' - \chi = \frac{\omega}{2}\left(3t_1 + 2T - t_0 - t_1 + t_0\right) = \omega\left(t_1 + T\right)$$

we get

$$\begin{aligned}\alpha_+(t_1') &= \frac{1}{2}e^{-i\omega(t_1+T)/2}\left(e^{-i(\omega_0-\omega)T/2} + e^{i(\omega_0-\omega)T/2}\right) \\ &= e^{-i\omega(t_1+T)/2}\cos\frac{(\omega_0-\omega)T}{2}\end{aligned} \tag{13.264}$$

Therefore, the probability that the neutron spin has flipped in the two-cavity system is

$$P_+ = \left|\alpha_+(t_1')\right|^2 = \cos^2\frac{(\omega_0-\omega)T}{2} \tag{13.265}$$

With the approximation $|\omega - \omega_0| \ll \omega_1$, the probability for a spin flip in a single cavity is independent of ω and equal to 1/2. In contrast, the two-cavity result exhibits strong modulation of the spin flip probability between 1 ($\omega = \omega_0$) and 0 ($(\omega_0 - \omega)T = \pi$). This modulation results from an interference process of the two quantum paths corresponding to:

1. a spin flip in 1^{st} cavity and no flip in 2^{nd} cavity
2. no flip in 1^{st} cavity and spin flip in 2^{nd} cavity

In practice, the velocities of the neutrons have some dispersion around the mean value v. This results in a dispersion in the time T to get from one cavity to the other. A typical experimental result giving the intensity of the outgoing beam in the + state as a function of the frequency $f = \omega/2\pi$ of the rotating field B_1 is shown in Figure 13.13 below.

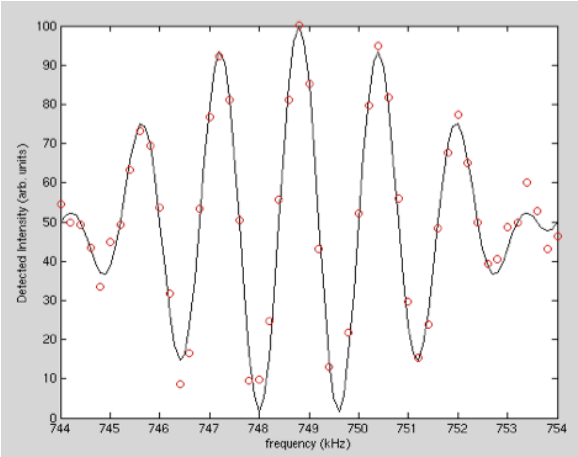


Figure 13.13: Experimental Results

Since $\cos^2 \varphi/2 = (1+\cos\varphi)/2$, the averaged probability distribution is

$$\left\langle \cos^2 \frac{(\omega-\omega_0)T}{2} \right\rangle = \left\langle \frac{1}{2} + \frac{1}{2}\cos((\omega-\omega_0)T) \right\rangle$$
$$= \frac{1}{2} + \frac{1}{2}\int_{-\infty}^{\infty} P(T)\cos((\omega-\omega_0)T)dT \qquad (13.266)$$

For

$$P(T) = \frac{1}{\tau\sqrt{2\pi}} e^{-(T-T_0)^2/2\tau^2} \qquad (13.267)$$

we get

$$\left\langle \cos^2 \frac{(\omega-\omega_0)T}{2} \right\rangle = \frac{1}{2} + \frac{1}{2} e^{-(\omega-\omega_0)^2\tau^2/2}\cos((\omega-\omega_0)T_0) \qquad (13.268)$$

This form agrees with the observed variation with frequency in Figure 14.13 of the experimental signal. The central maximum which is located at $\omega/2\pi = 748.8\,kHz$ corresponds to $\omega = \omega_0$. For that value a constructive interference appears whatever the neutron velocity. The lateral maxima and minima are less peaked, however, since the position of the lateral peak is velocity dependent. The first two lateral maxima correspond to $(\omega-\omega_0)T = \pm 2\pi$. Their amplitude is reduced compared to the central peak by the exponential factor.

This experiment can be compared to a Young's double slit interference experiment with polychromatic light.

Suppose that we insert between the two cavities a device which can measure the z component of the neutron spin (how this works will be discussed shortly). We

define P_{++} as the probability of detecting a neutron in the + state between the two cavities and in the + state when it leaves the second cavity. The probability P_{++} is the product of two probabilities, namely, the probability of finding the neutron in the state + when leaving the first cavity ($p = 1/2$) and, knowing that it is in the + state of finding it in the + state when it leaves the second cavity ($p = 1/2$). This gives $P_{++} = 1/4$. Similarly, $P_{-+} = 1/4$. The sum $P_{++} + P_{-+} = 1/2$ does not display any interference, since one has measured in which cavity the neutron spin has flipped. This is very similar to the electron double slit interference experiment if one measures through which slit the electron passes.

13.8.3 Detection of the Neutron Spin State

In order to measure the spin of the neutron, one lets it interact during a time τ with a spin 1/2 atom at rest. The atom's spin operator is $\hat{S}_a$. Let $|a : \pm z\rangle$ be the two eigenstates of the observable $\hat{S}_{az}$. After the interaction between the neutron and the atom, one measures the spin of the atom. Under certain conditions (which we will derive shortly) one can deduce the spin state of the neutron after the measurement.

Let $|a : \pm x\rangle$ be the eigenstates of $\hat{S}_{ax}$ and $|a : \pm y\rangle$ those of $\hat{S}_{ay}$. We can then write

$$|a : \pm x\rangle = \frac{1}{\sqrt{2}} \left(|a : +z\rangle \pm |a : -z\rangle \right) \tag{13.269}$$

$$|a : \pm y\rangle = \frac{1}{\sqrt{2}} \left(|a : +z\rangle \pm i\, |a : -z\rangle \right) \tag{13.270}$$

and

$$|a : \pm y\rangle = \frac{1}{2} \left((1 \pm i)\, |a : +x\rangle + (1 \mp i)\, |a : -x\rangle \right) \tag{13.271}$$

We assume that the neutron-atom interaction does not affect the neutron's trajectory. We represent the interaction between the neutron and the atom by a very simple model. This interaction is assumed to last a finite time τ during which the neutron-atom interaction Hamiltonian has the form

$$\hat{V} = \frac{2A}{\hbar} \hat{S}_{nz} \otimes \hat{S}_{ax} \tag{13.272}$$

where A is a constant. We neglect the action of any external field, including $\vec{B}_0$ during this time, i.e., we assume the atom-neutron interaction dominates for a short period of time.

The operators $\hat{S}_{nz}$ and $\hat{S}_{ax}$ commute since they act on two different Hilbert spaces. Therefore,

$$\left[\hat{S}_{nz}, \hat{V} \right] = 0 \tag{13.273}$$

The common eigenvectors of $\hat{S}_{nz}$ and $\hat{V}$ and the corresponding eigenvalues are

$$|n:+\rangle \otimes |a:\pm x\rangle \quad \hat{S}_{nz} = +\hbar/2 \quad V = \pm A\hbar/2 \tag{13.274}$$

$$|n:-\rangle \otimes |a:\pm x\rangle \quad \hat{S}_{nz} = -\hbar/2 \quad V = \mp A\hbar/2 \tag{13.275}$$

The operators $\hat{S}_{nz}$ and $\hat{V}$ form a complete set of commuting operators as far as the spin variables are concerned.

From now on we assume that $A\tau = \pi/2$. Suppose that the initial state of the system is

$$|\psi(0)\rangle = |n:+\rangle \otimes |a:+y\rangle \tag{13.276}$$

Expanding in terms of energy eigenstates, we get

$$\begin{aligned} |\psi(0)\rangle &= |n:+\rangle \otimes |a:+y\rangle \\ &= \frac{1}{2}|n:+\rangle \left((1+i)|a:+x\rangle + (1-i)|a:-x\rangle\right) \end{aligned} \tag{13.277}$$

and

$$|\psi(t)\rangle = \frac{1}{2}|n:+\rangle \otimes \left((1+i)|a:+x\rangle e^{-iAt/2} + (1-i)|a:-x\rangle e^{+iAt/2}\right)$$

so that

$$|\psi(\tau)\rangle = \frac{1}{2}|n:+\rangle \otimes \left((1+i)|a:+x\rangle e^{-iA\tau/2} + (1-i)|a:-x\rangle e^{+iA\tau/2}\right)$$

which for $A\tau = \pi/2$ gives

$$\begin{aligned} |\psi(\tau)\rangle &= \frac{1}{2}|n:+\rangle \otimes \left((1+i)|a:+x\rangle e^{-i\pi/4} + (1-i)|a:-x\rangle e^{+i\pi/4}\right) \\ &= \frac{1}{2}|n:+\rangle \otimes \left((1+i)|a:+x\rangle \frac{1}{\sqrt{2}}(1-i) + (1-i)|a:-x\rangle \frac{1}{\sqrt{2}}(1+i)\right) \\ &= \frac{1}{\sqrt{2}}|n:+\rangle \otimes (|a:+x\rangle + |a:-x\rangle) = |n:+\rangle \otimes |a:+\rangle \end{aligned} \tag{13.278}$$

Similarly, if $|\psi(0)\rangle = |n:-\rangle \otimes |a:-y\rangle$ then $|\psi(\tau)\rangle = |n:-\rangle \otimes |a:-\rangle$.

Physically, this means that the neutron's spin state does not change since it is an eigenstate of $\hat{V}$, while the atom's spin precesses around the $x-$axis with angular frequency A. At time $\tau = \pi/2A$ it lies along the $z-$axis.

We now suppose that the initial spin state is

$$|\psi(0)\rangle = (\alpha_+|n:+\rangle + \alpha_-|n:-\rangle) \otimes |a:+y\rangle \tag{13.279}$$

After the neutron-atom interaction described above, one measures the $z-$component $\hat{S}_{az}$ of the atom's spin. The state after the interaction is(using linearity)

$$|\psi(\tau)\rangle = \alpha_+|n:+\rangle \otimes |a:+\rangle + \alpha_-|n:-\rangle|a:-\rangle \tag{13.280}$$

The measurement of the z-component of the atoms spin gives $+\hbar/2$ with probability $|\alpha_+|^2$ and state $|n:+\rangle \otimes |a:+\rangle$ after the measurement or $-\hbar/2$ with probability $|\alpha_-|^2$ and state $|n:-\rangle \otimes |a:-\rangle$ after the measurement. In both cases, after measuring the $z-$component of the atom's spin, the neutron spin state is known - it is the same as that of the measured atom. It is not necessary to let the neutron interact with another measuring apparatus in order to know the value of $\hat{S}_{nz}$.

13.8.4 The Actual Quantum Eraser

We have seen above that if one measures the spin state of the atom between the two cavities, the interference signal disappears. We now want to show that it is possible to recover an interference if the information left by neutron on the detecting atom is *erased* by an appropriate measurement.

A neutron, initially in the spin state $-$, is sent into the two-cavity system. Immediately after the first cavity, there is a detecting atom of the type described above, prepared in the spin state $+y$. By assumption, the spin state of the atom evolves only during the time interval τ when it interacts with the neutron.

The successive states for the neutron are:

$$
\begin{aligned}
|\psi(t_1)\rangle &= |\text{after 1st cavity and before atom interaction}\rangle \\
&= \frac{1}{\sqrt{2}}\left(-ie^{-i\delta}\,|n:+\rangle \otimes |a:+y\rangle + e^{i\chi}\,|n:-\rangle \otimes |a:+y\rangle\right) \\
|\psi(t_1+\tau)\rangle &= |\text{just after atom interaction}\rangle \\
&= \frac{1}{\sqrt{2}}\left(-ie^{-i\delta}\,|n:+\rangle \otimes |a:+\rangle + e^{i\chi}\,|n:-\rangle \otimes |a:-\rangle\right) \\
|\psi(t_0')\rangle &= |\text{entering 2nd cavity}\rangle \\
&= \frac{1}{\sqrt{2}}\left(-ie^{-i(\delta+\omega_0 T/2)}\,|n:+\rangle \otimes |a:+\rangle + e^{i(\chi+\omega_0 T/2)}\,|n:-\rangle \otimes |a:-\rangle\right) \\
|\psi(t_1')\rangle &= |\text{after 2nd cavity}\rangle \\
&= \frac{1}{2}\left(\begin{array}{r} -ie^{-i(\delta+\omega_0 T/2)}\left(e^{-i\chi}\,|n:+\rangle - ie^{-i\delta'}\,|n:-\rangle\right) \otimes |a:+\rangle \\ +e^{i(\chi+\omega_0 T/2)}\left(-ie^{-i\delta'}\,|n:+\rangle + e^{i\chi}\,|n:-\rangle\right) \otimes |a:-\rangle \end{array}\right)
\end{aligned}
$$

The probability of finding the neutron in the state + at time t_1' (after the 2nd cavity) is the sum of the probabilities for finding

1. the neutron in state + and the atom in state +, i.e., the square modulus of the coefficient of $|n:+\rangle \otimes |a:+\rangle$, which = 1/4 in this case)

2. the neutron in state + and the atom in state $-$, which = 1/4 in this case also).

We therefore get $P_+ = 1/4+1/4 = 1/2$ - there is no interference since the quantum path leading in the end to a spin flip of the neutron can be determined from the state of the atom.

At time t_1', Bob measures the $z-$component of the neutron spin and Alice measures the $y-$component of the atom's spin. Assume that both measurements give $+\hbar/2$.

We can write

$$|\psi(t_1')\rangle = \frac{1}{2\sqrt{2}} \begin{pmatrix} -ie^{-i(\delta+\omega_0 T/2)} \left(e^{-i\chi}|n:+\rangle - ie^{-i\delta'}|n:-\rangle\right) \otimes (|a:+y\rangle + |a:-y\rangle) \\ +e^{i(\chi+\omega_0 T/2)} \left(-ie^{-i\delta'}|n:+\rangle + e^{i\chi}|n:-\rangle\right) \otimes (|a:+y\rangle - |a:-y\rangle) \end{pmatrix}$$

The probability amplitude that Bob finds $+\hbar/2$ along the $z-$axis while Alice finds $+\hbar/2$ along the $y-$axis is the coefficient of the term $|n:+\rangle \otimes |a:+y\rangle$ in the above state. Equivalently, the probability amplitude is found by projecting the state onto $|n:+\rangle \otimes |a:+y\rangle$ and squaring. We get

$$\begin{aligned} P\left(\hat{S}_{nz} = +\hbar/2, \hat{S}_{ay} = +\hbar/2\right) &= \frac{1}{8}\left|-ie^{-i(\delta+\chi+\omega_0 T/2)} - ie^{i(\chi-\delta'+\omega_0 T/2)}\right|^2 \\ &= \frac{1}{2}\cos^2\frac{(\omega-\omega_0)T}{2} \end{aligned} \tag{13.281}$$

which clearly exhibits a modulation reflecting an interference phenomenon. Similarly, one finds that

$$P\left(\hat{S}_{nz} = +\hbar/2, \hat{S}_{ay} = -\hbar/2\right) = \frac{1}{2}\sin^2\frac{(\omega-\omega_0)T}{2} \tag{13.282}$$

which is also modulated.

Let us now discuss the following three statements:

1. When Alice performs a measurement on the atom, Bob sees at once an interference appear in the signal he is measuring on the neutron.

2. Knowing the result obtained by Alice on each event, Bob can select a subsample of his own events which displays an interference phenomenon.

3. The experiment corresponds to an interference between two quantum paths for the neutron spin. By restoring the initial state of the atom, the measurement done by Alice erases the information concerning which quantum path is followed by the neutron spin and thus allows interference to reappear.

Statement (1) is clearly wrong. As seen earlier, if atom A is present, Bob no longer sees oscillations (in $\omega - \omega_0$) of the probability for detecting the neutron in the state +. This probability is equal to 1/2 whatever Alice does. Notice

that if the statement were correct, this would imply instantaneous transmission of information from Alice to Bob. By seeing interference reappear, Bob would know immediately that Alice is performing an experiment, even though she may be far away.

Statement (2) is correct. If Alice and Bob put together all their results, and if they select the subsample of events for which Alice finds $+\hbar/2$, then the number of events for which Bob also finds $+\hbar/2$ varies like

$$\cos^2 \frac{(\omega - \omega_0)T}{2} \tag{13.283}$$

Thus, they recover interference for this subset of events. In the complementary set where Alice found $-\hbar/2$, the number of Bob's results giving $+\hbar/2$ varies like

$$\sin^2 \frac{(\omega - \omega_0)T}{2} \tag{13.284}$$

This search for correlation between events occurring in different detectors is a common procedure in particle physics for example.

Statement (3), although less precise but more picturesque than than statement (2), is nevertheless acceptable. The

$$\cos^2 \frac{(\omega - \omega_0)T}{2} \tag{13.285}$$

signal found earlier can be interpreted as the interference of the amplitudes corresponding to two quantum paths for the neutron spin which is initially in the state $-$; either its spin flips in the 1^{st} cavity, or it flips in the 2^{nd} cavity. If there exists a possibility to determine which quantum path is followed by the system, interference cannot appear. It is necessary to *erase* this information, which is carried by the atom, in order to observe *some* interference. After Alice has measured the atom's spin along the $y-$axis, she has, in some sense *restored* the initial state of the system, and this enables Bob to see some interference. It is questionable to say that information has been erased - one may feel that, on the contrary, extra information has been acquired. Notice that the statement does not specify in which physical quantity the interference reappears. Notice also that the order of the measurements made by Alice and Bob has no importance, contrary to what this third statement seems to imply.

Chapter 14

Relativistic Wave Equations

Electromagnetic Radiation in Matter

14.1 Spin 0 particles: Klein-Gordon Equation

A classical nonrelativistic free particle has an energy-momentum relation $E = p^2/2m$. Under a Galilean transformation to a new coordinate system traveling with $-\vec{v}$ with respect to the first system, we have

$$\vec{r}\,' = \vec{r} + \vec{v}t \quad , \quad t' = t \tag{14.1}$$

which gives

$$\begin{aligned} \vec{u}\,' &= \frac{d\vec{r}\,'}{dt'} = \frac{d\vec{r}}{dt} + \vec{v} = \vec{u} + \vec{v} \\ m' &= m \\ \vec{p}\,' &= \vec{p} + m\vec{v} \end{aligned}$$

We then have

$$\vec{F}\,' = \frac{d\vec{p}\,'}{dt'} = \frac{d\vec{p}}{dt} = \vec{F} \tag{14.2}$$

which implies the invariance of the form $E = p^2/2m$, i.e.,

$$\begin{aligned} E' &= \int \vec{F}\,' \cdot d\vec{r}\,' = \frac{1}{m'} \int \vec{p}\,' \cdot d\vec{p}\,' = \frac{\vec{p}\,'^2}{2m'} \\ E &= \int \vec{F} \cdot d\vec{r} = \frac{1}{m} \int \vec{p} \cdot d\vec{p} = \frac{\vec{p}^2}{2m} \end{aligned}$$

and

$$E' = \frac{\vec{p}\,'^2}{2m'} = \frac{(\vec{p} + m\vec{v})^2}{2m} = E + \vec{p} \cdot \vec{v} + \frac{1}{2}mv^2 \tag{14.3}$$

Thus, the final Galilean transformation relations are

$$E' = E + \vec{p} \cdot \vec{v} + \frac{1}{2}mv^2 \quad , \quad \vec{p}\,' = \vec{p} + m\vec{v} \tag{14.4}$$

which, as shown, leaves the quadratic form $E = p^2/2m$ invariant, i.e., if $E = \vec{p}^2/2m$, then $E' = \vec{p}'^2/2m'$ and we derive the transformation rules for E and $\vec{p}$ from that condition.

The non-relativistic Schrodinger equation for the free particle then follows from the standard identifications

$$E \to i\hbar\frac{\partial}{\partial t} \quad , \quad \vec{p} \to \frac{\hbar}{i}\nabla$$
$$H\psi = E\psi \quad , \quad H = \frac{p^2}{2m}$$

which gives

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi \tag{14.5}$$

It is clear from the form of the Schrodinger equation for a free particle that the equation cannot be invariant under Lorentz transformation (Lorentz covariant), i.e., the time derivative is first order and the space derivatives are second order.

14.1.1 How to find correct form of relativistic wave equation?

Before proceeding let us recall some results from special relativity. Components of spacetime four-vectors will be labeled by Greek indices and the components of spatial three-vectors will be labeled by Latin indices and we will use Einstein summation convention.

Starting from $x^\mu(s) = (ct, \vec{x}) = (x^0, \vec{x})$, the contravariant 4-vector representation of the worldline as a function of the proper time s, we first obtain the 4-velocity, i.e.,

$$\dot{x}^\mu(s) = \frac{dx^\mu(s)}{ds} = \frac{dx^\mu(s)}{\frac{1}{\gamma}dx^0} = \gamma\left(\frac{dx^0}{dx^0}, \frac{d\vec{x}}{dx^0}\right) = \gamma\,(1, \vec{v}/c) \tag{14.6}$$

where

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad , \quad \vec{v} = \frac{d\vec{x}}{dt} = c\frac{d\vec{x}}{dx^0} \tag{14.7}$$

The 4-momentum vector is then given by

$$p^\mu = mc\dot{x}^\mu(s) = \gamma m\,(c, \vec{v}) = (E/c, \vec{p}) \tag{14.8}$$

where we have used the fact that

$$p^0 = \frac{E}{c} = \gamma mc \quad , \quad m = \text{ rest mass} \tag{14.9}$$

This says that the energy E and momentum $\vec{p}$ transform as the components of a contravariant 4-vector and we know that the *square* of any 4-vector is invariant.

The metric tensor defined by

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{14.10}$$

allows the construction of the covariant components(using Einstein summation convention)

$$p_\mu = g_{\mu\nu} p^\nu = \left(\frac{E}{c}, -\vec{p} \right) \tag{14.11}$$

We can then calculate the invariant *square* or the invariant scalar product of the 4-momentum with itself as

$$p_\mu p^\mu = g_{\mu\nu} p^\nu p^\mu = \frac{E^2}{c^2} - \vec{p}^2 = m^2 c^2 \tag{14.12}$$

We therefore have the relativistic energy momentum relation

$$E = \left(\vec{p}^2 c^2 + m^2 c^4 \right)^{1/2} \tag{14.13}$$

If we use this expression to construct a new wave equation by operator substitution we would have

$$i\hbar \frac{\partial \psi}{\partial t} = \left(-\hbar^2 c^2 \nabla^2 + m^2 c^4 \right)^{1/2} \psi \tag{14.14}$$

Although the energy formula is now relativistically correct, the time and space derivatives still do not appear symmetrically. In fact, a Taylor expansion of the square root gives infinitely high-order derivatives leading to a very difficult mathematical equation to deal with.

This fact, in itself, is not a valid reason for rejecting the equation.

There are, however, strong physical reasons for rejecting this equation. The equation says that the momentum space amplitude

$$\psi_{\vec{p}}(t) = \int d^3 r \, e^{-i\vec{p}\cdot\vec{r}/\hbar} \psi(\vec{r}, t) \tag{14.15}$$

obeys the equation

$$i\hbar \frac{\partial \psi_{\vec{p}}(t)}{\partial t} = \left(p^2 c^2 + m^2 c^4 \right)^{1/2} \psi_{\vec{p}}(t) \tag{14.16}$$

If we Fourier transform both sides back to position space we get

$$i\hbar \frac{\partial \psi(\vec{r}, t)}{\partial t} = \int d^3 r' \, K(\vec{r} - \vec{r}\,') \psi(\vec{r}\,', t) \tag{14.17}$$

where

$$K(\vec{r}-\vec{r}\,') = \int \frac{d^3p}{(2\pi\hbar)^3} e^{i\vec{p}\cdot(\vec{r}-\vec{r}\,')/h} \left(p^2c^2+m^2c^4\right)^{1/2} \tag{14.18}$$

This equation for $\psi(\vec{r},t)$ is *nonlocal*, which means that the value of the integral at *vecr* depends on the value of ψ at the other points *vecr′*. The function $K(\vec{r}-\vec{r}\,')$ is large as long as *vecr′* is within a distance

$$\approx \frac{\hbar}{mc} = \text{ Compton wavelength} \tag{14.19}$$

from *vecr*. As a consequence of the nonlocality, the rate of change in time of ψ at the spacetime point $(\vec{r},t)$ depends on the values of ψ at points $(\vec{r}\,',t)$ *outside the light cone* centered on $(\vec{r},t)$.

If we construct a wave packet localized well within a Compton wavelength of the origin, then the packet will be nonzero an arbitrarily short time later at points as distant as the Compton wavelength.

Thus, this equation leads to violations of relativistic causality when used to describe particles localized to within more than a Compton wavelength, which is unacceptable.

Instead, we will start from an equation involving E^2, i.e., we have

$$E^2\psi = \left(i\hbar\frac{\partial}{\partial t}\right)^2\psi = \left(\vec{p}^2c^2+m^2c^4\right)\psi = \left(-\hbar^2c^2\nabla^2+m^2c^4\right)\psi$$
$$\left(\nabla^2-\left(\frac{mc}{\hbar}\right)^2\right)\psi = \frac{1}{c^2}\frac{\partial^2\psi}{\partial^2 t} \tag{14.20}$$

which looks like a classical wave equation with an extra term of the form

$$\left(\frac{mc}{\hbar}\right)^2 \tag{14.21}$$

It is called the *Klein-Gordon* equation. In 4-vector notation it looks like

$$\left(\partial_\mu\partial^\mu+\left(\frac{mc}{\hbar}\right)^2\right)\psi = 0 \quad , \quad \partial_\mu = \frac{\partial}{\partial x^\mu} \tag{14.22}$$

The equation can be generalized in a relativistically invariant way to include the coupling of charged particles to the electromagnetic field by the substitutions (corresponding to minimal coupling we discussed earlier in Chapter 8)

$$i\hbar\frac{\partial}{\partial t} \to i\hbar\frac{\partial}{\partial t} - e\Phi \quad , \quad \frac{\hbar}{i}\nabla \to \frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A} \tag{14.23}$$

to obtain

$$\frac{1}{c^2}\left(i\hbar\frac{\partial}{\partial t} - e\Phi(\vec{r},t)\right)^2\psi(\vec{r},t) = \left(\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}(\vec{r},t)\right)^2 + m^2c^2\right)\psi(\vec{r},t) \tag{14.24}$$

The Klein-Gordon equation has several unusual features.

First, it is second-order in time (space and time derivatives are now the same order). This means we need to specify twice as much initial information (the function and its derivative) at one time to specify the relativistic solution as compared to the nonrelativistic solution which only required specification of the function at one time. .

This will mean that the equation has an extra degree of freedom. We will see shortly that this extra degree of freedom corresponds to specifying the *charge* of the particle and that the Klein-Gordon equation actually describes both a particle and its *antiparticle* together. .

Second, since the equation is second order in time, the functions

$$\psi = e^{i(\vec{p}\cdot\vec{r}-Et)/\hbar} \tag{14.25}$$

satisfy the free particle equation with either sign of E, i.e.,

$$E = \pm\left(\vec{p}^2c^2 + m^2c^4\right)^{1/2} \tag{14.26}$$

The Klein-Gordon equation has negative energy solutions for a free particle! For these solutions when we *increase* the magnitude of the momentum $\vec{p}$, then the energy of the particle *decreases*! As we will see later, these negative energy solutions are real and will correspond to antiparticles, while the positive energy solutions will be particles.

In nonrelativistic Schrodinger theory we were able to interpret $\psi^*\psi$ as a positive probability density that was conserved in time (no sinks or sources of probability in nonrelativistic Schrodinger theory). Let us see what happens in the case of the Klein-Gordon equation.

For the Klein-Gordon equation

$$\int \psi^*\psi d^3r \tag{14.27}$$

changes in time and thus, we cannot interpret $\psi^*(\vec{r},t)\psi(\vec{r},t)$ as being the probability of finding a particle at $\vec{r}$ at time t.

We can, however, construct a different conserved density as follows. We write

$$\psi^*\left(\partial_\mu\partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\psi = 0 \tag{14.28}$$

and

$$\psi\left(\partial_\mu\partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\psi^* = 0 \tag{14.29}$$

which give(subtracting)

$$\psi^* \partial_\mu \partial^\mu \psi - \psi \partial_\mu \partial^\mu \psi^* = 0 \tag{14.30}$$

$$\partial_\mu \left(\psi^* \partial^\mu \psi - \psi \partial^\mu \psi^* \right) = 0 \tag{14.31}$$

Expanding these expressions, we have the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = 0 \tag{14.32}$$

where

$$\rho(\vec{r}, t) = \frac{i\hbar}{2mc^2} \left(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t} \right) \tag{14.33}$$

$$\vec{j}(\vec{r}, t) = \frac{\hbar}{im} \left(\psi^* \nabla \psi - \psi \nabla \psi^* \right) \tag{14.34}$$

We have inserted a multiplicative constant so that the current density vector $\vec{j}(\vec{r}, t)$ is identical to the nonrelativistic case. Because this density $\rho(\vec{r}, t)$ satisfies a continuity equation, its integral over all space does not change in time. Clearly, however, it is not necessarily positive. In particular, $\rho < 0$ for a negative energy free particle eigenstate.

This means that we cannot interpret this new $\rho(\vec{r})$ as being the particle (probability) density at $\vec{r}$ and we cannot interpret $\vec{j}(\vec{r})$ as a particle current.

The interpretation that will eventually emerge is that for charged particles $e\rho(\vec{r})$ represents the charge density at $\vec{r}$, which can have either sign and $e\vec{j}(\vec{r})$ represents the electric current at $\vec{r}$.

14.1.2 Negative Energy States and Antiparticles

How do we interpret the Klein-Gordon equation and its solutions?

Consider a free particle at rest, i.e., $\vec{p} = 0$. The wave function for the positive energy solution is

$$\psi(\vec{r}, t) = e^{-imc^2 t/\hbar} \tag{14.35}$$

where the energy of a particle at rest is $E = mc^2$. The density for this state is $\rho(\vec{r}, t) = +1$.

Now make a Lorentz transformation to a new frame moving with velocity $-\vec{v}$ with respect to the particle at rest. The particle now appears to have a velocity $\vec{v}$ in this new frame. It, therefore, has

$$\text{momentum } = \vec{p} = \gamma m \vec{v} \text{ and energy } = E = \gamma mc^2 \tag{14.36}$$

where

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \tag{14.37}$$

This result follows because

$$\begin{aligned} p_\mu x'^\mu &= \text{ Lorentz scalar} \\ &= -mc^2 t \text{ in the rest frame} \\ &= \vec{p}\cdot\vec{r}\,' - Et' \text{ in the moving frame} \end{aligned}$$

The new wave function

$$\psi(\vec{r}\,', t') = e^{i(\vec{p}\cdot\vec{r}\,' - Et')/\hbar} = e^{i(\vec{p}\cdot\vec{r}\,' - E_{\vec{p}}t')/\hbar} \tag{14.38}$$

is the result we expect for a particle of momentum $\vec{p}$ and energy $E_{\vec{p}}$. If we calculate the density ρ for this wave function we get

$$\rho(\vec{r}\,', t') = \frac{E_{\vec{p}}}{mc^2} \tag{14.39}$$

and the current is

$$\vec{j}(\vec{r}\,', t') = \frac{\vec{p}}{m} = \frac{\vec{p}c^2}{E_{\vec{p}}}\rho(\vec{r}\,', t') = \vec{v}\rho(\vec{r}\,', t') \text{ (as expected)} \tag{14.40}$$

where

$$\vec{v} = \frac{\vec{p}c^2}{E_{\vec{p}}} \tag{14.41}$$

We see that $\rho(\vec{r}, t)$ transforms like $E_{\vec{p}}$ or as the time component of a 4-vector, which makes physical sense. Since a unit volume in the rest frame appears smaller by a factor

$$1/\gamma = \sqrt{1 - v^2/c^2} \tag{14.42}$$

when observed from the moving frame, a unit density in the rest frame will appear as a density

$$\frac{1}{\gamma} = \frac{E_{\vec{p}}}{mc^2} \tag{14.43}$$

in a frame in which the particle is moving.

What about the negative energy solutions? For a particle at rest we have, in this case,

$$\psi(\vec{r}, t) = e^{imc^2 t/\hbar} \tag{14.44}$$

where the energy of this particle at rest is $E = -mc^2$.

The density for this state is $\rho(\vec{r}, t) = -1$.

It turns out that one way to interpret a state with a negative particle density is to say that it is a state with a positive density of antiparticles.

We will make the interpretation that a particle at rest with energy $E = -mc^2$ is actually an antiparticle with positive energy $E = mc^2$. As we shall see,

this interpretation of negative energy states will lead to a consistent theoretical picture that is confirmed experimentally.

In a Lorentz frame traveling with velocity $-\vec{v}$ with respect to the antiparticle, the wave function is

$$\psi(\vec{r}\,', t') = e^{-i(\vec{p}\cdot\vec{r}\,' - Et')/\hbar} = e^{imc^2 t/\hbar} \tag{14.45}$$

where

$$\text{momentum } = \vec{p} = \gamma m\vec{v} \text{ and energy } = E = \gamma mc^2 \tag{14.46}$$

In this new frame the particle has velocity $\vec{v}$, momentum $\vec{p}$ and energy $E_{\vec{p}}$. The wave function, however, describes a particle of energy $-E_{\vec{p}}$ and momentum $-\vec{p}$.

The density in the moving frame is

$$\rho(\vec{r}\,', t') = -\frac{E_{\vec{p}}}{mc^2} \tag{14.47}$$

and the current is

$$\vec{j}(\vec{r}\,', t') = -\frac{\vec{p}}{m}\rho(\vec{r}\,', t') = \frac{\vec{p}c^2}{E_{\vec{p}}}\rho(\vec{r}\,', t') \tag{14.48}$$

Thus, an antiparticle moving with velocity $\vec{v}$ has associated with it a current moving in the opposite direction, i.e., a flow of antiparticles in one direction is equivalent to a flow of particles in the opposite direction.

For a charged particle $e\rho(\vec{r}, t)$ is the charge density. It is positive for a free particle with $e > 0$ and negative for a free antiparticle, which has opposite charge to the particle.

The quantity $e\vec{j}(\vec{r}, t)$ is the electric current of the state ψ. For a particle the electric current is in the direction of the particle velocity. For the antiparticle with $e < 0$, the electric current is opposite to the velocity.

This says that the interpretation of the negative energy solutions as antiparticles is consistent with the interpretation of the density ρ as a charge density and $\vec{j}$ as an electric current.

Is this interpretation consistent with the way charged particles interact with the electromagnetic field?

The Klein-Gordon equation with an electromagnetic field present is given by

$$\frac{1}{c^2}\left(i\hbar\frac{\partial}{\partial t} - e\Phi(\vec{r}, t)\right)^2 \psi(\vec{r}, t) = \left(\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}(\vec{r}, t)\right)^2 + m^2c^2\right)\psi(\vec{r}, t) \tag{14.49}$$

Taking the complex conjugate we have

$$\frac{1}{c^2}\left(i\hbar\frac{\partial}{\partial t}+e\Phi(\vec{r},t)\right)^2\psi^*(\vec{r},t)=\left(\left(\frac{\hbar}{i}\nabla+\frac{e}{c}\vec{A}(\vec{r},t)\right)^2+m^2c^2\right)\psi^*(\vec{r},t) \quad (14.50)$$

These equations say that if $\psi(\vec{r},t)$ is a solution to the Klein-Gordon equation with a certain sign of the charge, then $\psi^*(\vec{r},t)$ is a solution of the Klein-Gordon equation with the *opposite* sign of the charge and the *same* mass.

Thus, the relativistic theory of a spin zero particle predicts the existence of its antiparticle with the opposite charge and same mass, i.e., the theory contains solutions for both particles and antiparticles.

Relativistic invariance requires
the existence of antiparticles

The complex conjugate of a negative energy solution is a positive energy solution with the opposite sign of the charge. The operation of taking the complex conjugate of the wave function will be called *charge conjugation.* Charge conjugation changes particles into antiparticles and vice versa. If we label quantities calculated with the complex conjugate wave function by a subscript **c** , we find

$$\rho(\vec{r},t)=-\rho_c(\vec{r},t) \quad , \quad \vec{j}(\vec{r},t)=-\vec{j}_c(\vec{r},t) \quad (14.51)$$

as expected.

The solutions are normalized by the requirement that the total associated charge equals ±1 unit, i.e.,

$$\int d^3r\rho(\vec{r},t)=+1=-\int d^3r\rho_c(\vec{r},t) \quad (14.52)$$

This normalization is conserved in time and invariant under a Lorentz transformation.

14.2 Physics of the Klein-Gordon Equation

We first transform the Klein-Gordon equation into two equations, each first order in time.

We define

$$\psi^0(\vec{r},t)=\left(\frac{\partial}{\partial t}+\frac{ie}{\hbar}\Phi(\vec{r},t)\right)\psi(\vec{r},t) \qquad (1^{st}\text{ - order equation \#1}) \quad (14.53)$$

We then have

$$\left(\frac{\partial}{\partial t}+\frac{ie}{\hbar}\Phi(\vec{r},t)\right)\psi^0(\vec{r},t)=\left(\frac{\partial}{\partial t}+\frac{ie}{\hbar}\Phi(\vec{r},t)\right)^2\psi(\vec{r},t) \quad (14.54)$$

Now using the Klein-Gordon equation we have

$$\left(\frac{\partial}{\partial t}+\frac{ie}{\hbar}\Phi(\vec{r},t)\right)^2\psi(\vec{r},t)=c^2\left(\left(\nabla+\frac{ie}{\hbar c}\vec{A}(\vec{r},t)\right)^2-\frac{m^2c^4}{\hbar^2}\right)\psi(\vec{r},t) \qquad (14.55)$$

so we get

$$\left(\frac{\partial}{\partial t}+\frac{ie}{\hbar}\Phi(\vec{r},t)\right)\psi^0(\vec{r},t)$$
$$=c^2\left(\left(\nabla+\frac{ie}{\hbar c}\vec{A}(\vec{r},t)\right)^2-\frac{m^2c^4}{\hbar^2}\right)\psi(\vec{r},t) \qquad (1^{st}\text{ - order equation \#2}) \qquad (14.56)$$

These two new first-order equations involve the two functions $\psi^0(\vec{r},t)$ and $\psi(\vec{r},t)$.

Now define the linear combinations

$$\phi=\frac{1}{2}\left[\psi+\frac{i\hbar}{mc^2}\psi^0\right] \quad , \quad \chi=\frac{1}{2}\left[\psi-\frac{i\hbar}{mc^2}\psi^0\right] \qquad (14.57)$$

Substitution then gives the more symmetric equations

$$\left(i\hbar\frac{\partial}{\partial t}-e\Phi\right)\phi=\frac{1}{2m}\left[\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right]^2(\phi+\chi)+mc^2\phi \qquad (14.58)$$

$$\left(i\hbar\frac{\partial}{\partial t}-e\Phi\right)\chi=-\frac{1}{2m}\left[\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right]^2(\phi+\chi)+mc^2\chi \qquad (14.59)$$

Now define a two-component wave function

$$\Psi(\vec{r},t)=\begin{pmatrix}\phi(\vec{r},t)\\ \chi(\vec{r},t)\end{pmatrix} \qquad (14.60)$$

and three 2×2 matrices

$$\tau_1=\begin{pmatrix}0&1\\1&0\end{pmatrix} \quad , \quad \tau_2=\begin{pmatrix}0&-i\\i&0\end{pmatrix} \quad , \quad \tau_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix} \qquad (14.61)$$

The two symmetric equations can then be combined into the single equation

$$i\hbar\frac{\partial\Psi}{\partial t}=\left[\left[\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right]^2(\tau_3+i\tau_2)+mc^2\tau_3+e\Phi\right]\Psi \qquad (14.62)$$

This equation is completely equivalent to the original Klein-Gordon equation where

$$\psi=\phi+\chi \quad , \quad \psi^0=\frac{mc^2}{i\hbar}(\phi-\chi) \qquad (14.63)$$

The internal degree of freedom represented by these two components is the charge of the particle(one component represents the particle and the other the

antiparticle).

Using the two component equation, we can write the density as

$$\rho(\vec{r},t) = |\phi|^2 - |\chi|^2 = \Psi^+\tau_3\Psi \tag{14.64}$$

where $\Psi^+ = (\phi^*, \chi^*)$. A very simple expression. The current density, however, becomes less transparent

$$\begin{aligned}\vec{j}(\vec{r},t) = \frac{\hbar}{2im}\left[\Psi^+\tau_3(\tau_3 + i\tau_2)\nabla\Psi - (\nabla\Psi^+)\tau_3(\tau_3 + i\tau_2)\Psi\right] \\ - \frac{e\vec{A}}{mc}\Psi^+\tau_3(\tau_3 + i\tau_2)\Psi\end{aligned} \tag{14.65}$$

The normalization condition becomes

$$\int d^3r\, \Psi^+\tau_3\Psi = \pm 1 \tag{14.66}$$

The scalar product between two such wave functions Ψ and $|Psi'$ is defined by

$$\langle \Psi \mid \Psi' \rangle = \int d^3r\, \Psi^+(\vec{r},t)\tau_3\Psi'(\vec{r},t) \tag{14.67}$$

Finally, the wave equation is of the form

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi \tag{14.68}$$

where the Hamiltonian is

$$\hat{H} = \left[\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right]^2(\tau_3 + i\tau_2) + mc^2\tau_3 + e\Phi \tag{14.69}$$

Since $(\tau_3 + i\tau_2)^+ = \tau_3 - i\tau_2$, we find that $\hat{H}^+ \neq \hat{H}$, which seems to indicate that $\hat{H}$ is not Hermitian. It is Hermitian, however, when we use the proper scalar product definition of hermiticity, i.e.,

$$\langle\Psi|\left(\hat{H}\,|\Psi'\rangle\right) = \left[\langle\Psi'|\left(\hat{H}\,|\Psi\rangle\right)\right]^* \tag{14.70}$$

This relation requires that

$$\tau_3\hat{H}^+\tau_3 = \hat{H} \tag{14.71}$$

All required properties of Hermitian operators, i.e., real eigenvalues and so on, follow from the scalar product definition so that is all that is actually required.

Under the charge conjugation operation

$$\left.\begin{array}{l}\phi \to \chi^* \\ \chi \to \phi^*\end{array}\right\} \to \Psi_c = \tau_1\Psi^* \tag{14.72}$$

which is the form of the charge conjugation operation in two-component language.

We can now see the physical meaning of charge conjugation. Using

$$\vec{p}^{\,*} = \left(\frac{\hbar}{i}\nabla\right)^{*} = -\frac{\hbar}{i}\nabla = -\vec{p} \tag{14.73}$$

and $\tau_3 + i\tau_2$ = real matrix, we find that

$$\hat{H}^{*}(e) = \left[-\vec{p} - \frac{e}{c}\vec{A}\right]^2 (\tau_3 + i\tau_2) + mc^2\tau_3 + e\Phi \tag{14.74}$$

We then have

$$\tau_1\hat{H}^{*}(e)\tau_1 = -\left[\vec{p} + \frac{e}{c}\vec{A}\right]^2 (\tau_3 + i\tau_2) - mc^2\tau_3 + e\Phi = -\hat{H}(-e) \tag{14.75}$$

This means that, if Ψ solves the equation

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}(e)\Psi \tag{14.76}$$

we have

$$-i\hbar\frac{\partial\Psi^{*}}{\partial t} = \hat{H}^{*}(e)\Psi^{*} = -\tau_1\hat{H}(-e)\tau_1\Psi^{*} \tag{14.77}$$

Multiplying by τ_1 we get

$$i\hbar\frac{\partial\Psi_c}{\partial t} = \hat{H}(-e)\Psi_c \tag{14.78}$$

which is the two-component statement of the fact that Ψ_c solves the Klein-Gordon equation with the opposite sign of the charge.

What can we say about the two-component solutions for free particles and antiparticles?

The wave function of a free particle (positive energy solution) of momentum $\vec{p}$ (normalized to 1) is given by

$$\psi^{(+)}_{\vec{p}}(\vec{r},t) = \sqrt{\frac{mc^2}{E_{\vec{p}}}}\,e^{i(\vec{p}\cdot\vec{r} - E_{\vec{p}}t)/\hbar} \tag{14.79}$$

where

$$E_{\vec{p}} = \sqrt{p^2c^2 + m^2c^4} \tag{14.80}$$

Using

$$\psi^{0}(\vec{r},t) = \left(\frac{\partial}{\partial t} + \frac{ie}{\hbar}\Phi(\vec{r},t)\right)\psi(\vec{r},t) \tag{14.81}$$

and

$$\phi = \frac{1}{2}\left[\psi + \frac{i\hbar}{mc^2}\psi^{0}\right] \quad , \quad \chi = \frac{1}{2}\left[\psi - \frac{i\hbar}{mc^2}\psi^{0}\right] \tag{14.82}$$

we find (in two-component language)

$$\Psi_{\vec{p}}^{(+)}(\vec{r},t) = \Psi_{\vec{p}}^{(+)} e^{i(\vec{p}\cdot\vec{r}-E_{\vec{p}}t)/\hbar} \tag{14.83}$$

where the two-component vector $\Psi_{\vec{p}}^{(+)}$ is given by

$$\Psi_{\vec{p}}^{(+)} = \frac{1}{2\sqrt{E_{\vec{p}}mc^2}} \begin{pmatrix} mc^2 + E_{\vec{p}} \\ mc^2 - E_{\vec{p}} \end{pmatrix} \tag{14.84}$$

In a similar manner, we can write for the negative energy solutions (free antiparticles)

$$\psi_{\vec{p}}^{(-)}(\vec{r},t) = \sqrt{\frac{mc^2}{E_{\vec{p}}}} e^{-i(\vec{p}\cdot\vec{r}-E_{\vec{p}}t)/\hbar} \tag{14.85}$$

$$\Psi_{\vec{p}}^{(-)}(\vec{r},t) = \Psi_{\vec{p}}^{(-)} e^{-i(\vec{p}\cdot\vec{r}-E_{\vec{p}}t)/\hbar} \tag{14.86}$$

$$\Psi_{\vec{p}}^{(-)} = \frac{1}{2\sqrt{E_{\vec{p}}mc^2}} \begin{pmatrix} mc^2 - E_{\vec{p}} \\ mc^2 + E_{\vec{p}} \end{pmatrix} = \tau_1 \Psi_{\vec{p}}^{(+)} \tag{14.87}$$

We note that in the nonrelativistic limit

$$E_{\vec{p}} = \sqrt{p^2c^2 + m^2c^4} = mc^2 \left(1 + \frac{p^2}{m^2c^2}\right)^{1/2} \approx mc^2 \left(1 + \frac{p^2}{2m^2c^2}\right) \tag{14.88}$$

$$mc^2 \pm E_{\vec{p}} \approx \begin{cases} 2mc^2 \\ -p^2/2m \end{cases} \tag{14.89}$$

so that

$$\Psi_{\vec{p}}^{(+)} = \begin{pmatrix} 1 \\ -v^2/4c^2 \end{pmatrix} \quad , \quad \Psi_{\vec{p}}^{(-)} = \begin{pmatrix} -v^2/4c^2 \\ 1 \end{pmatrix} \tag{14.90}$$

This shows that in the nonrelativistic limit

$$\chi \text{ is } \approx v^2/c^2 \text{ times } \phi \text{ for a particle} \tag{14.91}$$

If we drop χ, then ϕ satisfies the nonrelativistic Schrodinger equation with the constant mc^2 included in the energy.

Similarly, dropping ϕ in the antiparticle solution shows that χ satisfies the nonrelativistic Schrodinger equation for the opposite charge with the constant mc^2 included in the energy.

The particle and antiparticle solutions are orthogonal in the sense that

$$\Psi_{\vec{p}}^{(+)} \tau_3 \Psi_{\vec{p}}^{(-)} = 0 = \Psi_{\vec{p}}^{(-)} \tau_3 \Psi_{\vec{p}}^{(+)} \tag{14.92}$$

which should be the case since they represent different energy eigenstates of the same Hamiltonian.

The free particle solutions form a complete set since any wave function Ψ can be expanded as a linear combination of the free particle and antiparticle solutions.

We first write Ψ as a Fourier transform

$$\Psi(\vec{r},t) = \int \frac{d^3p}{(2\pi\hbar)^3} e^{i\vec{p}\cdot\vec{r}/\hbar} \begin{pmatrix} \phi_{\vec{p}} \\ \chi_{\vec{p}} \end{pmatrix} \tag{14.93}$$

Since the two vectors $\Psi^{(+)}_{\vec{p}}$ and $\Psi^{(-)}_{\vec{p}}$ are linearly independent, we can write

$$\Psi_{\vec{p}}(t) = \begin{pmatrix} \phi_{\vec{p}} \\ \chi_{\vec{p}} \end{pmatrix} = u_{\vec{p}}(t)\Psi^{(+)}_{\vec{p}} + v^*_{-\vec{p}}(t)\Psi^{(-)}_{-\vec{p}} \tag{14.94}$$

Good reasons for the choice of $-\vec{p}$ and $*$ will appear shortly.

Substituting we get

$$\begin{aligned} \Psi(\vec{r},t) &= \int \frac{d^3p}{(2\pi\hbar)^3} e^{i\vec{p}\cdot\vec{r}/\hbar} \left[u_{\vec{p}}(t)\Psi^{(+)}_{\vec{p}} + v^*_{-\vec{p}}(t)\Psi^{(-)}_{-\vec{p}} \right] \\ &= \int \frac{d^3p}{(2\pi\hbar)^3} \left[u_{\vec{p}}(t)\Psi^{(+)}_{\vec{p}} e^{i\vec{p}\cdot\vec{r}/\hbar} + v^*_{\vec{p}}(t)\Psi^{(-)}_{\vec{p}} e^{-i\vec{p}\cdot\vec{r}/\hbar} \right] \end{aligned} \tag{14.95}$$

where a change of variables was made in the second term. From the form of this result, $u_{\vec{p}}(t)$ is the amplitude for a particle in the state Ψ to have momentum $\vec{p}$ and positive charge and $v_{\vec{p}}(t)$ is the amplitude for a particle in the state Ψ to have momentum $\vec{p}$ and negative charge.

Using the orthonormality of $\Psi^{(\pm)}_{\vec{p}}$ we get

$$u_{\vec{p}}(t) = \int d^3r \Psi^{(+)+}_{\vec{p}} e^{-i\vec{p}\cdot\vec{r}/\hbar} \tau_3 \Psi(\vec{r},t) \tag{14.96}$$

$$v^*_{\vec{p}}(t) = -\int d^3r \Psi^{(-)+}_{\vec{p}} e^{i\vec{p}\cdot\vec{r}/\hbar} \tau_3 \Psi(\vec{r},t) \tag{14.97}$$

The normalization integral for Ψ then becomes

$$\int \frac{d^3p}{(2\pi\hbar)^3} \left(|u_{\vec{p}}|^2 - |v_{\vec{p}}|^2 \right) = \pm 1 \tag{14.98}$$

This says that there is no restriction on the magnitude of either $u_{\vec{p}}$ or $v_{\vec{p}}$. Only the integral of the difference(above) is fixed.

Physically, we can then say that one can have a state with an arbitrarily large amplitude for finding a particle with a certain momentum, which is the first indication that we are dealing with *bosons* or that spin zero particles must be bosons.

We can write some expectation values in this formalism, i.e.,

$$\hat{H}_0 = \frac{p^2}{2m}(\tau_3 + i\tau_2)^2 + mc^2\tau_3 = \text{Kineticenergy} \tag{14.99}$$

$$\int \Psi^+(\vec{r})\tau_3\hat{H}_0\Psi(\vec{r})d^3r = \int \frac{d^3p}{(2\pi\hbar)^3}E_{\vec{p}}\left(|u_{\vec{p}}|^2 + |v_{\vec{p}}|^2\right) \tag{14.100}$$

and

$$\vec{p} = \frac{\hbar}{i}\nabla = \text{ momentum} \tag{14.101}$$

$$\int \Psi^+(\vec{r})\tau_3\left(\frac{\hbar}{i}\nabla\right)\Psi(\vec{r})d^3r = \int \frac{d^3p}{(2\pi\hbar)^3}\vec{p}\left(|u_{\vec{p}}|^2 + |v_{\vec{p}}|^2\right) \tag{14.102}$$

14.3 Free Particles as Wave Packets

A wave packet formed from the positive energy solutions is given by

$$\Psi^{(+)}(\vec{r},t) = \int \frac{d^3p}{(2\pi\hbar)^3}\, u_{\vec{p}}e^{i(\vec{p}\cdot\vec{r}-E_{\vec{p}}t)/\hbar}\Psi^{(+)}_{\vec{p}} \tag{14.103}$$

Let us assume that $u_{\vec{p}}$ is peaked about $\vec{p} = \vec{p}\,'$. Then, using arguments similar to our earlier discussions on stationary phase, the center of the wave packet moves with a group velocity

$$\vec{v}_g = (\nabla_{\vec{p}}E_{\vec{p}})_{\vec{p}=\vec{p}\,'} = \frac{\vec{p}\,'c^2}{E_{\vec{p}\,'}} \tag{14.104}$$

and similarly for a free wave packet made of the negative energy solutions for antiparticles.

Can we construct a free particle wave packet perfectly localized at the origin? It would have the form

$$\Psi(\vec{r}) = \begin{pmatrix} a \\ b \end{pmatrix}\delta(\vec{r}) \tag{14.105}$$

We then have

$$\begin{aligned} u_{\vec{p}} &= \int d^3r\Psi^{(+)+}_{\vec{p}}e^{-i\vec{p}\cdot\vec{r}/\hbar}\tau_3\Psi(\vec{r}) \\ &= \int d^3r\Psi^{(+)+}_{\vec{p}}e^{-i\vec{p}\cdot\vec{r}/\hbar}\tau_3\begin{pmatrix} a \\ b \end{pmatrix}\delta(\vec{r}) = \Psi^{(+)+}_{\vec{p}}\tau_3\begin{pmatrix} a \\ b \end{pmatrix} \\ &= \frac{1}{2\sqrt{E_{\vec{p}}mc^2}}\begin{pmatrix} mc^2 + E_{\vec{p}} \\ mc^2 - E_{\vec{p}} \end{pmatrix}^+\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \\ &= \frac{E_{\vec{p}}(a+b) + mc^2(a-b)}{2\sqrt{E_{\vec{p}}mc^2}} \end{aligned} \tag{14.106}$$

and similarly

$$v^*_{\vec{p}} = -\Psi^{(-)+}_{\vec{p}}\tau_3\begin{pmatrix} a \\ b \end{pmatrix} = \frac{E_{\vec{p}}(a+b) - mc^2(a-b)}{2\sqrt{E_{\vec{p}}mc^2}} \tag{14.107}$$

Looking at these results we can see that independent of the choice of a and b, the wave packet will always have both particle and antiparticle components. This means that it is impossible to construct a perfectly localized wave packet from positive energy solutions alone.

Suppose that we take a general wave packet made up of positive energy solutions and try to squeeze it(make it more localized) with real-world devices such as collimators. To see what might happen we multiply the wave packet by the position operator $\vec{r}$. We then have

$$\begin{aligned} \vec{r}\Psi^{(+)}(\vec{r},t) &= \int \frac{d^3p}{(2\pi\hbar)^3} u_{\vec{p}}(t)\Psi^{(+)}_{\vec{p}}\vec{r}e^{i\vec{p}\cdot\vec{r}/\hbar} \\ &= \int \frac{d^3p}{(2\pi\hbar)^3} u_{\vec{p}}(t)\Psi^{(+)}_{\vec{p}}\frac{\hbar}{i}\nabla_{\vec{p}}e^{i\vec{p}\cdot\vec{r}/\hbar} \end{aligned} \tag{14.108}$$

Integrating by parts we have

$$\begin{aligned} \vec{r}\Psi^{(+)}(\vec{r},t) = &\int \frac{d^3p}{(2\pi\hbar)^3}(i\hbar\nabla_{\vec{p}}u_{\vec{p}}(t))\Psi^{(+)}_{\vec{p}}e^{i\vec{p}\cdot\vec{r}/\hbar} \\ &+ \int \frac{d^3p}{(2\pi\hbar)^3}u_{\vec{p}}(t)(i\hbar\nabla_{\vec{p}}\Psi^{(+)}_{\vec{p}})e^{i\vec{p}\cdot\vec{r}/\hbar} \end{aligned} \tag{14.109}$$

Using

$$\nabla_{\vec{p}}\Psi^{(\pm)}_{\vec{p}} = -\frac{\vec{p}c^2}{2E^2_{\vec{p}}}\Psi^{(\mp)}_{\vec{p}} \tag{14.110}$$

we get

$$\vec{r}\Psi^{(+)}(\vec{r},t) = \vec{r}_+\Psi^{(+)}(\vec{r},t) + \vec{r}_-\Psi^{(+)}(\vec{r},t) \tag{14.111}$$

where

$$\vec{r}_+\Psi^{(+)}(\vec{r},t) = \int \frac{d^3p}{(2\pi\hbar)^3}(i\hbar\nabla_{\vec{p}}u_{\vec{p}}(t))\Psi^{(+)}_{\vec{p}}e^{i\vec{p}\cdot\vec{r}/\hbar} \tag{14.112}$$

$$\vec{r}_-\Psi^{(+)}(\vec{r},t) = -\int \frac{d^3p}{(2\pi\hbar)^3}u_{\vec{p}}(t)\frac{i\hbar\vec{p}c^2}{2E^2_{\vec{p}}}\Psi^{(-)}_{\vec{p}}e^{i\vec{p}\cdot\vec{r}/\hbar} \tag{14.113}$$

This says that multiplying a wave packet of positive energy states by the position operator mixes in negative energy solutions, i.e.,

$\vec{r}_+$ generates positive energy solutions while $\vec{r}_-$ generates negative energy solutions

or changes free particles in free antiparticles and vices versa

The same result occurs for any function of the position operator.

Suppose that $\vec{r}\Psi^{(+)}(\vec{r}) = \vec{r}_0\Psi^{(+)}(\vec{r})$, i.e., it is an eigenstate of $\vec{r}$ with eigenvalue $\vec{r}_0$. This says that

$$i\hbar\nabla_{\vec{p}}u_{\vec{p}} = \vec{r}_0 u_{\vec{p}} \rightarrow u_{\vec{p}} = e^{-i\vec{p}\cdot\vec{r}_0/\hbar} \tag{14.114}$$

and the state

$$\Psi^{+}_{\vec{r}_0}(\vec{r}) = \int \frac{d^3p}{(2\pi\hbar)^3} e^{i\vec{p}\cdot(\vec{r}-\vec{r}_0)/\hbar}\Psi^{(+)}_{\vec{p}} \tag{14.115}$$

is an eigenstate of $\vec{r}_+$.

This eigenstate is not normalizable. It is large over a region of space within $\hbar/mc$ (a Compton wavelength) of $\vec{r}_0$ or, in other words, the theory with positive energy solutions cannot describe particles localized to a region smaller than a Compton wavelength.

The presence of the $\vec{r}_-$ part in the position operator says that putting a wave packet made from positive energy solutions (a particle) through a potential $\Phi(\vec{r})$ (which multiplies by functions of r) causes the creation of antiparticles and because charge must be conserved, creates new particles also.

Thus, the relativistic spin-zero theory of the Klein-Gordon equation has built into it the mechanism of particle-antiparticle production by external potentials.

An example of this phenomenon is *Klein's paradox*. Suppose we have a beam of positively charged particles with momentum p hitting an electrostatic potential barrier of height $e\varphi$ from the left as shown in Figure 14.1 below.

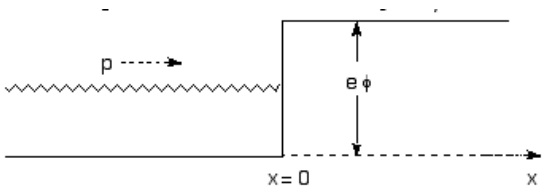


Figure 14.1: Electrostatic Potential Barrier - Klein Paradox

The solution follows the same lines as the nonrelativistic problem. For $x < 0$ we have

$$\psi(x) = ae^{ipx/\hbar} + be^{-ipx/\hbar} \quad , \text{ Energy } = E_p \tag{14.116}$$

This corresponds to incident and reflected waves. For $x > 0$, the Klein-Gordon equation is

$$(E_p - V)^2\psi(x) = -\hbar^2c^2\frac{\partial^2\psi(x)}{\partial x^2} + mc^2\psi(x) \quad , \quad V = e\varphi \tag{14.117}$$

The solution takes the form $\psi(x) = de^{ikx}$ where substitution gives

$$(E_p - V)^2 = \hbar^2 c^2 k^2 + m^2 c^4 \tag{14.118}$$

We have the boundary conditions at $x = 0$ (since potential only has a finite discontinuity)

$$\psi(x) \ , \ \frac{\partial \psi}{\partial x} \text{ continuous} \tag{14.119}$$

Note that ψ^0 is given by

$$\psi^0(\vec{r}, t) = \left(\frac{\partial}{\partial t} + \frac{ie}{\hbar} \Phi(\vec{r}, t) \right) \psi(\vec{r}, t) \tag{14.120}$$

is not continuous at $x = 0$.

We obtain

$$b = \frac{p - \hbar k}{p + \hbar k} a \quad , \quad d = \frac{2p}{p + \hbar k} a \tag{14.121}$$

We consider three cases:

1. If $E_p > V + mc^2$, then the particle can pass the over the barrier and the results are identical to the nonrelativistic case, i.e., part of the wave is reflected and part is transmitted.

2. If we have a stronger potential such that $E_p + mc^2 > V > E_p - mc^2$, then k must be imaginary so that the wave function goes to zero as $x \to \infty$. We then have

$$k = i\kappa \to \kappa = \frac{\sqrt{m^2 c^4 - (E_p - V)^2}}{\hbar c} \tag{14.122}$$

   and the wave is totally reflected at the barrier. The charge density on the right ($x > 0$) is given by

$$\rho(x) = \frac{E_p - V}{2mc^2} |d|^2 e^{-2\kappa x} \tag{14.123}$$

   For $E_p > V$, there exists a positive, exponentially decaying charge density to the right of the barrier. For $E_p < V$, however, the density is *negative* (remember it is a beam of positive particles). We reflect positively charged particles from the barrier and find negative particles inside the barrier.

3. We make the potential even stronger so that $V > E_p + mc^2$. Nonrelativistically it would be even more impossible for the particles to pass over the barrier. In the relativistic case, however, k is real again, This says that once again there is a particle current to the right of the barrier. The group velocity of the waves for $x > 0$ is

$$v_g = \frac{\partial E_p}{\partial(\hbar k)} \tag{14.124}$$

Using $(E_p - V)^2 = \hbar^2 c^2 k^2 + m^2 c^4$ we get

$$(E_p - V)\frac{\partial E_p}{\partial k} = \hbar c^2 k \rightarrow v_g = \frac{\hbar c^2 k}{E_p - V} \tag{14.125}$$

Since $E_p - V < 0$, we must have $k < 0$ (negative!) in order to have a wave(packet) traveling from the barrier towards positive x.

This says that the reflection coefficient b/a is greater than one, i.e., more wave is reflected than is incident! In addition, the charge density on the right is

$$\rho(x) = \frac{E_p - V}{2mc^2}|d|^2 < 0 \tag{14.126}$$

and the current on the right is negative.

One possible explanation is to say that the incident particle induces the creation of particle-antiparticle pairs at the barrier. The created antiparticles, having the opposite charge, find $x > 0$ a region of attractive potential and thus travel towards the right, which explains the negative current on the right. The created particles travel to the left and together with the incident particles (wave) which are(is) totally reflected, they add up to an outgoing current on the left that is larger than the incident current.

The total outgoing current on the left and right equals the incident current since total charge must be conserved.

This pair creation solution does not violate conservation of energy. The energy of a created particle on the left is E_p. The energy of a created antiparticle on the right is $\sqrt{\hbar^2 c^2 k^2 + m^2 c^4} - V$ since the electrostatic potential energy has the opposite sign for a particle of opposite charge.

Adding the two energies we get $E_p + \sqrt{\hbar^2 c^2 k^2 + m^2 c^4} - V = 0$, i.e., it takes zero energy to create a particle-antiparticle pair. This happens because the potential V is so large that the energy of the antiparticle on the right is not only less than mc^2 but is negative.

14.4 Bound State Problems

We now study the bound states of spin zero relativistic particles in a static potential $\Phi(\vec{r})$. For a positively charged particle with energy E the bound state wave function is

$$\psi(\vec{r}, t) = e^{-iEt/\hbar}\psi(\vec{r}) \tag{14.127}$$

and the charge density of the bound state particle is

$$e\rho(\vec{r}) = \frac{e\left[E - e\Phi(\vec{r})\right]}{mc^2}|\psi(\vec{r})|^2 \tag{14.128}$$

This says that in regions where $E > e\Phi(\vec{r})$, which includes classically accessible regions, the charge density is positive. But, in regions where $E < e\Phi(\vec{r})$, the charge density is negative. The way to think about this is to say the particle in the potential is a linear combination of free particle and free antiparticle states.

Because we are considering an electrostatic potential, the positively charged parts will be found mainly in regions of smaller $e\Phi(\vec{r})$ while the negatively charged parts will be found in regions of larger $e\Phi(\vec{r})$. A relativistic particle seems to have an internal structure that can be polarized by and electric field.

Alternatively, we might say that the potential produces particle-antiparticle pairs in the vacuum, The positively charged particles are attracted to regions of smaller $e\Phi(\vec{r})$ while the negatively charged particles are attracted to regions of larger $e\Phi(\vec{r})$. We say that the electric potential has *polarized* the vacuum. This polarization modifies the effective potential felt by the bound particle.

This interaction cannot be taken into account in the present one-particle relativistic theory(requires quantum field theory).

We now turn to the problem of a spin zero particle bound in a Coulomb potential. An example is a π^- bound to a nucleus. We have

$$e\Phi(\vec{r}) = -\frac{Ze^2}{r} \tag{14.129}$$

which leads to the Klein-Gordon equation

$$\left[\left(E + \frac{Ze^2}{r}\right)^2 + \hbar^2 c^2 \nabla^2 - mc^2\right]\psi(\vec{r}) = 0 \tag{14.130}$$

Since this is a central potential, we can assume that the eigenstates have definite values of total orbital angular momentum. We then have

$$\left[\left(\frac{E^2}{c^2} - m^2c^2\right) + \hbar^2\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\ell(\ell+1) - (Z\alpha)^2}{r^2}\right) + \frac{2Ze^2}{r}\frac{E}{c}\right]\psi(r) = 0 \tag{14.131}$$

or

$$\left[-\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{\ell(\ell+1) - (Z\alpha)^2}{r^2} - \frac{2Z\alpha E}{\hbar c r} - \left(\frac{E^2 - m^2c^4}{\hbar^2 c^2}\right)\right]\psi(r) = 0 \tag{14.132}$$

where

$$\alpha = \frac{e^2}{\hbar c} = \text{ fine structure constant} \tag{14.133}$$

Now we define

$$\gamma = Z\alpha \quad , \quad \ell'(\ell'+1) = \ell(\ell+1) - \gamma^2$$
$$\lambda = \frac{2E\gamma}{\hbar c \sigma} \quad , \quad \frac{4(m^2c^4 - E^2)}{\hbar^2 c^2} = \sigma^2 \quad , \quad \rho = \sigma r$$

and we get

$$\left[\frac{d^2}{d(\rho/2)^2}+\frac{2\lambda}{\rho/2}-1+\frac{\ell'(\ell'+1)}{(\rho/2)^2}\right]\rho\psi(\rho)=0 \tag{14.134}$$

which is identical to the radial equation for the nonrelativistic Coulomb problem for the function $u=\rho\psi(\rho)$. The difference is that ℓ' is not necessarily an integer (remember that it is an integer in the nonrelativistic problem), which causes the orbits of the relativistic Coulomb (Kepler) problem to no longer be closed, i.e., the orbits precess. This also means that the extra degeneracy of the nonrelativistic problem which causes the energy to be independent of ℓ is broken in the relativistic problem. We now solve this equation in the standard way. For

$$\begin{aligned}&\rho\to 0 \qquad \psi(\rho)\to\rho^{\ell'}\\ &\rho\to\infty \qquad \psi(\rho)\to e^{-\rho/2}\end{aligned}$$

Therefore, we guess a solution of the form

$$u=\rho\psi(\rho)=\left(\frac{\rho}{2}\right)^{\ell'+1}e^{-\rho/2}w(\rho/2) \tag{14.135}$$

The solution method is identical to the nonrelativistic hydrogen atom. We get a power series which must terminate (so that the solution is normalizable) when

$$\begin{aligned}&\lambda=N+\ell'+1 \quad , \quad N=0,1,2,3,.....\\ &E=mc^2\left(1+\frac{\gamma^2}{\lambda^2}\right)^{-1/2}\rightarrow E=\frac{mc^2}{\sqrt{1+\frac{\gamma^2}{\left[N+\frac{1}{2}+\sqrt{(\ell+\frac{1}{2})^2-\gamma^2}\right]^2}}}\end{aligned} \tag{14.136}$$

If we define the principal quantum number $n=N+\ell+1=$ integer, then we have

$$E_{n\ell}=\frac{mc^2}{\sqrt{1+\frac{\gamma^2}{\left[n-(\ell+\frac{1}{2})+\sqrt{(\ell+\frac{1}{2})^2-\gamma^2}\right]^2}}} \tag{14.137}$$

The principal quantum number has the possible values $n=1,2,3,.....$ For a given n the possible values of the total orbital angular momentum are $\ell=0,1,2,3,....,n-1$.

The degeneracy that was present in the nonrelativistic theory with respect to orbital angular momentum ℓ is clearly removed.

If we expand the energy in a power series in the fine structure constant α (or γ) we get

$$E_{n\ell}=mc^2-Ry\frac{1}{n^2}-Ry\frac{\gamma^2}{n^3}\left(\frac{1}{\ell+\frac{1}{2}}-\frac{3}{4n}\right)+O(Ry\gamma^4) \tag{14.138}$$

The first term is the rest energy. The second term is the nonrelativistic Rydberg formula. The third term is the relativistic correction due to using the relativistic form of the kinetic energy, which as we saw earlier in Chapter 10 took the form

$$\hat{H}_{rel} = -\frac{p^4}{8m^3c^2} \tag{14.139}$$

It is this correction that removes the degeneracy in ℓ, i.e.,

$$E_{n,\ell=0} - E_{n,\ell=n-1} = Ry\frac{4\gamma^2}{n^3}\frac{n-1}{2n-1} \tag{14.140}$$

As we shall see later when we derive the Dirac equation, there are more corrections to this formula due to the fact that the electron has spin = 1/2.

14.4.1 Nonrelativistic Limit

The Klein-Gordon equation in the presence of an electromagnetic field is

$$\frac{1}{c^2}\left(i\hbar\frac{\partial}{\partial t} - e\Phi(\vec{r},t)\right)^2\psi(\vec{r},t) = \left(\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}(\vec{r},t)\right)^2 + m^2c^2\right)\psi(\vec{r},t) \tag{14.141}$$

and using

$$\psi(\vec{r},t) = \begin{pmatrix} \phi(\vec{r},t) \\ \chi(\vec{r},t) \end{pmatrix} \tag{14.142}$$

we have as earlier

$$\left(i\hbar\frac{\partial}{\partial t} - e\Phi\right)\phi = \frac{1}{2m}\left[\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right]^2\phi + mc^2\phi + \frac{1}{2m}\left[\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right]^2\chi \tag{14.143}$$

$$\left(i\hbar\frac{\partial}{\partial t} - e\Phi\right)\chi = -\frac{1}{2m}\left[\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right]^2(\phi+\chi) - mc^2\chi \tag{14.144}$$

Remember that in the nonrelativistic limit the dominant term in the energy will be mc^2 so that we expect the zeroth order equation for χ to be

$$i\hbar\frac{\partial\chi}{\partial t} = mc^2\chi \tag{14.145}$$

which then implies in the next approximation that

$$\chi = -\frac{1}{4m^2c^2}\left[\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right]^2\phi \tag{14.146}$$

and that ϕ satisfies the equation

$$\left(i\hbar\frac{\partial}{\partial t} - e\Phi\right)\phi = \frac{1}{2m}\left[\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right]^2\phi + mc^2\phi - \frac{1}{8m^3c^2}\left[\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right]^4\phi \tag{14.147}$$

The operator on the right-hand side is just the kinetic energy operator

$$\sqrt{m^2c^4 + c^2\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)^2} \tag{14.148}$$

expanded to second order in $1/mc^2$. This agrees with our earlier result that the first relativistic correction for a spinless particle is entirely due to the relativistic modification of the kinetic energy.

For a weak magnetic field $\vec{B}$ this becomes (to order $(v/c)^3$) after much algebra

$$i\hbar\frac{\partial\phi}{\partial t} = -\frac{\hbar^2\nabla^2}{2m}\left[1+\frac{\hbar^2\nabla^2}{2m^2c^2}\right]\phi + (mc^2+e\Phi)\phi - \frac{e}{2mc}\vec{B}\cdot\vec{L}\left[1+\frac{\hbar^2\nabla^2}{2m^2c^2}\right]\phi \quad (14.149)$$

where $\vec{L}$ is the orbital angular momentum of the particle. The term

$$\left[1+\frac{\hbar^2\nabla^2}{2m^2c^2}\right] \to \left[1-\frac{p^2}{2m^2c^2}\right] \quad (14.150)$$

represents the relativistic correction to the magnetic moment.

14.5 Relativistic Spin 1/2 Particles - The Dirac Equation

14.5.1 Lorentz Transformation of Spin

The contravariant and covariant components of the position 4-vector in spacetime are:

$$x^\mu: \quad x^0 = ct \quad , \quad x^1 = x \quad , \quad x^2 = y \quad , \quad x^3 = z$$
$$x_\mu: \quad x^0 = ct \quad , \quad x_1 = -x \quad , \quad x_2 = -y \quad , \quad x_3 = -z$$

The flat spacetime metric tensor is defined by

$$g = (g_{\mu\nu}) = (g^{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (14.151)$$

The metric tensor relates covariant and contravariant components by

$$x_\mu = g_{\mu\nu}x^\nu \quad , \quad x^\mu = g^{\mu\nu}x_\nu \quad (14.152)$$

We also have

$$g^\mu_\nu = g^{\mu\alpha}g_{\alpha\nu} \equiv \delta^\mu_\nu \quad , \quad (\delta^\mu_\nu) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (14.153)$$

Under the action of a Lorentz transformation along the z-axis with velocity $v = \beta c$, a 4-vector (any type) since it is a first-rank tensor, transforms as

$$V'^\mu = \Lambda^\mu_\nu V^\nu \quad (14.154)$$

where

$$(\Lambda^{\mu}_{\nu}) = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \quad , \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} \quad , \quad \beta = \frac{v}{c} \tag{14.155}$$

This corresponds to the standard transformation relations for the position and momentum 4-vectors

$$ct' = \gamma(ct - \beta z) \quad , \quad x' = x \quad , \quad y' = y \quad , \quad z' = \gamma(z - \beta ct)$$
$$\frac{E'}{c} = \gamma\left(\frac{E}{c} - \beta p_z\right) \quad , \quad p'_x = p_x \quad , \quad p'_y = p_y \quad , \quad p'_z = \gamma\left(p_z - \beta\frac{E}{c}\right)$$

Spin is an angular momentum corresponding to internal degrees of freedom of the system. This means, as we showed earlier, that spin must have the same transformation properties as any other angular momentum.

Nonrelativistically, we think of angular momentum as a vector and, in fact, under a simple spatial rotation it does transform as a vector (as we saw earlier). Consider, however, the behavior of an orbital angular momentum

$$\vec{L} = \vec{r} \times \vec{p} \rightarrow L_i = \varepsilon_{ijk} x_j p_k \tag{14.156}$$

under the action of a Lorentz transformation along the z–direction. We find that

$$L'_z = x' p'_y - y' p'_x = x p_y - y p_x = L_z \tag{14.157}$$

since the components of vectors orthogonal to the z–axis are unchanged. This is clearly not the transformation property of a vector.

In fact, $\vec{L}$ is the product of two vectors and therefore should have the transformation properties of a second-rank tensor, i.e., as

$$Q'^{\,\mu\nu} = \Lambda^{\mu}_{\alpha}\Lambda^{\nu}_{\beta} Q^{\alpha\beta} \tag{14.158}$$

Relativistic electrodynamics can be written in terms of a second-rank field tensor

$$(\,F^{\mu\nu}) = \begin{pmatrix} 0 & \varepsilon_1 & \varepsilon_2 & \varepsilon_3 \\ -\varepsilon_1 & 0 & B_3 & -B_2 \\ -\varepsilon_2 & -B_3 & 0 & B_1 \\ -\varepsilon_3 & B_2 & -B_1 & 0 \end{pmatrix} \tag{14.159}$$

as Maxwell's equations

$$\frac{\partial F^{\mu\nu}}{\partial x^{\mu}} = \frac{4\pi}{c} J^{\nu}$$
$$\frac{\partial F^{\mu\nu}}{\partial x^{\alpha}} + \frac{\partial F^{\nu\alpha}}{\partial x^{\mu}} + \frac{\partial F^{\alpha\mu}}{\partial x^{\nu}} = 0$$

where the current density 4-vector is

$$J^\mu = (c\rho, J_x, J_y, J_z) \tag{14.160}$$

and the Lorentz force law is

$$\frac{dp^\mu}{d\tau} = \frac{q}{m} p_\nu F^{\mu\nu} \tag{14.161}$$

The transformation rule then says that the fields transform according to the relations

$$\begin{aligned} F'^{01} &= \varepsilon'_1 = \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta} = \Lambda^0_\alpha \Lambda^1_\beta F^{\alpha\beta} = \Lambda^0_\alpha \Lambda^1_1 F^{\alpha 1} = \Lambda^0_\alpha F^{\alpha 1} \\ &= \Lambda^0_0 F^{01} + \Lambda^0_3 F^{31} = \gamma\varepsilon_1 - \beta\gamma B_2 = \gamma(\varepsilon_1 - ((\vec{v}/c) \times \vec{B})_1) \end{aligned}$$

or

$$\varepsilon'_1 = \gamma(\varepsilon_1 + ((\vec{v}/c) \times \vec{B})_1) \text{ where } \vec{v} = v\hat{e}_z \tag{14.162}$$

Similarly, we find

$$\begin{aligned} &\varepsilon'_2 = \gamma(\varepsilon_2 + ((\vec{v}/c) \times \vec{B})_2) \quad , \quad \varepsilon'_3 = \varepsilon_3 \\ &B'_1 = \gamma(B_1 - ((\vec{v}/c) \times \vec{\varepsilon})_1) \quad , \quad B'_2 = \gamma(B_2 - ((\vec{v}/c) \times \vec{\varepsilon})_2) \quad , \quad B'_3 = B_3 \end{aligned}$$

We can summarize these results for an arbitrary (direction) Lorentz transformation applied to a second-rank tensor by

$$\begin{aligned} &B'_\parallel = B_\parallel \quad , \quad \varepsilon'_\parallel = \varepsilon_\parallel \; \parallel = \text{ component parallel to } \vec{v} \\ &\vec{B}'_\perp = \gamma\left(\vec{B}_\perp - (\vec{v}/c) \times \vec{\varepsilon}\right) \\ &\vec{\varepsilon}'_\perp = \gamma\left(\vec{\varepsilon}_\perp + (\vec{v}/c) \times \vec{B}\right) \; \perp = \text{ component perpendicular to } \vec{v} \end{aligned}$$

Thus, a pure magnetic field in one frame is a mixture of magnetic and electric fields in the new frame.

Now, under a spatial inversion transformation, we have

$$\begin{aligned} &\vec{\varepsilon} \to -\vec{\varepsilon} \quad , \quad \vec{B} \to \vec{B} \\ &\vec{r} \to -\vec{r} \quad , \quad \vec{p} \to -\vec{p} \Rightarrow \vec{L} \to \vec{L} \end{aligned}$$

Therefore, an angular momentum(including spin) has the same transformation properties as the magnetic field.

Since spin,

$$\bar{S} = \frac{1}{2}\vec{\sigma} \tag{14.163}$$

must transform as an angular momentum, which transforms like a magnetic field and the magnetic field is part of second-rank tensor with the electric field, we must conclude that there exists another set of dynamical variables generated

by the internal degrees of freedom of the particle that will be analogous to the electric field. Do not think of the operator $\vec{\sigma}$ as the standard 2×2 Pauli matrices; we shall see later that $\vec{\sigma}$ will need to be represented by 4×4 matrices relativistically.

We define these new variables as $i\vec{\alpha}/2$ where the $i/2$ factor is chosen for later convenience. We then have that $\vec{S}$ and $i\vec{\alpha}/2$ or $\vec{\sigma}$ and $i\vec{\alpha}$ transform as $\vec{B}$ and $\vec{\varepsilon}$, i.e.,

$$\begin{aligned} \sigma'_{\parallel} &= \sigma_{\parallel} \quad , \quad \vec{\sigma}'_{\perp} = \gamma\left(\vec{\sigma}_{\perp} - (\vec{v}/c) \times i\vec{\alpha}\right) \\ i\alpha'_{\parallel} &= i\alpha_{\parallel} \quad , \quad i\vec{\alpha}'_{\perp} = \gamma\left(i\vec{\alpha}_{\perp} + (\vec{v}/c) \times \vec{\sigma}\right) \end{aligned}$$

and they form a second-rank tensor $\sigma^{\mu\nu}$ analogous to $F^{\mu\nu}$, i.e.,

$$(\sigma^{\mu\nu}) = \begin{pmatrix} 0 & i\alpha_1 & i\alpha_2 & i\alpha_3 \\ -i\alpha_1 & 0 & \sigma_3 & -\sigma_2 \\ -i\alpha_2 & -\sigma_3 & 0 & \sigma_1 \\ -i\alpha_3 & \sigma_2 & -\sigma_1 & 0 \end{pmatrix} \tag{14.164}$$

We must now investigate the dynamical properties of the new variables $\vec{\alpha}$ and also ask this question - where have these objects been hiding in all of previous discussions?

Since spin is an angular momentum, we know its algebraic properties (commutators). In addition, spin generates rotations of the internal degrees of freedom. Spin commutes with spatial degrees of freedom like $\vec{r}$ and $\vec{p}$ and, thus, so does $\vec{\alpha}$.

Since $\vec{\alpha}$ behaves like a vector under spatial rotations (it is like the electric field vector), it must have the standard commutation relations with $\vec{S}$

$$[\alpha_i, S_j] = i\varepsilon_{ijk}\alpha_k \rightarrow [\alpha_i, \sigma_j] = 2i\varepsilon_{ijk}\alpha_k \tag{14.165}$$

Since $\vec{\sigma}$ is angular momentum, it satisfies the relations

$$\sigma_i\sigma_j = i\varepsilon_{ijk}\sigma_k + \delta_{ij} \tag{14.166}$$

which must be true in all Lorentz frames, i.e., since $\sigma_1^2 = 1$, we must have $\sigma_1'^2 = 1$.

Let us now determine all the properties of $\vec{\alpha}$.

For a Lorentz transformation along the $z-$direction we have

$$\begin{aligned} \sigma'_x &= \gamma\left(\sigma_x + iv\alpha_y/c\right) \\ \sigma'_y &= \gamma\left(\sigma_y - iv\alpha_x/c\right) \end{aligned}$$

Squaring σ'_x we get

$$\begin{aligned} \sigma_x'^2 = 1 &= \gamma^2\left(\sigma_x^2 + iv(\sigma_x\alpha_y + \alpha_y\sigma_x)/c - \left(v/c\right)^2\right)\alpha_y^2\Big) \\ &= \gamma^2\left(1 - \left(v/c\right)^2\right)\alpha_y^2\Big) + \gamma^2\left(iv(\sigma_x\alpha_y + \alpha_y\sigma_x)/c\right) \end{aligned}$$

Since this must be true for all v, the coefficient of v/c must vanish. Thus,

$$\sigma_x\alpha_y + \alpha_y\sigma_x = 0 \tag{14.167}$$

We then have

$$1 = \gamma^2\left(1 - \left(v/c\right)^2\right)\alpha_y^2\Big) \tag{14.168}$$

Since

$$1 = \gamma^2\left(1 - \left(v/c\right)^2\right)\alpha_y^2\Big) \tag{14.169}$$

we must also have

$$\alpha_y^2 = 1 \tag{14.170}$$

These results generalize to the following:

$$i \neq j \qquad \sigma_i\alpha_j = -\alpha_j\sigma_i \tag{14.171}$$

$$i = j \qquad [\sigma_i, \alpha_i] = 0 \tag{14.172}$$

Multiplying σ'_y by σ'_x we get

$$\sigma'_x\sigma'_y = i\sigma'_z = i\sigma_z = \gamma^2\left(\sigma_x + iv\alpha_y/c\right)\left(\sigma_y - iv\alpha_x/c\right)$$
$$i\sigma_z = \gamma^2\left(\sigma_x\sigma_y + (v/c)^2\alpha_y\alpha_x + i(v/c)(\alpha_y\sigma_y - \sigma_x\alpha_x\right)$$

Multiplying σ'_x by σ'_y we get

$$\sigma'_y\sigma'_x = -i\sigma'_z = -i\sigma_z = \gamma^2\left(\sigma_y - iv\alpha_x/c\right)\left(\sigma_x + iv\alpha_y/c\right)$$
$$-i\sigma_z = \gamma^2\left(\sigma_y\sigma_x + (v/c)^2\alpha_x\alpha_y + i(v/c)(\sigma_y\alpha_y - \alpha_x\sigma_x\right)$$

Adding, we have

$$(\sigma_y\sigma_x + \sigma_x\sigma_y) + (v/c)^2(\alpha_x\alpha_y + \alpha_y\alpha_x) = 0$$
$$(\alpha_x\alpha_y + \alpha_y\alpha_x) = 0 \rightarrow (\alpha_i\alpha_j + \alpha_j\alpha_i) = 0 \quad i \neq j$$

Continuing, we find these other relations

$$\alpha_y\alpha_x = -i\sigma_z = -\alpha_x\alpha_y \rightarrow \alpha_i\alpha_j - \alpha_j\alpha_i = 2i\varepsilon_{ijk}\sigma_k \tag{14.173}$$

or summarizing we have

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij} \quad , \quad [\alpha_i, \alpha_j] = 2i\varepsilon_{ijk}\sigma_k \tag{14.174}$$

$$[\alpha_i, \sigma_j] = 2i\varepsilon_{ijk}\alpha_k \quad , \quad \{\alpha_i, \sigma_j\} = 0 \quad , \quad i \neq j \tag{14.175}$$

So α obeys exactly the same algebraic relations as σ. How do we know that α is not equal to σ? If we apply a parity transformation, we find that $\vec{\sigma} \rightarrow \vec{\sigma}$ since angular momentum is unchanged by spatial inversion, i.e., the space-space components of a second-rank tensor do not change sign under parity. On the other hand, the time-space components such as the electric field or $i\vec{\alpha}$ do change sign, i.e., $\vec{\alpha} \rightarrow -\vec{\alpha}$. So they cannot be the same operator!

Let β be the operator that corresponds to the parity transformation in spin space. Now two successive inversions brings us back to the starting configuration. Remember, however, that the spin representation of rotations is doubled valued, i.e., a rotation by 2π produces a minus sign. We have a choice of letting the square of the parity operation include a 2π rotation(about any axis) or not. This means that we can have $\beta^2 = +1 \; or \; -1$.

In the first case, the eigenvalues of β are ± 1 and in the second case $\pm i$. We choose $\beta^2 = +1 \rightarrow \beta^{-1} = \beta$. The properties under parity become

$$\beta^{-1}\vec{\sigma}\beta = \vec{\sigma} \rightarrow \beta\vec{\sigma} = \vec{\sigma}\beta \tag{14.176}$$

$$\beta^{-1}\vec{\alpha}\beta = -\vec{\alpha} \rightarrow \beta\vec{\alpha} = -\vec{\alpha}\beta \tag{14.177}$$

We can now construct an explicit matrix representation for the operators $\vec{\alpha}$, β and $\vec{\sigma}$ similar to the two-dimensional Pauli matrix representation in the nonrelativistic case.

Consider the determinant of the matrix for $\beta^{-1}\alpha_i\beta$. Using $\beta\vec{\alpha} = -\vec{\alpha}\beta$ we have

$$\det(\beta^{-1}\alpha_i\beta) = \det(-\beta^{-1}\beta\alpha_i) = \det(-\alpha_i) = (-1)^N \det(\alpha_i) \tag{14.178}$$

where N is the dimension of the matrix representation. However, using the cyclic property of determinants, i.e.,

$$\det ABC = \det BCA = \det CAB \tag{14.179}$$

we get

$$\det(\beta^{-1}\alpha_i\beta) = \det(\beta\beta^{-1}\alpha_i) = \det(\alpha_i) \tag{14.180}$$

Putting these results together we get

$$(-1)^N \det(\alpha_i) = \det(\alpha_i) \rightarrow (-1)^N = 1 \rightarrow N = 2, 4, 6, ... \tag{14.181}$$

We have used the fact that $\det(\alpha_i) \neq 0$, since $\alpha_i^2 = 1$.

Now, $N = 2$ is not possible as we show below.

All 2×2 matrices can be constructed from the set $\{I, \vec{\sigma}\}$ and $[\beta, \vec{\sigma}] = 0$. This means that β would have to commute with all 2×2 matrices. Since $\vec{\alpha}$ would then have to commute with β, we would then violate the relation $\beta\vec{\alpha} = -\vec{\alpha}\beta$.

This means N must be at least as large as 4. This says that a relativistic spin 1/2 particle would have 4 internal states (the nonrelativistic case has 2). This is similar to the Klein-Gordon case and it will turn out here also that this doubling signals the appearance of antiparticles.

An explicit representation(not unique) using the 2×2 matrices

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad \tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

where the the last three matrices are the standard Pauli matrices. It is given by

$$\vec{\sigma} = \begin{pmatrix} \vec{\tau} & 0 \\ 0 & \vec{\tau} \end{pmatrix} \quad , \quad \vec{\alpha} = \begin{pmatrix} 0 & \vec{\tau} \\ \vec{\tau} & 0 \end{pmatrix} \quad , \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{14.182}$$

Note that the trace of each of these matrices is zero, which is a general property of matrices that obey anticommutation relations.

We can make the following physical interpretations of the components of the tensor $\sigma^{\mu\nu}$.

It follows from earlier discussions that the space-space components, i.e., the spin operators σ_j, generate (in the spin degrees of freedom) a rotation of the coordinate system.

This implies that the operators $\vec{\sigma}\,'$, $\vec{\alpha}\,'$, β' in a spatially rotated frame are given by the operator relations

$$\vec{\sigma}\,' = R_\varphi \vec{\sigma} R_\varphi^{-1} \ , \quad \vec{\alpha}\,' = R_\varphi \vec{\alpha} R_\varphi^{-1} \ , \quad \beta' = R_\varphi \beta R_\varphi^{-1} \tag{14.183}$$

where

$$R_\varphi = e^{-i\vec{\sigma}\cdot\hat{n}\varphi/2} \tag{14.184}$$

and

$\hat{n} =$ unit vector in direction of axis of rotation

$\varphi =$ angle of rotation

We then assume that time-space components generate a rotation of the space axes with the time axis, which is a Lorentz transformation and that the operators in the new frame are given by

$$\vec{\sigma}\,' = L_v \vec{\sigma} L_v^{-1} \ , \ \vec{\alpha}\,' = L_v \vec{\alpha} L_v^{-1} \ , \ \beta' = L_v \beta L_v^{-1} \tag{14.185}$$

where

$$L_v = e^{-i(i\vec{\alpha})\cdot\vec{\omega}/2} = e^{\vec{\alpha}\cdot\vec{\omega}/2} \rightarrow L_v^{-1} = e^{-\vec{\alpha}\cdot\vec{\omega}/2} \tag{14.186}$$

and

$\vec{\omega} =$ vector in direction of velocity of primed frame
with respect to unprimed frame and of magnitude

$\tanh(\omega) = \dfrac{v}{c}$

Proof: First,

$$\sigma'_{\|} = L\sigma_{\|}L^{-1} = e^{\alpha_{\|}\omega/2}\sigma_{\|}e^{-\alpha_{\|}\omega/2} = e^{\alpha_{\|}\omega/2}e^{-\alpha_{\|}\omega/2}\sigma_{\|} = \sigma_{\|} \tag{14.187}$$

where we have used

$$[\alpha_i, \sigma_j] = 2i\varepsilon_{ijk}\alpha_k \rightarrow \left[\alpha_{\|}, \sigma_{\|}\right] = 0 \tag{14.188}$$

This agrees with our earlier result.

Second,

$$\sigma'_\perp = L\sigma_\perp L^{-1} = e^{\alpha_\parallel \omega/2}\sigma_\perp e^{-\alpha_\parallel \omega/2} = e^{\alpha_\parallel \omega/2}e^{\alpha_\parallel \omega/2}\sigma_\perp = e^{\alpha_\parallel \omega}\sigma_\perp = e^{\vec{\alpha}\cdot\vec{\omega}}\sigma_\perp \quad (14.189)$$

where we have used

$$\{\alpha_i, \sigma_j\} = 0 \quad , \quad i \neq j \rightarrow \{\alpha_\parallel, \sigma_\perp\} = 0 \quad (14.190)$$

Now using $\vec{\omega} = \omega\hat{\omega}$, $\hat{\omega}\cdot\hat{\omega} = 1$ and $(\vec{\alpha}\cdot\hat{\omega})^2 = 1$ we get

$$e^{\vec{\alpha}\cdot\vec{\omega}} = \cosh\omega + \vec{\alpha}\cdot\hat{\omega}\sinh\omega \quad (14.191)$$

This derivation is the analog of

$$e^{i\vec{\sigma}\cdot\hat{n}\theta} = \cos\theta + i\vec{\sigma}\cdot\hat{n}\sin\theta \quad (14.192)$$

Therefore,

$$\sigma'_\perp = e^{\vec{\alpha}\cdot\vec{\omega}}\sigma_\perp = \cosh\omega\left[1 + \vec{\alpha}\cdot\hat{\omega}\tanh\omega\right]\sigma_\perp = \gamma\left[1 + \vec{\alpha}\cdot\hat{\omega}\tanh\omega\right]\sigma_\perp \quad (14.193)$$

where we have used

$$\cosh^2\omega = \frac{1}{1-\tanh^2\omega} = \frac{1}{1-(v/c)^2} = \gamma^2 \quad (14.194)$$

Using $\hat{\omega}\tanh\omega = \vec{v}/c$ we then get

$$\sigma'_\perp = \gamma\left[1 + \vec{\alpha}\cdot\hat{\omega}\tanh\omega\right]\sigma_\perp = \gamma\left[1 + \vec{\alpha}\cdot\vec{v}/c\right]\sigma_\perp \quad (14.195)$$

Finally, assuming $\vec{v} = v\hat{e}_z$ and using $\alpha_i\sigma_j = i\varepsilon_{ijk}\alpha_k,\ i \neq j$ we get

$$(\vec{\alpha}\cdot\vec{v}/c)\sigma_\perp = i\frac{\vec{v}}{c}\times\vec{\alpha} \quad (14.196)$$

so that

$$\sigma'_\perp = \gamma\left[\sigma_\perp + i\frac{\vec{v}}{c}\times\vec{\alpha}\right] \quad (14.197)$$

which agrees with our earlier result. Thus, $\vec{\sigma}$ transforms correctly. A similar calculation shows that $\vec{\alpha}$ transforms correctly also and, thus, our interpretation is correct.

What about the operator β? In a new Lorentz frame we get

$$\beta' = L_v\beta L_v^{-1} = e^{\vec{\alpha}\cdot\vec{\omega}}\beta \quad (14.198)$$

since β anticommutes with $\vec{\alpha}$. We then get (as above)

$$\beta' = \gamma\left[\beta - (\vec{v}/c)\cdot\beta\vec{\alpha}\right] \quad (14.199)$$

where we have used $\beta\,\vec{\alpha} = -\,\vec{\alpha}\beta$. From the form of this transformation relation, it looks like β transforms as the time-component of a 4-vector of which $\beta\vec{\alpha}$ is the space part.

Some algebra shows that

$$\beta'\alpha'_{\perp} = L\beta\alpha_{\perp}L^{-1} = \beta\alpha_{\perp} \tag{14.200}$$

since both β and $\alpha_{\perp}$ anticommute with L and therefore $\beta\alpha_{\perp}$ commutes with L. In addition, we can show that

$$\beta'\alpha'_{\|} = \gamma\left[\beta\alpha_{\|} - (v/c)\beta\right] \tag{14.201}$$

Therefore, $(\beta, \beta\,\vec{\alpha})$ does transform like a 4-vector.

This 4-vector is called $\gamma^{\mu} = (\beta, \beta\,\vec{\alpha})$. In our earlier notation the space part is

$$\vec{\gamma} = \beta\vec{\alpha} = \begin{pmatrix} I & \vec{\tau} \\ -\vec{\tau} & I \end{pmatrix} \tag{14.202}$$

$\vec{\gamma}$ is anti-Hermitian and γ^0 is Hermitian.

Some properties

$$(\gamma^0)^2 = 1 \quad , \quad (\gamma^i)^2 = -1 \quad , \quad i = 1,2,3 \tag{14.203}$$

$$\{\gamma^{\mu}, \gamma^{\nu}\} = 0 \quad , \quad \mu \neq \nu \tag{14.204}$$

which are summarized by the relation

$$\{\gamma^{\mu}, \gamma^{\nu}\} = 2g^{\mu\nu} \tag{14.205}$$

We also have

$$\sigma^{\mu\nu} = \frac{i}{2}\left[\gamma^{\mu}, \gamma^{\nu}\right] \tag{14.206}$$

In fact, any 4×4 matrix can be written as a unique linear combination of the γ^{μ}. The set of 16 matrices

$$I \ , \ \gamma^{\mu} \ , \ \sigma^{\mu\nu} \ , \ \gamma_5\gamma^{\mu} \ , \ \gamma_5 \quad , \qquad \text{where } \gamma_5 = \gamma^0\gamma^1\gamma^2\gamma^3 \tag{14.207}$$

are linearly independent and complete. All are traceless except for the identity matrix.

The new operator γ_5 commutes with the γ^{μ}. This implies that it commutes with $\alpha_i = \gamma^0\gamma^i$ and is invariant under a Lorentz transformation. It is not a scalar, however, since under parity

$$\beta\gamma_5\beta = -\gamma_5 \tag{14.208}$$

This means that it is a *pseudoscalar*.

Similarly, $\gamma_5\gamma^\mu$ is a *pseudovector* (or *axial vector*), which is a 4-vector whose space part does not change sign under parity and whose time component does.

We see that the set I , γ^μ , $\sigma^{\mu\nu}$, $\gamma_5\gamma^\mu$, γ_5 transforms as a scalar or zeroth-rank tensor, a vector or first-rank tensor, a second-rank tensor, a pseudovector or axial vector or a first-rank pseudotensor and a pseudoscalar or a zeroth-rank pseudotensor.

This is a clear indication that they are linearly independent.

14.6 The Dirac Equation

The scalar product of two 4-vectors is a Lorentz invariant. We have identified two 4-vectors, namely,

$$\gamma^\mu = (\beta, \beta\vec{\alpha}) \quad , \quad p^\mu = (E/c, \vec{p}) \tag{14.209}$$

Their scalar product is

$$\beta\frac{E}{c} - \beta\vec{\alpha}\cdot\vec{p} \tag{14.210}$$

Since it is an invariant, it has the same value in all frames. What is that value? We can find out by looking at its square

$$\begin{aligned}\left(\beta\frac{E}{c} - \beta\vec{\alpha}\cdot\vec{p}\right)^2 &= \beta^2\left(\frac{E}{c}\right)^2 + (\beta\vec{\alpha}\cdot\vec{p})^2 - \beta\left(\beta\vec{\alpha} + \vec{\alpha}\beta\right)\cdot\frac{\vec{p}E}{c} \\ &= (1)\left(\frac{E}{c}\right)^2 + \sum_i (\beta\alpha_i p_i)^2 - \beta\,(0)\cdot\frac{\vec{p}E}{c} = \left(\frac{E}{c}\right)^2 - \sum_i (\beta\alpha_i)^2 p_i^2 \\ &= \left(\frac{E}{c}\right)^2 - \sum_i (1)^2 p_i^2 = \left(\frac{E}{c}\right)^2 - p^2\end{aligned} \tag{14.211}$$

This says that

$$\beta\frac{E}{c} - \beta\vec{\alpha}\cdot\vec{p} = \sqrt{\left(\frac{E}{c}\right)^2 - p^2} = \pm mc \tag{14.212}$$

The sign depends on the sign we choose for β. If we had interpreted $\beta^2 = 1$ to mean $\beta = -1$ instead of $+1$, which is equivalent to choosing the parity operator as $-\beta$, no physics would have changed. This means we are free to choose the sign. We choose

$$\beta\frac{E}{c} - \beta\vec{\alpha}\cdot\vec{p} = +mc \tag{14.213}$$

or

$$\beta E - \beta c\vec{\alpha}\cdot\vec{p} = mc^2 \tag{14.214}$$

This operator equation involves 4×4 matrices which implies that any physical state vectors must be 4-component spinors.

We make the standard operator correspondence

$$E \to i\hbar \frac{\partial}{\partial t} \quad , \quad \vec{p} \to i\hbar \nabla \tag{14.215}$$

and obtain a wave equation

$$i\hbar\beta \frac{\partial \psi}{\partial t} = \beta c \vec{\alpha} \cdot \frac{\hbar}{i} \nabla \psi + mc^2 \psi \tag{14.216}$$

or multiplying by β we get

$$i\hbar \frac{\partial \psi}{\partial t} = \left[c\vec{\alpha} \cdot \frac{\hbar}{i} \nabla + \beta mc^2 \right] \psi \tag{14.217}$$

which is the *Dirac equation* for a relativistic spin 1/2 particle.

The form of the result says that the Hamiltonian of a relativistic spin 1/2 particle is

$$\hat{H} = c\vec{\alpha} \cdot \vec{p} + \beta mc^2 \tag{14.218}$$

In the presence of an electromagnetic field we use minimal coupling to get

$$\left(i\hbar \frac{\partial}{\partial t} - e\Phi \right) \psi(\vec{r}, t) = \left[c\vec{\alpha} \cdot \left(\frac{\hbar}{i} \nabla - \frac{e}{c} \vec{A} \right) + \beta mc^2 \right] \psi(\vec{r}, t) \tag{14.219}$$

We note that the vector potential $\vec{A}$ (corresponding to spatial degrees of freedom) is directly coupled to $\vec{\alpha}$ corresponding to internal degrees of freedom).

14.6.1 Nonrelativistic Limit

First, we separate time using

$$\psi(t) = \psi e^{-iEt/\hbar} \tag{14.220}$$

to get

$$(E - e\Phi)\, \psi = \left[c\vec{\alpha} \cdot \left(\frac{\hbar}{i} \nabla - \frac{e}{c} \vec{A} \right) + \beta mc^2 \right] \psi \tag{14.221}$$

We then write

$$\psi = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} \tag{14.222}$$

where ψ_A and ψ_B are still two-component functions and use the explicit Dirac matrices to obtain

$$(E - e\Phi) \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} = \left[\begin{pmatrix} 0 & \vec{\tau} \\ \vec{\tau} & 0 \end{pmatrix} \cdot \left(c\vec{p} - e\vec{A} \right) + \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} mc^2 \right] \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} \tag{14.223}$$

This is equivalent to two coupled equations

$$\vec{\tau} \cdot \left(c\vec{p} - e\vec{A} \right) \psi_B + mc^2 \psi_A = (E - e\Phi)\, \psi_A \tag{14.224}$$

$$\vec{\tau} \cdot \left(c\vec{p} - e\vec{A} \right) \psi_A - mc^2 \psi_B = (E - e\Phi)\, \psi_B \tag{14.225}$$

Letting $E = E' + mc^2$ the second equation of the pair becomes

$$\psi_B = \frac{1}{E' - e\Phi + 2mc^2}\vec{\tau}\cdot\left(c\vec{p} - e\vec{A}\right)\psi_A \tag{14.226}$$

Inserting this result into the first equation of the pair we get

$$\frac{1}{2mc^2}\vec{\tau}\cdot\left(c\vec{p} - e\vec{A}\right)\frac{1}{1+\frac{E'-e\Phi}{2mc^2}}\vec{\tau}\cdot\left(c\vec{p} - e\vec{A}\right)\psi_A = \left(E' - e\Phi\right)\psi_A \tag{14.227}$$

These last two equations are exact and very useful substitutes for the Dirac equation.

We now make some approximations relevant to the nonrelativistic case. We assume that

$$E' << mc^2 \quad , \quad e\Phi << mc^2$$
$$\text{eigenvalues of } \vec{p} \text{ are of order } mv << mc$$

This says that the components satisfy

$$\psi_B \approx \frac{v}{c}\psi_A \tag{14.228}$$

Or that the 4-component wavefunction ψ has two large components ψ_A and two small components ψ_B.

If we ignore terms of order $(v/c)^2$ the equation for ψ_A becomes

$$\frac{1}{2mc^2}\left(\vec{\tau}\cdot\left(c\vec{p} - e\vec{A}\right)\right)^2\psi_A + e\Phi\psi_A = E'\psi_A \tag{14.229}$$

Now, earlier we derived the identity

$$(\vec{\tau}\cdot\vec{a})(\vec{\tau}\cdot\vec{b}) = \vec{a}\cdot\vec{b} + i\vec{\tau}\cdot\left(\vec{a}\times\vec{b}\right) \tag{14.230}$$

We then have

$$\left(\vec{\tau}\cdot\left(c\vec{p} - e\vec{A}\right)\right)^2 = \left(c\vec{p} - e\vec{A}\right)^2 + i\vec{\tau}\cdot\left(\left(c\vec{p} - e\vec{A}\right)\times\left(c\vec{p} - e\vec{A}\right)\right) \tag{14.231}$$

Now

$$\left(\vec{\tau}\cdot\left(c\vec{p} - e\vec{A}\right)\right)^2 = \left(c\vec{p} - e\vec{A}\right)^2 \tag{14.232}$$

and

$$\left(c\vec{p} - e\vec{A}\right)\times\left(c\vec{p} - e\vec{A}\right) = -ec\left(\vec{p}\times\vec{A} + \vec{A}\times\vec{p}\right) = +ie\hbar c\left(\nabla\times\vec{A} + \vec{A}\times\nabla\right) \tag{14.233}$$

Now

$$\begin{aligned}\left(\nabla\times\vec{A} + \vec{A}\times\nabla\right)^i\psi_A &= \varepsilon^{ijk}\left(\partial_j A^k - A^k\partial_j\right)\psi_A \\ &= \varepsilon^{ijk}\left((\partial_j A^k)\psi_A + A^k(\partial_j\psi_A) - A^k(\partial_j\psi_A)\right) \\ &= \varepsilon^{ijk}(\partial_j A^k)\psi_A = (\nabla\times\vec{A})\psi_A = \vec{B}\psi_A\end{aligned} \tag{14.234}$$

Putting everything together we get

$$\frac{1}{2m}\left(\vec{p}-\frac{e}{c}\vec{A}\right)^2\psi_A-\frac{e}{mc}\frac{\hbar}{2}\vec{\tau}\cdot\vec{B}+e\Phi\psi_A=E'\psi_A \tag{14.235}$$

This the *Pauli equation*. The term involving the magnetic field has the form of a magnetic dipole interaction energy

$$-\frac{e}{mc}\vec{S}\cdot\vec{B} \tag{14.236}$$

with a gyromagnetic ratio

$$2\times\frac{e}{mc}\rightarrow g=2 \tag{14.237}$$

The full time-dependent form of the nonrelativistic limit is given by

$$\frac{1}{2m}\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)^2\psi_A-\frac{e}{mc}\frac{\hbar}{2}\vec{\tau}\cdot\vec{B}+(e\Phi+mc^2)\psi_A=i\hbar\frac{\partial\psi_A}{\partial t} \tag{14.238}$$

14.6.2 Currents and Continuity Equations

Going back to the full equation

$$\left(i\hbar\frac{\partial}{\partial t}-e\Phi\right)\psi(\vec{r},t)=\left[c\vec{\alpha}\cdot\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)+\beta mc^2\right]\psi(\vec{r},t) \tag{14.239}$$

we take the Hermitian conjugate to get

$$\left(-i\hbar\frac{\partial}{\partial t}-e\Phi\right)\psi^+(\vec{r},t)=c\left(-\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)\psi^+(\vec{r},t)\cdot\vec{\alpha}+\beta mc^2\psi^+(\vec{r},t) \tag{14.240}$$

Note that the Hermitian conjugate operation reverses matrix order. Now multiply the first equation by $\psi^+(\vec{r},t)$ on the left and the second equation by $\psi(\vec{r},t)$ on the right and subtracting we get the continuity-type equation

$$\frac{\partial(\psi^+\psi)}{\partial t}+\nabla\cdot(\psi^+c\vec{\alpha}\psi)=0 \tag{14.241}$$

This says that the quantity $\psi^+\psi$ is a positive conserved quantity that can be interpreted as a probability density and then

$$\vec{j}=\psi^+c\vec{\alpha}\psi \tag{14.242}$$

is the corresponding probability current. The operator $c\vec{\alpha}$ corresponds to the velocity operator, which is the derivative of the Hamiltonian with respect to $\vec{p}$.

What happens to the Dirac equation under a Lorentz transformation?

In one frame we have

$$i\hbar\beta\frac{\partial\psi}{\partial t}=\beta c\vec{\alpha}\cdot\frac{\hbar}{i}\nabla\psi+mc^2\psi \tag{14.243}$$

and in a new frame we have

$$i\hbar\beta'\frac{\partial\hat{\psi}(\vec{r}\,',t')}{\partial t'} = \beta' c\vec{\alpha}\,'\cdot\frac{\hbar}{i}\nabla'\hat{\psi}(\vec{r}\,',t') + mc^2\hat{\psi}(\vec{r}\,',t') \tag{14.244}$$

where $\hat{\psi}(\vec{r}\,',t')$ is the wave function in the new frame.

We already determined, however, that

$$i\hbar\beta'\frac{\partial}{\partial t'} - \beta' c\vec{\alpha}\,'\cdot\frac{\hbar}{i}\nabla' = i\hbar\beta\frac{\partial}{\partial t} - \beta c\vec{\alpha}\cdot\frac{\hbar}{i}\nabla \tag{14.245}$$

since the scalar product of two 4-vectors is an invariant. This implies that

$$\hat{\psi}(\vec{r}\,',t') = \psi(\vec{r},t) = \text{ Lorentz scalar} \tag{14.246}$$

i.e., they both satisfy the same equation when $\vec{r}\,',t'$ and $\vec{r},t$ are the same space-time point.

It turns out, however, that a more convenient equation to use in the new frame is one that still involves the old β and $\vec{\alpha}$ matrices, i.e., β' and $\vec{\alpha}\,'$ are represented by the same matrices as β and $\vec{\alpha}$. We can find this other equation as follows. We have

$$\begin{aligned}
i\hbar\beta'\frac{\partial\hat{\psi}(\vec{r}\,',t')}{\partial t'} &= \beta' c\vec{\alpha}\,'\cdot\frac{\hbar}{i}\nabla'\hat{\psi}(\vec{r}\,',t') + mc^2\hat{\psi}(\vec{r}\,',t') \\
i\hbar L_v\beta L_v^{-1}\frac{\partial\hat{\psi}(\vec{r}\,',t')}{\partial t'} &= L_v\beta c\vec{\alpha}L_v^{-1}\cdot\frac{\hbar}{i}\nabla'\hat{\psi}(\vec{r}\,',t') + mc^2\hat{\psi}(\vec{r}\,',t') \\
i\hbar\beta L_v^{-1}\frac{\partial\hat{\psi}(\vec{r}\,',t')}{\partial t'} &= \beta c\vec{\alpha}L_v^{-1}\cdot\frac{\hbar}{i}\nabla'\hat{\psi}(\vec{r}\,',t') + mc^2 L_v^{-1}\hat{\psi}(\vec{r}\,',t')
\end{aligned}$$

If we define

$$\psi'(\vec{r}\,',t') = L_v^{-1}\hat{\psi}(\vec{r}\,',t') = L_v^{-1}\psi(\vec{r},t) \tag{14.247}$$

we have the equation

$$i\hbar\beta\frac{\partial\psi'(\vec{r}\,',t')}{\partial t'} = \beta c\vec{\alpha}\cdot\frac{\hbar}{i}\nabla'\psi'(\vec{r}\,',t') + mc^2\psi'(\vec{r}\,',t') \tag{14.248}$$

This form of the equation has the same matrices β and $\vec{\alpha}$ in all frames with the wave function in the new frame related to the wave function in the old frame by the Lorentz transformation.

Alternatively, we can write the Dirac equation in covariant form. The Dirac equation is

$$i\hbar\frac{\partial\psi}{\partial t} = \left[c\vec{\alpha}\cdot\frac{\hbar}{i}\nabla + \beta mc^2\right]\psi \tag{14.249}$$

which we can rewrite as

$$-i\hbar\beta\partial_0\psi - i\hbar\beta\alpha^i\partial_i\psi + mc\psi = 0 \tag{14.250}$$

using the definition of the Dirac gamma matrices we have

$$\left(-i\gamma^\mu\partial_\mu + \frac{mc}{\hbar}\right)\psi = 0 \tag{14.251}$$

which is clearly covariant.

14.6.3 Free Particle Solutions

We start off by constructing solutions for a particle at rest. In this case, we have

$$\psi(\vec{r},t) = e^{-iEt/\hbar}u \tag{14.252}$$

where u is a spinor independent of space and time. Substituting into the Dirac equation we have

$$Eu = \beta mc^2 u \tag{14.253}$$

The eigenvalues of β are ± 1. If u is an eigenstate of β with eigenvalue $+1$, then $E = +mc^2$ and if u is an eigenstate of β with eigenvalue -1, then $E = -mc^2$. So we find negative energy solutions again and we will associate them with particles and antiparticles as before. Their properties in the spin 1/2 case will be different, however.

We choose to write four linearly independent solutions to the free particle Dirac equation as:

$$u_{0\uparrow}^{(+)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad , \quad u_{0\downarrow}^{(+)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad u_{0\downarrow}^{(-)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad u_{0\uparrow}^{(-)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

where the upper index $(\pm)$ denotes the eigenvalue of β, the 0 denotes that the particle is at rest $\vec{p} = 0$, and the arrow denotes the value of the spin associated physically with these states.

The spinors $u_{0\uparrow}^{(+)}$ and $u_{0\downarrow}^{(-)}$ are eigenstates of σ_z with eigenvalue $+1$ and $u_{0\downarrow}^{(+)}$ and $u_{0\uparrow}^{(-)}$ are eigenstates of σ_z with eigenvalue -1.

We are saying here that while $u_{0\uparrow}^{(-)}$ is the spinor of a negative energy particle with spin up, we will associate it with a positive energy antiparticle with spin down.

The states with $\beta = +1$ vary in time as $e^{-imc^2t/\hbar}$ and those with $\beta = -1$ vary in time as $e^{+imc^2t/\hbar}$. The positive and negative states have opposite parity(intrinsic).

We can now construct states for a particle with momentum $\vec{p}$ by starting with

the particle at rest and applying a Lorentz transformation to take us to a frame moving with velocity

$$\vec{v} = -\frac{\vec{p}c^2}{E_p} \text{ where } E_p = +\sqrt{p^2c^2 + m^2c^4} \tag{14.254}$$

We showed earlier that

$$\psi'(\vec{r}\,', t') = L_v^{-1}\psi(\vec{r}, t) = e^{-\vec{\alpha}\cdot\vec{\omega}/2}\psi(\vec{r}, t) = e^{-\vec{\alpha}\cdot\vec{\omega}/2}e^{\mp imc^2t/\hbar}u_{0,\sigma}^{(\pm)} \tag{14.255}$$

Now $E_p't' - \vec{p}\,'\cdot\vec{r}\,'$ is a Lorentz scalar (scalar product of two 4-vectors). In the rest frame it is equal to mc^2t. Therefore we can write

$$e^{\mp imc^2t/\hbar} = e^{\pm i(\vec{p}\,'\cdot\vec{r}\,' - E_p't')/\hbar} \tag{14.256}$$

Dropping the superfluous primes we then have

$$\psi(\vec{r}, t) = e^{\pm i(\vec{p}\cdot\vec{r} - E_pt)/\hbar}e^{-\vec{\alpha}\cdot\vec{\omega}/2}u_{0,\sigma}^{(\pm)} \tag{14.257}$$

as the wave function for nonzero momentum. The new spinors are given by

$$u_{\vec{p},\sigma}^{(\pm)} = e^{-\vec{\alpha}\cdot\vec{\omega}/2}u_{0,\sigma}^{(\pm)} = \left[\cosh\frac{\omega}{2} - \vec{\alpha}\cdot\hat{v}\sinh\frac{\omega}{2}\right]u_{0,\sigma}^{(\pm)} \tag{14.258}$$

Using

$$\vec{v} = -\frac{\vec{p}c^2}{E_p} \tag{14.259}$$

we get

$$\cosh\frac{\omega}{2} = \sqrt{\frac{E_p + mc^2}{2mc^2}} \quad , \quad \hat{v}\tanh\frac{\omega}{2} = -\frac{\vec{p}c}{E_p + mc^2} \tag{14.260}$$

so that

$$u_{\vec{p},\sigma}^{(\pm)} = \sqrt{\frac{E_p + mc^2}{2mc^2}}\left[1 + \frac{c\vec{p}\cdot\vec{\alpha}}{E_p + mc^2}\right]u_{0,\sigma}^{(\pm)} \tag{14.261}$$

We then have (in the standard representation)

$$u_{\vec{p},\uparrow}^{(+)} = \sqrt{\frac{E_p + mc^2}{2mc^2}}\left[1 + \frac{c}{E_p + mc^2}\vec{p}\cdot\vec{\alpha}\right]u_{0,\uparrow}^{(+)} \tag{14.262}$$

Now

$$\vec{p}\cdot\vec{\alpha} = p_x\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} + p_y\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$
$$+ p_z\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad , \quad u_{0,\uparrow}^{(+)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and we get

$$u^{(+)}_{\vec{p},\uparrow} = \sqrt{\frac{E_p+mc^2}{2mc^2}}\left[\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + \frac{c}{E_p+mc^2}\left(p_x\begin{pmatrix}0\\0\\0\\1\end{pmatrix} + p_y\begin{pmatrix}0\\0\\0\\i\end{pmatrix} + p_z\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right)\right]$$

$$= \sqrt{\frac{E_p+mc^2}{2mc^2}}\begin{pmatrix}1\\0\\\frac{cp_z}{E_p+mc^2}\\\frac{c(p_x+ip_y)}{E_p+mc^2}\end{pmatrix} \tag{14.263}$$

and similarly

$$u^{(+)}_{\vec{p},\downarrow} = \sqrt{\frac{E_p+mc^2}{2mc^2}}\begin{pmatrix}0\\1\\\frac{c(p_x-ip_y)}{E_p+mc^2}\\-\frac{cp_z}{E_p+mc^2}\end{pmatrix} \tag{14.264}$$

$$u^{(-)}_{\vec{p},\downarrow} = \sqrt{\frac{E_p+mc^2}{2mc^2}}\begin{pmatrix}\frac{cp_z}{E_p+mc^2}\\\frac{c(p_x+ip_y)}{E_p+mc^2}\\1\\0\end{pmatrix} \tag{14.265}$$

$$u^{(-)}_{\vec{p},\uparrow} = \sqrt{\frac{E_p+mc^2}{2mc^2}}\begin{pmatrix}\frac{c(p_x-ip_y)}{E_p+mc^2}\\-\frac{cp_z}{E_p+mc^2}\\0\\1\end{pmatrix} \tag{14.266}$$

Remember that the arrow refers to the spin associated with the state in the rest frame, which is minus the σ_z eigenvalue for the $(-)$ spinors. We see that a particle in a σ_z eigenstate in its rest frame appears to be in a σ_z eigenstate to an observer moving with respect to the particle only if the observer is moving along the $z-$direction, i.e., if $p_x = p_y = 0$ we have

$$u^{(+)}_{\vec{p},\uparrow} = \sqrt{\frac{E_p+mc^2}{2mc^2}}\begin{pmatrix}1\\0\\\frac{cp_z}{E_p+mc^2}\\0\end{pmatrix} = \sqrt{\frac{E_p+mc^2}{2mc^2}}\left[u^{(+)}_{0,\uparrow} + \frac{cp_z}{E_p+mc^2}u^{(-)}_{0,\downarrow}\right] \tag{14.267}$$

which is a sum of a particle and an antiparticle where both have spin up!

The positive energy solutions $u^{(+)}_{\vec{p}\sigma}e^{i(\vec{p}\cdot\vec{r}-E_pt)/\hbar}$ correspond to particles with momentum $\vec{p}$, energy E_p and spin orientation σ. The negative energy solutions

$u^{(-)}_{\vec{p}\sigma} e^{-i(\vec{p}\cdot\vec{r}-E_p t)/\hbar}$ correspond to particles with momentum $-\vec{p}$, energy $-E_p$ and spin orientation $-\sigma$ which we will soon associate with antiparticles with momentum $\vec{p}$, energy E_p and spin orientation σ.

The nonzero momentum spinors are orthogonal but not normalized to one (as is the case with the zero momentum spinors). Since $L^+ \neq L^{-1}$, in general, the Lorentz transformations are not represented by a unitary operator and hence the lengths of vectors or normalizations change. In particular. The normalization is given by

$$u^{(\pm)+}_{\vec{p}\sigma} u^{(\pm)}_{\vec{p}\sigma} = \frac{E_p}{mc^2} \tag{14.268}$$

Since $\vec{\alpha}$ is Hermitian, we have $L^+ = L$.

Thus, if $u_{\vec{p}} = L^{-1}u_0$, then $(u_{\vec{p}})^+ = (u_0)^+(L^{-1})^+ = (u_0)^+L^{-1}$ and $u^+_{\vec{p}}u_{\vec{p}} = u^+_0(L^{-1})^2u_0$.

It is possible to define a normalization that is invariant under a Lorentz transformation. Since β anticommutes with $\vec{\alpha}$, we can write

$$(L^{-1})^+\beta = L^{-1}\beta = e^{\vec{\alpha}\cdot\vec{\omega}/2}\beta = \beta e^{-\vec{\alpha}\cdot\vec{\omega}/2} = \beta L \tag{14.269}$$

Now, if the spinor u transforms as $u' = L^{-1}u$, then the spinor $\bar{u} = u^+\beta$ is given in the new frame by

$$\bar{u}' = u'^+\beta' = u'^+\beta = u^+(L^{-1})^+\beta = u^+L^{-1}\beta = u^+\beta L = \bar{u}L \tag{14.270}$$

This means that the product $\bar{u}_1u_2$ of any two spinors is a Lorentz invariant, i.e.,

$$\bar{u}'_1u'_2 = (\bar{u}_1L)(L^{-1}u_2) = \bar{u}_1u_2 \tag{14.271}$$

In the rest frame

$$u^{(b)+}_{0\sigma}u^{(b')}_{0\sigma'} = b\delta_{bb'}\delta_{\sigma\sigma'} \quad , \quad b = \pm \tag{14.272}$$

which says that the same relation is true for all momentum $\vec{p}$

$$u^{(b)+}_{\vec{p}\sigma}u^{(b')}_{\vec{p}\sigma'} = b\delta_{bb'}\delta_{\sigma\sigma'} \quad , \quad b = \pm \tag{14.273}$$

The spinors $u^{(\pm)}_{\vec{p}\sigma}$ obey the completeness relation that says that the 4×4 identity matrix can be written as the sum of the outer products of the four spinors, i.e.,

$$\sum_{b,\sigma} bu^{(b)}_{\vec{p}\sigma}\bar{u}^{(b)}_{\vec{p}\sigma} = 1 \tag{14.274}$$

The spinors $u^{(\pm)}_{\vec{p}\sigma}$ obey

$$(\beta E_p - c\beta\vec{\alpha}\cdot\vec{p})u^{(\pm)}_{\vec{p}\sigma} = \pm mc^2u^{(\pm)}_{\vec{p}\sigma} \tag{14.275}$$

and

$$u^{(\pm)+}_{\vec{p}\sigma}(\beta E_p - c\vec{\alpha}\cdot\vec{p}\beta) = \pm mc^2u^{(\pm)+}_{\vec{p}\sigma} \tag{14.276}$$

Multiplying the last equation on the right by β we have the equation satisfied by $\bar{u}^{(\pm)}_{\vec{p}\sigma}$

$$\bar{u}^{(\pm)}_{\vec{p}\sigma}(\beta E_p - c\beta\vec{\alpha}\cdot\vec{p}) = \pm mc^2\bar{u}^{(\pm)}_{\vec{p}\sigma} \tag{14.277}$$

14.6.4 More About Currents

As we have seen, the solution of the Dirac equation $\psi(\vec{r},t)$ has the following behavior under a Lorentz transformation

$$\psi(\vec{r},t) \to \psi'(\vec{r}',t') = L^{-1}\psi(\vec{r},t) \tag{14.278}$$

The spinor $\bar{\psi}(\vec{r},t) = \psi^+(\vec{r},t)\beta$ transforms like

$$\bar{\psi}(\vec{r},t) \to \bar{\psi}'(\vec{r}',t') = [L^{-1}\psi(\vec{r},t)]^+\beta = \bar{\psi}(\vec{r},t)L \tag{14.279}$$

This says that the product $\bar{\psi}(\vec{r},t)\psi(\vec{r},t)$ transforms like a Lorentz scalar.

Now the γ^μ transform as the components of a 4-vector, i.e., $\gamma^\mu \to L\gamma^\mu L^{-1}$. Therefore, the product

$$\bar{\psi}(\vec{r},t)\gamma^\mu\psi(\vec{r},t) = \left(\rho(\vec{r},t), \frac{1}{c}\vec{j}(\vec{r},t)\right) \tag{14.280}$$

transforms like a 4-vector under a Lorentz transformation. It is the particle 4-current multiplied by $1/c$. In the same manner,

$$\begin{aligned}
&\bar{\psi}\sigma^{\mu\nu}\psi \to \text{ second -rank tensor}\\
&\bar{\psi}\gamma_5\gamma^\mu\psi \to \text{ axial vector}\\
&\bar{\psi}\gamma_5\psi \to \text{ pseudo - scalar}
\end{aligned}$$

The positive density $\rho(\vec{r},t) = \psi^+(\vec{r},t)\psi(\vec{r},t)$ and the current $\vec{j}(\vec{r},t) = c\psi^+(\vec{r},t)\vec{\alpha}\psi(\vec{r},t)$ satisfy the continuity equation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\vec{j} = 0 \tag{14.281}$$

which implies that the quantity

$$\int \rho(\vec{r},t)d^3r \tag{14.282}$$

is a constant of the motion.

In this case, we can interpret $\rho(\vec{r},t)$ as a probability density (same as in non-relativistic case). Remember in the spin zero case this was not so since the corresponding conserved density needed to be interpreted as a charge density which could be positive and negative.

One important consequence in the spin zero case was that it is impossible for a particle to make a transition from a state normalized to +1 to a state normalized to −1 since the normalization remains constant in time. We associated the negative energy states with particles and the negatively normalized states with antiparticles. We then see that the impossibility of a transition between

positive and negative energy states just corresponds to charge conservation.

In the spin−1/2 case, however, both positive and negative energy states have positive normalization so that there is nothing in the theory (so far) that prevents a particle in a positive energy state from making a transition to a negative energy state radiating away several high energy photons in the process. A difficulty in the theory that we must return to later!

Let us say some more about the position and velocity operators in the Dirac theory.

The position operator has strange features similar to those of the Klein-Gordon theory. If we apply the position operator to a wave packet made up of positive energy free particle states we get

$$\begin{aligned}\vec{r}\psi^{(+)}(\vec{r}) &= \vec{r}\left(\sum_\sigma \int \frac{d^3p}{(2\pi\hbar)^3} a_{\vec{p}\sigma} u^{(+)}_{\vec{p}\sigma} e^{i\vec{p}\cdot\vec{r}/\hbar}\right) = \sum_\sigma \int \frac{d^3p}{(2\pi\hbar)^3} a_{\vec{p}\sigma} u^{(+)}_{\vec{p}\sigma} \frac{\hbar}{i}\nabla_{\vec{p}} e^{i\vec{p}\cdot\vec{r}/\hbar} \\ &= \sum_\sigma \int \frac{d^3p}{(2\pi\hbar)^3} (i\hbar\nabla_{\vec{p}} a_{\vec{p}\sigma}) u^{(+)}_{\vec{p}\sigma} e^{i\vec{p}\cdot\vec{r}/\hbar} + \sum_\sigma \int \frac{d^3p}{(2\pi\hbar)^3} a_{\vec{p}\sigma} (i\hbar\nabla_{\vec{p}} u^{(+)}_{\vec{p}\sigma}) e^{i\vec{p}\cdot\vec{r}/\hbar}\end{aligned}$$

where we have integrated by parts to get the last two terms. The first term contains only positive energy components. The second term, however, contains the factor $i\hbar\nabla_{\vec{p}} u^{(+)}_{\vec{p}\sigma}$, which generates both positive and negative components (explicitly do the derivatives on the column vectors we derived earlier). If we define, as before

$$\vec{r} = \vec{r}_{(+)} + \vec{r}_{(-)} \tag{14.283}$$

then, as before, the even part $\vec{r}_{(+)}$ acting on the wave packet of positive energy free particle states produces only positive energy free particle states and acting on the wave packet of negative energy free particle states produces only negative energy free particle states, while the odd part $\vec{r}_{(-)}$ turns positive positive energy states to negative energy states and vice versa.

As in the Klein-Gordon case, both positive and negative energy free particle solutions are needed to produce a localized wave packet.

Looking at the current expression $\vec{j}(\vec{r},t) = c\psi^+(\vec{r},t)\vec{\alpha}\psi(\vec{r},t)$ we see that the operator $c\vec{\alpha}$ acts as a velocity operator. This interpretation also agrees with the commutator relation

$$-i\hbar\left[\vec{r},\hat{H}\right] = c\vec{\alpha} \tag{14.284}$$

which leads to the Heisenberg representation operator equation

$$\frac{d\vec{r}}{dt} = c\vec{\alpha} \tag{14.285}$$

If, however, we consider the z−component of this velocity operator we get $(c\alpha_z)^2 = c^2\alpha_z^2 = c^2$. Thus the eigenvalues of each component of the velocity

operator are $\pm c$, which says that a particle in an eigenstate of the velocity operator travels at the speed of light!

This means that the velocity operator is not simply related to the momentum operator relativistically. The eigenstates of any component of $\vec{\alpha}$ are linear combinations of positive and negative energy free particle states and thus cannot be realized in any physical situation! For any arbitrary state the expectation value of $c\vec{\alpha}$ has a magnitude between 0 and c.

14.6.5 Non-relativistic Limit

We now derive corrections to the Pauli equation. Earlier we had

$$\vec{\tau}\cdot\left(c\vec{p}-e\vec{A}\right)\psi_B + mc^2\psi_A = (E-e\Phi)\,\psi_A \tag{14.286}$$

$$\vec{\tau}\cdot\left(c\vec{p}-e\vec{A}\right)\psi_A - mc^2\psi_B = (E-e\Phi)\,\psi_B \tag{14.287}$$

or

$$\vec{\tau}\cdot\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)\psi_B + mc\psi_A = \frac{1}{c}\left(i\hbar\frac{\partial}{\partial t}-e\Phi\right)\psi_A \tag{14.288}$$

$$\vec{\tau}\cdot\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)\psi_A - mc\psi_B = \frac{1}{c}\left(i\hbar\frac{\partial}{\partial t}-e\Phi\right)\psi_B \tag{14.289}$$

The second equation of the above pair gives (an exact equation)

$$\psi_B = \frac{1}{2mc}\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)\cdot\vec{\tau}\psi_A - \frac{1}{2mc^2}\left(i\hbar\frac{\partial}{\partial t}-mc^2-e\Phi\right)\psi_B \tag{14.290}$$

Now the ψ_A term is much larger than the ψ_B term on the right. Thus, we get the first correction by iterating once, i.e.

$$\begin{aligned}\psi_B = {} & \frac{1}{2mc}\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)\cdot\vec{\tau}\psi_A \\ & - \frac{1}{4m^2c^3}\left(i\hbar\frac{\partial}{\partial t}-mc^2-e\Phi\right)\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)\cdot\vec{\tau}\psi_A\end{aligned} \tag{14.291}$$

Substituting this expression into the first equation of the pair we get the first relativistic correction term to the Pauli equation

$$-\frac{1}{4m^2c^3}\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)\cdot\vec{\tau}\left(i\hbar\frac{\partial}{\partial t}-mc^2-e\Phi\right)\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)\cdot\vec{\tau}\psi_A \tag{14.292}$$

which is $\approx (v/c)^2$ smaller than the kinetic energy term $p^2/2m$.

The correction term can be rewritten as

$$\begin{aligned}& -\frac{1}{4m^2c^3}\left(\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)\cdot\vec{\tau}\right)^2\left(i\hbar\frac{\partial}{\partial t}-mc^2-e\Phi\right)\psi_A \\ & \quad -\frac{ie\hbar}{4m^2c^3}\left(\left(\frac{\hbar}{i}\nabla-\frac{e}{c}\vec{A}\right)\cdot\vec{\tau}\right)(\vec{\varepsilon}\cdot\vec{\tau})\,\psi_A\end{aligned}$$

where

$$\vec{\varepsilon} = -\nabla\Phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} = \text{ electric field} \tag{14.293}$$

To lowest order in (v/c) we have

$$\left(i\hbar\frac{\partial}{\partial t} - mc^2 - e\Phi\right)\psi_A = \frac{p^2}{2m}\psi_A \tag{14.294}$$

Using this relation with the identity

$$(\vec{\tau}\cdot\vec{a})(\vec{\tau}\cdot\vec{b}) = \vec{a}\cdot\vec{b} + i\vec{\tau}\cdot\left(\vec{a}\times\vec{b}\right) \tag{14.295}$$

the correction becomes

$$-\left[\frac{p^4}{8m^3c^2} + \frac{e\hbar}{4m^2c^2}\vec{\tau}\cdot(\vec{\varepsilon}\times\vec{p}) + \frac{ie\hbar}{4m^2c^2}\vec{p}\cdot\vec{\varepsilon}\right] \tag{14.296}$$

The first term is the relativistic correction to the kinetic energy. The second term is the spin-orbit coupling. The third term is new and is not even Hermitian!

The reason for this non-Hermitian term is that we are only working to order $(v/c)^2$. Such a non-Hermitian term in the wave equation means that the normalization integral

$$\int \psi_A^+\psi_A d^3r$$

can change in time. Now the full Dirac equation obeys the normalization condition

$$\int \psi^+\psi d^3r = \int [\psi_A^+\psi_A + \psi_B^+\psi_B]d^3r = 1 \tag{14.297}$$

To lowest order, however,

$$\psi_B = \frac{\hbar}{2imc}\nabla\cdot\vec{\tau}\psi_A \rightarrow \psi_B^+\psi_B = \psi_A^+\frac{p^2}{4m^2c^2}\psi_A \tag{14.298}$$

Thus, the integral stays constant to order $(v/c)^2$. It is the integral

$$\int \psi_A^+\left[1 + \frac{p^2}{4m^2c^2}\right]\psi_A d^3r = \int \left[\left[1 + \frac{p^2}{8m^2c^2}\right]\psi_A\right]^+\left[\left[1 + \frac{p^2}{8m^2c^2}\right]\psi_A\right]d^3r$$

to order $(v/c)^2$, that remains constant and equal to 1. This implies that the correct nonrelativistic limit of the Dirac wave function (the limit whose normalization remains constant in time) is

$$\psi(\vec{r},t) = \left[1 + \frac{p^2}{8m^2c^2}\right]\psi_A(\vec{r},t) \rightarrow \int \psi^+\psi d^3r = 1 \tag{14.299}$$

The equation for this form of the wave function will not have any non-Hermitian terms. A large amount of algebra gives the equation for $\psi(\vec{r},t)$ as

$$i\hbar\frac{\partial\psi}{\partial t} = \left[mc^2 + \frac{1}{2m}\left(\vec{p} - \frac{e}{c}\vec{A}\right)^2 - \frac{p^4}{8m^3c^2}\right]\psi \tag{14.300}$$
$$-\left[\frac{e\hbar}{2mc}\vec{\tau}\cdot\vec{B} + \frac{e\hbar}{4m^2c^2}\vec{\tau}\cdot(\vec{\varepsilon}\times\vec{p})\right]\psi + \left[e\Phi + \frac{\hbar^2}{8m^2c^2}(\nabla^2 e\Phi)\right]$$

This is the correct nonrelativistic limit of the Dirac equation. All terms are Hermitian.

The terms on the right-hand side are

[rest energy + kinetic energy(to order $(v/c)^2$)]
– [Pauli magnetic moment energy + spin-orbit energy]
+ [correction to the potential energy term]

Spin-Orbit Term - Letting $\vec{A} = 0$ for simplicity we have

$$\frac{e\hbar}{4m^2c^2}\vec{\tau}\cdot(\vec{\varepsilon}\times\vec{p}) = -\frac{e\hbar}{4m^2c^2}\vec{\tau}\cdot(\nabla\Phi\times\vec{p}) \tag{14.301}$$

If we assume the potential is spherically symmetric, then

$$\nabla\Phi = \frac{1}{r}\frac{d\Phi}{dr}\vec{r} \tag{14.302}$$

and we get

$$\frac{e\hbar}{4m^2c^2}\vec{\tau}\cdot(\vec{\varepsilon}\times\vec{p}) = -\frac{e\hbar}{4m^2c^2r}\frac{d\Phi}{dr}\vec{\tau}\cdot(\vec{r}\times\vec{p}) = -\frac{e}{2m^2c^2r}\frac{d\Phi}{dr}\vec{S}\cdot\vec{L} \tag{14.303}$$

which is the spin-orbit energy. It correctly contains the Thomas precession correction! We do not have to add any terms in an ad hoc manner!

Correction to the Potential - This is called the *Darwin term*. Now, from Poisson's equation we have

$$\nabla^2 e\Phi(\vec{r}) = -4\pi e Q(\vec{r}) \quad , \quad Q(\vec{r}) = \text{ charge density producing } \Phi(\vec{r})$$

For a Coulomb potential we get

$$\frac{\hbar^2}{8m^2c^2}(\nabla^2 e\Phi) = \frac{\pi\hbar^2}{2m^2c^2}Ze^2\delta(\vec{r}) \tag{14.304}$$

This term tends to raise the energy of s-states since they do not vanish at the origin.

14.6.6 The Dirac Hydrogen Atom

We start with the equations

$$\vec{\sigma}\cdot\left(c\vec{p}-e\vec{A}\right)\psi_B+mc^2\psi_A=\left(E-e\Phi\right)\psi_A \tag{14.305}$$

$$\vec{\sigma}\cdot\left(c\vec{p}-e\vec{A}\right)\psi_A-mc^2\psi_B=\left(E-e\Phi\right)\psi_B \tag{14.306}$$

where we have substituted $\vec{\sigma}$ for $\vec{\tau}$. The potential function is $e\Phi=-Ze^2/r$ and we let $\vec{A}=0$. Writing

$$\psi_A=\begin{pmatrix}u_1\\u_2\end{pmatrix}\quad,\quad\psi_B=\begin{pmatrix}u_3\\u_4\end{pmatrix} \tag{14.307}$$

we get

$$-\frac{i}{\hbar c}\left[E+\frac{Ze^2}{r}-mc^2\right]u_1+\frac{\partial u_4}{\partial x}-i\frac{\partial u_4}{\partial y}+\frac{\partial u_3}{\partial z}=0 \tag{14.308}$$

$$-\frac{i}{\hbar c}\left[E+\frac{Ze^2}{r}-mc^2\right]u_2+\frac{\partial u_3}{\partial x}+i\frac{\partial u_3}{\partial y}-\frac{\partial u_4}{\partial z}=0 \tag{14.309}$$

$$-\frac{i}{\hbar c}\left[E+\frac{Ze^2}{r}+mc^2\right]u_3+\frac{\partial u_2}{\partial x}-i\frac{\partial u_2}{\partial y}+\frac{\partial u_2}{\partial z}=0 \tag{14.310}$$

$$-\frac{i}{\hbar c}\left[E+\frac{Ze^2}{r}+mc^2\right]u_4+\frac{\partial u_1}{\partial x}+i\frac{\partial u_1}{\partial y}-\frac{\partial u_2}{\partial z}=0 \tag{14.311}$$

We now use another clever trick I learned from Professor Hans Bethe at Cornell University to find a solution.

If we consider only large components, i.e., set the small components to zero, then $\left[\vec{L},\hat{H}\right]$, which is proportional to $\vec{\alpha}\times\vec{p}$, will be zero, since $\vec{\alpha}$ connects the small and large components. This means that ψ_A will be an eigenfunction of $\vec{L}$. In addition, it must contain one spin component with spin up and another with spin down.

Of course, $\vec{j}$ and j_z are constants of the motion. Hence, for $j=\ell+1/2$ we can set

$$u_1=g(r)\sqrt{\frac{\ell+m+\frac{1}{2}}{2\ell+1}}Y_\ell^{m-\frac{1}{2}}(\Omega) \tag{14.312}$$

$$u_2=-g(r)\sqrt{\frac{\ell-m+\frac{1}{2}}{2\ell+1}}Y_\ell^{m+\frac{1}{2}}(\Omega) \tag{14.313}$$

where the unknown function $g(r)$ will be the solution of some relativistic radial equation.

To get the small components we recall the equation

$$\psi_B=\frac{1}{E'-e\Phi+2mc^2}\vec{\tau}\cdot\left(c\vec{p}-e\vec{A}\right)\psi_A \tag{14.314}$$

and note that the operator which gives the small component from the large component has odd parity ($\vec{p}$ is odd, $\vec{A}$ = 0 and everything else is even) and commutes with $\vec{j}$. Hence, ψ_B must belong to the same j value as ψ_A but must have a different ℓ.

Corresponding to $j = \ell + 1/2$ the only other possible value of the orbital angular momentum is $\ell' = \ell+1$. Therefore, we set (remembering the appropriate Clebsch-Gordon coefficients)

$$u_3 = if(r)\sqrt{\frac{\ell - m + \frac{3}{2}}{2\ell + 3}}Y_{\ell+1}^{m-\frac{1}{2}}(\Omega) \tag{14.315}$$

$$u_4 = -if(r)\sqrt{\frac{\ell + m + \frac{3}{2}}{2\ell + 3}}Y_{\ell+1}^{m+\frac{1}{2}}(\Omega) \tag{14.316}$$

where the unknown function $f(r)$ will be the solution of some relativistic radial equation. Inserting these solution guesses into the 4 coupled equations we find that for $j = \ell + 1/2$ the connection between f and g is given by

$$\frac{1}{\hbar c}\left[E + \frac{Ze^2}{r} + mc^2\right]f = \frac{dg}{dr} - \ell\frac{g}{r} \tag{14.317}$$

$$\frac{1}{\hbar c}\left[E + \frac{Ze^2}{r} - mc^2\right]g = -\frac{df}{dr} - (\ell + 2)\frac{f}{r} \tag{14.318}$$

In an analogous way for $j = \ell - 1/2$ we have

$$u_1 = g(r)\sqrt{\frac{\ell - m + \frac{1}{2}}{2\ell + 1}}Y_{\ell}^{m-\frac{1}{2}}(\Omega) \tag{14.319}$$

$$u_2 = g(r)\sqrt{\frac{\ell + m + \frac{1}{2}}{2\ell + 1}}Y_{\ell}^{m+\frac{1}{2}}(\Omega) \tag{14.320}$$

$$u_3 = -if(r)\sqrt{\frac{\ell + m - \frac{1}{2}}{2\ell - 1}}Y_{\ell-1}^{m-\frac{1}{2}}(\Omega) \tag{14.321}$$

$$u_4 = if(r)\sqrt{\frac{\ell - m - \frac{1}{2}}{2\ell - 1}}Y_{\ell-1}^{m+\frac{1}{2}}(\Omega) \tag{14.322}$$

and

$$\frac{1}{\hbar c}\left[E + \frac{Ze^2}{r} + mc^2\right]f = \frac{dg}{dr} + (\ell + 1)\frac{g}{r} \tag{14.323}$$

$$\frac{1}{\hbar c}\left[E + \frac{Ze^2}{r} - mc^2\right]g = -\frac{df}{dr} + (\ell - 1)\frac{f}{r} \tag{14.324}$$

We now define

$$k = \begin{cases} -(\ell + 1) & \text{if} \quad j = \ell + 1/2 \\ \ell & \text{if} \quad j = \ell + 1/2 \end{cases} \tag{14.325}$$

i.e.,

$$k = \begin{cases} -1, -2, ... & \text{if} \quad j = \ell + 1/2 \\ 1, 2, ... & \text{if} \quad j = \ell + 1/2 \end{cases} \tag{14.326}$$

We can then combine the 4 equations for f and g into 2 equations as

$$\frac{1}{\hbar c}\left[E + \frac{Ze^2}{r} + mc^2\right] f - \left(\frac{dg}{dr} + (1+k)\frac{g}{r}\right) = 0 \tag{14.327}$$

$$\frac{1}{\hbar c}\left[E + \frac{Ze^2}{r} - mc^2\right] g + \left(\frac{df}{dr} + (1-k)\frac{f}{r}\right) = 0 \tag{14.328}$$

Setting

$$F = rf \quad , \quad G = rg$$
$$\alpha_1 = \frac{mc^2 + E}{\hbar c} \quad , \quad \alpha_2 = \frac{mc^2 - E}{\hbar c}$$
$$\alpha = (\alpha_1 \alpha_2)^{1/2} \quad , \quad \gamma = \frac{Ze^2}{\hbar c} \quad , \quad \rho = \alpha r$$

we get

$$\left(\frac{d}{d\rho} + \frac{k}{\rho}\right) G - \left(\frac{\alpha_1}{\alpha} + \frac{\gamma}{\rho}\right) F = 0 \tag{14.329}$$

$$\left(\frac{d}{d\rho} - \frac{k}{\rho}\right) F - \left(\frac{\alpha_2}{\alpha} - \frac{\gamma}{\rho}\right) G = 0 \tag{14.330}$$

We now solve these coupled equations using the standard series method to obtain the positive energy bound state solutions.

We substitute

$$F = \phi(\rho) e^{-\rho} \quad , \quad F = \chi(\rho) e^{-\rho} \tag{14.331}$$

and obtain

$$\chi' - \chi + \frac{k}{\rho}\chi - \left(\frac{\alpha_1}{\alpha} + \frac{\gamma}{\rho}\right)\phi = 0 \tag{14.332}$$

$$\phi' - \phi - \frac{k}{\rho}\phi - \left(\frac{\alpha_2}{\alpha} - \frac{\gamma}{\rho}\right)\chi = 0 \tag{14.333}$$

We now substitute the series

$$\phi = \rho^s \sum_{m=0}^{\infty} a_m \rho^m \quad , \; a_0 \neq 0 \quad , \quad \chi = \rho^s \sum_{m=0}^{\infty} b_m \rho^m \quad , \; b_0 \neq 0 \tag{14.334}$$

the requirement that f and g be finite everywhere turns out to be impossible to satisfy. Instead, we require that the integrated probability density be finite, i.e.,

$$\int_0^{\infty} \left[|F(\rho)|^2 + |G(\rho)|^2\right] d\rho < \infty \tag{14.335}$$

This makes sure that $s \neq -\infty$. Substituting the series and equating coefficients of the same power of ρ we get the recursion relations

$$(s+\nu+k)b_\nu - b_{\nu-1} - \gamma a_\nu - \frac{\alpha_1}{\alpha}a_{\nu-1} = 0 \tag{14.336}$$

$$(s+\nu-k)a_\nu - a_{\nu-1} + \gamma b_\nu - \frac{\alpha_2}{\alpha}b_{\nu-1} = 0 \tag{14.337}$$

For $\nu = 0$ we get

$$(s+k)b_0 - \gamma a_0 = 0 = (s-k)a_0 + \gamma b_0 \tag{14.338}$$

These equations have a nontrivial solution if and only if

$$s = \pm(k^2 - \gamma^2)^{1/2} \tag{14.339}$$

First we look at the negative root. For small ρ the integrand for the integrated probability density is $\sim \rho^{2s}$ and we must have $2s > -1$ or $(k^2-\gamma^2)^{1/2} > 1/2$. The minimum s occurs when $k^2 = 1$. This corresponds to $Z \geq 109$. For $k^2 > 1$, no value of Z will permit the negative root.

Restricting ourselves to $Z < 109$, we choose the positive root $s = (k^2-\gamma^2)^{1/2}$. For $k = 1$, $s < 1$, f and g diverge at the origin. The probability density integral converges, however.

The recursion relations lead to function of the order $e^{2\rho}$ (the probability density integral would diverge) unless the series terminate. Suppose the series terminate for $\nu = n'$, i.e., $a_{n'+1} = b_{n'+1} = 0$. We then have from the recursion relations that

$$\alpha_1 a_{n'} = -\alpha b_{n'} \quad , \quad n' = 0, 1, 2, ... \tag{14.340}$$

We now multiply the first recursion relation by α and the second by α_1 and subtract them to get

$$b_\nu[\alpha(s+\nu+k) - \alpha_1\gamma] = a_\nu[\alpha_1(s+\nu-k) + \alpha\gamma] \tag{14.341}$$

Inserting $\nu = n'$ and using $\alpha_1 a_{n'} = -\alpha b_{n'}$ we get

$$2\alpha(s+n') = \gamma(\alpha_1 - \alpha_2) = \frac{2E\gamma}{\hbar c} \tag{14.342}$$

Putting everything together we get

$$E = mc^2\left[1 + \frac{\gamma^2}{(s+n')^2}\right]^{-1/2} = mc^2\left[1 + \frac{\gamma^2}{\left(n' + \sqrt{k^2-\gamma^2}\right)^2}\right]^{-1/2} \tag{14.343}$$

Since $|k| = j + \frac{1}{2}$ we get

$$E = mc^2\left[1 + \frac{\gamma^2}{\left(n' + \sqrt{(j+\frac{1}{2})^2-\gamma^2}\right)^2}\right]^{-1/2} \quad , \; n' = 0, 1, 2... \quad j + \tfrac{1}{2} = 1, 2, 3, ... \tag{14.344}$$

where $\gamma = Ze^2/\hbar c$.

Before looking at the physics in this result let us investigate an alternative approach involving a second-order Dirac Equation. The first-order Dirac equation is

$$\beta\left(i\hbar\frac{\partial}{\partial t} - \hat{H}\right)\psi(\vec{r},t) = 0 \quad , \quad \hat{H} = c\vec{\alpha}\cdot\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right) + \beta mc^2 + e\Phi \tag{14.345}$$

We now define the projection operator $\hat{P}$ as

$$\hat{P} = \frac{\beta\left(i\hbar\frac{\partial}{\partial t} - \hat{H}\right) + 2mc^2}{2mc^2} \tag{14.346}$$

and operate on the Dirac equation from the left. After some algebra we get

$$\left[\frac{1}{c^2}\left(i\hbar\frac{\partial}{\partial t} - e\Phi\right)^2 - \left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)^2 - m^2c^2 + \frac{e\hbar}{c}\left(\vec{\sigma}\cdot\vec{B} - i\vec{\alpha}\cdot\vec{\varepsilon}\right)\right]\psi = 0 \tag{14.347}$$

where we have used the relations

$$\left[\vec{\alpha}\cdot\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)\right]^2 = \left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}\right)^2 - \frac{e\hbar}{c}\vec{\sigma}\cdot\vec{B} \tag{14.348}$$

and

$$\left[\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}, i\hbar\frac{\partial}{\partial t} - e\Phi\right] = -i\hbar e\vec{\varepsilon} \tag{14.349}$$

The new second-order equation is just the Klein-Gordon equation with an additional term $\left(\vec{\sigma}\cdot\vec{B} - i\vec{\alpha}\cdot\vec{\varepsilon}\right)$, which represents the direct coupling of the electromagnetic fields to the magnetic(and electric) moments of the particle.

Every solution of the Dirac equation is a solution of this new second-order equation, but every solution of the second-order equation in not necessarily a solution of the Dirac equation.

If, however, ψ is a solution of the second-order equation, then $\phi = \hat{P}\psi$ is a solution of the Dirac equation. We can see this as follows. The second-order equation can be written as

$$\hat{P}\beta\left(i\hbar\frac{\partial}{\partial t} - \hat{H}\right)\psi(\vec{r},t) = \beta\left(i\hbar\frac{\partial}{\partial t} - \hat{H}\right)\hat{P}\psi(\vec{r},t) = 0 \tag{14.350}$$

or the second order equation is equivalent to

$$\left(i\hbar\frac{\partial}{\partial t} - \hat{H}\right)\psi(\vec{r},t) = \left(i\hbar\frac{\partial}{\partial t} - \hat{H}\right)\phi(\vec{r},t) = 0 \tag{14.351}$$

This says that $\hat{P}$ acts as a projection operator, which reduces solutions of the second-order equation to solutions of the first-order Dirac equation.

Let us now use the second-order equation to find the energy levels of the Dirac hydrogen atom (Glauber, et al PR **109**,1307(1958)). For a stationary state of energy E in the Coulomb potential the second-order equation becomes

$$\left[\frac{1}{c^2}\left(E+\frac{Ze^2}{r}\right)^2-\left(\frac{\hbar}{i}\nabla\right)^2-m^2c^2+\frac{i\hbar Ze^2}{r^2c}\alpha_r\right]\psi=0 \tag{14.352}$$

where $\alpha_r=\vec{\alpha}\cdot\hat{r}$. We now write

$$\left(\frac{\hbar}{i}\nabla\right)^2=-\frac{\hbar^2}{r^2}\frac{\partial^2}{\partial r^2}r^2+\frac{\hat{L}^2}{r^2} \tag{14.353}$$

and get the equation

$$\left[\frac{E^2-m^2c^4}{c^2}+\frac{2EZe^2}{rc^2}+\frac{\hbar^2}{r^2}\frac{\partial^2}{\partial r^2}r^2-\frac{\hat{L}^2-\left(\frac{Ze^2}{c}\right)^2-i\hbar\left(\frac{Ze^2}{c}\right)\alpha_r}{r^2}\right]\psi=0 \tag{14.354}$$

We now use a few tricks to change this equation, which is almost in the same form as the Klein-Gordon equation for the Coulomb potential, into exactly the same form.

We first define the operator

$$\hat{K}=\beta\left(1+\vec{\sigma}\cdot\frac{\vec{L}}{\hbar}\right) \tag{14.355}$$

with these properties

$$\left[\hat{K},\vec{\alpha}\cdot\vec{p}\right]=0\quad,\quad\left[\hat{K},\vec{\alpha}\cdot\vec{r}\right]=0\quad,\quad\left[\hat{K},r^2\right]=0 \tag{14.356}$$

$$\left[\hat{K},\vec{J}\right]=0\quad,\quad\vec{J}=\vec{L}+\frac{\hbar}{2}\vec{\sigma} \tag{14.357}$$

These imply that $\hat{K}$ commutes with the Hamiltonian

$$\hat{H}=c\vec{\alpha}\cdot\vec{p}+\beta mc^2-\frac{Ze^2}{r} \tag{14.358}$$

for the relativistic hydrogen atom.

This says that $\hat{K}$ is a constant of the motion and since it also commutes with the total angular momentum we can label the common eigenstates or energy levels of the hydrogen atom by the eigenvalues of $\hat{K}$, $\hat{J}^2$ and $\hat{J}_z$. $\hat{K}$ is a constant of the motion for any spherically symmetric, spin-independent potential and physically it measures the degree to which the spin and the orbital angular momentum of the particle are aligned.

Let us find the eigenvalues k of $\hat{K}$. We note that

$$\begin{aligned}\hat{K}^2 &= \left(1+\vec{\sigma}\cdot\frac{\vec{L}}{\hbar}\right)^2 = 1+\left(\vec{\sigma}\cdot\frac{\vec{L}}{\hbar}\right)^2+2\vec{\sigma}\cdot\frac{\vec{L}}{\hbar}\\ &= 1+\left(\frac{\vec{L}\cdot\vec{L}}{\hbar^2}+\frac{i}{\hbar^2}\vec{\sigma}\cdot(\vec{L}\times\vec{L})\right)+2\vec{\sigma}\cdot\frac{\vec{L}}{\hbar}\\ &= 1+\left(\frac{\vec{L}\cdot\vec{L}}{\hbar^2}+\frac{i}{\hbar^2}\vec{\sigma}\cdot(i\hbar\vec{L})\right)+2\vec{\sigma}\cdot\frac{\vec{L}}{\hbar}\\ &= 1+\frac{L^2}{\hbar^2}+\vec{\sigma}\cdot\frac{\vec{L}}{\hbar} = \frac{1}{\hbar^2}\left(\vec{L}+\frac{\hbar}{2}\vec{\sigma}\right)^2+\frac{1}{4}\\ &= \frac{\hat{J}^2}{\hbar^2}+\frac{1}{4}\end{aligned} \tag{14.359}$$

where we have used

$$\begin{aligned}&(\vec{\sigma}\cdot\vec{A})(\vec{\sigma}\cdot\vec{B}) = \vec{A}\cdot\vec{B}+i\vec{\sigma}\cdot(\vec{A}\times\vec{B})\\ &\vec{L}\times\vec{L} = i\hbar\vec{L}\quad,\quad \vec{J}=\vec{L}+\vec{S}\quad,\quad \vec{S}=\frac{\hbar}{2}\vec{\sigma}\end{aligned}$$

Therefore, we have

$$k^2 = j(j+1)+\frac{1}{4} = (j+\tfrac{1}{2})^2 \tag{14.360}$$

Now, since $\left\{\hat{K},\gamma_5\right\}=0$ we find that, if k is an eigenvalue of $\hat{K}$, i.e.,

$$\hat{K}\,|k\rangle = k\,|k\rangle \tag{14.361}$$

then

$$\hat{K}\gamma_5\,|k\rangle = -\gamma_5\hat{K}\,|k\rangle = -k\gamma_5\,|k\rangle \tag{14.362}$$

which says that $-k$ is also an eigenvalue of $\hat{K}$. The eigenvalues are then

$$k = \pm1\ ,\ \pm2\ ,\ \pm3\ ,\ldots\ldots. \tag{14.363}$$

since $j = 1/2,\ 3/2,\ 5/2,\ \ldots\ldots$. Note that zero is not an eigenvalue of $\hat{K}$. In addition, an eigenstate of $\hat{K}$ with eigenvalue k is an eigenstate of $\hat{J}^2$ with eigenvalue $j = |k| - 1/2$.

We now find the energy eigenvalues. Define the operator

$$\hat{\Lambda} = -\beta\hat{K} - i\frac{Ze^2}{\hbar c}\alpha_r \tag{14.364}$$

with the properties

$$\left[\hat{\Lambda},\hat{K}\right]=0\quad,\quad \left[\hat{\Lambda},\vec{J}\right]=0\quad,\quad \Lambda^2 = K^2-\left(\frac{Ze^2}{\hbar c}\right)^2 \tag{14.365}$$

A little algebra then shows that

$$\hbar^2\hat{\Lambda}(\hat{\Lambda}+1) = \hat{L}^2 - \left(\frac{Ze^2}{c}\right)^2 - i\hbar\left(\frac{Ze^2}{c}\right)\alpha_r \tag{14.366}$$

which is the operator in the last term of the second-order equation. We can then write

$$\left[\frac{E^2 - m^2c^4}{c^2} + \frac{2EZe^2}{rc^2} + \frac{\hbar^2}{r^2}\frac{\partial^2}{\partial r^2}r^2 - \frac{\hbar^2\hat{\Lambda}(\hat{\Lambda}+1)}{r^2}\right]\psi = 0 \tag{14.367}$$

This is exactly the same form as the Klein-Gordon equation except that

$$\ell'(\ell'+1) \to \hat{\Lambda}(\hat{\Lambda}+1) \tag{14.368}$$

or a number has been replaced by an operator. Now if $\psi(\vec{r})$ is an eigenstate of $\hat{\Lambda}(\hat{\Lambda}+1)$ then the operator $\hat{\Lambda}(\hat{\Lambda}+1)$ in the equation is replaced by its eigenvalue which we can write as $\ell'(\ell'+1)$. This says that the energy eigenvalues are given by the same formula as in the spin zero case, i.e.,

$$E = \frac{mc^2}{\left[1 + \left(\frac{Ze^2}{\hbar cn'}\right)^2\right]^{1/2}} \quad , \quad n' = \ell' + 1 + \nu \quad , \quad \nu = 0, 1, 2, ... \tag{14.369}$$

Since $\hat{\Lambda}$, $\hat{K}$, $\hat{J}^2$ and $\hat{J}_z$ all commute, we can construct solutions which are eigenstates of $\hat{K}$, $\hat{J}^2$ and $\hat{J}_z$ as well as $\hat{\Lambda}$.

$\hat{\Lambda}$ does not commute with $\hat{H}$ however. This means that the solutions we have found for the second order equation cannot directly be eigenfunctions of $\hat{H}$. Instead, since

$$\hat{H}(\hat{P}\psi) = E(\hat{P}\psi) \tag{14.370}$$

i.e., the energy eigenvalues from the second-order equation are also the eigenvalues of $\hat{H}$, we can find eigenfunctions of $\hat{H}$ by using the projection operator $\hat{P}$. Since $\hat{P}$ and $\hat{\Lambda}$ do not commute, the eigenfunction of $\hat{H}$, namely $\hat{P}\psi$, will generally be a linear combination of different $\hat{\Lambda}$ eigenfunctions.

To find the energy eigenvalues we need to know the eigenvalues of $\hat{\Lambda}$. Consider an eigenstate of $\hat{\Lambda}$ and $\hat{K}$ with eigenvalues k. We then have

$$\hat{\Lambda}^2|k\rangle = \lambda^2|k\rangle = \left(\hat{K}^2 - \left(\frac{Ze^2}{\hbar c}\right)^2\right)|k\rangle = \left(k^2 - \left(\frac{Ze^2}{\hbar c}\right)^2\right)|k\rangle \tag{14.371}$$

or

$$\lambda = \left(k^2 - \left(\frac{Ze^2}{\hbar c}\right)^2\right)^{1/2} \tag{14.372}$$

and the possible eigenvalues of $\hat{\Lambda}$ are $\pm\lambda$. When $\hat{\Lambda}(\hat{\Lambda}+1)$ acts on a $(\hat{\Lambda}\ ,\hat{K})$ eigenstate it has the eigenvalue $\pm\lambda(\pm\lambda+1) = \ell'(\ell'+1)$. This leads to two possible ℓ' values for each eigenvalue of $\hat{\Lambda}$:

$$\ell' = \begin{cases} \lambda,\ -\lambda-1 & \text{for } \Lambda = \lambda \\ -\lambda,\ \lambda-1 & \text{for } \Lambda = -\lambda \end{cases} \tag{14.373}$$

For each eigenvalue of $\hat{\Lambda}$, the two ℓ' solutions add up to -1. The smaller of the two solutions $-lambda$ and $-\lambda-1$ are eliminated because they are not normalizable (behavior near the origin). This leave two cases to consider:

$$\begin{aligned} &(1) \quad \Lambda = \lambda \quad , \quad \ell' = \lambda \\ &(2) \quad \Lambda = -\lambda \quad , \quad \ell' = \lambda - 1 \end{aligned}$$

The possible energy eigenvalues are given by

$$E = \frac{mc^2}{\left[1+\left(\frac{Ze^2}{\hbar c n'}\right)^2\right]^{1/2}} \quad , \quad n' = \ell' + 1 + \nu \quad , \quad \nu = 0, 1, 2, ... \tag{14.374}$$

or redefining some quantities

$$n' = n - |k| + \lambda = n - j - \tfrac{1}{2} + \sqrt{\left(j+\tfrac{1}{2}\right)^2 - \left(\frac{Ze^2}{\hbar c}\right)^2} \tag{14.375}$$

so that n takes on the values

$$\begin{aligned} &n = |k|\ ,\ |k|+1\ ,\ |k|+2\ ,\ .....\text{for}\Lambda = -\lambda \\ &n = |k|+1\ ,\ |k|+2\ ,\ |k|+3\ ,\ .....\text{for}\Lambda = \lambda \end{aligned}$$

the energy levels are then given by

$$E = mc^2\left[1 + \frac{\left(\frac{Ze^2}{\hbar c}\right)^2}{\left[n - j - \frac{1}{2} - \sqrt{\left(j+\frac{1}{2}\right)^2 - \left(\frac{Ze^2}{\hbar c}\right)^2}\right]^2}\right]^{-1/2} \tag{14.376}$$

The quantum number n is just the principal quantum number of the hydrogen atom.

The energy level structure looks like Figure 20.2 below.

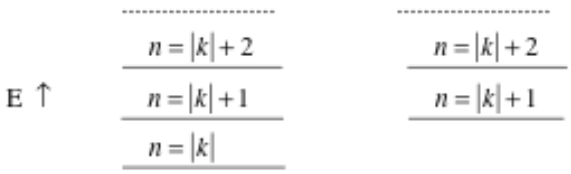


Figure 14.2: Dirac hydrogen energy level structure

where the left sequence corresponds to $\Lambda = -\lambda$ and the right sequence to $\Lambda = \lambda$.

Some Features

The energy levels for the spin 1/2 particle are the same as those found for the spin 0 particle with $\ell \to j$.

The energy is real only if

$$j + \tfrac{1}{2} < \frac{Ze^2}{\hbar c} \tag{14.377}$$

which corresponds to $Z < 137$ for $j = 1/2$.

The Dirac theory leads to an accidental degeneracy in ℓ, i.e., states with the same j but different ℓ have the same energy. This degeneracy is removed by the Lamb shift, which is due to the interaction of the electron with its own field. As we shall see later, for $j = 1/2$, the effect is one order of magnitude smaller then the fine structure splitting. For $j \geq 3/2$, it is two orders of magnitude smaller.

An expansion in powers of $Z\alpha$, where

$$\alpha = \frac{e^2}{\hbar c} = \text{ fine structure constant} \tag{14.378}$$

looks like

$$E_{n,j} = mc^2 \left[1 - \frac{Z^2\alpha^2}{2n^2} - \frac{(Z\alpha)^4}{2n^3} \left(\frac{1}{j + \frac{1}{2}} - \frac{3}{4n} \right) + O((Z\alpha)^6) \right] \tag{14.379}$$

which agrees with the perturbation calculations we carried out earlier.

Some Details about the Energy Levels

The solutions of the Dirac equation are not $\hat{\Lambda}$ eigenstates but they are $\hat{K}$ eigenstates and $\hat{K}$ is a constant of the motion (it commutes with the Hamiltonian). The total orbital angular momentum $\vec{L}$ is not a constant of the motion and neither is $\hat{L}^2$. We need to come up with some way to classify the energy levels in the relativistic hydrogen atom using the eigenvalues k.

To get a handle on how to proceed we look at the nonrelativistic limit where

$$\hat{\Lambda} \to -\beta\hat{K} \to -\hat{K} \text{ for positive energy states} \tag{14.380}$$

We expect the solutions of the second order equation with one sign of $\hat{\Lambda}$ to correspond to solutions of the first-order equation with the opposite sign of $\hat{K}$. This means that

$$\begin{aligned} n &= |k| \ , \ |k|+1 \ , \ |k|+2 \ , \ \ldots\ldots\ldots \text{ for } \Lambda = -\lambda \to k > 0 \\ n &= |k|+1 \ , \ |k|+2 \ , \ |k|+3 \ , \ \ldots\ldots\ldots \text{ for } \Lambda = \lambda \to k < 0 \end{aligned}$$

It turns out to be convenient to still label the solutions by the ℓ value that they would have in the nonrelativistic limit. To find this ℓ value we use

$$\begin{aligned} \hat{K}^2 &= 1 + \frac{\hat{L}^2}{\hbar^2} + \vec{\sigma}\cdot\frac{\vec{L}}{\hbar} = \beta\hat{K} + \frac{\hat{L}^2}{\hbar^2} \\ \hat{K}(\hat{K}-\beta) &= \frac{\hat{L}^2}{\hbar^2} \end{aligned}$$

In the nonrelativistic limit, $\beta \to 1$ and we have

$$\hat{K}(\hat{K}-1) = \frac{\hat{L}^2}{\hbar^2} \to k(k-1) = \ell(\ell+1) \tag{14.381}$$

so that ℓ becomes the total orbital angular momentum quantum number in the nonrelativistic limit. Solving for ℓ in terms of k we get

$$\ell = \begin{cases} k-1 = j-1/2 & \text{for } k > 0 \\ |k| = j+1/2 & \text{for } k < 0 \end{cases} \tag{14.382}$$

Now K measures the alignment of the spin and the orbital angular momentum. The above results say that for $k > 0$, they are essentially parallel and so $j = \ell + 1/2$ and for $k < 0$ they are essentially antiparallel so $j = \ell - 1/2$.

A detailed calculation of the wave functions shows that the upper two components of the wave function (the large components) are eigenstates of total orbital angular momentum with eigenvalue ℓ, while the lower two components (the small components) are eigenstates of total orbital angular momentum with eigenvalue $\ell + 1$ for $k > 0$ and with $\ell - 1$ for $k < 0$.

The complete energy level scheme for the relativistic hydrogen atom for $n = 1$, 2, and 3 looks like Figure 14.3 below.

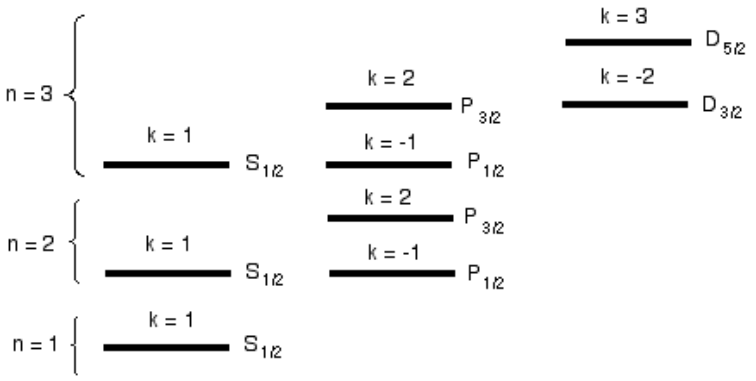


Figure 14.3: Energy level structure for relativistic hydrogen

The complete degeneracy of a given n in the nonrelativistic case is lifted by relativistic effects. The degeneracy between states like $1S_{1/2}$, $2P_{3/2}$, $3D_{5/2}$, $4F_{7/2}$, etc is now broken. The degeneracy still remains between states like $2S_{1/2}$ and $2P_{1/2}$, $3S_{1/2}$ and $3P_{1/2}$, $3P_{3/2}$ and $3D_{3/2}$, etc., levels.

All levels except $1S_{1/2}$, $2P_{3/2}$, $3D_{5/2}$, etc., are 2-fold degenerate because they are the eigenstates of of K with opposite eigenvalues, i.e., $2P_{3/2} \to k = 2$, $2D_{3/2} \to k = -2$.

Hyperfine Structure

The are two corrections that modify the energy level results from the Dirac equation. The two-fold degeneracy is removed by the interaction of the electron with vacuum fluctuations of the electromagnetic radiation field. This effect is called the Lamb shift. In addition, there is also a hyperfine interaction which splits every level into two, It is due to the interaction of the electron with the magnetic moment of the proton. We consider hyperfine splitting first.

As an example we derive the hyperfine splitting of an $s-$state using nonrelativistic first-order perturbation theory. The interaction of the electron spin with the magnetic moment of the proton is given by

$$\hat{H}' = \frac{|e|\hbar}{2mc}\vec{\sigma}\cdot\vec{B}(\vec{r}) = \mu_B\vec{\sigma}\cdot\vec{B}(\vec{r}) \tag{14.383}$$

where $\vec{B}(\vec{r})$ is the magnetic field due to the magnetic moment of the proton. This magnetic moment is given by

$$\vec{M}_p = \frac{|e|\hbar g_p}{4m_p c}\vec{\sigma}_p = \frac{1}{2}g_p\mu_p\vec{\sigma}_p \tag{14.384}$$

where g_p is the gyromagnetic ratio of the proton, m_p is the proton mass and $\hbar\vec{\sigma}_p/2$ is the spin of the proton. The magnetic field from this magnetic moment

(assuming the proton is fixed at the origin) is given by the relations

$$\vec{A}(\vec{r}) = -\vec{M}_p \times \nabla\left(\frac{1}{r}\right) = \text{ vector potential} \tag{14.385}$$

$$\vec{B}(\vec{r}) = \nabla \times \vec{A}(\vec{r}) \tag{14.386}$$

$$\vec{B}(\vec{r}) = -\nabla \times (\vec{\sigma}_p \times \nabla)\frac{g_p \mu_p}{2r} \tag{14.387}$$

This gives

$$\begin{aligned}
\hat{H}' &= -g_p \mu_B \mu_p \vec{\sigma} \cdot \left(\nabla \times (\vec{\sigma}_p \times \nabla)\frac{1}{2r}\right) \\
&= -g_p \mu_B \mu_p \vec{\sigma} \cdot (\vec{\sigma}_p(\nabla \cdot \nabla) - \nabla(\vec{\sigma}_p \cdot \nabla))\,\frac{1}{2r} \\
&= -g_p \mu_B \mu_p \left((\vec{\sigma} \cdot \vec{\sigma}_p)(\nabla \cdot \nabla) - (\vec{\sigma} \cdot \nabla)(\vec{\sigma}_p \cdot \nabla)\right)\frac{1}{2r}
\end{aligned} \tag{14.388}$$

The first-order shift of the level is

$$\left\langle \hat{H}' \right\rangle = -g_p \mu_B \mu_p \int d^3 r\, |\psi(\vec{r})|^2 \left[\left(\langle \vec{\sigma} \cdot \vec{\sigma}_p \rangle\right)(\nabla \cdot \nabla) - \langle (\vec{\sigma} \cdot \nabla)(\vec{\sigma}_p \cdot \nabla) \rangle \right) \frac{1}{2r} \Bigg] \tag{14.389}$$

where the brackets $\langle ..... \rangle$ denote the expectation value in the relative spin state of the electron and proton and $\psi(\vec{r})$ is the nonrelativistic wave function of the level. If we only consider $s-$states, which are spherically symmetric, then

$$\langle (\vec{\sigma} \cdot \nabla)(\vec{\sigma}_p \cdot \nabla) \rangle = \frac{1}{3} \langle (\vec{\sigma} \cdot \vec{\sigma}_p) \rangle \nabla^2 \tag{14.390}$$

and we get

$$\begin{aligned}
\left\langle \hat{H}' \right\rangle &= -\frac{1}{3} g_p \mu_B \mu_p \int d^3 r\, |\psi(\vec{r})|^2 \langle (\vec{\sigma} \cdot \vec{\sigma}_p) \rangle \left(\nabla^2 \frac{1}{2r}\right) \\
&= \frac{4\pi}{3} g_p \mu_B \mu_p \langle (\vec{\sigma} \cdot \vec{\sigma}_p) \rangle \int d^3 r\, |\psi(\vec{r})|^2 \delta(\vec{r}) \\
&= \frac{4\pi}{3} g_p \mu_B \mu_p \langle (\vec{\sigma} \cdot \vec{\sigma}_p) \rangle |\psi(0)|^2
\end{aligned} \tag{14.391}$$

where we have used

$$\nabla^2 \frac{1}{r} = -4\pi \delta(\vec{r}) \tag{14.392}$$

For the hydrogen atom $s-$state

$$|\psi(0)|^2 = \frac{1}{\pi (n a_0)^3} \tag{14.393}$$

and we get

$$\left\langle \hat{H}' \right\rangle = \frac{2}{3}\left(\frac{e^2}{2a_0}\right) g_p \frac{m}{m_p} \frac{\alpha^2}{n^3} \langle (\vec{\sigma} \cdot \vec{\sigma}_p) \rangle \tag{14.394}$$

We then have

$$\vec{F} = \vec{S} + \vec{I} = \text{ total spin} \tag{14.395}$$

For $S = \frac{1}{2}$, $I = \frac{1}{2}$, we have $F = 0(\text{singlet})$, $1(\text{triplet})$. But

$$\vec{F} = \vec{S} + \vec{I} \rightarrow \vec{F}^2 = \vec{S}^2 + \vec{I}^2 + 2\vec{S} \cdot \vec{I} \tag{14.396}$$

$$\vec{S} \cdot \vec{I} = \frac{\hbar^2}{4} \vec{\sigma} \cdot \vec{\sigma}_p = \frac{\hbar^2}{2} (F(F+1) - 3/2) \tag{14.397}$$

We then have for a relative triplet state $\langle (\vec{\sigma} \cdot \vec{\sigma}_p) \rangle = 1$ and for a relative singlet state $\langle (\vec{\sigma} \cdot \vec{\sigma}_p) \rangle = -3$. This says that the singlet state lies lower than the triplet.

The total splitting of the *ground state* is

$$\Delta E = \frac{8}{3} \left(\frac{e^2}{2a_0} \right) g_p \frac{m}{m_p} \alpha^2 \tag{14.398}$$

between the triplet and singlet. The transition between these two levels generates radiation with a frequency of $1420\, MHz$ and a wavelength of $21.4\, cm$. This radiation is very important in astronomy. From its intensity, Doppler broadening, and Doppler shift, one obtains information concerning the density, temperature, and motion of interstellar and intergalactic hydrogen clouds.

The Lamb Shift

The coupling

$$\hat{H}_{\text{int}} = -\frac{e}{c} \int d^3 r \vec{j}(\vec{r}) \cdot \vec{A}(\vec{r}) \tag{14.399}$$

of the electron to the quantum mechanical radiation field causes a shift in the energy levels of the hydrogen atom. Although not an exact calculation, we can get some idea of the fundamental difficulties in quantum electrodynamics by doing a nonrelativistic second-order perturbation calculation.

We consider an electron in the state $|n\rangle$ with energy ε_n. Because of the above interaction (see last part of this chapter) the electron is able to spontaneously emit a photon thereby going to some state $|n'\rangle$. This produces a second-order shift in the energy given by

$$\Delta E_n = \sum_{n'} \sum_{\vec{k}\vec{\lambda}} \frac{\left| \left\langle n', \vec{k}\vec{\lambda} \right| \hat{H}_{\text{int}} |n, 0\rangle \right|^2}{\varepsilon_n - \varepsilon_{n'} - ck} \tag{14.400}$$

where $|n, 0\rangle$ is the initial state with the electron in $|n\rangle$ with no photons present, and $\left|n', \vec{k}\vec{\lambda}\right\rangle$ is the intermediate state with an electron in $|n'\rangle$ and on photon of momentum $\vec{k}$ and polarization $\vec{\lambda}$ present. The energy of this intermediate state is $\varepsilon_{n'} + ck$.

From the quantum theory of electromagnetic radiation (see end of this chapter) we have that

$$\left\langle n', \vec{k}\vec{\lambda}\right| \hat{H}_{\text{int}} \left| n, 0 \right\rangle = -\frac{e}{c}\sqrt{\frac{2\pi\hbar^2 c^2}{\omega_k V}} \left\langle n' \right| \vec{j}_{\vec{k}} \cdot \vec{\lambda}^* \left| n \right\rangle \tag{14.401}$$

where $\vec{j}_{\vec{k}}$ is the k^{th} Fourier component of the current $\vec{j}(\vec{r})$. Therefore,

$$\begin{aligned} \Delta E_n &= \int \frac{d^3k}{(2\pi\hbar)^3} \frac{2\pi\hbar^2 e^2}{ck} \sum_{n'} \frac{\sum_{\vec{\lambda}} \left| \left\langle n' \right| \vec{j}_{\vec{k}} \cdot \vec{\lambda}^* \left| n \right\rangle \right|^2}{\varepsilon_n - \varepsilon_{n'} - ck} \\ &= \int \frac{k^2 dk}{4\pi^2\hbar} \frac{e^2}{ck} \sum_{n'} \frac{\int d\Omega \sum_{\vec{\lambda}} \left| \left\langle n' \right| \vec{j}_{\vec{k}} \cdot \vec{\lambda}^* \left| n \right\rangle \right|^2}{\varepsilon_n - \varepsilon_{n'} - ck} \end{aligned} \tag{14.402}$$

In the dipole approximation, we can use

$$\vec{j}_{\vec{k}} \to \vec{j}_0 = \frac{\vec{p}}{m} \quad , \quad \vec{p} = \text{ electron momentum operator} \tag{14.403}$$

The angular integration over the polarizations is given by

$$\int d\Omega \sum_{\vec{\lambda}} \left| \left\langle n' \right| \vec{j}_{\vec{k}} \cdot \vec{\lambda}^* \left| n \right\rangle \right|^2 = \frac{1}{m} \int d\Omega \sum_{\vec{\lambda}} \left| \left\langle n' \right| \vec{p} \cdot \vec{\lambda}^* \left| n \right\rangle \right|^2 = \frac{4\pi}{m} \frac{2}{3} \left| \left\langle n' \right| \vec{p} \left| n \right\rangle \right|^2$$

where the factor 2/3 comes from the fact that there are only 2 independent polarizations for each $\vec{k}$ value. This gives

$$\Delta E_n = \frac{2e^2}{3\pi\hbar c^3 m^2} \int_0^\infty \omega d\omega \sum_{n'} \frac{\left| \left\langle n' \right| \vec{p} \left| n \right\rangle \right|^2}{\varepsilon_n - \varepsilon_{n'} - \omega} \quad , \quad \omega = ck \tag{14.404}$$

The first problem we encounter is that the ω integral diverges!! This means that the interaction with the radiation field produces an *infinite shift downward* in the energy of the electron.

This result presented theoretical physics with a great difficulty for many years. In the late 1940's it was resolved due to the work of Feynman, Schwinger and Tomonaga in producing new calculation rules within the context of quantum electrodynamics and by Bethe and Weisskopf who actually carried out the calculation using the new rules and got a finite number agreeing with experiment.

Let us try to understand some aspects of what happened.

If we do a similar calculation for a free electron, then one gets an infinite result again. In the dipole approximation, we can evaluate the energy shift for a free electron in a momentum state $|\vec{p}\rangle$. We get

$$\Delta E_{\vec{p}} = \frac{2e^2}{3\pi\hbar c^3 m^2} \int_0^\infty \omega d\omega \sum_{\vec{q}} \frac{\left| \left\langle \vec{q} \right| \vec{p} \left| \vec{p} \right\rangle \right|^2}{\varepsilon_q - \varepsilon_p - \omega} \tag{14.405}$$

Since this is a free particle $\varepsilon_q - \varepsilon_p = 0$ and we have

$$\Delta E_{\vec{p}} = -\frac{2e^2}{3\pi\hbar c^3 m^2}\int_0^\infty \omega d\omega \frac{\left|\langle\vec{p}|\,\vec{p}\,|\vec{p}\rangle\right|^2}{\omega} = -\frac{2e^2}{3\pi\hbar c^3 m^2}p^2\int_0^\infty d\omega \tag{14.406}$$

which is infinite. What Bethe and Weisskopf noticed was that this expression is proportional to p^2. In their development of quantum electrodynamics, Feynman, Schwinger and Tomonaga had similar problems which they were able to deal with be redefining the electron parameters that appeared in the theory (like mass and charge). The process is called *renormalization*. In this process all infinite expression are consistently incorporated into the mass or charge parameters and then these are defined to have the known experimental values.

In our case, we can interpret the infinite result as redefining the mass, i.e., as representing a shift of the mass of the electron. In terms of the mathematics, this means the following. If we say that m_0 is the mass and $p^2/2m_0$ is the kinetic energy of a free electron of momentum $\vec{p}$ neglecting the electromagnetic interactions, then the energy including the effects of the electromagnetic interactions is given by

$$\frac{p^2}{2m_0} + \Delta E_{\vec{p}} = \left(\frac{1}{m_0} - \frac{2e^2}{3\pi\hbar c^3 m^2}\int_0^\infty d\omega\right)\frac{p^2}{2} = \frac{1}{m}\frac{p^2}{2} \tag{14.407}$$

i.e., we have *renormalized* the electron mass. The so-called electromagnetic self-energy of the electron can thus be interpreted as giving a shift of the mass of the electron from its *bare* (no electromagnetic interactions) value m_0 to its *observed* (measured in the laboratory where all interactions are present) value m.

We then argue that the reason the interacting electron has an infinite energy shift is that it includes the infinite energy change that we already have counted once when we use the observed mass m rather than the bare mass in the calculation and, thus, we are double counting. In other words, we should really start out with the Hamiltonian for the hydrogen atom in the presence of the radiation field given by

$$\hat{H} = \frac{p^2}{2m_0} - \frac{e^2}{r} + \hat{H}_{\text{int}} \tag{14.408}$$

Then using the corrected expression for m we get

$$\hat{H} = \frac{p^2}{2m} - \frac{e^2}{r} + \left(\hat{H}_{\text{int}} + \frac{2e^2}{3\pi\hbar c^3 m^2}\int_0^\infty d\omega\right) \tag{14.409}$$

This means that if we write the observed free particle mass in the kinetic energy (which we always do) we should not count that part of $\hat{H}_{\text{int}}$ that produces the infinite mass shift, i.e., we should regard

$$\hat{H}_{\text{int}} + \frac{2e^2}{3\pi\hbar c^3 m^2}\int_0^\infty d\omega \tag{14.410}$$

as the effective interaction of an electron of renormalized mass m with the radiation field. It is now finite to second-order of our calculation. Feynman, et al, showed that this could be done to all orders of perturbation theory!!

We therefore modify our calculation by adding in the required term. We get

$$\Delta E_n' = \frac{2e^2}{3\pi\hbar c^3 m^2} \int_0^\infty \omega d\omega \left(\sum_{n'} \frac{|\langle n'|\vec{p}\,|n\rangle|^2}{\varepsilon_n - \varepsilon_{n'} - \omega} + \frac{|\langle n|\vec{p}\,|n\rangle|^2}{\omega} \right) \tag{14.411}$$

Using completeness, we have

$$\langle n|p^2|n\rangle = \sum_{n'} \langle n|p|n'\rangle \langle n'|p|n\rangle = \sum_{n'} |\langle n'|p|n\rangle|^2 \tag{14.412}$$

so that

$$\Delta E_n' = \frac{2e^2}{3\pi\hbar c^3 m^2} \sum_{n'} |\langle n'|p|n\rangle|^2 \int_0^\infty d\omega \frac{\varepsilon_{n'} - \varepsilon_n}{\varepsilon_n - \varepsilon_{n'} - \omega} \tag{14.413}$$

The integral is still divergent but only logarithmically and, in fact, not at all in more sophisticated relativistic calculations. We can imagine that the correct calculation would yield a similar result but with a convergent integral. We can simulate this result by integrating to some cutoff value (and not to infinity) say at $\hbar\omega = mc^2$. We then get

$$\Delta E_n' = \frac{2e^2}{3\pi\hbar c^3 m^2} \sum_{n'} |\langle n'|p|n\rangle|^2 (\varepsilon_{n'} - \varepsilon_n) \ell n \left| \frac{mc^2}{\varepsilon_{n'} - \varepsilon_n} \right| \tag{14.414}$$

where we have neglected quantities the size of $\varepsilon_{n'} - \varepsilon_n$ in comparison to mc^2.

Bethe evaluated this result numerically and obtained $\Delta E_n' = +1040$ megacycles (the $2P_{1/2}$ level turns out to be shifted downward) and the observed value equals +1057 megacycles, which is remarkable agreement!

Taking into account both the Lamb shift and the hyperfine splitting we have the level scheme shown in Figure 14.4 below for $n = 2$:

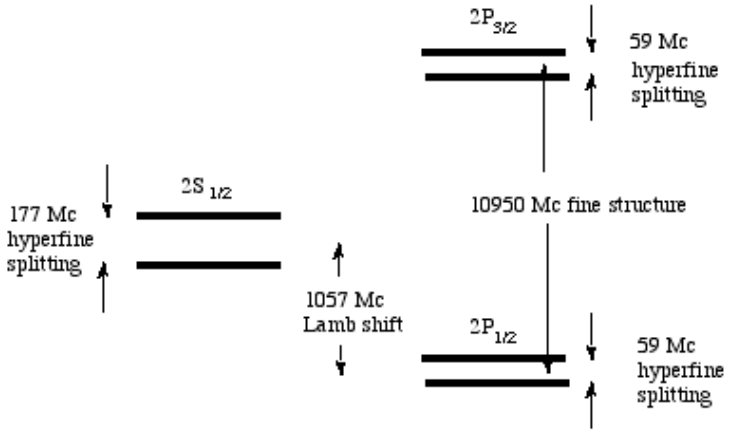


Figure 14.4: $n = 2$ Energy level structure for relativistic hydrogen

Dirac Hole Theory

Finally, we tackle the problem of the negative energy states in the Dirac theory.

As we said earlier, there is no simple conservation law that prevents an electron or any other spin 1/2 particle in a positive energy state from making a radiative transition to a negative energy state. This means all atoms must be unstable! An energy diagram is shown in Figure 14.5 below.

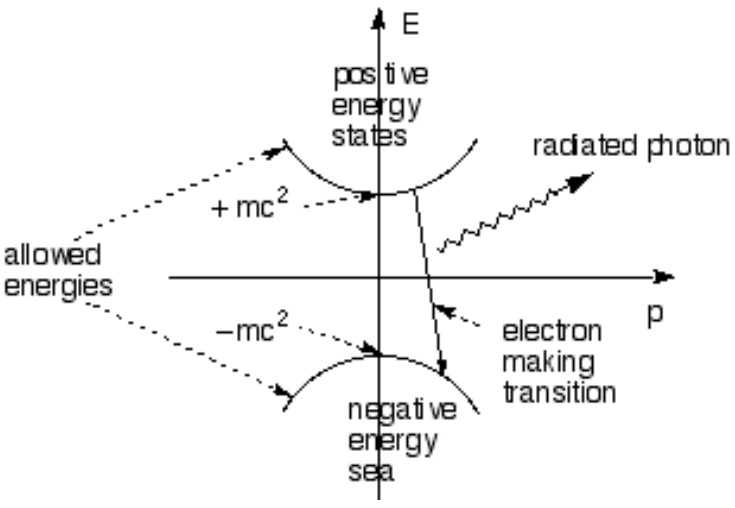


Figure 14.5: $n = 2$ Energy level structure for relativistic hydrogen

The properties of the positive energy states show remarkable agreement with

experiment. Can we simply ignore the negative energy states? The answer is no because an arbitrary wave packet, as we saw earlier will always contain negative energy components via interactions even if we start off only with positive energy components.

Dirac proposed a clever way out of this dilemma : since spin 1/2 particles obey the exclusion principle, all one needs to do to insure stability is to say that the negative energy states are completely filled. Then a particle cannot make a transition from a positive to a negative energy state for this would put two particles into the same (negative energy) state. The vacuum state in this picture consists of an *infinite sea* of particles in negative energy states. The particle and charge density at every point is infinite. This is not a problem for the physical theory since Dirac contended that we only measure deviations from the vacuum. In the absence of any potential, the charge density of the negative sea is uniform and Dirac argued that this charge density can produce no forces, since by isotropy, the forces have no special direction to point!

Now this theory has some very useful special property. Suppose that we remove a negative energy electron from the vacuum. What is left behind is a *hole* in the negative energy sea. Measured with respect to the vacuum, the hole would appear to have positive charge and positive energy,i.e., since it is the absence of negative charge and negative energy. Dirac interpreted it as a positron, which is the electron antiparticle.

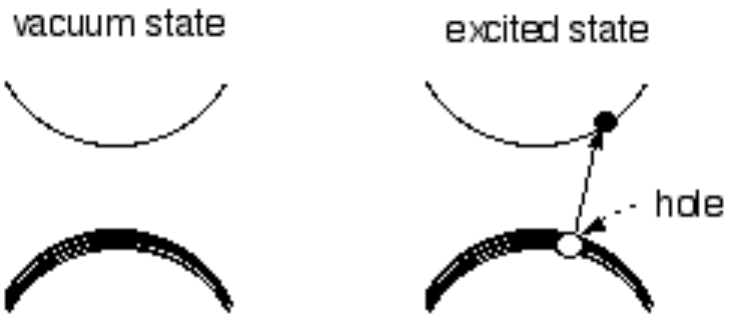


Figure 14.6: A hole appears

Let me say that again..... an excited state of the vacuum arises as shown in the figure. A negative energy electron is excited into a positive energy state, leaving behind a hole with charge $-(-e) = +e$ and the same mass as the electron, which is the antiparticle. It looks like a positive charge since if we apply an electric field the infinite sea of electron translates opposite to the field direction, which is unobservable since the sea is infinite. However, the hole seems to be traveling in the direction of the field like a positive charge!

In this way, antiparticles appear in the Dirac theory as unoccupied negative energy states, which is very different from the way they appear in the spin zero theory.

This Dirac hole theory gives a simple description for pair production. Suppose that a photon of energy $> 2mc^2$ traveling through the vacuum is absorbed by a negative energy electron and the negative energy electron gets excited to a positive energy state. What remains, as we have said, is a hole in the negative energy sea, i.e., a positive energy positron and a positive energy electron. This says that pair production is simply the excitation of a particle from a negative to a positive energy state.

Since we could exchange the roles of positrons and electrons in the entire Dirac theory, electrons would appear as holes in a positron sea. This forces us to conclude that negative energy seas cannot have any physical reality. The *hole theory* is simply a mathematical model that allows us to do the correct *book-keeping* within the framework of a single-particle Dirac theory.

With a filled negative energy sea, the Dirac theory would become a many-particle theory in which we are unable to take into account the interactions between these particles. The Dirac theory gives valid results only when these interactions can be neglected. For example, in the hydrogen atom, the modification of the Coulomb potential by vacuum polarization accounts for about 2.5% of the Lamb shift.

If we second-quantized the Dirac theory, we can treat both particles and antiparticles on the same basis.

The full relativistic quantum field theory of the electrons and positrons and their interactions with photons was carried out by Feynman, et al in a theory which is beyond the scope of these volumes.

14.7 Electromagnetic Radiation and Matter

14.7.1 Interacting with the Classical Radiation Field

We assume *classical* EM radiation in the *transverse gauge*, where

$$\phi(\vec{r},t) = 0 \quad , \quad \nabla \cdot \vec{A}(\vec{r},t) = 0 \tag{14.415}$$

The electric and magnetic fields are given in terms of the vector potential (in this gauge) by

$$\vec{\varepsilon}(\vec{r},t) = -\frac{1}{c}\frac{\partial \vec{A}(\vec{r},t)}{\partial t} \quad , \quad \vec{B}(\vec{r},t) = \nabla \times \vec{A}(\vec{r},t) \tag{14.416}$$

The electromagnetic energy is given by

$$E = \int d^3\vec{r} \frac{\varepsilon^2(\vec{r},t) + B^2(\vec{r},t)}{8\pi} \tag{14.417}$$

and the rate and direction of energy transport is given by the Poynting vector

$$\vec{\wp}(\vec{r},t) = \frac{c}{4\pi}\vec{\varepsilon}(\vec{r},t) \times \vec{B}(\vec{r},t) \tag{14.418}$$

The radiation field generated by a classical current $\vec{j}(\vec{r},t)$ is given by

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\vec{A}(\vec{r},t) = -\frac{4\pi}{c} j_\perp(\vec{r},t) \tag{14.419}$$

where $\perp$ means the transverse/divergence-free part.

We first consider the *monochromatic plane wave* solution of this equation. It takes the form

$$\vec{A}(\vec{r},t) = \alpha\vec{\lambda}e^{i\vec{k}\cdot\vec{r}-i\omega t} + \alpha^*\vec{\lambda}^* e^{-i\vec{k}\cdot\vec{r}+i\omega t} \tag{14.420}$$

where

$$\begin{aligned}
&\omega = ck \\
&\vec{\lambda} = \text{ polarization vector with } \left|\vec{\lambda}\right|^2 = 1 \\
&\alpha = \text{ amplitude} = \text{constant}
\end{aligned}$$

To insure that $\nabla\cdot\vec{A}(\vec{r},t) = 0$ we require $\vec{\lambda}\cdot\vec{k} = 0$ which corresponds to transverse polarizations only.

The energy per unit volume in the electromagnetic wave is

$$\frac{\varepsilon^2 + B^2}{8\pi} = \frac{\omega^2}{2\pi c^2}\left[|\alpha|^2 - \mathrm{Re}\left(\alpha^2\lambda^2 e^{2i\vec{k}\cdot\vec{r}-2i\omega t}\right)\right] \tag{14.421}$$

The quantity Re(..) oscillates in time and averages to zero so that the average energy density is

$$\frac{E}{volume} = \frac{\omega^2}{2\pi c^2}|\alpha|^2 \tag{14.422}$$

In a similar way the time average of the Poynting vector is

$$\frac{\omega^2}{2\pi c}|\alpha|^2\,\hat{k} \tag{14.423}$$

Any general wave solution is a *linear superposition* of these monochromatic wave solutions.

$$\vec{A}(\vec{r},t) = \sum_{\vec{k}\vec{\lambda}}\left[A_{\vec{k}\vec{\lambda}}\vec{\lambda}\frac{e^{i\vec{k}\cdot\vec{r}-i\omega t}}{\sqrt{V}} + A^*_{\vec{k}\vec{\lambda}}\vec{\lambda}^*\frac{e^{-i\vec{k}\cdot\vec{r}+i\omega t}}{\sqrt{V}}\right] \tag{14.424}$$

where the sum is over all allowed $\vec{k}$ values and over the two orthogonal $\vec{\lambda}$ polarizations for each $\vec{k}$ such that $\vec{\lambda} \cdot \vec{k} = 0$ and we have assumed that the universe is a very large box of volume V. The total energy in this wave solution is

$$E = \sum_{\vec{k}\vec{\lambda}} \frac{\omega^2}{2\pi c^2} \left| A_{\vec{k}\vec{\lambda}} \right|^2 \tag{14.425}$$

How does this classical electromagnetic field interact with a quantum mechanical particle?

In general (no transverse gauge at this point), the classical Hamiltonian is

$$\hat{H} = \frac{\left(\vec{p} - \frac{e}{c}\vec{A}(\vec{r},t)\right)^2}{2m} + e\phi(\vec{r},t) + V(\vec{r},t) \tag{14.426}$$

where $V(\vec{r},t)$ represents all the other potentials seen by the particle.

We get to quantum mechanics via the standard substitutions

$$\vec{r} \to \vec{r}_{op} \quad , \quad \vec{p} \to \vec{p}_{op} = \frac{\hbar}{i}\nabla \tag{14.427}$$

Substituting, we get the Schrodinger equation for an electron in an electromagnetic field

$$i\hbar\frac{\partial\psi(\vec{r},t)}{\partial t} = \left[\frac{1}{2m}\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\vec{A}(\vec{r},t)\right)^2 + e\phi(\vec{r},t) + V(\vec{r},t)\right]\psi(\vec{r},t) \tag{14.428}$$

14.7.2 Relation to Gauge Invariance

In order to have the Schrodinger equation invariant under a gauge transformation, the wave function has to change by a phase factor,i.e.,

$$\psi'(\vec{r},t) = e^{i\frac{e}{\hbar c}\chi(\vec{r},t)}\psi(\vec{r},t) \tag{14.429}$$

where $\chi(\vec{r},t)$ is the some scalar function.

This means that the solutions of the gauge-transformed Schrodinger equation will still describe the same physical states.

The wave functions or state vectors differ by a phase factor that depends on space and time and thus, the invariance is *LOCAL* rather than *GLOBAL* (a phase factor independent of space and time).

It is then clear that it is *NOT* the canonical momentum $\hat{p} \to -i\hbar\nabla$ (whose expectation value is *NOT* gauge invariant), but the genuine kinetic momentum

$$\hat{p} - \frac{q}{c}\vec{A}(\hat{r},t) \tag{14.430}$$

(whose expectation value *IS* gauge invariant), that represents a *measurable* quantity.

In any physical system, if the momentum operator $\hat{p}$ appears, then it must always be replaced by

$$\psi'(\vec{r},t) = e^{i\frac{e}{\hbar c}\chi(\vec{r},t)}\psi(\vec{r},t) \tag{14.431}$$

if we turn on an electromagnetic fields. This is the only way to guarantee gauge invariance in quantum mechanics.

Quantum mechanics + electromagnetism requires *minimal coupling* for gauge invariance to be valid.

14.7.3 Interactions

We now write

$$\hat{H} = \hat{H}_0 + \hat{H}_{int} \tag{14.432}$$

where

$$\hat{H}_0 = \frac{\vec{p}^2}{2m} + V(\vec{r},t) \tag{14.433}$$

is the Hamiltonian in the absence of electromagnetic fields and

$$\hat{H}_{\text{int}} = -\frac{e}{2mc}\left(\vec{p}\cdot\vec{A}(\vec{r},t) + \vec{A}(\vec{r},t)\cdot\vec{p}\right) + \frac{e^2}{2mc^2}\vec{A}^2(\vec{r},t) + e\phi(\vec{r},t) \tag{14.434}$$

is the operator giving the *interaction between matter and radiation.*

One must treat the term $\vec{p}\cdot\vec{A}(\vec{r},t) + \vec{A}(\vec{r},t)\cdot\vec{p}$ with care since $[x_i, p_j] = i\hbar\delta_{ij}$. In general, we can show that

$$\vec{p}\cdot\vec{A}(\vec{r},t) - \vec{A}(\vec{r},t)\cdot\vec{p} = -i\hbar\left(\nabla\cdot\vec{A}(\vec{r},t)\right) \tag{14.435}$$

which says that

$$\vec{p}\cdot\vec{A}(\vec{r},t) = \vec{A}(\vec{r},t)\cdot\vec{p} \text{ only when } \nabla\cdot\vec{A} = 0 \tag{14.436}$$

or when we are operating in the transverse gauge.

For multi-electron atoms we have

$$\hat{H} = \sum_{i=1}^{N}\frac{\left(\vec{p}_i - \frac{e}{c}\vec{A}(\vec{r}_i,t)\right)^2}{2m} + e\sum_{i=1}^{N}\phi(\vec{r}_i,t) + V \tag{14.437}$$

and

$$\hat{H}_{int} = \sum_{i=1}^{N}\left\{-\frac{e}{2mc}\left(\vec{p}_i\cdot\vec{A}(\vec{r}_i,t) + \vec{A}(\vec{r}_i,t)\cdot\vec{p}_i\right) + \frac{e^2}{2mc^2}\vec{A}^2(\vec{r}_i,t) + e\phi(\vec{r}_i,t)\right\} \tag{14.438}$$

We now define a particle number density

$$\rho(\vec{r}) = \sum_i \delta(\vec{r} - \vec{r}_i) \tag{14.439}$$

and a current density

$$\vec{j}(\vec{r}) = \frac{1}{2}\sum_i \left(\frac{\vec{p}_i}{m}\delta(\vec{r} - \vec{r}_i) + \delta(\vec{r} - \vec{r}_i)\frac{\vec{p}_i}{m} \right) \tag{14.440}$$

where we constructed a symmetric combination of the terms so that the operator would be Hermitian.

These quantities imply that

$$\sum_i e\phi(\vec{r}_i, t) = \sum_i \int d^3\vec{r}\, e\delta(\vec{r} - \vec{r}_i)\phi(\vec{r}, t) = \int d^3\vec{r}\, e\rho(\vec{r})\phi(\vec{r}, t) \tag{14.441}$$

where $\phi(\vec{r}, t) \neq$ operator (all operators are in $\rho(\vec{r})$) and

$$\int d^3\vec{r}\rho(\vec{r}) = N = \text{ total number of particles} \tag{14.442}$$

Finally, we have

$$\sum_{i=1}^{N} \left\{ -\frac{e}{2mc}\left(\vec{p}_i \cdot \vec{A}(\vec{r}_i, t) + \vec{A}(\vec{r}_i, t) \cdot \vec{p}_i \right) \right\} = -\frac{e}{c}\int d^3\vec{r}\vec{j}(\vec{r}) \cdot \vec{A}(\vec{r}, t) \tag{14.443}$$

Since

$$\vec{v}_i = \frac{\vec{p}_i}{m} - \frac{e}{mc}\vec{A} \tag{14.444}$$

when an electromagnetic field is present, the true current operator is

$$\vec{J}(\vec{r}) = \vec{j}(\vec{r}) - \frac{e}{mc}\vec{A}(\vec{r}, t)\rho(\vec{r}) = (\text{paramagnetic} + \text{diamagnetic}) \text{ currents} \tag{14.445}$$

and therefore,

$$\hat{H}_{\text{int}} = \int d^3\vec{r}\left[-\frac{e}{c}\vec{j}(\vec{r}) \cdot \vec{A}(\vec{r}, t) + \frac{e^2}{2mc^2}\rho(\vec{r})\vec{A}^2(\vec{r}, t) + e\phi(\vec{r}, t)\rho(\vec{r}) \right] \tag{14.446}$$

14.7.4 Induced Absorption and Emission

We will now use the transverse gauge, which says that the $\phi(\vec{r}, t)$ term is zero. We also assume that the radiation fields are small compared to the fields inside the atom, i.e., $\left|\vec{A}\right| << e^2/a_0$, which implies that we can neglect the $\vec{A}^2$ term compared to the $\vec{j} \cdot \vec{A}$ term. Therefore, we have

$$\hat{H}_{int} = -\frac{e}{c}\int d^3\vec{r}\,\vec{j}(\vec{r}) \cdot \vec{A}(\vec{r}, t) \tag{14.447}$$

For $\vec{A}$ as a linear superposition of monochromatic plane waves we then have

$$\begin{aligned}\hat{H}_{int} &= -\frac{e}{c}\int d^3\vec{r}\left(\frac{1}{2}\sum_i\left(\frac{\vec{p}_i}{m}\delta(\vec{r}-\vec{r}_i)+\delta(\vec{r}-\vec{r}_i)\frac{\vec{p}_i}{m}\right)\right)\\ &\qquad\times\left(\sum_{\vec{k}\vec{\lambda}}\left[A_{\vec{k}\vec{\lambda}}\vec{\lambda}\frac{e^{i\vec{k}\cdot\vec{r}-i\omega t}}{\sqrt{V}}+A^*_{\vec{k}\vec{\lambda}}\vec{\lambda}^*\frac{e^{-i\vec{k}\cdot\vec{r}+i\omega t}}{\sqrt{V}}\right]\right)\\ &= -\frac{e}{2c\sqrt{V}}\sum_{\vec{k}\vec{\lambda}}\sum_i\left(\begin{array}{l}A_{\vec{k}\vec{\lambda}}\vec{\lambda}\frac{\vec{p}_i}{m}e^{i\vec{k}\cdot\vec{r}_i-i\omega t}+A^*_{\vec{k}\vec{\lambda}}\vec{\lambda}^*\frac{\vec{p}_i}{m}e^{-i\vec{k}\cdot\vec{r}_i+i\omega t}\\ \quad+A_{\vec{k}\vec{\lambda}}\vec{\lambda}e^{i\vec{k}\cdot\vec{r}_i-i\omega t}\frac{\vec{p}_i}{m}+A^*_{\vec{k}\vec{\lambda}}\vec{\lambda}^*e^{-i\vec{k}\cdot\vec{r}_i+i\omega t}\frac{\vec{p}_i}{m}\end{array}\right)\\ &= -\frac{e}{c\sqrt{V}}\sum_{\vec{k}\vec{\lambda}}\left[A_{\vec{k}\vec{\lambda}}\vec{j}_{-\vec{k}}\cdot\vec{\lambda}e^{-i\omega t}+A^*_{\vec{k}\vec{\lambda}}\vec{j}_{\vec{k}}\cdot\vec{\lambda}^*e^{i\omega t}\right]\end{aligned}\tag{14.448}$$

where

$$\vec{j}_{\vec{k}} = \frac{1}{2}\sum_i\left(\frac{\vec{p}_i}{m}e^{-i\vec{k}\cdot\vec{r}_i}+e^{-i\vec{k}\cdot\vec{r}_i}\frac{\vec{p}_i}{m}\right) = \int d^3\vec{r}e^{-i\vec{k}\cdot\vec{r}}\vec{j}(\vec{r})\tag{14.449}$$

As we saw in the discussion of time-dependent perturbation theory for a harmonic perturbation, the $e^{-i\omega t}$ term implies an absorption of radiation process and the $e^{i\omega t}$ term implies and emission of radiation process. Following the same steps as that case, we have for the absorption transition rate

$$\Gamma^{abs}_{0\to n;\vec{k}\vec{\lambda}} = \frac{2\pi}{\hbar}\delta(\varepsilon_n-\varepsilon_0-\hbar\omega)\frac{e^2}{Vc^2}\left|A_{\vec{k}\vec{\lambda}}\right|^2\left|\langle n|\vec{j}_{\vec{k}}\cdot\vec{\lambda}|0\rangle\right|^2\tag{14.450}$$

To find the total rate of transition we must sum over $\vec{k}$ and $\vec{\lambda}$ (2 polarizations for each $\vec{k}$) to get

$$\Gamma^{abs}_{0\to n} = \frac{2\pi}{\hbar V}\sum_{\vec{k}\vec{\lambda}}\delta(\varepsilon_n-\varepsilon_0-\hbar\omega)\frac{e^2}{c^2}\left|A_{\vec{k}\vec{\lambda}}\right|^2\left|\langle n|\vec{j}_{-\vec{k}}\cdot\vec{\lambda}|0\rangle\right|^2\tag{14.451}$$

Now we can write

$$\frac{1}{V}\sum_{\vec{k}}\to\int\frac{k^2dkd\Omega}{(2\pi)^3} = \int\frac{\omega^2d\omega d\Omega}{(2\pi c)^3}\tag{14.452}$$

so that

$$\Gamma^{abs}_{0\to n} = \frac{2\pi e^2}{\hbar^2c^2}\frac{\omega^2}{2\pi c^3}\int d\Omega\sum_{\vec{\lambda}}\left|A_{\vec{k}\vec{\lambda}}\right|^2\left|\langle n|\vec{j}_{\vec{k}}\cdot\vec{\lambda}|0\rangle\right|^2\tag{14.453}$$

where

$$\omega = \frac{\varepsilon_n-\varepsilon_0}{\hbar}\ \text{(from the }\delta-\text{ function)}\tag{14.454}$$

If the incident light beam subtends a solid angle $d\Omega$ and it is polarized with polarization vector $\vec{\lambda}$, then the total rate of energy transport in the beam is the time average of the Poynting vector which is given by

$$\frac{1}{V}\sum_{\vec{k}}\frac{\omega^2}{2\pi c}\left|A_{\vec{k}\vec{\lambda}}\right|^2 = d\Omega\int d\omega\frac{\omega^4}{(2\pi c)^4}\left|A_{\vec{k}\vec{\lambda}}\right|^2\tag{14.455}$$

Now

$$I(\omega) = d\Omega \frac{\omega^4}{(2\pi c)^4} \left|A_{\vec{k}\vec{\lambda}}\right|^2$$
$$= \text{ intensity of the incident beam per unit frequency} \tag{14.456}$$

In a similar way

$$\Gamma_{n\to 0}^{ind\ emis} = \frac{4\pi^2 e^2}{\hbar^2 c\omega^2} I(\omega) \left|\langle n| \vec{j}_{\vec{k}} \cdot \vec{\lambda}^* |0\rangle\right|^2 \tag{14.457}$$

Since

$$\langle n| \vec{j}_{\vec{k}} \cdot \vec{\lambda}^* |0\rangle = \langle n| \vec{j}_{-\vec{k}} \cdot \vec{\lambda} |0\rangle^* \tag{14.458}$$

we have

$$\Gamma_{0\to n}^{abs} = \Gamma_{n\to 0}^{ind\ emis} \tag{14.459}$$

(this is the origin of the Einstein A and B coefficients).

In the absorption process, the absorption of one photon of energy $\hbar\omega = \varepsilon_n - \varepsilon_0$ causes an upward transition. The electron gains energy and the electromagnetic field loses energy. Induced emission is just the opposite.

Now a photon of frequency ω and energy $\hbar\omega$ and therefore, the total energy in the incident beam is

$$E = \sum_{\vec{k}\vec{\lambda}} \hbar\omega N_{\vec{k}\vec{\lambda}} \tag{14.460}$$

where $N_{\vec{k}\vec{\lambda}}$ = the number of photons in the $(\vec{k}, \vec{\lambda})$ mode in the beam. But we already have

$$E = \sum_{\vec{k}\vec{\lambda}} \frac{\omega^2}{2\pi c^2} \left|A_{\vec{k}\vec{\lambda}}\right|^2 \tag{14.461}$$

which says that

$$\left|A_{\vec{k}\vec{\lambda}}\right|^2 = \frac{2\pi\hbar c^2}{\omega} N_{\vec{k}\vec{\lambda}} \tag{14.462}$$

and thus

$$\Gamma_{0\to n}^{abs} = \Gamma_{n\to 0}^{ind\ emis} = \sum_{\vec{k}\vec{\lambda}} \frac{4\pi^2 e^2}{\omega V} \delta(\varepsilon_n - \varepsilon_0 - \hbar\omega) \left|\langle n| \vec{j}_{-\vec{k}} \cdot \vec{\lambda} |0\rangle\right|^2 N_{\vec{k}\vec{\lambda}} \tag{14.463}$$

14.7.5 Quantized Radiation Field and Spontaneous Emission

Up to this point we have been treating the electromagnetic field classically as a wave. We have mentioned the idea of photons, but have not created any formal quantum mechanical structure to describe them, i.e., we have been considering what happens to the atom and ignoring what is happening to the EM field during these processes.

To bring out the structure of the theory in terms of photons, we must now describe these processes in terms of state vectors, such that, in the absorption process the atom makes a transition from $|0\rangle \rightarrow |n\rangle$ while the electromagnetic field makes a transition from an initial state to a state with *one less photon* (it has been absorbed).

All of our development so far has involved what is physically called an incoherent beam of light.

We related $\left|A_{\vec{k}\vec{\lambda}}\right|$ and $N_{\vec{k}\vec{\lambda}}$ so that knowledge of the $N_{\vec{k}\vec{\lambda}}$ clearly does not imply any information about the relative phases of the $A_{\vec{k}\vec{\lambda}}$ which is the meaning of the term incoherent.

An incoherent beam, therefore, is completely specified by the photon numbers, i.e., the $N_{\vec{k}\vec{\lambda}}$. It is in this sense that we can write the initial state(normalized) of the electromagnetic field as

$$\left|N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right\rangle \tag{14.464}$$

where, as before, the $N_{\vec{k}\vec{\lambda}}$ = the number of photons in the mode $(\vec{k}, \vec{\lambda})$.

Any two of these states are orthogonal if they differ in the number of photons in any mode.

The final state of the electromagnetic field after photon absorption of a photon in the mode $(\vec{k}, \vec{\lambda})$ is

$$\left|N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}} - 1, ...\right\rangle \tag{14.465}$$

We assume that there exists some $\hat{H}_{int}$ that causes both transitions (atom and electromagnetic field) as it couples the electromagnetic field to matter. We define

$$\text{initial state } = |0\rangle \left|N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right\rangle \tag{14.466}$$

$$\text{final state } = |n\rangle \left|N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}} - 1, ...\right\rangle \tag{14.467}$$

so that

$$E_{initial} = \varepsilon_0 + \sum_{\vec{k}'\vec{\lambda}'} \hbar c k' N_{\vec{k}'\vec{\lambda}'} \tag{14.468}$$

$$E_{final} = \varepsilon_n + \sum_{\vec{k}'\vec{\lambda}'} \hbar c k' N_{\vec{k}'\vec{\lambda}'} - \hbar c k \tag{14.469}$$

The transition rate between the two states is given by Fermi's golden rule as

$$\frac{2\pi}{\hbar}\delta(\varepsilon_n - \varepsilon_0 - \hbar\omega)\left|\langle final|\, \hat{H}_{\text{int}}\, |initial\rangle\right|^2 \tag{14.470}$$

This must be the same as our earlier result (20.463) which implies that we must have

$$\begin{aligned}\left|\langle final|\,\hat{H}_{int}\,|initial\rangle\right|^2 &= \frac{e^2}{Vc^2}\left|A_{\vec{k}\vec{\lambda}}\right|^2\left|\langle n|\,\vec{j}_{-\vec{k}}\cdot\vec{\lambda}\,|0\rangle\right|^2 \\ &= \frac{e^2}{Vc^2}\frac{2\pi\hbar c^2}{\omega}N_{\vec{k}\vec{\lambda}}\left|\langle n|\,\vec{j}_{-\vec{k}}\cdot\vec{\lambda}\,|0\rangle\right|^2 \end{aligned} \tag{14.471}$$

This implies that as yet undetermined operator $\hat{H}_{int}$ must have the following properties:

1. it must include a part $\vec{j}_{-\vec{k}}\cdot\vec{\lambda}$ that acts on the atom
2. it must have a part that decreases the number of photons in the $(\vec{k},\vec{\lambda})$ mode by 1
3. it must be Hermitian

One way of doing this is to write

$$\hat{H}_{int} = \frac{e}{c\sqrt{V}}\sum_{\vec{k}'\vec{\lambda}'}\left(\vec{j}_{-\vec{k}'}\cdot\vec{\lambda}'A^{(op)}_{\vec{k}'\vec{\lambda}'} + \vec{j}_{\vec{k}'}\cdot\vec{\lambda}'^{*}A^{(op)+}_{\vec{k}'\vec{\lambda}'}\right) \tag{14.472}$$

where $A^{(op)}_{\vec{k}\vec{\lambda}}$ *reduces the number of photons* in the $(\vec{k},\vec{\lambda})$ mode by 1. It is a photon in mode $(\vec{k},\vec{\lambda})$ *annihilation operator.*

The second term is required to make $\hat{H}_{int}$ Hermitian. Using this model we then have

$$\begin{aligned}&\langle final|\,\hat{H}_{int}\,|initial\rangle \\ &\quad= \left\langle n; N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}-1, ...\right|\hat{H}_{\text{int}}\left|0; N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right\rangle \\ &\quad= -\frac{e}{c}\langle n|\,\vec{j}_{-\vec{k}}\cdot\vec{\lambda}\,|0\rangle \\ &\qquad\times\left\langle N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}-1, ...\right|A^{(op)}_{\vec{k}\vec{\lambda}}\left|N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right\rangle\end{aligned} \tag{14.473}$$

For agreement with the earlier result we must have

$$\begin{aligned}&\left\langle N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}-1, ...\right|A^{(op)}_{\vec{k}\vec{\lambda}}\left|N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right\rangle \\ &\qquad= \sqrt{\frac{2\pi\hbar c^2}{\omega}}\sqrt{N_{\vec{k}\vec{\lambda}}}\end{aligned} \tag{14.474}$$

This matrix element of $A^{(op)}_{\vec{k}\vec{\lambda}}$ corresponds to the $A_{\vec{k}\vec{\lambda}}$ term in the classical field picture.

The matrix element implies that

$$\begin{aligned}
&\left\langle N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}-1, ...\right| A^{(op)}_{\vec{k}\vec{\lambda}} \left| N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right\rangle^* \\
&\quad = \left\langle N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right| A^{(op)+}_{\vec{k}\vec{\lambda}} \left| N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}-1, ...\right\rangle \\
&\quad = \sqrt{\frac{2\pi\hbar c^2}{\omega}}\sqrt{N_{\vec{k}\vec{\lambda}}}
\end{aligned} \tag{14.475}$$

which says that $A^{(op)+}_{\vec{k}\vec{\lambda}}$ is an operator that increases the number of photons in the $(\vec{k},\vec{\lambda})$ mode by 1. It is a photon in mode $(\vec{k},\vec{\lambda})$ *creation operator*. We thus have

$$\begin{aligned}
&A^{(op)}_{\vec{k}\vec{\lambda}} \left| N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right\rangle \\
&\qquad = \sqrt{\frac{2\pi\hbar c^2}{\omega}}\sqrt{N_{\vec{k}\vec{\lambda}}} \left| N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}-1, ...\right\rangle
\end{aligned} \tag{14.476}$$

$$\begin{aligned}
&A^{(op)+}_{\vec{k}\vec{\lambda}} \left| N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right\rangle \\
&\qquad = \sqrt{\frac{2\pi\hbar c^2}{\omega}}\sqrt{N_{\vec{k}\vec{\lambda}}+1} \left| N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}+1, ...\right\rangle
\end{aligned} \tag{14.477}$$

This behavior is identical (aside from the $\sqrt{2\pi\hbar c^2/\omega}$ factor) to that of the $\hat{a}$ and $\hat{a}^+$ operators in the harmonic oscillator problem.

This model gives a quantum mechanical picture of the electromagnetic radiation field as an infinite number of *harmonic oscillators* - one per mode and the quanta associated with these oscillators are photons.

If we define a Hermitian electromagnetic field operator as

$$\vec{A}^{(op)}(\vec{r}) = \sum_{\vec{k}\vec{\lambda}} \left[A^{(op)}_{\vec{k}\vec{\lambda}} \vec{\lambda} \frac{e^{i\vec{k}\bullet\vec{r}}}{\sqrt{V}} + A^{(op)+}_{\vec{k}\vec{\lambda}} \vec{\lambda}^* \frac{e^{-i\vec{k}\bullet\vec{r}}}{\sqrt{V}} \right] \tag{14.478}$$

we have

$$\hat{H}_{int} = \int d^3\vec{r} \left[-\frac{e}{c}\vec{j}(\vec{r}) \cdot \vec{A}^{(op)}(\vec{r}) + \frac{e^2}{2mc^2}\rho(\vec{r}) \left(\vec{A}^{(op)}(\vec{r})\right)^2 \right] \tag{14.479}$$

In the interaction representation $\vec{A}^{(op)}(\vec{r},t)$ has the time dependence

$$\vec{A}^{(op)}(\vec{r},t) = e^{\frac{i}{\hbar}\hat{H}_{em}t} \vec{A}^{(op)}(\vec{r}) e^{-\frac{i}{\hbar}\hat{H}_{em}t} \tag{14.480}$$

where $\hat{H}_{em}$ = Hamiltonian for free radiation. We then have

$$\hat{H}_{em} = \frac{1}{8\pi}\int d^3\vec{r} \left(\vec{\varepsilon}^2 + \vec{B}^2\right) = \sum_{\vec{k}\vec{\lambda}} \hbar ck \left(A^{(op)+}_{\vec{k}\vec{\lambda}} A^{(op)}_{\vec{k}\vec{\lambda}} + \frac{1}{2} \right) \tag{14.481}$$

The operator algebra similarity to the $\hat{a}$ and $\hat{a}^+$ problem then allows us to write

$$\left[A^{(op)}_{\vec{k}\vec{\lambda}}, A^{(op)+}_{\vec{k}'\vec{\lambda}'}\right] = \frac{2\pi\hbar c^2}{\omega}\delta_{\vec{k}\vec{k}'}\delta_{\vec{\lambda}\vec{\lambda}'} \quad , \quad \left[A^{(op)}_{\vec{k}\vec{\lambda}}, A^{(op)}_{\vec{k}'\vec{\lambda}'}\right] = 0 \tag{14.482}$$

and

$$\left|N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right\rangle = \frac{1}{\sqrt{N_{\vec{k}\vec{\lambda}}!}}\left(A^{(op)+}_{\vec{k}\vec{\lambda}}\right)^{N_{\vec{k}\vec{\lambda}}}\left|N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., 0, ...\right\rangle \tag{14.483}$$

and $\hat{H} = \hat{H}_0 + \hat{H}_{em} + \hat{H}_{int}$ where $\hat{H}_0$ = Hamiltonian for the electrons.

We then have

$$\begin{aligned} & e^{\frac{i}{\hbar}\hat{H}_{em}t}A^{(op)}_{\vec{k}\vec{\lambda}}e^{-\frac{i}{\hbar}\hat{H}_{em}t}\left|....., N_{\vec{k}\vec{\lambda}}, ...\right\rangle \\ & \quad = e^{\frac{i}{\hbar}\hat{H}_{em}t}A^{(op)}_{\vec{k}\vec{\lambda}}e^{-i(N_{\vec{k}\vec{\lambda}}+\frac{1}{2})t}\left|....., N_{\vec{k}\vec{\lambda}}, ...\right\rangle \\ & \quad = e^{\frac{i}{\hbar}\hat{H}_{em}t}e^{-i(N_{\vec{k}\vec{\lambda}}+\frac{1}{2})t}\sqrt{N_{\vec{k}\vec{\lambda}}}\left|....., N_{\vec{k}\vec{\lambda}}-1, ...\right\rangle \\ & \quad = e^{i(N_{\vec{k}\vec{\lambda}}-1+\frac{1}{2})t}e^{-i(N_{\vec{k}\vec{\lambda}}+\frac{1}{2})t}\sqrt{N_{\vec{k}\vec{\lambda}}}\left|....., N_{\vec{k}\vec{\lambda}}-1, ...\right\rangle \\ & \quad = e^{-ickt}A^{(op)}_{\vec{k}\vec{\lambda}}\left|....., N_{\vec{k}\vec{\lambda}}, ...\right\rangle \end{aligned} \tag{14.484}$$

or

$$e^{\frac{i}{\hbar}\hat{H}_{em}t}A^{(op)}_{\vec{k}\vec{\lambda}}e^{-\frac{i}{\hbar}\hat{H}_{em}t} = e^{-ickt}A^{(op)}_{\vec{k}\vec{\lambda}} \tag{14.485}$$

and similarly

$$e^{\frac{i}{\hbar}\hat{H}_{em}t}A^{(op)+}_{\vec{k}\vec{\lambda}}e^{-\frac{i}{\hbar}\hat{H}_{em}t} = e^{ickt}A^{(op)+}_{\vec{k}\vec{\lambda}} \tag{14.486}$$

Putting this all together we have

$$\vec{A}^{(op)}(\vec{r}, t) = \sum_{\vec{k}\vec{\lambda}}\left[A^{(op)}_{\vec{k}\vec{\lambda}}\vec{\lambda}\frac{e^{i\vec{k}\cdot\vec{r}-i\omega t}}{\sqrt{V}} + A^{(op)+}_{\vec{k}\vec{\lambda}}\vec{\lambda}^*\frac{e^{-i\vec{k}\cdot\vec{r}+i\omega t}}{\sqrt{V}}\right] \tag{14.487}$$

By construction, we have forced the quantum mechanical description of absorption of the electromagnetic field in terms of the photon to be identical to the description in terms of the classical electromagnetic field for the induced absorption process.

We now apply the formalism to the emission process. This corresponds to the transition between the states

$$\text{initial state } = |0\rangle\left|N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}, ...\right\rangle \tag{14.488}$$

$$\text{final state } = |n\rangle\left|N_{\vec{k}_1\vec{\lambda}_1}, N_{\vec{k}_2\vec{\lambda}_2}, ..., N_{\vec{k}\vec{\lambda}}+1, ...\right\rangle \tag{14.489}$$

so that

$$E_{initial} = \varepsilon_0 + \sum_{\vec{k}'\vec{\lambda}'}\hbar ck' N_{\vec{k}'\vec{\lambda}'} \tag{14.490}$$

$$E_{final} = \varepsilon_n + \sum_{\vec{k}'\vec{\lambda}'}\hbar ck' N_{\vec{k}'\vec{\lambda}'} + \hbar ck \tag{14.491}$$

The transition rate is

$$\frac{2\pi}{\hbar}\delta(\varepsilon_n - \varepsilon_0 - \hbar ck)\left|\left\langle 0;.....N_{\vec{k}\vec{\lambda}}+1,.....\right|\hat{H}_{\rm int}\left|n;.....N_{\vec{k}\vec{\lambda}},....\right\rangle\right|^2 \qquad (14.492)$$

where

$$\begin{aligned}
&\left\langle 0;.....N_{\vec{k}\vec{\lambda}}+1,.....\right|\hat{H}_{int}\left|n;.....N_{\vec{k}\vec{\lambda}},....\right\rangle \\
&\quad = -\frac{e}{c\sqrt{V}}\left\langle 0\right|\vec{j}_{\vec{k}}\cdot\vec{\lambda}^*\left|n\right\rangle\left\langle .....N_{\vec{k}\vec{\lambda}}+1,.....\right|A^{(op)+}_{\vec{k}\vec{\lambda}}\left|.....N_{\vec{k}\vec{\lambda}},....\right\rangle \\
&\quad = -\frac{e}{c}\sqrt{\frac{2\pi\hbar c}{\omega V}}\left\langle 0\right|\vec{j}_{\vec{k}}\cdot\vec{\lambda}^*\left|n\right\rangle\sqrt{N_{\vec{k}\vec{\lambda}}+1} \qquad (14.493)
\end{aligned}$$

We no longer have any features that are unknown and hence adjustable. Forcing agreement with induced absorption makes the results for the emission process a prediction!

We get

$$\Gamma^{emis}_{n\to 0;\vec{k}\vec{\lambda}} = \frac{4\pi^2e^2}{\omega V}\delta(\varepsilon_n - \varepsilon_0 - \hbar ck)\left|\left\langle 0\right|\vec{j}_{\vec{k}}\cdot\vec{\lambda}^*\left|n\right\rangle\right|^2\left(N_{\vec{k}\vec{\lambda}}+1\right) \neq \Gamma^{abs}_{0\to n;\vec{k}\vec{\lambda}} \qquad (14.494)$$

which disagrees with the classical field result but agrees with experiment.

The $N_{\vec{k}\vec{\lambda}}$ part corresponds to the classical result.

The $+1$ part is a purely quantum mechanical effect.

This term implies that there is an emission process that can take place even if there is *no external field* present.

This process is called *spontaneous emission*. A clear victory for the quantum approach.

14.8 Problems

14.8.1 Dirac Spinors

The Dirac spinors are (with $E = \sqrt{\vec{p}^2 + m^2}$)

$$u(p,s) = \frac{\not{p}+m}{\sqrt{E+m}}\begin{pmatrix}\varphi_s\\0\end{pmatrix} \quad , \quad v(p,s) = \frac{-\not{p}+m}{\sqrt{E+m}}\begin{pmatrix}0\\\chi_s\end{pmatrix}$$

where $\not{p} = \gamma^\mu p_\mu$, $\varphi_s (s = \pm 1/2)$ are orthonormalized 2-spinors and similarly for χ_s. Prove(using $\bar{u} = u^+\gamma^0$, etc):

(a) $\bar{u}(p.s)u(p.s') = -\bar{v}(p,s)v(p,s') = 2m\delta_{ss'}$

(b) $\bar{v}(p,s)u(p,s') = 0$

(c) $\bar{u}(p,s)\gamma^0 u(p,s') = 2E\delta_{ss'}$

(d) $\sum_s u(p.s)\bar{u}(p,s) = \not{p} + m$

(e) $\sum_s v(p.s)\bar{v}(p,s) = \not{p} - m$

(f) $\bar{u}(p,s)\gamma^\mu u(p',s') = 2E\delta_{ss'} = \frac{1}{2m}\bar{u}(p,s)\left[(p+p')^\mu + i\sigma^{\mu\nu}(p-p')_\nu\right]u(p',s')$ $(The\,Gordon\,Identity)$

14.8.2 Lorentz Transformations

In a Lorentz transformation $x' = \Lambda x$ the Dirac wave function transforms as $\psi'(x') = S(\Lambda)\psi(x)$, where $S(\Lambda)$ is a 4×4 matrix.

(a) Show that the Dirac equation is invariant in form, i.e., $\left(i\gamma^\mu\partial'_\mu - m\right)\psi'(x') = 0$, provided
$$S^{-1}(\Lambda)\gamma^\mu S(\Lambda) = \Lambda^\mu{}_\nu\gamma^\nu$$

(b) For an infinitesimal transformation $\Lambda^\mu{}_\nu = g^\mu{}_\nu + \delta\omega^\mu{}_\nu$, where $\delta\omega_{\mu\nu} = -\delta\omega_{\nu\mu}$. The spin dependence of $S(\Lambda)$ is given by $I - i\sigma_{\mu\nu}\delta\omega^{\mu\nu}/4$. Show that $\sigma_{\mu\nu} = i[\gamma_\mu,\gamma_{nu}]$ satisfies the equation in part (a). For finite transformations we then have $S(\Lambda) = e^{-i\sigma_{\mu\nu}\omega^{\mu\nu}/4}$.

14.8.3 Dirac Equation in $1+1$ Dimensions

Consider the Dirac equation in $1+1$ Dimensions (i.e., one space and one time dimension):
$$\left(i\gamma^0\frac{\partial}{\partial x^0} + i\gamma^1\frac{\partial}{\partial x^1} - m\right)\psi(x) = 0$$

(a) Find a 2×2 matrix representation of γ^0 and γ^1 which satisfies $\{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}$ and has correct hermiticity. What is the physical reason that ψ can have only two components in $1+1$ dimensions?

(b) Find the representation of $\gamma_5 = \gamma^0\gamma^1$, $\gamma_5\gamma^\mu$ and $\sigma^{\mu\nu} = \frac{1}{2}i\left[\gamma^\mu,\gamma^\nu\right]$. Are they independent? Define a minimal set of matrices which form a complete basis.

(c) Find the plane wave solutions $\psi_+(x) = u(p^1)e^{-ip\cdot x}$ and $\psi_-(x) = v(p^1)e^{ip\cdot x}$ in $1+1$ dimensions, normalized to $\bar{u}u = -\bar{v}v = 2m$ (where $\bar{u} = u^+\gamma^0$).

14.8.4 Trace Identities

Prove the following trace identities for Dirac matrices using only their property $\{\gamma^\mu,\gamma^\nu\} = g^{\mu\nu}$ (i.e., do not use a specific matrix representation)

(a) $Tr(\gamma^\mu) = 0$

(b) $Tr(\gamma^\mu\gamma^\nu) = 4g^{\mu\nu}$

(c) $Tr(\gamma^\mu\gamma^\nu\gamma^\rho) = 0$

(d) $Tr(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) = 4g^{\mu\nu}g^{\rho\sigma} - 4g^{\mu\rho}g^{\nu\sigma} + 4g^{\mu\sigma}g^{\nu\rho}$

(e) $Tr(\gamma_5) = 0$ where $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$

14.8.5 Right- and Left-Handed Dirac Particles

The right (R) and left (L) -handed Dirac particles are defined by the projections

$$\psi_R(x) = \frac{1}{2}(1+\gamma_5)\psi(x) \quad , \quad \psi_L(x) = \frac{1}{2}(1-\gamma_5)\psi(x)$$

In the case of a massless particle (m=0):

(a) Show that the Dirac equation $(i\not{\partial} - e\not{A})\psi = 0$ does not couple $\psi_R(x)$ to $\psi_L(x)$, i.e., they satisfy independent equations. Specifically, show that in the chiral representation of the Dirac matrices

$$\gamma^0 = \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix} \quad , \quad \boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}$$

we have

$$\psi = \begin{pmatrix} \phi_R \\ \phi_L \end{pmatrix} e^{-ip\cdot x}$$

i.e., that the lower(upper) two components of ψ_R (ψ_L) vanish.

(b) For the free Dirac equation ($A^\mu = 0$) show that ϕ_R and ϕ_L are eigenstates of the helicity operator $\frac{1}{2}\boldsymbol{\sigma}\cdot\boldsymbol{p}$ with positive and negative helicity, respectively, for plane wave states with $p^0 > 0$.

14.8.6 Gyromagnetic Ratio for the Electron

(a) Reduce the Dirac equation $(i\not{\partial} - e\not{A} - m)\psi = 0$ by multiplying it with $(i\not{\partial} - e\not{A} + m)\psi = 0$ to the form

$$\left[(i\partial - eA)^2 - \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu} - m^2\right]\psi = 0$$

where $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$ and the field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

(b) Show that the dependence in the magnetic field $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ in the spin-dependent term $\sigma^{\mu\nu}F_{\mu\nu}$ is of the form $-(ge/2m)\frac{1}{2}\boldsymbol{\Sigma}\cdot\boldsymbol{B}$ when the kinetic energy is normalized to $-\nabla^2/2m$ ($\boldsymbol{\Sigma} = \gamma_5\gamma^0\boldsymbol{\gamma}$ is the spin matrix). Determine the value of the gyromagnetic ration g for the electron.

14.8.7 Dirac → Schrodinger

Reduce tyhe Dirac equation $(i\partial\!\!\!/ - eA\!\!\!/ - m)\psi = 0$ for the Hydrogen atom ($A^0 = -Ze/4\pi r$, $\boldsymbol{A} = 0$) to the standard Schrodinger equation

$$i\frac{\partial}{\partial t}\Psi(t, boldsymbolx) = \left(-\frac{\nabla^2}{2m} + eA^0\right)\Psi(t, boldsymbolx)$$

in the non-relativistic limit, where $|\boldsymbol{p}|$, $A^0 \ll m$. **HINT**: You may start from the reduced form of the Dirac equation in Problem 20.6(a). Extract the leading time dependence by writing $\psi(x) = \Psi(t, \boldsymbol{x})e^{-imt}$.

14.8.8 Positive and Negative Energy Solutions

Positive energy solutions of the Dirac equation correspond to the 4-vector current $\boldsymbol{J}^\mu = 2\boldsymbol{p}^\mu = 2(E, \vec{p})$, $E > 0$. Show that the negative energy solutions correspond to the current $\boldsymbol{J}^\mu = -2(E, \vec{p}) = -2(|E|, -\vec{p}) = -2\boldsymbol{p}^\mu$, $E < 0$.

14.8.9 Helicity Operator

(1) Show that the helicity operator commutes with the Hamiltonian:

$$\left[\vec{\Sigma}\cdot\hat{\boldsymbol{p}}, \boldsymbol{H}\right] = 0$$

(2) Show explicitly that the solutions to the Dirac equation are eigenvectors of the helicity operator:

$$\left[\vec{\Sigma}\cdot\hat{\boldsymbol{p}}\right]\Psi = \pm\Psi$$

14.8.10 Non-Relativisitic Limit

Consider

$$\Psi = \begin{pmatrix}\boldsymbol{u}_A \\ \boldsymbol{u}_B\end{pmatrix}$$

to be a solution of the Dirac equation where $\boldsymbol{u}_A$ and $\boldsymbol{u}_B$ are two-component spinors. Show that in the non-relativistic limit $\boldsymbol{u}_B \sim \beta = v/c$.

14.8.11 Gyromagnetic Ratio

Show that in the non-relativisitc limit the motion of a spin 1/2 fermion of charge e in the presence of an electromagnetic field $A^\mu = (A^0, \vec{A})$ is described by

$$\left[\frac{(\vec{p} - e\vec{A})^2}{2m} - \frac{e}{2m}\vec{\sigma}\cdot\vec{B} + eA^0\right]\chi = E\chi$$

where $\vec{B}$ is the magnetic field, σ^i are the Pauli matrices and $E = p^0 - m$. Identify the g-factor of the fermion and show that the Dirac equation predicts the correct gyromagnetic ratio for the fermion. To write down the Dirac equation in the presence of an electromagnetic field substitute: $p^\mu \to p^\mu - eA^\mu$.

14.8.12 Properties of γ_5

Show that:

(a) $\bar{\Psi}\gamma_5\Psi$ is a pseudoscalar.

(b) $\bar{\Psi}\gamma_5\gamma^\mu\Psi$ is an axial vector.

14.8.13 Lorentz and Parity Properties

Comment on the Lorentz and parity properties of the quantities:

(a) $\bar{\Psi}\gamma_5\gamma^\mu\Psi\bar{\Psi}\gamma_\mu\Psi$

(b) $\bar{\Psi}\gamma_5\Psi\bar{\Psi}\gamma_5\Psi$

(c) $\bar{\Psi}\Psi\bar{\Psi}\gamma_5\Psi$

(d) $\bar{\Psi}\gamma_5\gamma^\mu\Psi\bar{\Psi}\gamma_5\gamma_\mu\Psi$

(e) $\bar{\Psi}\gamma^\mu\Psi\bar{\Psi}\gamma_\mu\Psi$

14.8.14 A Commutator

Explicitly evaluate the commutator of the Dirac Hamiltonian with the orbital angular momentum operator $\hat{\boldsymbol{L}}$ for a free particle.

14.8.15 Solutions of the Klein-Gordon equation

Let $\phi(\vec{r},t)$ be a solution of the free Klein-Gordon equation. Let us write

$$\phi(\vec{r},t) = \psi(\vec{r},t)e^{-imc^2t/\hbar}$$

Under what conditions will $\psi(\vec{r},t)$ be a solution of the non-relativistic Schrodinger equation? Interpret your condition physically when ϕ is given by a plane-wave solution.

14.8.16 Matrix Representation of Dirac Matrices

The Dirac matrices must satisfy the anti-commutator relationships:

$$\{\alpha_i,\alpha_j\} = 2\delta_{ij} \ , \ \{\alpha_i,\beta\} = 0 \text{ with } \beta^2 = 1$$

(1) Show that the α_i, β are Hermitian, traceless matrices with eigenvalues ±1 and even dimensionality.

(2) Show that, as long as the mass term mis not zero and the matrix β is needed, there is no 2×2 set of matrices that satisfy all the above relationships. Hence the Dirtac matrices must be of dimension 4 or higher. First show that the set of matrices $\{I,\vec{\sigma}\}$ can be used to express any 2×2

matrix, i.e., the coefficients c_0, c_i always exist such that any 2×2 matrix can be written as:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = c_0 I + c_i \sigma_i$$

Having shown this, you can pick an intelligent choice for the α_i in terms of the Pauli matrices, for example $\alpha_i = \sigma_i$ which automatically obeys $\{\alpha_i, \alpha_j\} = 2\delta_{ij}$, and express β in terms of $\{I, \vec{\sigma}\}$ using the relation above. Show then that there is no 2×2 β matrix that satisfies $\{\alpha_i, \beta\} = 0$.

14.8.17 Weyl Representation

(1) Show that the Weyl matrices:

$$\vec{\alpha} = \begin{pmatrix} -\vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \quad , \quad \beta = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

satisfy all the Dirac conditions of Problem 20.16. Hence, they form just another representation of the Dirac matrices, the Weyl representation, which is different than the standard Pauli-Dirac representation.

(2) Show that the Dirac matrices in the Weyl representation are

$$\vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix} \quad , \quad \gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

(3) Show that in the Weyl representation $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}$

(4) Solve the Dirac equation $[\vec{\alpha} \cdot \vec{p} + \beta m]\Psi = E\Psi$ in the particle rest frame using the Weyl representation.

(5) Compute the result of the chirality operators

$$\frac{1 \pm \gamma_5}{2}$$

when they are acting on the Dirac solutions in the Weyl representation.

14.8.18 Total Angular Momentum

Use the Dirac Hamiltonian in the standard Pauli-Dirac representation

$$H = \vec{\alpha} \cdot \vec{p} + \beta m$$

to compute $[H, \hat{L}]$ and $[H, \hat{\Sigma}]$ and show that they are zero. Use the results to show that:

$$[H, \hat{L} + \hat{\Sigma}/2] = 0$$

where the components of the angular momentum operator are given by:

$$\hat{L}_i = \varepsilon_{ijk}\hat{x}_j\hat{p}_k$$

and the components of the spin operator are given by:

$$\hat{\Sigma}_i = \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}$$

Recall that the Pauli matrices satisfy $\sigma^i\sigma^j = \delta^{ij} + i\varepsilon^{ijk}\sigma^k$.

14.8.19 Dirac Free Particle

The Dirac equation for a free particle is

$$i\hbar\frac{\partial|\psi\rangle}{\partial t} = \left(c\alpha_x p_x + c\alpha_y p_y + c\alpha_z p_z + \beta mc^2\right)|\psi\rangle$$

Find all solutions and discuss their meaning. Using the identity

$$(\vec{\sigma}\cdot\vec{A})(\vec{\sigma}\cdot\vec{B}) = \vec{A}\cdot\vec{B} + i\vec{\sigma}\cdot(\vec{A}\times\vec{B})$$

will be useful.