Paul E. Dimotakis

Lectures on Quantum Physics

and

Applications

California Institute of Technology

CONTENTS

I. CLASSICAL MECHANICS

1.1 The principle of stationary action.

Consider a system whose configuration at any time t can be described in terms of N coordinates

$$q_i(t) \quad \text{where} \quad i = 1, 2, 3, \ldots, N \; .$$

These can be thought of as the components of an N-dimensional vector*

$$\underline{q}(t) = [q_1(t),\; q_2(t),\; \ldots,\; q_N(t)] \; , \tag{1.1}$$

which is called the configuration of the system. The solution of the equations of motion for the coordinates $q_i(t)$ describes the evolution of the configuration vector $\underline{q}(t)$ in time, i.e.

$$\begin{aligned} \underline{q}(t+\delta t) &= [q_1(t+\delta t),\; q_2(t+\delta t),\; \ldots,\; q_N(t+\delta t)] \\ &= [q_1(t) + \dot{q}_1(t)\,\delta t,\; \ldots\ldots,\; q_N(t) + \dot{q}_N(t)\,\delta t] \\ &= \underline{q}(t) + \dot{\underline{q}}(t)\,\delta t \; , \end{aligned} \tag{1.2}$$

where

$$\dot{\underline{q}}(t) = [\dot{q}_1(t),\; \dot{q}_2(t),\; \ldots,\; \dot{q}_N(t)] \; . \tag{1.3}$$

It is a remarkable fact that all of classical mechanics is derivable from a unifying principle, the principle of stationary action (Hamilton's principle). The principle of stationary action states that it is possible to define a scalar function of the configuration vector $\underline{q}(t)$ and its time derivative $\dot{\underline{q}}(t)$, i.e.

$$L = L(\underline{q}, \dot{\underline{q}}, t) \tag{1.4}$$

called the Lagrangian of the system, such that the integral

$$I\{\underline{q}\} = \int_{t_1}^{t_2} L(\underline{q}, \dot{\underline{q}}, t)\, dt \tag{1.5}$$

is an extremum (minimum or maximum), or stationary along the particular path $\underline{q}(t)$, in the time interval $t_1 < t < t_2$, which describes the evolution of the system. The integral of equation 1.5 is called the action and is a functional of the path $\underline{q}(t)$, i.e., it assumes a particular value $I\{\underline{q}\}$ for a given (vector) function of time $\underline{q}(t)$ in the interval $t_1 < t < t_2$. A functional is richer in content than a function. A function assumed a particular value for a given value (or set of values) of its independent variables. A functional assumes a particular value for a given function (or set of functions), which

---

*Vectors will be denoted with underbars in these lectures.

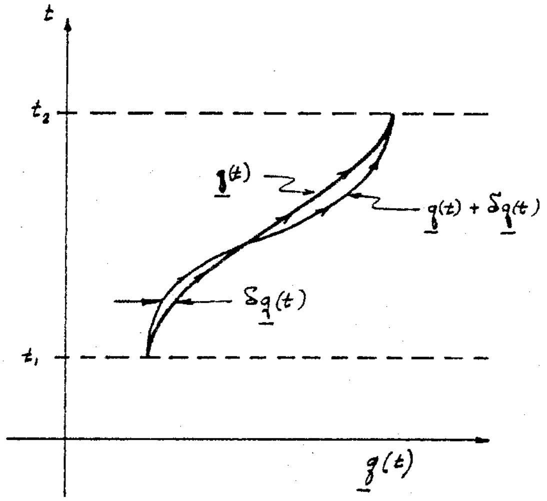
serves as its independent variable.

The condition that the functional $I\{\underline{q}\}$ is stationary along the particular path $\underline{q}(t)$ implies that if we alter $\underline{q}(t)$ by small amount $\delta\underline{q}(t)$ to obtain a neighboring path, i.e.

$$\underline{q}(t) \to \underline{q}(t) + \delta\underline{q}(t) \ , \tag{1.6a}$$

which goes through the endpoints of the original path, i.e.

$$\delta\underline{q}(t_1) = \delta\underline{q}(t_2) = 0 \ , \tag{1.6b}$$



the functional $I\{\underline{q}\}$ remains unchanged to first order in $\delta\underline{q}$. In other words,

$$\delta I\{\underline{q}\} = I\{\underline{q}+\delta\underline{q}\} - I\{\underline{q}\} = 0(\delta\underline{q}^2) \ . \tag{1.7}$$

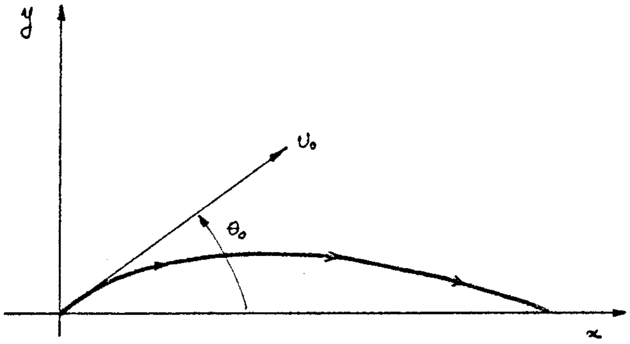
The requirement that the functional $I\{\underline{q}\}$, the action in this case, be stationary along the path $\underline{q}(t)$ is, in fact, sufficient to specify the path in the time interval $t_1 < t < t_2$, given the initial conditions

$$\underline{q}(t_1) = \underline{q}_1 \tag{1.8a}$$

$$\dot{\underline{q}}(t_1) = \dot{\underline{q}}_1 \ . \tag{1.8b}$$

The particular path that solves the problem of extremizing the functional is called the <u>extremal</u>.

Example 1.1. The ballistic trajectory



The configuration vector for this system is two dimensional, i.e.

$$\underline{q}(t) = [x(t),\ y(t)] \tag{1.9}$$

The initial conditions specify

$$\underline{q}(0) = [x(0),\ y(0)] = (0,0) \tag{1.10a}$$

and

$$\dot{\underline{q}}(0) = [\dot{x}(0),\ \dot{y}(0)] = (v_0 \cos\theta_0,\ v_0 \sin\theta_0)\ . \tag{1.10b}$$

The Lagrangian function in this case is given by

$$L(\underline{q}, \underline{q}, t) = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - mgy\ , \tag{1.11}$$

where m is the mass of the ballistic object and g is the accelleration of gravity, and the action integral is therefore given by

$$I\{\underline{q}\} = m \int_0^{t_2} [\tfrac{1}{2}(\dot{x}^2 + \dot{y}^2) - gy]\, dt\ . \tag{1.12}$$

The end time t_2 is defined implicitly by the condition

$$y(t_2) = 0\ . \tag{1.13}$$

You can convince yourself that the parabolic trajectory

$$x(t) = (v_0 \cos\theta_0)t \tag{1.14a}$$

$$y(t) = (v_0 \sin\theta_0)t - \tfrac{1}{2}gt^2 \tag{1.14b}$$

is the extremal of the action integral of equation 1.13. Try a few other paths!

Given the Lagrangian function $L(\underset{\sim}{q}, \dot{\underset{\sim}{q}}, t)$, the problem of determining the extremal, i.e. the particular path $\underset{\sim}{q}(t)$ that extremizes the action integral of equation 1.5, is a topic of the Calculus of Variations. See Appendix A. In particular, requiring that the variation vanish to first order, we have

$$\delta I\{\underset{\sim}{q}\} = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial \underset{\sim}{q}} \cdot \delta \underset{\sim}{q} + \frac{\partial L}{\partial \dot{\underset{\sim}{q}}} \cdot \delta \dot{\underset{\sim}{q}} \right) dt = 0 , \tag{1.15}$$

where

$$\frac{\partial L}{\partial \underset{\sim}{q}} \equiv \left(\frac{\partial L}{\partial q_1}, \frac{\partial L}{\partial q_2}, \ldots \quad , \frac{\partial L}{\partial q_N} \right) \tag{1.16}$$

and

$$\frac{\partial L}{\partial \dot{\underset{\sim}{q}}} \equiv \left(\frac{\partial L}{\partial \dot{q}_1}, \frac{\partial L}{\partial \dot{q}_2}, \ldots \quad , \frac{\partial L}{\partial \dot{q}_N} \right) . \tag{1.17}$$

We can now integrate the second term in the integrand of equation 1.15 by parts to obtain

$$\delta I = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial \underset{\sim}{q}} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\underset{\sim}{q}}} \right) \right] \cdot \delta \underset{\sim}{q}(t) \, dt \quad . \tag{1.18}$$

Since we now require that the action $I\{\underset{\sim}{q}\}$ is stationary <u>for an arbitrary variation</u> of the path, we must have

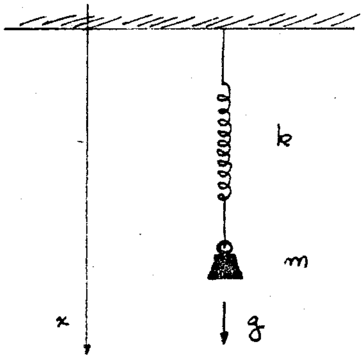
$$\boxed{\frac{\partial L}{\partial \underset{\sim}{q}} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\underset{\sim}{q}}} \right) = 0} \quad . \tag{1.19}$$

Equation 1.19 is a vector equation which must hold for each component $q_i(t)$ of the N-dimensional vector $\underset{\sim}{q}(t)$, i.e.,

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = 0 \quad , \; i = 1, N \tag{1.19'}$$

<u>Example 1.2</u> <u>One-dimensional harmonic oscilator</u>.

Consider a mass m suspended by a spring with a spring constant k in a gravitational field with an acceleration $\underset{\sim}{g}$.



The Lagrangian for this system is given by

$$L(x, \dot{x}, t) = \frac{1}{2} m \dot{x}^2 - \frac{1}{2} k x^2 + mgx \quad .$$

Note that, in this example, the $\underline{q}$ vector is one-dimensional, i.e.,

$$\underline{q}(t) = [q_1(t)] = [x(t)]$$

and the equation of motion is given by 1.19'

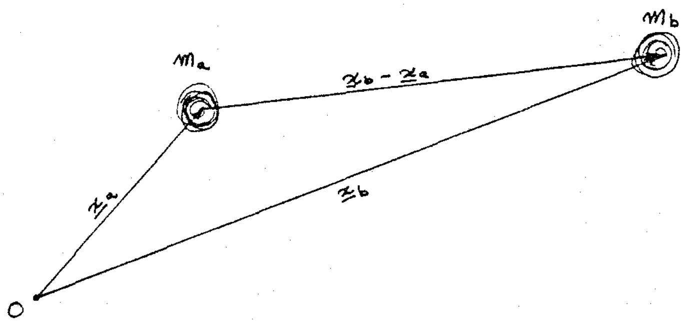
$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = mg - kx - m\ddot{x} = 0 \ ,$$

or

$$m\ddot{x} = - kx + mg \ ,$$

recognized as Newton's law ($\underline{F} = m\underline{a}$).

Example 1.3 Motion of two bodies subject to mutual gravitational force.



The Lagrangian for this system is given by

$$L = \frac{1}{2} m_a |\dot{\underline{x}}_a|^2 + \frac{1}{2} m_b |\dot{\underline{x}}_b|^2 + \frac{G\, m_a m_b}{|\underline{x}_a - \underline{x}_b|} \quad , \tag{1.20}$$

where G is the gravitational constant and m_a, m_b are the masses of the two bodies. The $\underline{q}$ vector in this example is six-dimensional, i.e.,

$$\underline{q} = (q_1, q_2, q_3, q_4, q_5, q_6)$$

$$= (x_a, y_a, z_a, x_b, y_b, z_b) = (\underline{x}_a, \underline{x}_b)$$

and the equations of motion are given by 1.19', for i = 1, 6, which can be written as two 3-vector equations, i.e.,

$$\frac{\partial L}{\partial \underline{x}_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\underline{x}}_a}\right) = 0 \tag{1.21a}$$

and

$$\frac{\partial L}{\partial \underline{x}_b} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\underline{x}}_b}\right) = 0 \;, \tag{1.21b}$$

or

$$m_a \ddot{\underline{x}}_a = -G\, m_a m_b \frac{\underline{x}_a - \underline{x}_b}{|\underline{x}_a - \underline{x}_b|^3} \tag{1.22a}$$

and

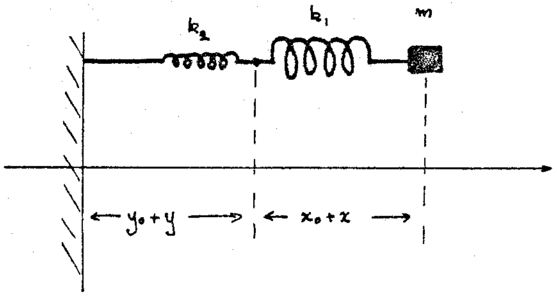
$$m_b \ddot{\underline{x}}_b = -\,G\, m_a m_b \frac{\underline{x}_b - \underline{x}_a}{|\underline{x}_b - \underline{x}_a|} \;. \tag{1.22b}$$

The Lagrangian for non-relativistic systems whose forces are derivable from a potential U which is a function of $\underline{q} = (q_1, q_2, \ldots, q_N)$ only, and not of $\dot{\underline{q}}$, can be expressed as the difference between the kinetic energy K and the potential energy U, i.e.,

$$L = K - U \tag{1.23}$$

and can usually be written down by inspection. In fact, this is how the appropriate Lagrangian was found for the systems described in examples 1.1, 1.2 and 1.3.

Example 1.4 Find the frequency of oscillation of a mass m constrained to move in one dimension and connected to a rigid boundary by means of two springs in series, with spring constants k_1 and k_2.



If we define x and y as the elongations of springs k_1 and k_2 respectively, we have

$$K = \frac{1}{2} m(\dot{x} + \dot{y})^2$$

$$U = \frac{1}{2}(k_1 x^2 + k_2 y^2)$$

and therefore

$$L = \frac{1}{2} m(\dot{x} + \dot{y})^2 - \frac{1}{2}(k_1 x^2 + k_2 y^2) \;. \tag{1.24}$$

The $\underline{q}$ vector for this system is two-dimensional and given by

$$\underline{q}(t) = [x(t), y(t)] \ .$$

The equations of motion are

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = -k_1 x - m(\ddot{x} + \ddot{y}) = 0 \tag{1.25a}$$

and

$$\frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) = -k_2 y - m(\ddot{x} + \ddot{y}) = 0 \ . \tag{1.25b}$$

Subtracting the two equations we have

$$k_1 x = k_2 y \tag{1.26}$$

which if we substitute in 1.25a yields

$$m\left(1 + \frac{k_1}{k_2}\right)\ddot{x} + k_1 x = 0$$

or

$$m\ddot{x} + kx = 0 \tag{1.27a}$$

where

$$k = \frac{k_1 k_2}{k_1 + k_2} \ . \tag{1.27b}$$

The system therefore behaves as if the mass was connected by means of a single spring with a spring constant k given by 1.27b. The frequency of oscillation is then given by

$$\omega = (k/m)^{1/2}$$

as can easily be verified by substituting $x(t) = x(0)\, e^{-i\omega t}$ into equation 1.27a.

The form of the equations of motion, i.e., 1.19', suggests that we should consider the quantity $\partial L/\partial \dot{q}_i$ as the momentum associated with the coordinate q_i. In particular, since

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} \tag{1.28}$$

and, in the cases where

$$L(\underline{q}, \dot{\underline{q}}, t) = K(\dot{\underline{q}}, t) - U(\underline{q}, t) \ , \tag{1.29}$$

the quantity on the right of 1.28 is identifiable as the force affecting the coordinate q_i, i.e.,

$$\frac{\partial L}{\partial q_i} = -\frac{\partial}{\partial q_i} U(\underline{q}, t) = F_i(\underline{q}, t) \ , \tag{1.30}$$

and since, from Newton's laws

$$\frac{d}{dt} p_i = F_i(\underline{q}, t) \qquad (1.31)$$

we have

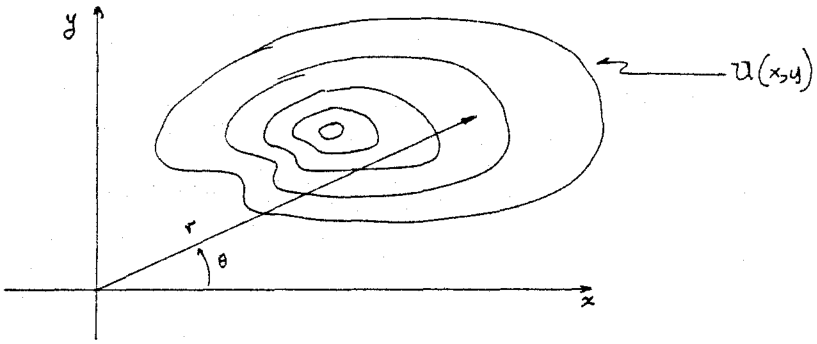
$$\frac{\partial L}{\partial \dot{q}_i} = p_i \ .$$

The identification of $\partial L/\partial \dot{q}_i$ as the momentum associated with the coordinate q_i survives generalizations of equation 1.28 and we are led to define as the generalized momentum p_i, associated with the coordinate q_i, as

$$\boxed{p_i \equiv \frac{\partial L}{\partial \dot{q}_i}} \ , \qquad (1.32)$$

in general.

Example 1.5 Motion of a particle of mass m in a two-dimensional potential.



Kinetic energy: $K = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2)$

Potential energy: $U(x, y)$

Consider now a transformation to polar coordinates

$$x = r \cos \theta$$
$$y = r \sin \theta$$

in terms of which we have

$$\dot{x} = \dot{r} \cos \theta - r\dot{\theta} \sin \theta$$
$$\dot{y} = \dot{r} \sin \theta + r\dot{\theta} \cos \theta$$

and therefore

Kinetic energy: $K = \frac{1}{2} m[\dot{r}^2 + (r\dot{\theta})^2]$ (1.33a)

Potential energy: $U = U(r \cos \theta, \ r \sin \theta) \equiv V(r, \theta)$. (1.33b)

The Lagrangian is then given by

$$L(r, \theta, \dot{r}, \dot{\theta}, t) = \frac{1}{2} m[\dot{r}^2 + (r\dot{\theta})^2] - V(r, \theta) \ .$$

The $\underset{\sim}{q}$ vector for this system is two dimensional, i.e.,

$$q_1(t) = r(t) \qquad\qquad q_2(t) = \theta(t) \ .$$

The equations of motion are then

$$\frac{\partial L}{\partial q_1} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_1}\right) = \frac{\partial L}{\partial r} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) = 0 \ .$$

or

$$\frac{d}{dt}(m\dot{r}) = -\frac{\partial V}{\partial r} + mr\dot{\theta}^2 \tag{1.34}$$

and

$$\frac{\partial L}{\partial q_2} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_2}\right) = \frac{\partial L}{\partial \theta} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = 0$$

or

$$\frac{d}{dt}(mr^2\dot{\theta}) = -\frac{\partial V}{\partial \theta} \ . \tag{1.35}$$

We see that the quantity $m\dot{r}$ in equation 1.34 must be associated with the (linear) momentum along the radial direction, i.e.,

$$p_r = m\dot{r} = \frac{\partial L}{\partial \dot{r}} \ .$$

Note that $-\partial V/\partial r$ is the proper expression of the radial force, while the term $mr\dot{\theta}^2$ properly accounts for the centripetal acceleration (centrifugal force). Similarly, the quantity $-\partial V/\partial \theta$ can be identified as the torque applied to the particle by the potential, since

$$-\frac{\partial V}{\partial \theta} = \left(-\frac{\partial U}{\partial x}\right)\frac{\partial x}{\partial \theta} + \left(-\frac{\partial U}{\partial y}\right)\frac{\partial y}{\partial \theta} = -r\sin\theta\, F_x + r\sin\theta\, F_y = \underline{r} \times \underline{F} \ ,$$

where

$$\underline{F} = -\frac{\partial U}{\partial \underline{x}} = -\frac{\partial U}{\partial x}\hat{e}_x - \frac{\partial U}{\partial y}\hat{e}_y$$

is the (two-dimensional) force derivable as the (two-dimensional) gradient of the potential. Therefore the quantity

$$p_\theta \equiv mr^2\dot{\theta} = \frac{\partial L}{\partial \dot{\theta}}$$

of equation 1.35 can now be identified with the angular momentum about the origin. Note that in this example the kinetic energy, as given by equation 1.33a in polar coordinates, is a function of both $\underset{\sim}{\dot{q}}$ and $\underset{\sim}{q}$.

Example 1.6 Motion of a charged particle in an electromagnetic field.

Recall that the force on a charge q_c[#] moving in an electromagnetic field is given by

---

[#] There is a problem with symbols because q is reserved for the generalized coordinates in the Lagrangian.

$$\underline{F} = q_c[\underline{\mathcal{E}}(\underline{x}, t) + \underline{\dot{x}} \times \underline{\mathcal{B}}(\underline{x}, t)] = \underline{F}(\underline{x}, \underline{\dot{x}}, t) \ , \tag{1.36}$$

where $\underline{\mathcal{E}}(\underline{x}, t)$ and $\underline{\mathcal{B}}(\underline{x}, t)$ are the electric and magnetic fields respectively at $(\underline{x}, t)$. Note that the electric and magnetic fields are not constituted by six independent scalar functions $(\mathcal{E}_x, \mathcal{E}_y, \mathcal{E}_z, \mathcal{B}_x, \mathcal{B}_y, \mathcal{B}_z)$, since they must satisfy Maxwell's equations. In particular, we can define a scalar field $\varphi(\underline{x}, t)$ and a vector field $\underline{A}(\underline{x}, t)$, such that

$$\underline{\mathcal{E}}(\underline{x}, t) = -\frac{\partial}{\partial \underline{x}} \varphi(\underline{x}, t) - \frac{\partial}{\partial t} \underline{A}(\underline{x}, t) \tag{1.37a}$$

and

$$\underline{\mathcal{B}}(\underline{x}, t) = \frac{\partial}{\partial \underline{x}} \times \underline{A}(\underline{x}, t) \tag{1.37b}$$

see, for example, reference 1., chapter 21.

The appropriate Lagrangian for this system is given by

$$L(\underline{x}, \underline{\dot{x}}, t) = \frac{1}{2} m|\underline{\dot{x}}|^2 - q_c[\varphi(\underline{x}, t) - \underline{\dot{x}} \cdot \underline{A}(\underline{x}, t)] \ , \tag{1.38}$$

as can be shown directly. In particular

$$\frac{\partial L}{\partial \underline{x}} = q_c \left\{ -\frac{\partial \varphi}{\partial \underline{x}} + \frac{\partial}{\partial \underline{x}} [\underline{\dot{x}} \cdot \underline{A}(\underline{x}, t)] \right\}$$

$$= q_c \left\{ -\frac{\partial \varphi}{\partial \underline{x}} + \underline{\dot{x}} \times \left[\frac{\partial}{\partial \underline{x}} \times \underline{A}(\underline{x}, t) \right] + \left(\underline{\dot{x}} \cdot \frac{\partial}{\partial \underline{x}} \right) \underline{A}(\underline{x}, t) \right\}$$

Using equations 1.37a and 1.37b we then have

$$\frac{\partial L}{\partial \underline{x}} = q_c \left\{ \underline{\mathcal{E}}(\underline{x}, t) + \underline{\dot{x}} \times \underline{\mathcal{B}} + \frac{\partial A}{\partial t} + \left(\underline{\dot{x}} \cdot \frac{\partial}{\partial \underline{x}} \right) \underline{A}(\underline{x}, t) \right\} . \tag{1.39}$$

Taking the partial with respect to $\underline{\dot{x}}$, we now have

$$\frac{\partial L}{\partial \underline{\dot{x}}} = m\underline{\dot{x}} + q_c \underline{A}(\underline{x}, t) \ , \tag{1.40}$$

and therefore

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \underline{\dot{x}}} \right) = m\underline{\ddot{x}} + q_c \left[\frac{\partial}{\partial t} \underline{A}(\underline{x}, t) + \left(\underline{\dot{x}} \cdot \frac{\partial}{\partial \underline{x}} \right) \underline{A}(\underline{x}, t) \right] . \tag{1.41}$$

We now have, from the condition for stationary action

$$\frac{\partial L}{\partial \underline{x}} - \frac{d}{dt} \left(\frac{\partial L}{\partial \underline{\dot{x}}} \right) = 0 \ ,$$

and therefore, from equations 1.39 and 1.41, we obtain Newton's law of motion, corresponding to the Coulomb force, i.e.,

$$m\underline{\ddot{x}} = q_c[\underline{\mathcal{E}}(\underline{x}, t) + \underline{\dot{x}} \times \underline{\mathcal{B}}(\underline{x}, t)] \tag{1.42}$$

[#] note that

$$\frac{\partial}{\partial \underline{x}} [\underline{a} \cdot \underline{A}(\underline{x})] \equiv \underline{a} \times \left[\frac{\partial}{\partial \underline{x}} \times \underline{A}(\underline{x}) \right] + \left(\underline{a} \cdot \frac{\partial}{\partial \underline{x}} \right) \underline{A}(\underline{x})$$

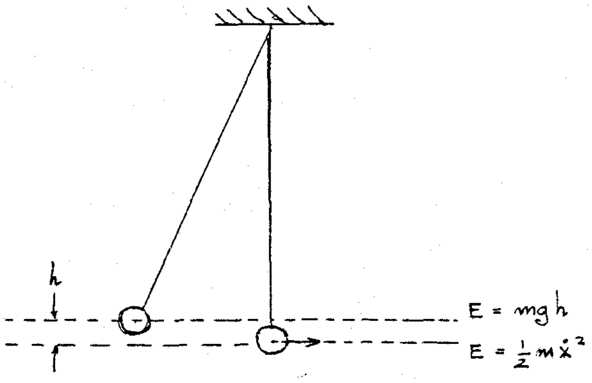
Note now, that the appropriate momentum associated with the motion of the charged particle through the electromagnetic field is given by (equation 1.40

$$\underline{p} = m\dot{\underline{x}} + q_c \, \underline{A}(\underline{x}, t) \; . \tag{1.43}$$

This result may seem a little strange, because we usually consider the momentum of a particle as something intrinsic to the motion of the particle, i.e., $p = m\dot{\underline{x}}$. Equation 1.43, however, should be viewed in the same way as the statement in Newtonian mechanics that the energy of a particle is the sum of a kinetic and potential part, i.e.,

$$E = \frac{1}{2} m |\dot{\underline{x}}|^2 + U(\underline{x}) \; . \tag{1.44}$$

If the particle moves subject to forces derivable from the potential $U(\underline{x})$ (i.e., conservative), E remains constant even though an exchange may be taking place between the kinetic part of the energy [i.e., $\frac{1}{2} m|\dot{\underline{x}}|^2$] and the potential part of the energy [i.e., $U(\underline{x})$]. A picture that may be helpful at this point is that of a (lossless) pendulum in a gravitational field. The system (pendulum) moves, continuously exchanging kinetic for potential energy in a periodic manner.



Note that in the absence of a potential, the energy possesses only a kinetic part and is equal to $E = \frac{1}{2} m|\dot{\underline{x}}|^2$. We should think of the momentum $\underline{p}$, as given by equation 1.43, in much the same manner, as the sum of two parts, a kinetic and an electromagnetic. In the absence of an electromagnetic field, the momentum is only kinetic and given by $\underline{p} = m\dot{\underline{x}}$. In the presence of an electromagnetic field, however, the momentum is composed of two parts which must be considered together in describing the motion. In particular, by analogy to mechanics where, if the energy is conserved, the sum of the kinetic and potential parts is a constant, also in the case of the motion of a charge in an electromagnetic field, if the momentum is conserved, the (vector) sum of the kinetic and electromagnetic parts is a constant. See problem 1.5.

It should also be noted that in writing the Lagrangian of equation 1.38 we have assumed that the scalar field $\varphi(\underline{x}, t)$ and vector field $\underline{A}(\underline{x}, t)$ are externally imposed. In particular, we have not taken into account the effect of the motion of the charged particle itself on $\varphi(\underline{x}, t)$ and $\underline{A}(\underline{x}, t)$, which you will recall (reference 1, chapter 21) are given by

$$\varphi(\underline{x}, t) = \frac{1}{4\pi\varepsilon_0} \int_{\underline{x}'} \frac{\rho_c(\underline{x}', t')}{|\underline{x} - \underline{x}'|} d^3 \underline{x}' \qquad (1.45a)$$

and

$$\underline{A}(\underline{x}, t) = \frac{1}{4\pi\varepsilon_0 c^2} \int_{\underline{x}'} \frac{\underline{j}_c(\underline{x}', t')}{|\underline{x} - \underline{x}'|} d^3 \underline{x}' \ , \qquad (1.45b)$$

where $\rho_c(\underline{x}, t)$ and $\underline{j}_c(\underline{x}, t)$ are the external charge density and current density, connected through the charge conservation equation

$$\frac{\partial}{\partial t} \rho_c(\underline{x}, t) + \frac{\partial}{\partial \underline{x}} \cdot \underline{j}_c(\underline{x}, t) = 0 \ , \qquad (1.45c)$$

and t' is the time in the past, spaced from the present t by the time it takes light to travel the distance $|x - \underline{x}'|$, i.e.,

$$t' = t - \frac{1}{c} |\underline{x} - \underline{x}'| \ , \qquad (1.45d)$$

where c is the speed of light, not to be confused (hopefully!) with the subscript c in q_c, ρ_c, $\underline{j}_c$ which denotes charge.#

1.2 Conservation laws. An immediate benefit can be derived by expressing the equations of motion of the coordinates of the system by means of the condition of stationary action. In particular, equation 1.19' can be re-written as

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = \frac{\partial L}{\partial q_i} \ , \quad i = 1, N \qquad (1.46)$$

From this equation, we see that if the Lagrangian does not explicitly contain a particular coordinate q_i, the corresponding momentum $p_i \equiv \partial L / \partial \dot{q}_i$

#Have you ever wondered why the fields $\varphi(\underline{x}, t)$ and $\underline{A}(\underline{x}, t)$ must wait for light (!?) to come from the charges and currents?

is concerned. In other words, if $L(\underset{\sim}{q}, \underset{\sim}{\dot{q}}, t)$ does not explicitly depend on q_j, i.e. if

$$\frac{\partial L}{\partial q_j} = 0 ,$$

then

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) = 0$$

and therefore

$$\frac{\partial L}{\partial \dot{q}_j} = p_j = \text{const.}$$

It is clear that the discovery of the conserved momenta depends in a rather crucial way on the coordinate system that is chosen to express the Lagrangian. A very dramatic illustration of this effect is afforded by our previous example 1.5 of the pair of gravitating masses. In that case the Lagrangian (which we rewrite here)

$$L = \tfrac{1}{2}m_a|\underset{\sim}{\dot{x}}_a|^2 + \tfrac{1}{2}m_b|\underset{\sim}{\dot{x}}_b|^2 + \frac{G m_a m_b}{|\underset{\sim}{x}_a - \underset{\sim}{x}_b|}$$

depends on all six of the coordinates $(x_a, y_a, z_a, x_b, y_b, z_b)$ and it would appear that no momenta are conserved. While it is true that no momenta associated with any of these six coordinates are conserved, this Lagrangian in fact conceals six conserved momenta, three linear momenta and three angular momenta. See problem 1.4.

In the preceding discussion we saw that if the Lagrangian does not depend on a particular coordinate q_i, the associated momentum p_i is conserved. We shall now show that if the Lagrangian does not depend on time, the energy of the system is conserved. Now if the Lagrangian does not have an explicit time dependence, we must have

$$\frac{\partial L}{\partial t} = 0$$

and therefore

$$\frac{dL}{dt} = \frac{\partial L}{\partial \underset{\sim}{q}} \cdot \underset{\sim}{\dot{q}} + \frac{\partial L}{\partial \underset{\sim}{\dot{q}}} \cdot \underset{\sim}{\ddot{q}} . \tag{1.47}$$

Along the path of stationary action, however, we have that

$$\frac{\partial L}{\partial \underset{\sim}{q}} = \frac{d}{dt}\left(\frac{\partial L}{\partial \underset{\sim}{\dot{q}}}\right)$$

and therefore

$$\frac{dL}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \underline{\dot{q}}}\right) \cdot \underline{\dot{q}} + \frac{\partial L}{\partial \underline{\dot{q}}} \cdot \underline{\ddot{q}}$$

or

$$\frac{dL}{dt} = \frac{d}{dt}\left(\underline{\dot{q}} \cdot \frac{\partial L}{\partial \underline{\dot{q}}}\right)$$

and finally, combining the two, we have

$$\frac{d}{dt}\left(\underline{\dot{q}} \cdot \frac{\partial L}{\partial \underline{\dot{q}}} - L\right) = 0. \qquad (1.48)$$

Consequently, if $\partial L/\partial t = 0$, we have along the path of stationary action,

$$\underline{\dot{q}} \cdot \frac{\partial L}{\partial \underline{\dot{q}}} - L = \text{const.} \qquad (1.49)$$

Now note that the kinetic energy must always be a quadratic function of the $\dot{q}_i$, i.e.

$$K = \frac{1}{2} \sum_{i,j} a_{ij}(\underline{q})\, \dot{q}_i \dot{q}_j \qquad (1.50)$$

(see for example reference 2, §5), and we must therefore have

$$\underline{\dot{q}} \cdot \frac{\partial L}{\partial \underline{\dot{q}}} = 2K , \qquad (1.51)$$

where we have assumed that the potential energy does not depend on $\underline{\dot{q}}$, i.e. $U = U(\underline{q})$ only. With that restriction we also have

$$L = K(\underline{q}, \underline{\dot{q}}) - U(\underline{q})$$

and therefore

$$\underline{\dot{q}} \cdot \frac{\partial L}{\partial \underline{\dot{q}}} - L = K(\underline{q}, \underline{\dot{q}}) + U(\underline{q}) . \qquad (1.52)$$

We can now identify the constant of equation 1.49 with the energy E.

Therefore, if $\partial L/\partial t = 0$, then along the path of stationary action

$$\boxed{\underline{\dot{q}} \cdot \frac{\partial L}{\partial \underline{\dot{q}}} - L = E = \text{const.}} \qquad (1.53)$$

Again, as in the case of the generalized momentum $\partial L/\partial \dot{q}_i$, the validity of equation 1.53 is quite general.

1.3 The Hamiltonian. The Lagrangian that was used to derive the equations of motion in the form 1.19 is a function of $\underline{q}$, $\dot{\underline{q}}$ and t. These equations are second order differential equations (see examples 1.2 - 1.5) because K, the kinetic energy, and as a consequence L, the Lagrangian, are quadratic in the components of $\dot{\underline{q}}$. This is often undesirable and we are motivated to derive a formalism that yields equations of motion that are first order differential equations. Alternatively we might like to express the evolution of the system in terms of the coordinates q_i and and the momenta $p_i \equiv \partial L/\partial \dot{q}_i$, instead of the coordinates q_i and the velocities $\dot{q}_i$ which must be specified as initial conditions to solve the problem. Both of these objectives can be realized by means of a Legendre transformation from a function of the independent variables $(\underline{q}, \dot{\underline{q}}, t)$ to the variables $(\underline{q}, \underline{p}, t)$ as follows.

The total differential of the Lagrangian, which is a function of $\underline{q}$, $\dot{\underline{q}}$ and t, can be written as

$$\begin{aligned} dL &= \sum_i \frac{\partial L}{\partial q_i}\, dq_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\, d\dot{q}_i + \frac{\partial L}{\partial t}\, dt \\ &= \frac{\partial L}{\partial \underline{q}} \cdot d\underline{q} + \frac{\partial L}{\partial \dot{\underline{q}}} \cdot d\dot{\underline{q}} + \frac{\partial L}{\partial t}\, dt \; . \end{aligned} \tag{1.54}$$

Substituting the momenta $p_i = \partial L/\partial \dot{q}_i$ we have

$$dL = \sum_i \frac{\partial L}{\partial q_i}\, dq_i + \sum_i p_i d\dot{q}_i + \frac{\partial L}{\partial t}\, dt \; .$$

The second term can be re-written as

$$\sum_i p_i\, d\dot{q}_i \equiv d\left(\sum_i p_i \dot{q}_i\right) - \sum_i \dot{q}_i\, dp_i \quad ,$$

so that

$$d\left(-L + \sum_i p_i \dot{q}_i\right) = \sum_i \left(-\frac{\partial L}{\partial q_i}\right) dq_i + \sum_i \dot{q}_i dp_i - \frac{\partial L}{\partial t}\, dt \; . \tag{1.55}$$

We therefore see that this is the total differential of a function, which we will denote by

$$H \equiv -L + \sum_i p_i \dot{q}_i \tag{1.56}$$

which is a function of $(\underline{q}, \underline{p}, t)$ and which satisfies the equations

$$\frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i} \tag{1.57a}$$

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \tag{1.57b}$$

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \; . \tag{1.57c}$$

Now note that along the path of stationary action

$$\frac{\partial L}{\partial q_i} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \dot{p}_i \tag{1.58}$$

so that the equations for the extremal can be written as

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{1.59a}$$

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \; . \tag{1.59b}$$

The function $H = H(\underline{q}, \underline{p}, t)$ is called the Hamiltonian of the system and can be seen to be a constant of the motion if $\partial H/\partial t = -\partial L/\partial t = 0$. See equation 1.53 and related discussion.

It should be emphasized that the (Legendre) transformation from a function of $(\underline{q}, \dot{\underline{q}}, t)$ to a function of $(\underline{q}, \underline{p}, t)$, i.e. from the Lagrangian to the Hamiltonian does not depend on the assumption that $q(t)$ is any particular path or, for that matter, the path of stationary action. It is a simultaneous transformation of the dependent and independent variables. Equations 1.57 were derived without the benefit of Euler's equations (equations 1.58). Now, if in addition we impose the conditions of Euler's equations, we obtain equations 1.59, called the canonical equations, for $q(t)$ and $p(t)$ along the classical path of stationary action. These are first order equations which we can solve to obtain the system evolution, given $\underline{q}(o)$ and $\underline{p}(o)$, i.e. the coordinates and associated momenta at an **instant** in time.

1.4 Poisson brackets and constants of the motion.

From the preceding discussion (section 1.2), we have seen that if the Lagrangian does not depend explicitly on a particular coordinate q_i, then the corresponding momentum p_i is conserved along the path of stationary action, i.e. is a constant of the motion. Similarly if the Lagrangian is not an explicit function of time, then the energy is a constant of the motion. These two statements provide us with a very powerful means of knowing ahead of time (without solving for the path of stationary action) whether the momenta associated with the corresponding coordinates are conserved, and whether the energy is conserved. They do not allow, however, for an explicit scheme that would tell us if a particular quantity is conserved. In fact we have already seen that it may well be that the momenta associated with an alternative choice of coordinates are conserved, in a manner that might not be obvious from some original choice of coordinates for the Lagrangian. We therefore seek a general method for determining whether a particular quantity is conserved or not.

Consider a quantity $F = F(\underline{q}, \underline{p}, t)$, which is some function of the coordinates, the momenta and time. Then

$$\frac{dF}{dt} = \sum_i \left(\frac{\partial F}{\partial q_i} \dot{q}_i + \frac{\partial F}{\partial p_i} \dot{p}_i \right) + \frac{\partial F}{\partial t} , \tag{1.60}$$

and along the path of stationary action (see equations 1.59), we have

$$\frac{dF}{dt} = \sum_i \left(\frac{\partial H}{\partial p_i} \frac{\partial F}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial F}{\partial p_i} \right) + \frac{\partial F}{\partial t} ,$$

or, in vector notation

$$\frac{dF}{dt} = \left\{ \frac{\partial H}{\partial \underline{p}} \cdot \frac{\partial F}{\partial \underline{q}} - \frac{\partial H}{\partial \underline{q}} \cdot \frac{\partial F}{\partial \underline{p}} \right\} + \frac{\partial F}{\partial t} . \tag{1.61}$$

The quantity in the braces is called the Poisson bracket of the quantity $F(\underline{q}, \underline{p}, t)$ and the Hamiltonian, and denoted by

$$\{H, F\} \equiv \frac{\partial H}{\partial \underline{p}} \cdot \frac{\partial F}{\partial \underline{q}} - \frac{\partial H}{\partial \underline{q}} \cdot \frac{\partial F}{\partial \underline{p}} . \tag{1.62}$$

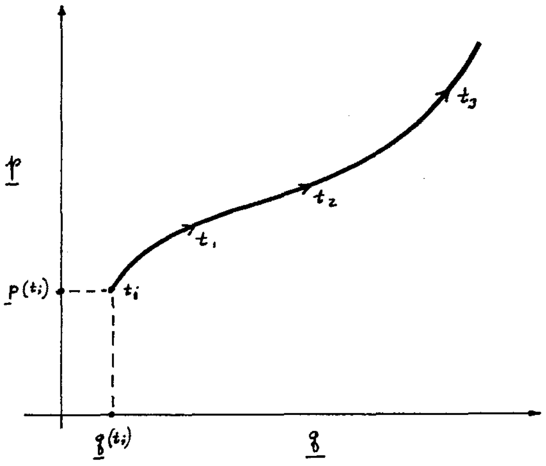
The total time derivative along the path of stationary action is then given by

$$\frac{dF}{dt} = \{H, F\} + \frac{\partial F}{\partial t} . \tag{1.63}$$

This equation provides the means of determining whether $F(\underline{q}, \underline{p}, t)$ is a constant of the motion. In particular, if F is not an explicit function of time (i.e. $\partial F/\partial t = 0$) and if the Poisson bracket $\{H, F\}$ is zero, F is a constant of the motion.

1.5 The assumptions of classical mechanics

The preceding formalism provides a means by which the evolution of any mechanical system can be determined. From the canonical equations, for example, the coordinate vector $\underline{q}(t)$ and the associated momentum vector $\underline{p}(t)$ can be obtained for all times t, given their initial values $\underline{q}(t_i)$, $\underline{p}(t_i)$ at some time $t = t_i$. The resulting solution describes a path through the 2N-dimensional $\underline{q}$ - $\underline{p}$ space, called phase space, expressed parametrically in terms of time.



Implicit in this description, you will recognize, is that,

(i) it is possible to specify both the coordinates and momenta of the system at some time t_i (or any time for that matter),

or

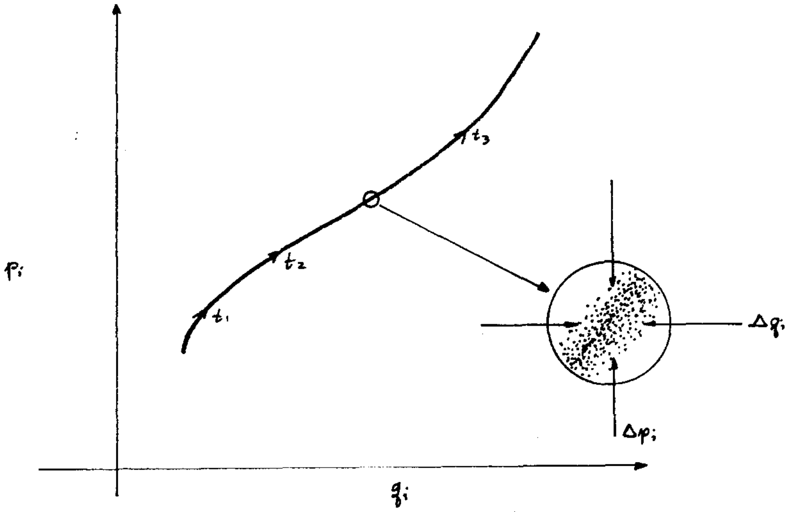
(ii) that the concept of a path through phase space, along which the system evolves, is a valid one.#

Interestingly enough, the validity of these assumptions is in fact not supported by experimental evidence, even though they may appear intuitively self-evident. Our current views of the workings of nature are consistent with the notion that the evolution of a system proceeds through a region in phase space (as opposed to a well-defined trajectory) whose joint extent (e.g. root mean square) along each coordinate-momentum conjugate pair, at any one time, can never be less than a certain minimum. If we

---

[#] These two assumptions are evidently not independent, for if we could specify $\underline{q}(t)$ and $\underline{p}(t)$ at the same time then a path would be definable.

were to look closely enough at a portion of the system evolution, in the part of phase space along, say, a particular conjugate coordinate-momentum pair, we might see something like the sketch below



where $\Delta q_i \Delta p_i \gtrsim$ fixed constant.

Classical mechanics is then to be considered as applicable to the range of phenomena for which the finite extent of the region in phase space along the evolution of the system is negligible. This is almost always the case in the macroscopic world, where classical mechanics and the concept of a path are valid approximations. It is found to be inadequate in the microscopic world of nuclei, atoms, electrons, photons, etc., for which we need a different formalism, the formalism of Quantum Mechanics.

References

1.1 R. P. Feynman, R. B. Leighton, M. Sands; The Feynman Lectures in Physics, volume II, Chapter 19.

1.2 L. D. Landau and E. M. Lifshitz, Mechanics, § 1-5.

1.3 H. Goldstein, Classical Mechanics, Chapter 2.

Problems

1.1 Compute the value of the action integral of equation 1.12 for,

- a. The parabolic trajectory of equation 1.14.
- b. A circular arc that satisfies the initial conditions. Let the radius be a variable and compute the integral as a function of the radius of curvature.
- c. A half-cycle sine wave, as a function of the period.
- d. Use any other curve of your choosing.
- e. Choose a particular value for v_0 and θ_0, in example 1.1, plot the four curves and compute the numerical value of the action in each case.

1.2 Show that the Lagrangian $L(\underline{q}, \dot{\underline{q}}, t)$ is not uniquely defined through the equations of motion and the principle of stationary action. In particular show that the alternative Lagrangian $L'(\underline{q}, \dot{\underline{q}}, t)$, given by

$$L'(\underline{q}, \dot{\underline{q}}, t) \equiv \alpha \, L(\underline{q}, \dot{\underline{q}}, t) + \frac{d}{dt} \, G(\underline{q}, t) \ ,$$

where α is any constant and G is any function of $\underline{q}$ and t, yields the same equations of motion.

1.3 Express the kinetic energy of a single particle, i.e.,

$$K = \frac{1}{2} m \, |\dot{\underline{x}}|^2$$

in (i) Cartesian coordinates, i.e.,

$$\underline{x} = (x, y, z)$$

(ii) Cylindrical coordinates, i.e.,

$$\underline{x} = (r, \phi, z) \ ,$$

where

$$x = r \cos\phi \ , \qquad y = r \sin\phi \ ,$$

and

(iii) spherical polar coordinates

$$\underline{x} = (r, \phi, \theta) \ ,$$

where

$$x = r \cos\phi \sin\theta, \ \ y = r \sin\phi \sin\theta, \ \ z = r \cos\theta \ .$$

1.4 Express the Lagrangian of example 1.3 in terms of the separation between the two masses

$$\underline{x} \equiv \underline{x}_b - \underline{x}_a ,$$

the position of the center of mass

$$\underline{X} \equiv \frac{m_b \underline{x}_b + m_a \underline{x}_a}{m_b + m_a} ,$$

and the <u>total</u> and <u>reduced</u> masses

$$M = m_b + m_a , \quad m = \frac{m_a m_b}{m_a + m_b} .$$

(i) Show that the three components of the momentum of the center of mass are conserved, i.e.,

$$\underline{P} = M\dot{\underline{X}} = \text{const.}$$

(ii) Express the vector $\underline{x}$ in terms of spherical polar coordinates and show that the angular momentum about the z-axis is conserved. Can you show that the other two components of angular momentum are also conserved?

1.5 Express the Lagrangian corresponding to the motion of a charged particle in an electromagnetic field (equation 1.38) in spherical polar coordinates.

(i) Compute the angular momentum p_ϕ, corresponding to the azimuthal angle ϕ.

(ii) Show that if the scalar potential $\varphi(\underline{x}, t)$ and the vector potential $\underline{A}(\underline{x}, t)$ are azimuthally symmetric, p_ϕ of part (i) is conserved.

(iii) Discuss the "paradox" described in reference 1, section 17-4.

1.6 Assuming that

$$L = K(\underline{q}, \dot{\underline{q}}) - U(\underline{q})$$

and that K is a quadratic function of $\dot{\underline{q}}$, i.e.,

$$K = \frac{1}{2} \sum_{ij} a_{ij}(\underline{q})\, \dot{q}_i \dot{q}_j ,$$

show that

$$\dot{\underline{q}} \cdot \frac{\partial L}{\partial \dot{q}} = 2K .$$

1.7 Derive the Hamiltonian for the motion of a charged particle in a constant electromagnetic field.

1.8 From the defining equation for action (equation 1.5)

a. Compute $\frac{dI}{dt}$

b. Express the action integral in terms of the Hamiltonian and its variables.

c. Compute dI as a total differential and from it the partial derivatives

$$\frac{\partial I}{\partial q_i} \quad , \quad \frac{\partial I}{\partial t} \quad .$$

d. Show that action is a relativistic invariant (Hint: consider dI) .

1.9 Using the definition of the Poisson bracket, show

a. $\{f, g\} = - \{g, f\}$

b. if c = constant $\{f, c\} = 0$

c. $\{f_1 + f_2 , g\} = f_1 \{f_2 , g\} + f_2 \{f_1 , g\}$

d. $\frac{\partial}{\partial t}\{f, g\} = \{\frac{\partial f}{\partial t} , g\} + \{f, \frac{\partial g}{\partial t}\}$

e. $\{f, q_i\} = \frac{\partial f}{\partial p_i}$, $\{f, p_i\} = - \frac{\partial f}{\partial q_i}$

f. $\{q_i , q_j\} = 0$, $\{p_i , p_j\} = 0$, $\{p_i , q_j\} = \delta_{ij}$.

II. WAVE MECHANICS

2.1 An interference experiment.

Consider a monochromatic light source at the origin (a laser for example), of wavelength λ, whose output is incident on a plate P_1 at $z = z_1$ which has two holes at $x = \pm \ell/2$ and $y = 0$. Consider also a second plate P_2 at $z = z_2 = z_1 + L$, parallel to P_1. See figure 2.1.

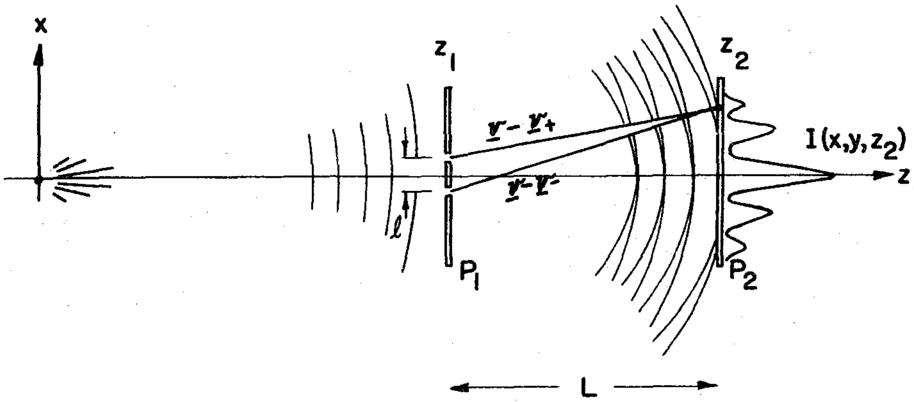


Figure 2.1. An interference apparatus.

If we expose a photographic plate placed on P_2 we will record a complicated pattern $I(x, y, z_2)$ arising from the interference, at each point (x, y) on P_2, between the waves emanating from each of the two orifices on P_1. More specifically, the total amplitude (electric field) at a location (x, y) on P_2 is given by

$$E(x, y, z_2) = E_+(x, y, z_2) + E_-(x, y, z_2) \tag{2.1}$$

where E+ and E- are the amplitudes of the waves from each of the orifices at $x = + \ell/2$ and $x = - \ell/2$ respectively. The resulting intensity $I_\ell(x, y, z_2)$, or for low light levels, the probability of detection $P(x, y, z_2)$, is then given by

$$I_\ell(x, y, z_2) \propto P(x, y, z_2) \propto E^*(x, y, z_2)\ E(x, y, z_2)\ , \tag{2.2}$$

where a superscript * denotes the complex conjugate function.

These results can be derived from Maxwell's equations which also predict that the shape of $I_\ell(x, y, z_2)$ is independent of the light intensity; a result also borne out by experiment. In other words, if we define

$$f(x, y, z_2) \equiv \frac{1}{W} I_\ell(x, y, z_2)\ , \tag{2.3}$$

where W is the power output of the source, the function $f(x, y, z_2)$ is independent of W.

Example 2.1

Compute the light interference pattern arising in the apparatus of figure 2.1.

Both $E_+(\underline{r}, t)$ and $E_-(\underline{r}, t)$, for $z_1 < z < z_2$, are spherical waves which, for distances $|\underline{r} - \underline{r}_+|$ and $|\underline{r} - \underline{r}_-|$ much larger than the dimension (say diameter) of the orifices, can be written as

$$E_+(\underline{r}, t) = E_0 a_+ \frac{e^{i(k|\underline{r} - \underline{r}_+| - \omega t)}}{|\underline{r} - \underline{r}_+|}$$

and

$$E_-(\underline{r}, t) = E_0 a_- \frac{e^{i(k|\underline{r} - \underline{r}_-| - \omega t)}}{|\underline{r} - \underline{r}_-|}$$

where a_+ and a_- are complex constants (with units of length) and E_0 is the amplitude of the electric field at some reference point to the left of P_1 [say at $(0, 0, z_1^-)$]. Then the intensity at P_2 is given by equations 2.1 and 2.2, or

$$I_\ell(x, y, z_2) \propto |E_+(x, y, z_2) + E_-(x, y, z_2)|^2 ,$$

and expanding the square we have

$$I_\ell(x, y, z_2) \propto E_+^*(x, y, z_2)\ E_+(x, y, z_2) + E_-^*(x, y, z_2)\ E_-(x, y, z_2)$$
$$+ E_+^*(x, y, z_2)\ E_-(x, y, z_2) + E_-^*(x, y, z_2)\ E_+(x, y, z_2)$$

The first two terms can be identified as the intensities of the waves from the $x = +\ell/2$ and $x = -\ell/2$ orifices respectively that we would observe separately, i.e., for each one as if the other one was plugged up. The cross products in the third and fourth terms give rise to the interference, i.e.,

$$I_\ell(x, y, z_2) = I_+(x, y, z_2) + I_-(x, y, z_2) + I_{int}(x, y, z_2) . \tag{2.4a}$$

Substituting for the expressions of E_+ and E_-, we have

$$I_+(x, y, z_2) \propto \frac{E_0^2 |a_+|^2}{(x - \ell/2)^2 + y^2 + L^2} \tag{2.4b}$$

$$I_-(x, y, z_2) \propto \frac{E_0^2 |a_-|^2}{(x + \ell/2)^2 + y^2 + L^2} \tag{2.4c}$$

and, for small x/L, y/L, ℓ/L,

$$I_{int}(x, y, z_2) \propto E_0^2 \left\{ \frac{a_+^* a_-\ e^{-ikx(\ell/L)} + a_-^* a_+\ e^{ikx(\ell/L)}}{L^2 + (\ell/2)^2 + x^2 + y^2} \right\} \tag{2.4d}$$

which holds to cubic order in x/L, y/L and ℓ/L. If we now make the substitutions

$$a_+ = |a_+| e^{i\phi_+} , \qquad a_- = |a_-| e^{i\phi_-}$$

we have,

$$I_{int}(x, y, z_2) \propto \frac{2 E_0^2 |a_+||a_-| \cos [kx(\ell/L) + (\phi_+ - \phi_-)]}{L^2 + (\ell/2)^2 + x^2 + y^2} \tag{2.5}$$

Since E_0^2 is proportional to the source strength W, the linearity of the interference phenomenon (equation 2.3 and related discussion) can be seen.

This behavior of light is well documented and does not appear surprising being attributable to the accepted wave nature of light. What is perhaps less amenable to intuitive explanation is that if we replace the monochromatic light source with a source of monoenergetic electrons, the result is essentially the same. A similar interference pattern $I_e(x, y, z_2)$ will appear on a suitable detector plate, with a possible change in scale depending on the energy of the electrons and the dimensions ℓ and L of the interference apparatus. This interference pattern, properly normalized as we did for light in equation 2.3, remains unchanged as we decrease the intensity (electrons per second) of the source, even to the point when at any one time only one electron is in flight between the source and the detector plate P_2. Of course, as we decrease the source electron current we will start detecting individual electrons arriving at P_2.# Nevertheless, the count rate as a function of x and y would be proportional to the interference function $I_e(x, y, z_2)$. The inescapable conclusion is that, just like light, a single electron behaves as a wave with an amplitude $\psi(\underline{x}, t)$ such that the probability of detecting an electron at a location $\underline{x}$ is given by

$$P(\underline{x},t) = \psi^*(\underline{x}, t)\ \psi(\underline{x}, t) \ . \tag{2.6}$$

The interference pattern then arises because, as in the case of light, the amplitude on the detector plate is a superposition of the amplitudes of the waves from each of the orifices, i.e.

$$\psi(x, y, z_2) = \psi_+(x, y, z_2) + \psi_-(x, y, z_2) \ . \tag{2.7}$$

#The situation is, of course, no different from that of light where, if we decrease the light intensity, we would replace the photographic plate with a photomultiplier tube and count individual photons.

2.2 Wave packets.

A plane wave of wavelength λ can be represented by an amplitude $u(\underline{x}, t)$ given by

$$u(\underline{x}, t) = U(\underline{k}) e^{i[\underline{k} \cdot \underline{x} - \omega(\underline{k}) t]} \tag{2.8}$$

where $\underline{k}$ is the wavevector along the direction of propagation, whose magnitude is given by

$$k = |\underline{k}| = \frac{2\pi}{\lambda} , \tag{2.9}$$

and $\omega(\underline{k})$ is the angular frequency associated with the wavenumber $\underline{k}$.

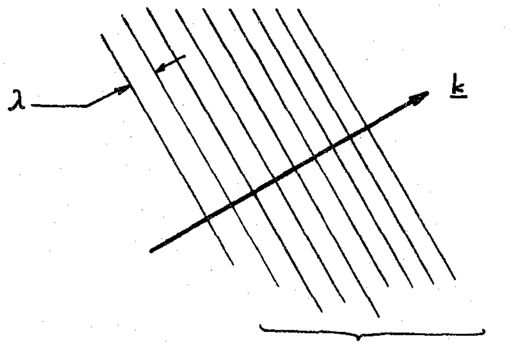


Figure 2.2. Plane wave of wavenumber $\underline{k}$.

The constant $U(\underline{k})$ can be complex and has a modulus whose square is proportional to the intensity of the wave, i.e.,

$$I \propto U^*(\underline{k}) \cdot U(\underline{k}) = |U(\underline{k})|^2 . \tag{2.10}$$

Using a superposition of waves of this type it is possible to synthesize any distribution $f(\underline{x}, t_0)$ of wave amplitude in space, at an instant in time $t = t_0$, i.e.

$$f(\underline{x}, t_0) = \left(\frac{1}{2\pi}\right)^3 \iiint_{-\infty}^{\infty} F(\underline{k}, t_0)\, e^{i\underline{k} \cdot \underline{x}}\, d^3\underline{k} \tag{2.11}$$

where, from the properties of the Fourier transform (see Appendix B) the function $F(\underline{k}, t_0)$ is given by[#]

$$F(\underline{k}, t_0) = \iiint_{-\infty}^{\infty} f(\underline{x}, t_0)\, e^{-i\underline{k} \cdot \underline{x}}\, d^3\underline{x} . \tag{2.12}$$

[#] There is a variety of conventions that different authors use concerning the $(1/2\pi)$ factor in the Fourier integrals. The convention adopted here is that the 2π is used to cancel the "units" of radians in the wavenumber k or the angular frequency ω.

Moreover, if the frequencies $\omega(\underline{k})$ are known, the evolution in time of the wave amplitude $f(\underline{x}, t)$ is now completely described by

$$f(\underline{x}, t) = \left(\frac{1}{2\pi}\right)^3 \iiint_{-\infty}^{\infty} F(\underline{k}, t_0)\, e^{i[\underline{k}\cdot\underline{x} - \omega(\underline{k})(t - t_0)]}\, d^3\underline{k} \,. \qquad (2.13)$$

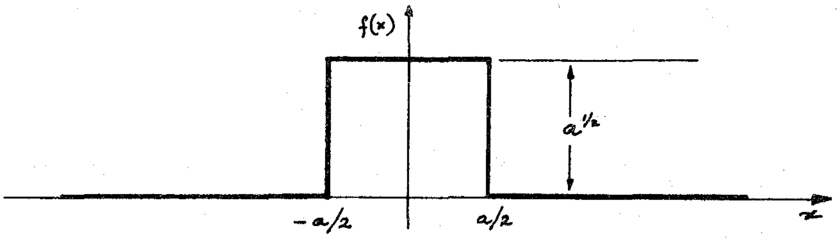
It is noteworthy to observe that even though waves of the type described by equation 2.8 are uniformly distributed in $\underline{x}$, they can be superimposed to synthesize distributions that are localized in some region of space

Example 2.2

A wave amplitude at $t = 0$ is given by (one dimension)

$$f(x) = \begin{cases} a^{-1/2} & \text{for } |x| < a/2 \\ 0 & \text{for } |x| > a/2 \,. \end{cases} \qquad (2.14)$$

Find the distribution of wavenumbers contributing to that function.



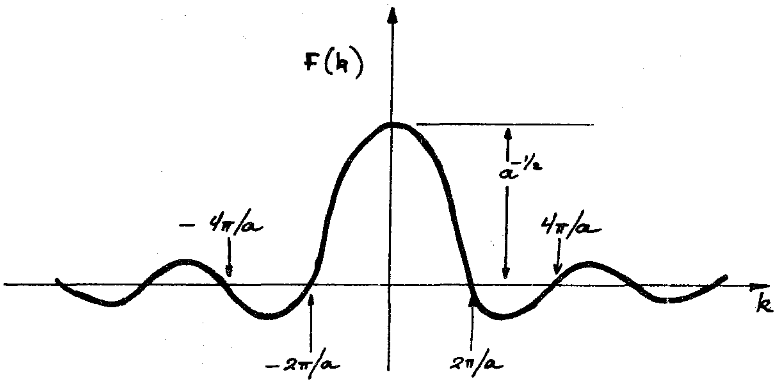
Using the one dimensional form of equation 2.12, we have

$$F(k) = \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx$$

$$= \frac{1}{a}^{1/2} \int_{-a/2}^{a/2} e^{-ikx}\, dx$$

or

$$F(k) = a^{1/2}\, \frac{\sin(ka/2)}{(ka/2)} \,. \qquad (2.15)$$



Consequently, using the one dimensional form of equation 2.11, we have that

$$f(x) = \frac{1}{\pi a^{1/2}} \int_{-\infty}^{\infty} \frac{1}{k} \sin \frac{ka}{2} \, e^{ikx} \, dk \qquad (2.16)$$

The linear superposition of waves e^{ikx} of equation 2.16 interferes constructively for $|x| < a/2$ to yield $a^{1/2}$ and destructively for $|x| > a/2$ to result in complete cancellation outside.

The symmetry in the pair of equations 2.11 and 2.12 should be noted. It suggests that it is equivalent to think of $F(\underline{k}, t)$ as the distribution of wavenumbers required to synthesize $f(\underline{x}, t)$, and to think of $f(\underline{x}, t)$ as the amplitude distribution in $\underline{x}$-space required to synthesize a particular superposition $F(\underline{k}, t)$ of wavenumbers $\underline{k}$ at time t.

Parseval's theorem (normalization). The square of the modulus of $f(\underline{x}, t)$ is a measure of the intensity of the wave at $(\underline{x}, t)$, i.e.,

$$I_x(\underline{x}, t) = f^*(\underline{x}, t) \, f(\underline{x}, t) \equiv |f(\underline{x}, t)|^2 \; . \qquad (2.17)$$

Conversly, the square of the modulus of $F(\underline{k}, t)$ is a measure of the intensity of the wave to be associated with the wavenumber $\underline{k}$ at time t, i.e.,

$$I_k(\underline{k}, t) = F^*(\underline{k}, t) \, F(\underline{k}, t) \equiv |F(\underline{k}, t)|^2 \; . \qquad (2.18)$$

As a consequence of Parseval's theorem (see Appendix B, §14), we have

$$\int_{-\infty}^{\infty} |f(\underline{x}, t)|^2 \, d^3\underline{x} = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} |F(\underline{k}, t)|^2 \, d^3\underline{k} \; , \qquad (2.19)$$

which is a plausible statement since the total intensity in $\underline{x}$-space must be equal to the total intensity in $\underline{k}$-space. We are talking about the same wave!

If the integrals in equation 2.19 exist, we can normalize $f(\underline{x}, t)$ and $F(\underline{k}, t)$ such that

$$\tilde{f}(\underline{x}, t) = \frac{f(\underline{x}, t)}{\left[\int_{-\infty}^{\infty} |f(\underline{x}, t)|^2 \, d^3\underline{x} \right]^{1/2}} \qquad (2.20)$$

and

$$\widetilde{F}(\underline{k},t) = \frac{F(\underline{k},t)}{\left[\left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} |F(\underline{k},t)|^2 \, d^3\underline{k}\right]^{1/2}} \qquad (2.21)$$

where the two denominators are, of course the same. We then have

$$\int_{-\infty}^{\infty} |\widetilde{f}(\underline{x},t)|^2 \, d^3\underline{x} = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} |\widetilde{F}(\underline{k},t)|^2 \, d^3\underline{k} = 1 \qquad (2.22)$$

<u>Expectation values</u>. If the distributions $f(\underline{x},t)$ and $F(\underline{k},t)$ are localized in $\underline{x}$-space and $\underline{k}$-space respectively, we can compute the mean position (center of "mass" if you will) of a superposition $f(\underline{x},t)$ of wave amplitudes by weighting each coordinate $\underline{x}$ with the intensity $I_x(\underline{x},t)$, i.e.,

$$\langle x \rangle = \int_{-\infty}^{\infty} \underline{x} \, I_x(\underline{x},t) \, d^3\underline{x} \qquad (2.23)$$

or

$$\langle \underline{x} \rangle = \int_{-\infty}^{\infty} f^*(\underline{x},t) \; \underline{x} \; f(\underline{x},t) \, d^3\underline{x} \ , \qquad (2.23')$$

which, in general, can be a function of time. The interchange of $\underline{x}$ and $f^*(\underline{x},t)$ is legal since $\underline{x} \, f^*(\underline{x},t) = f^*(\underline{x},t) \, \underline{x}$. It was done for reasons that will be apparent in a moment. The wave distribution $f(\underline{x},t)$ in equation 2.23 and 2.23' have been normalized such that

$$\int_{-\infty}^{\infty} |f(\underline{x},t)|^2 \, d^3\underline{x} = 1 \ .$$

If not, the expectation value of $\underline{x}$ can be computed by means of

$$\langle \underline{x} \rangle = \frac{\int_{-\infty}^{\infty} f^*(\underline{x},t) \; \underline{x} \; f(\underline{x},t) \, d^3\underline{x}}{\int_{-\infty}^{\infty} |f(\underline{x},t)|^2 \, d^3\underline{x}} \qquad (2.23'')$$

Similarly, we can compute the mean wavenumber of the corresponding superposition $F(\underline{k},t)$ of wave amplitudes, i.e.,

$$\langle \underline{k} \rangle = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} \underline{k} \; I_k(\underline{k},t) \, d^3\underline{k} \qquad (2.24)$$

or

$$\langle \underline{k} \rangle = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} F^*(\underline{k}, t)\ \underline{k}\ F(\underline{k}, t)\, d^3\underline{k}\ , \tag{2.24'}$$

which, in general, can also be a function of time. The interchange of $F^*(\underline{k}, t)$ and $\underline{k}$ is, again, obviously legal. The quantities $\langle \underline{x} \rangle$ and $\langle \underline{k} \rangle$ are called <u>expectation values</u> of $\underline{x}$ and $\underline{k}$ respectively. In fact, the expectation values of any function of $\underline{x}$ or $\underline{k}$ can be computed in a similar fashion, i.e.,

$$\langle Q(\underline{x}) \rangle = \int_{-\infty}^{\infty} f^*(\underline{x}, t)\ Q(\underline{x}, t)\ f(\underline{x}, t)\, d^3\underline{x} \tag{2.25}$$

and similarly

$$\langle R(\underline{k}) \rangle = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} F^*(\underline{k}, t)\ R(\underline{k})\ F(\underline{k}, t)\, d^3\underline{k}\ . \tag{2.26}$$

The functions $f(\underline{x}, t)$ and $F(\underline{k}, t)$ in equations 2.24, 2.25 and 2.26 are assumed normalized to unity in the sense of equation 2.22.

An interesting symmetry arises between the expectation values of $\underline{x}$ and $\underline{k}$ from the fact that $f(\underline{x}, t)$ and $F(\underline{k}, t)$ are Fourier transform pairs. Consider equation 2.24', which we re-write below,

$$\langle \underline{k} \rangle = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} F^*(\underline{k}, t)\ [\underline{k}\ F(\underline{k}, t)]\, d^3\underline{k}\ . \tag{2.24'}$$

The quantity in the brackets can be computed in terms of $f(\underline{x}, t)$ using the differentiation property of the Fourier transform (Appendix B, problem B.4). In particular

$$\underline{k}\ F(\underline{k}, t) = \int_{-\infty}^{\infty} \left[-i\,\frac{\partial}{\partial \underline{x}}\ f(\underline{x}, t)\right] e^{-i\underline{k}\cdot\underline{x}}\, d^3\underline{x}\ , \tag{2.27}$$

where

$$\frac{\partial}{\partial \underline{x}} \equiv \hat{e}_x \frac{\partial}{\partial x} + \hat{e}_y \frac{\partial}{\partial y} + \hat{e}_z \frac{\partial}{\partial z} \tag{2.28}$$

denotes the gradient in $\underline{x}$-space. Substituting equation 2.27 into 2.24' we have

$$\langle \underline{k} \rangle = \int_{-\infty}^{\infty} \left\{\left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} F^*(\underline{k}, t)\, e^{-i\underline{k}\cdot\underline{x}}\, d^3\underline{k}\right\}\left[-i\,\frac{\partial}{\partial \underline{x}}\ f(\underline{x}, t)\right] d^3\underline{x}\ ,$$

where we have interchanged the order of integration. The quantity in the braces we recognize as $f^*(\underline{x}, t)$ and therefore

$$\langle \underline{k} \rangle = \int_{-\infty}^{\infty} f^*(\underline{x}, t)\ \left(-i\,\frac{\partial}{\partial \underline{x}}\right) f(\underline{x}, t)\, d^3\underline{x}\ . \tag{2.29}$$

The quantity $-i\ \partial/\partial\underline{x}$ is an <u>operator</u> that operates on the function to its right. In other words we first compute

$$g_x = -i \frac{\partial f}{\partial x} \qquad (2.30a)$$

$$g_y = -i \frac{\partial f}{\partial y} \qquad (2.30b)$$

$$g_z = -i \frac{\partial f}{\partial z} \qquad (2.30c)$$

and then the expectation values of

$$\langle k_x \rangle = \int_{-\infty}^{\infty} f^*(\underline{x},t) \; g_x(\underline{x},t)\, d^3\underline{x} \qquad (2.31a)$$

$$\langle k_y \rangle = \int_{-\infty}^{\infty} f^*(\underline{x},t) \; g_y(\underline{x},t)\, d^3\underline{x} \qquad (2.31b)$$

$$\langle k_z \rangle = \int_{-\infty}^{\infty} f^*(\underline{x},t) \; g_z(\underline{x},t)\, d^3\underline{x} \qquad (2.31c)$$

to obtain

$$\langle \underline{k} \rangle = \hat{e}_x \langle k_x \rangle + \hat{e}_y \langle k_y \rangle + \hat{e}_z \langle k_z \rangle \, . \qquad (2.32)$$

Equation 2.26 is shorthand for equations 2.30a, 2.30b, 2.30c, 2.31a, 2.31b, 2.31c, 2.32.

The result of equation 2.29 can be generalized so that the expectation value of any function of $\underline{k}$ can be computed from the distribution of wave amplitudes in $\underline{x}$-space. In particular, the expectation value of equation 2.26 can also be computed in terms of

$$\langle R(\underline{k}) \rangle = \int_{-\infty}^{\infty} f^*(\underline{x},t) \; R\left(-i \frac{\partial}{\partial \underline{x}}\right) f(\underline{x},t)\, d^3\underline{x} \, . \qquad (2.33)$$

Corresponding formulae also apply when computing expectation values of $\underline{x}$ using wave amplitudes in $\underline{k}$-space. By similar arguments we can show that if

$$F(\underline{k},t) = \int_{-\infty}^{\infty} f(\underline{x},t)\, e^{-i\underline{k}\cdot\underline{x}}\, d^3\underline{x} \, ,$$

then

$$\langle \underline{x} \rangle = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} F^*(\underline{k},t) \left(i \frac{\partial}{\partial \underline{k}}\right) F(\underline{k},t)\, d^3\underline{k} \qquad (2.34)$$

and, in general,

$$\langle Q(\underline{x}) \rangle = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty} F^*(\underline{k},t)\; Q\left(i \frac{\partial}{\partial \underline{k}}\right) F(\underline{k},t)\, d^3\underline{k} \, , \qquad (2.35)$$

where

$$\frac{\partial}{\partial \underline{k}} = \hat{e}_x \frac{\partial}{\partial k_x} + \hat{e}_y \frac{\partial}{\partial k_y} + \hat{e}_z \frac{\partial}{\partial k_z} \qquad (2.36)$$

denotes the gradient in $\underline{k}$-space.

Note that

$$\begin{array}{c|c|c} & \underline{x}\text{-space} & \underline{k}\text{-space} \\ \hline \hat{\underline{x}} & = \underline{x} & = i\,\dfrac{\partial}{\partial \underline{k}} \\ \hline \hat{\underline{k}} & = -i\,\dfrac{\partial}{\partial \underline{x}} & = \underline{k}\ , \end{array} \qquad (2.37)$$

where the caret ($\hat{}$) denotes the corresponding operator.

Localization. In addition to asking where a wave packet is, in $\underline{x}$-space or $\underline{k}$-space, by means of the expectation value of $\underline{x}$ or $\underline{k}$ respectively (equations 2.23, 2.24 or 2.29, 2.34), we can find out something about the degree of localization of the packet in $\underline{x}$-space or $\underline{k}$-space by computing the mean square deviation from $\langle \underline{x} \rangle$ or $\langle \underline{k} \rangle$ respectively. To simplify the thinking, we will restrict the discussion to one dimension. The extension to higher dimensions is straight forward.

Assume that a wave packet has a wave amplitude in position space given by f(x) at some initial time t = 0, such that

$$\langle x \rangle = \int_{-\infty}^{\infty} f^*(x)\ x f(x)\, dx = x_0 \qquad (2.38a)$$

and

$$\langle k \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} F^*(k)\ k\ F(k)\, dk = k_0\ , \qquad (2.38b)$$

where

$$F(k) \equiv \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx \qquad (2.39)$$

and where f(x) has been normalized such that

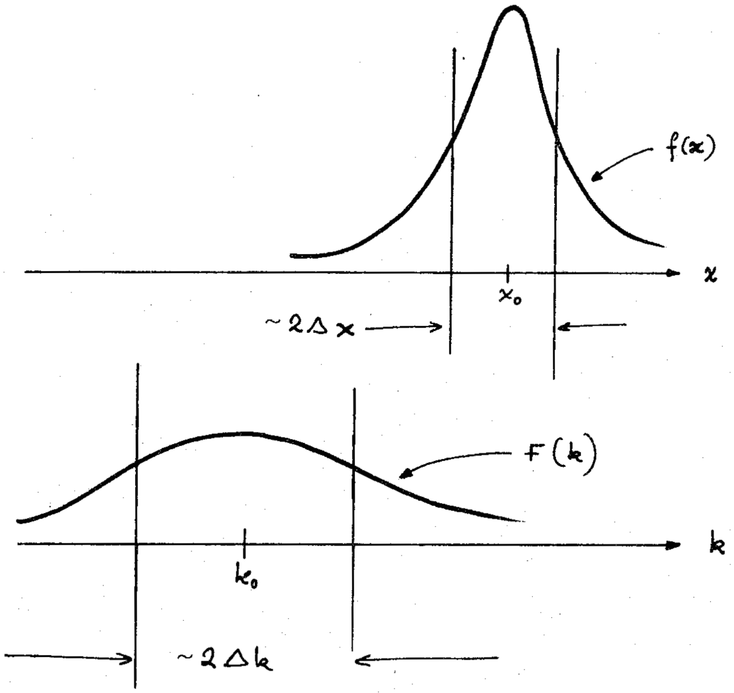
$$\int_{-\infty}^{\infty} |f(x)|^2\, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(k)|^2\, dk = 1\ . \qquad (2.40)$$

To measure the spread of the wave packet we now compute the mean square deviations about x_0 and k_0 respectively, i.e., (see equations 2.25 and 2.26)

$$(\Delta x)^2 \equiv \langle (x - x_0)^2 \rangle = \int_{-\infty}^{\infty} f^*(x)\ (x - x_0)^2\ f(x)\, dx \qquad (2.41a)$$

and

$$(\Delta k)^2 \equiv \langle (k - k_0)^2 \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} F^*(k)\ (k - k_0)^2\ F(k)\, dk\ . \qquad (2.41b)$$



If we now define ($x_0 = \langle x \rangle$, $k_0 = \langle k \rangle$)

$$g(x - x_0)\, e^{ik_0 x} \equiv f(x) \; , \tag{2.42}$$

we have (see Appendix B, §10)

$$F(k) = e^{-i(k - k_0)x_0}\, G(k - k_0) \tag{2.43}$$

where

$$G(k) \equiv \int_{-\infty}^{\infty} g(x)\, e^{-ikx}\, dx \; , \tag{2.44}$$

and therefore

$$(\Delta x)^2 = \int_{-\infty}^{\infty} x^2 \; |g(x)|^2\, dx \tag{2.45}$$

and

$$(\Delta k)^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} k^2 \; |G(k)|^2\, dk \; . \tag{2.46}$$

Making use of the operator $\hat{k}$ for x-space, we can also write equation 2.46 as

$$(\Delta k)^2 = -\int_{-\infty}^{\infty} g^*(x)\, \frac{d^2 g(x)}{dx^2}\, dx \tag{2.47}$$

or, integrating by parts assuming that $g(\pm\infty) = 0$, we have $[g'(x) \equiv dg/dx]$

$$(\Delta k)^2 = \int_{-\infty}^{\infty} |g'(x)|^2 dx \; , \tag{2.48}$$

subject to the normalization condition

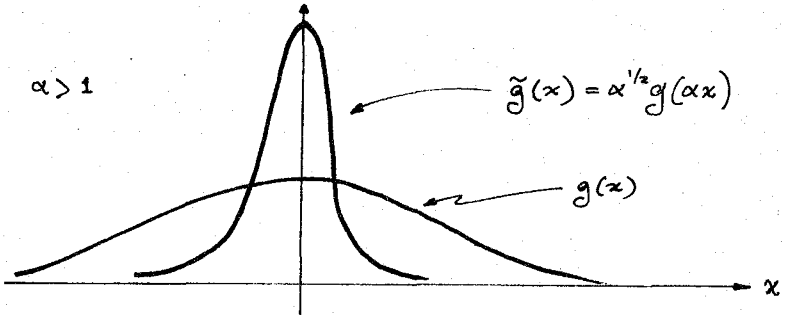
$$\int_{-\infty}^{\infty} |g(x)|^2 dx = 1 . \qquad (2.49)$$

Suppose we now make the transformation

$$\tilde{g}(x) = \alpha^{1/2} g(\alpha x) , \qquad (2.50)$$

where α is a real constant. If $\alpha > 1$, the effect of this transformation is to generate a function $\tilde{g}(x)$ which is narrower (more localized) but taller such that

$$\int_{-\infty}^{\infty} |\tilde{g}(x)|^2 dx = \int_{-\infty}^{\infty} |g(x)|^2 dx = 1 . \qquad (2.51)$$



The spread of $\tilde{g}(x)$ in x-space is given by

$$(\Delta\tilde{x})^2 = \int_{-\infty}^{\infty} x^2 |\tilde{g}(x)|^2 dx = \frac{1}{\alpha^2} \int_{-\infty}^{\infty} x^2 |g(x)|^2 dx = \frac{1}{\alpha^2} (\Delta x)^2 ,$$

and similarly

$$(\Delta\tilde{k})^2 = \int_{-\infty}^{\infty} \left|\frac{d\tilde{g}(x)}{dx}\right|^2 dx = \alpha^2 (\Delta k)^2 ,$$

or, equivalently

$$\Delta\tilde{x} = \frac{1}{\alpha} \Delta x \qquad (2.52a)$$

$$\Delta\tilde{k} = \alpha \, \Delta k . \qquad (2.52b)$$

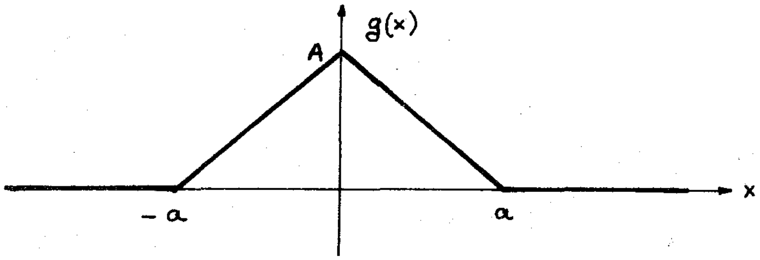
Consequently, we can localize a wave packet g(x) to a smaller region in x-xpace, but only at the expense of localization in k-space. What is interesting, however, is that the joint spread $\Delta x \cdot \Delta k$ remains unaltered by such a transformation, i.e.,

$$\Delta x \cdot \Delta k = \Delta\tilde{x} \cdot \Delta\tilde{k} \qquad (2.53)$$

and is therefore only a function of the shape of g(x).

Example 2..3. Consider the wave packet

$$g(x) = \begin{cases} 0 & ; \ x > -a \\ A\ (1 - |\frac{x}{a}|) & ; \ |x| < a \\ 0 & ; \ x > a \ . \end{cases}$$



Find the root mean square deviations in x-space and k-space.

Normalization:

$$\int_{-\infty}^{\infty} g^2(x)\,dx = 2A^2 \int_0^a (1 - x/a)^2\,dx = \frac{2A^2 a}{3} = 1$$

and therefore

$$A = \left(\frac{3}{2a}\right)^{1/2} .$$

Note that $\langle x \rangle = \langle k \rangle = 0$. Therefore

$$(\Delta x)^2 = \frac{3}{2a} \int_{-a}^{a} x^2 \ (1 - |\frac{x}{a}|)^2\,dx = a^2/10 \ ,$$

and

$$(\Delta k)^2 = \frac{3}{2a} \int_{-a}^{a} \left[\frac{d}{dx}\ (1 - |\frac{x}{a}|)\right]^2 dx = 3/a^2$$

and consequently, $\Delta x = a/\sqrt{10}$ and $\Delta k = \sqrt{3}/a$. Note that the joint spread $\Delta x \cdot \Delta k = (3/10)^{1/2}$, is indeed independent of a.

Does there exist a minimum joint spread? In other words, does there exist a function g(x) such that

$$\Pi^2\,\{g(x)\} \equiv (\Delta x)^2 \cdot (\Delta k)^2 = \min ?$$

Substituting from equations 1.45 and 1.48, we have

$$\Pi^2\,\{g(x)\} = \int_{-\infty}^{\infty} x^2\,|g(x)|^2\,dx \cdot \int_{-\infty}^{\infty} |g'(x)|^2\,dx \ , \qquad (2.54)$$

subject to

$$g(\pm\infty) = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} |g(x)|^2 \, dx = 1 \; . \tag{2.55}$$

Now $g(x)$ is in general a complex function which we can write as

$$g(x) = u(x)\, e^{i\phi(x)} \; . \tag{2.56}$$

It is easy to see, however, that for any function $u(x)$, the functional $\Pi^2\{u, \varphi\}$ attains its minimum if $\varphi'(x) = 0$, or $\varphi(x) = \varphi_0 = \text{const.}$ The problem then becomes the minimization of

$$\Pi^2\{u, \varphi_0\} = \int_{-\infty}^{\infty} x^2 u^2(x)\, dx \cdot \int_{-\infty}^{\infty} u'^2(x)\, dx \; , \tag{2.57}$$

subject to the constraints

$$u(\pm\infty) = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} u^2(x)\, dx = 1 \; , \tag{2.58}$$

with respect to all real functions $u(x)$. Now, the minimization of $\Pi^2\{u,\varphi_0\}$, with respect to $u(x)$, subject to the normalization integral, is equivalent to the minimization of

$$J_\lambda\{u\} \equiv \Pi^2\{u,\varphi_0\} + \lambda \, [1 - \int_{-\infty}^{\infty} u^2(x)\, dx] \tag{2.59}$$

with respect to $u(x)$ <u>and λ</u>. This is evidently so, since the condition for a minimum with respect to λ, i.e.,

$$\frac{\partial}{\partial\lambda} J_\lambda\{u\} = 0 \; ,$$

is in fact the normalization condition.

This problem can be solved using the methods of the calculus of variations. See appendix A. We cannot use the Euler equation, however, since our functional does not involve a single integral but rather the product of two integrals. Nevertheless, the same ideas apply and we seek the solution $u(x)$, such that a small variation $u(x) + \epsilon\eta(x)$ leads to a second order change in $J\{u\}$, i.e.

$$\delta J = J\{u+\epsilon\eta\} - J\{u\} = 0(\epsilon^2) \; .$$

The function $\epsilon\eta(x)$ must be small, i.e. $\epsilon\,|\eta(x)/u(x)| \ll 1$, and $u(x) + \epsilon\eta(x)$ must also satisfy the boundary conditions at $\pm\infty$, i.e.

$$\eta(\pm\infty) = 0 \; . \tag{2.60}$$

The resulting functional J_λ is then given by

$$\begin{aligned}
J_\lambda\{u+\epsilon\eta\} &= \int_{-\infty}^{\infty} x^2(u+\epsilon\eta)^2\,dx \cdot \int_{-\infty}^{\infty} (u'+\epsilon\eta')^2\,dx + \lambda\,[1 - \int_{-\infty}^{\infty} (u+\epsilon\eta)^2\,dx] \\
&= \int_{-\infty}^{\infty} x^2u^2\,dx \cdot \int_{-\infty}^{\infty} u'^2\,dx + \lambda(1 - \int_{-\infty}^{\infty} u^2\,dx) \\
&\quad + 2\left\{ \int_{-\infty}^{\infty} u'^2\,dx \cdot \int_{-\infty}^{\infty} x^2u\eta\epsilon\,dx + \int_{-\infty}^{\infty} x^2u^2\,dx \cdot \int_{-\infty}^{\infty} u'\eta'\epsilon\,dx \right. \\
&\quad \left. - \lambda \int_{-\infty}^{\infty} u\eta\epsilon\,dx \right\} + \text{H.O.T.}
\end{aligned}$$

Note that the sum of the first two terms are equal to $J_\lambda\{u\}$ and that,

$$\int_{-\infty}^{\infty} u'\eta'\,dx = \cancelto{0}{u'\eta\Big|_{-\infty}^{\infty}} - \int_{-\infty}^{\infty} u''\eta\,dx \;.$$

Therefore

$$\begin{aligned}
\delta J &= J_\lambda\{u+\epsilon\eta\} - J_\lambda\{u\} \\
&= -2\int_{-\infty}^{\infty} \{(\Delta x)^2u''(x) + [\lambda - (\Delta k)^2x^2]\,u(x)\}\,\eta(x)\,\epsilon\,dx + \text{H.O.T.},
\end{aligned}$$

where we have substituted

$$(\Delta x)^2 = \int_{-\infty}^{\infty} x^2u^2(x)\,dx \tag{2.61a}$$

and

$$(\Delta k)^2 = \int_{-\infty}^{\infty} u'^2(x)\,dx \quad . \tag{2.61b}$$

Since δJ must be zero to first order in ϵ, for all $\eta(x)$, we must have

$$(\Delta x)^2u''(x) + [\lambda - (\Delta k)^2x^2]\,u(x) = 0 \;, \tag{2.62}$$

where the boundary conditions on $u(x)$ are given by equation 2.58.

Equation 2.62 can be simplified by a change of variables

$$\xi = \alpha x \quad \text{and} \quad v(\xi) = \alpha^{-1/2}u(x) \;, \tag{2.63}$$

[note that $v^2(\xi)\,d\xi = u^2(x)\,dx$] to yield

$$\alpha^2(\Delta x)^2v''(\xi) + \left[\lambda - \frac{(\Delta k)^2}{\alpha^2}\,\xi^2\right]v(\xi) = 0 \;.$$

Choosing

$$\alpha = \left(\frac{\Delta k}{\Delta x}\right)^{1/2} \tag{2.64}$$

we have

$$v''(\xi) + (\beta - \xi^2)\,v(\xi) = 0 \quad , \tag{2.65}$$

with the boundary conditions

$$v(\pm\infty) = 0 \quad \text{and} \quad \int_{-\infty}^{\infty} v^2(\xi)\,d\xi = 1 \,, \tag{2.66}$$

and where the substitution was made

$$\beta = \frac{\lambda}{\Delta x \cdot \Delta k} \,. \tag{2.67}$$

It is perhaps a trivial point but may be worth mentioning that the transformation given by equation 2.50 actually describes a <u>different wave packet</u>, whereas the transformation of equations 2.63 describes <u>the same wave packet</u> by means of different variables.

Rewriting equation 2.65 we have

$$- v''(\xi) + \xi^2 v(\xi) = \beta\, v(\xi)$$

which, when we multiply with $v(\xi)$ and integrate from $-\infty$ to $+\infty$, yields

$$- \int_{-\infty}^{\infty} v(\xi)\, v''(\xi)\, d\xi + \int_{-\infty}^{\infty} \xi^2 v^2(\xi)\, d\xi = \beta \int_{-\infty}^{\infty} v^2(\xi)\, d\xi \,.$$

The integral on the right hand side is just the normalization integral and equal to unity. The integral on the left can be integrated by parts to yield

$$\int_{-\infty}^{\infty} v'^2(\xi)\, d\xi + \int_{-\infty}^{\infty} \xi^2 v^2(\xi)\, d\xi = \beta > 0 \,. \tag{2.68}$$

The two integrals are given by

$$\int_{-\infty}^{\infty} v'^2(\xi)\, d\xi = (\Delta k)^2/\alpha^2 = \Delta x \cdot \Delta k = \Pi \tag{2.69a}$$

and

$$\int_{-\infty}^{\infty} \xi^2 v^2(\xi)\, d\xi = \alpha^2 (\Delta x)^2 = \Delta x \cdot \Delta k = \Pi \,, \tag{2.69b}$$

where we have substituted for α in terms of equation 2.64. Therefore, from equation 2.68 we have

$$\beta = 2\Pi \,. \tag{2.70}$$

It would appear that since $\Pi = \beta/2$, the lowest value of the joint spread $\Pi = \Delta x \cdot \Delta K$ can be obtained by setting β equal to zero. This, however, is not so because the problem posed by equation 2.65, <u>combined</u> with the conditions 2.66, is an <u>eigenvalue</u> problem. Perhaps a more recognizable form might be

$$\hat{\mathcal{L}}\, v(\xi) = \beta\, v(\xi) \ , \qquad (2.71a)$$

where

$$\hat{\mathcal{L}} \equiv -\frac{d^2}{d\xi^2} + \xi^2 \ , \qquad (2.71b)$$

with $v(\pm\infty) = 0$ and $v(\xi)$ normalizable. Such problems have solutions for selected values of the constant $\beta = \beta_n$ called the eigenvalues of the problem and to each of these there corresponds a unique function $v_n(\xi)$ which is called the eigenfunction of the n^{th} eigenvalue.

The solution to this particular problem is given by

$$\beta_n = 2n + 1 \ ; \qquad n = 0, 1, \ldots \qquad (2.72a)$$

and

$$v_n(\xi) = A_n h_n(\xi)\, e^{-\xi^2/2} \qquad (2.72b)$$

where the $h_n(\xi)$ are the Hermite polynomials, defined by

$$h_n(\xi) \equiv (-1)^n e^{\xi^2} \frac{d^n}{d\xi^n} e^{-\xi^2} \qquad (2.73)$$

and A_n is the normalization constant given by

$$A_n = (\pi^{1/2}\, 2^n\, n!)^{-1/2} \qquad (2.74)$$

(see Appendix C).

From this solution we therefore have that the minimum joint spread is equal to $\beta_0/2$, or, equivalently,

$$\boxed{\Delta x \cdot \Delta k \geq 1/2} \ . \qquad (2.75)$$

The wave packet corresponding to this joint spread is then given by $v_0(\xi)$, or

$$v_0(\xi) = \pi^{-1/4} e^{-\xi^2/2} \ , \qquad (2.76)$$

$[h_0(\xi) \equiv 1]$ and is seen to be a gaussian. Any other shape will necessarily be characterized by a larger joint spread.

There is a relatively simple way, using a trick, to prove that $\Delta x \cdot \Delta k \geq 1/2$ and that the wave packet that does it is a gaussian (see problem 2. 6). The method of calculus of variations was used, however, because it is instructive and leads to results which will be directly related to subsequent material in quite a different context.

Suppose now that we ask for the superposition of wavenumbers corresponding to $v_o(\xi)$? Well we have that if

$$v_o(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} V_o(\kappa) e^{i\kappa\xi} d\kappa \quad ,$$

then

$$V_o(\kappa) = \int_{-\infty}^{\infty} v_o(\xi) e^{-i\kappa\xi} d\xi$$

$$= \pi^{-1/4} \int_{-\infty}^{\infty} e^{-(\xi^2/2 + i\kappa\xi)} d\xi$$

$$= 2\pi^{-1/4} e^{-\kappa^2/2} \int_{-\infty}^{\infty} e^{-t^2/2} dt \qquad \#$$

or

$$V_o(\kappa) = (2\pi)^{1/2} [\pi^{-1/4} e^{-\kappa^2/2}] \, , \qquad (2.77)$$

also a gaussian. Note that for the minimum joint spread wave packet (see equations 2.69a and 2.69b) we have

$$\int_{-\infty}^{\infty} \xi^2 v^2(\xi) d\xi = \int_{-\infty}^{\infty} v'^2(\xi) d\xi$$

What about u(x)? To transform from ξ to x (and κ to k) we need $\alpha = (\Delta k/\Delta x)^{1/2}$, which has units of reciprocal length (see eq. 2.63 and 2.64). Note, however, that the joint spread $\Pi = \Delta x \cdot \Delta k$ is dimensionless and, as we in fact showed, can depend only on the shape of u(x), being invariant under a transformation that scales the x-axis (or the k-axis for that matter). See equations 2.50, 2.52 and 2.53 and related discussion. Consequently, the solution to the minimization of Π can only determine the wave packet to within such a transformation. If Δx is given, however, we have (see equation 2.63)

$$u_o(x) = \alpha^{1/2} v_o(\alpha x)$$

$$= \left(\frac{\Delta k}{\pi \Delta x}\right)^{1/4} e^{-\left(\frac{\Delta k}{\Delta x}\right)\frac{\xi^2}{2}}$$

and since $\Delta k = \Pi/\Delta x$ and, corresponding to u_o, $\Pi = 1/2$, we have

$$u_o(x) = [2\pi(\Delta x)^2]^{-1/4} e^{-x^2/4(\Delta x)^2} \quad . \qquad (2.78)$$

Similarly (see Appendix B, §8), we have for the Fourier transform

---

$\#$

$$-(ax^2 + bx) = -a(x^2 + \frac{b}{a} x) = -a(x + \frac{b}{2a})^2 + \frac{b^2}{4a}$$

$$U_0(k) = a^{-1/2}\; V_0(k/a) \quad ,$$

and therefore

$$U_0(k) = (2\pi)^{1/2}\; [2\pi(\Delta k)^2]^{-1/4}\; e^{-k^2/4(\Delta k)^2} \quad . \tag{2.79}$$

The symmetry of equations 2.78 and 2.79 is noteworthy.

Wave packets of definite x or k. What happens to the wave packet as we let $\Delta x \to 0$? Well it will get narrower, since it will have a smaller second moment, but it will also get taller since it must remain normalized. See equation 2.50 and related discussion. The packet in k-space, however, is getting wider and lower since $\Delta x \cdot \Delta k$ = constant. By way of example, compare the wave packet described by equations 2.78 and 2.79 where the dependence on Δx, Δk is explicit. As we pass to the limit of $\Delta x = 0^+$, $u_0^2(x)$ has become infinitely tall and infinitely narrow, i.e.,

$$\lim_{\Delta x \to 0^+} \left\{ u_0^2(x) \right\} = \begin{cases} \infty & \text{for } x = 0 \\ 0 & \text{for } x \neq 0 \ , \end{cases} \tag{2.80a}$$

remaining, however, normalized, i.e.,

$$\int_{\infty}^{\infty} \lim_{\Delta x \to 0^+} \left\{ u_0^2(x) \right\} dx = 1 \quad . \tag{2.80b}$$

The pair of equations, 2.80a and 2.80b, define a generalized function $\delta(x)$,[#] which has the properties of these equations, namely

$$\delta(x) = \begin{cases} \infty & \text{for } x = 0 \\ 0 & \text{for } x \neq 0 \end{cases} \tag{2.81a}$$

and

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1 \quad . \tag{2.81b}$$

We can imagine $\delta(x)$ as an infinitely tall, infinitely thin, unit-area spike at the origin. It is called the Dirac delta function and described in Appendix D. Corresponding to this limiting u(x), the modulus squared of the original wave packet f(x) is given by (see equations 2.42 and 2.56)

$$|f(x)|^2 \to \delta(x - x_0) \quad \text{as} \quad \Delta x \to 0^+ \ , \tag{2.82}$$

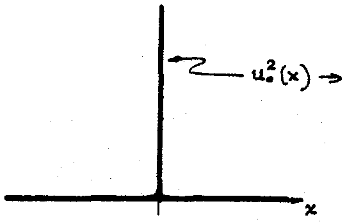
[#] The function $\delta(x)$ is called a "generalized function" because, strictly speaking, a function that is non-zero at only one point has a measure of zero and must therefore have an integral which is zero in conventional analysis.

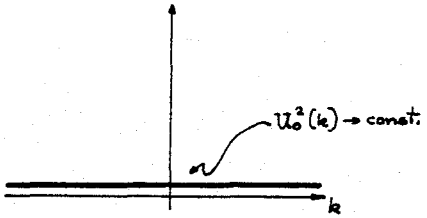
where

$$x_0 \equiv \langle x \rangle = \int_{-\infty}^{\infty} f^*(x)\ x\ f(x)\ dx\ , \qquad (2.83)$$

or, in other words, it is all concentrated at $x = \langle x \rangle = x_0$.

What happens to $U_0^2(k)$ as $\Delta x \to 0^+$? From equation 2.79, we see that since $\Delta k = 2/\Delta x$, $\Delta k \to \infty$ and $U_0^2(k) \to$ constant.





Consequently a wave packet that is completely localized in x-space is uniformly spread out in k-space. Conversly, a wave packet with a definite wavenumber (localized in k-space) is uniformly spread in x-space.

These results are quite general and hold for any shape f(x) or F(k). To illustrate this behavior, study example 2.2 on page 2.5 and problem 2.3 as $a \to 0^+$.

Commuting operators and simultaneous localization. Our inability to localize the wave packet in x and k simultaneously (i.e., $\Pi = \Delta x \cdot \Delta k \neq 0$), is related in an important way to the fact that the operators $\hat{x}$ and $\hat{k}$ do not commute. This can be shown as follows.

Note, first of all, that the expectation values for both $\hat{x}$ and $\hat{k}$ are real i.e.,

$$\langle \hat{x} \rangle = \int f^*(x)\ \hat{x}\ f(x)\ dx = x_0$$

and

$$\langle \hat{k} \rangle = \int f^*(x)\ \hat{k}\ f(x)\ dx = k_0\ .$$

Now because x_0 is real, we must have

$$x_0 = \int f^*(\hat{x} f)\ dx = \int f\ (\hat{x} f)^* dx$$

and therefore the operator $\hat{x}$ must satisfy the relation

$$\int f^*(x)\ [\hat{x}\ f(x)]\ dx = \int [\hat{x}\ f(x)]^*\ f(x)\ dx\ , \qquad (2.84)$$

for any function f(x). Equation 2.84 is easily verified since in x-space, $\hat{x} = x$ and $f^* x = (x f)^*$. Similarly, because k_0 is also real, we must have

$$\int f^*(x)\ [\hat{k} f(x)]\ dx = \int [\hat{k} f(x)]^*\ f(x)\ dx \tag{2.85}$$

for any function f(x). Equation 2.85 can also be directly verified by integration by parts recalling that in x-space $\hat{k} = -i\frac{\partial}{\partial x}$. In general, an operator $\hat{a}$ that satisfies the relationship, for any pair of functions $f(\underline{x})$ and g(x),

$$\int g^*(\hat{a} f)\ d^3\underline{x} = \int (\hat{a} g)^*\ f\ d^3\underline{x} \tag{2.86}$$

is called Hermitian. From the preceding discussion we see that the expectation values of a Hermitian operator are always real.

Consider now two Hermitian operators $\hat{a}$ and $\hat{b}$. These two operators will in general not commute, i.e.,

$$\hat{a}\hat{b} \neq \hat{b}\hat{a} \quad .$$

By way of example, consider the operators $\hat{x}$ and $\hat{k}_x$ and the action of the operator products $\hat{x}\hat{k}_x$ and $\hat{k}_x\hat{x}$ on a wave packet $f(\underline{x})$. In particular

$$\hat{x}\,\hat{k}_x f(\underline{x}) = -i\,x\,\frac{\partial f}{\partial x}\ ,$$

whereas

$$\hat{k}_x\hat{x}\ f(\underline{x}) = -i\,\frac{\partial}{\partial x}\ (xf) = -if - ix\,\frac{\partial f}{\partial x}$$

or

$$(\hat{x}\hat{k}_x - \hat{k}_x\hat{x})\ f\ (\underline{x}) = i f(\underline{x}) \quad . \tag{2.87}$$

The quantity in parentheses is called the commutator of $\hat{x}$ and $\hat{k}_x$ and denoted by $[\hat{x}, \hat{k}_x]$. For any pair of operators $\hat{a}$ and $\hat{b}$ we denote

$$[\hat{a}, \hat{b}] \equiv \hat{a}\hat{b} - \hat{b}\hat{a} \quad . \tag{2.88}$$

See problem 2. From equation 2.87 we then see that

$$\boxed{[\hat{x}, \hat{k}_x] = i} \quad . \tag{2.89}$$

The commutator for a pair of operators $\hat{a}$ and $\hat{b}$ will, in general, be an operator itself, i.e.

$$[a, b] \equiv \hat{a}\hat{b} - \hat{b}\hat{a} = i\hat{c} \tag{2.90}$$

where if $\hat{a}$ and $\hat{b}$ are Hermitian, $\hat{c}$ is also Hermitian. See problem 2.

We can compute the expectation values of $\hat{a}$ and $\hat{b}$ and the mean square deviations Δa and Δb corresponding to any wave packet $\psi(\underline{x})$, i.e.,

$$(\Delta a)^2 \equiv \langle(\hat{a} - \langle\hat{a}\rangle)^2\rangle = \int \psi^*(\hat{a} - \langle\hat{a}\rangle)^2\,\psi\, d^3\underline{x} \tag{2.91}$$

and

$$(\Delta b)^2 \equiv \langle(\hat{b} - \langle\hat{b}\rangle)^2\rangle = \int \psi^*(\hat{b} - \langle\hat{b}\rangle)^2 \psi \, d^3\underline{x} \,, \tag{2.92}$$

or, if we define the operators

$$\hat{a}' = \hat{a} - \langle\hat{a}\rangle \quad , \qquad \hat{b}' = \hat{b} - \langle\hat{b}\rangle \tag{2.93}$$

we have

$$(\Delta a)^2 = \int \psi^* \hat{a}'^2 \psi \, d^3\underline{x} \quad , \quad (\Delta b)^2 = \int \psi^* \hat{b}'^2 \psi \, d^3\underline{x} \quad . \tag{2.94}$$

Now if the operators $\hat{a}$ and $\hat{b}$ are Hermitian, the operators $\hat{a}'$ and $\hat{b}'$ are also Hermitian and we have

$$(\Delta a)^2 = \int (\hat{a}'\psi)^*(\hat{a}'\psi) \, d^3\underline{x} = \int |\hat{a}'\psi|^2 \, d^3\underline{x} \tag{2.95a}$$

and similarly

$$(\Delta b)^2 = \int |\hat{b}'\psi|^2 \, d^3\underline{x} \quad . \tag{2.95b}$$

Combining 2.95a and 2.95b we have

$$(\Delta a)^2(\Delta b)^2 = \int |\hat{a}'\psi|^2 \, d^3\underline{x} \cdot \int |\hat{b}'\psi|^2 d^3\underline{x} \quad . \tag{2.96}$$

Now for any two functions $f(\underline{x})$ and $g(\underline{x})$ we must have (Schwartz inequality)

$$\int |f|^2 d^3\underline{x} \cdot \int |g|^2 d^3\underline{x} \geq \left| \int (f^* g) \, d^3\underline{x} \right|^2 \,, \tag{2.97}$$

where the equality holds if $f(\underline{x}) = \gamma \cdot g(\underline{x})$. Substituting $f(\underline{x}) = \hat{a}'\psi(\underline{x})$ and $g(\underline{x}) = \hat{b}'\psi(\underline{x})$ we have

$$(\Delta a)^2(\Delta b)^2 \geq \left| \int (\hat{a}'\psi)^*(\hat{b}'\psi) \, d^3\underline{x} \right|^2$$

or, since $\hat{a}'$ has been assumed Hermitian,

$$(\Delta a)^2(\Delta b)^2 \geq \left| \int \psi^*(\hat{a}'\hat{b}')\psi \, d^3\underline{x} \right|^2 \quad . \tag{2.98}$$

Now,

$$\begin{aligned} \hat{a}'\hat{b}' &= \frac{1}{2}(\hat{a}'\hat{b}' + \hat{b}'\hat{a}') + \frac{1}{2}(\hat{a}'\hat{b}' - \hat{b}'\hat{a}') \\ &= \frac{1}{2}(\hat{a}'\hat{b}' + \hat{b}'\hat{a}') + \frac{1}{2}[\hat{a}', \hat{b}'] \\ &= \frac{1}{2}(\hat{a}'\hat{b}' + \hat{b}'\hat{a}') + \frac{1}{2}[\hat{a}, \hat{b}] \quad , \end{aligned}$$

or, substituting from equation 2.90 for the commutator, we have

$$\hat{a}'\hat{b}' = \frac{1}{2}(\hat{a}'\hat{b}' + \hat{b}'\hat{a}') + \frac{i}{2}\hat{c} \tag{2.99}$$

and therefore, substituting in equation 2.98,

$$(\Delta a)^2(\Delta b)^2 \geq \frac{1}{4}|\langle(\hat{a}'\hat{b}' + \hat{b}'\hat{a}')\rangle|^2 + \frac{1}{4}|\langle\hat{c}\rangle|^2$$

or

$$\boxed{\Delta a \cdot \Delta b \geq \frac{1}{2}|\langle\hat{c}\rangle|} \quad . \tag{2.100}$$

Substituting $\hat{a} = \hat{x}$ and $\hat{b} = \hat{k}_x$ we have

$$\hat{c} = -i\,[\hat{x}, \hat{k}_x] = 1$$

and obtain the previously established result,

$$\Delta x \cdot \Delta k_x \geq \frac{1}{2} .$$

The equality in equation 2.100 holds if the Schwartz inequality 2.97 is an equality, i.e., if

$$(\hat{a} - \langle\hat{a}\rangle)\psi = \gamma\,(\hat{b} - \langle\hat{b}\rangle)\psi \ , \tag{2.101a}$$

for some constant γ, <u>and</u> if

$$\langle[(\hat{a} - \langle\hat{a}\rangle)(\hat{b} - \langle\hat{b}\rangle) + (\hat{b} - \langle\hat{b}\rangle)(\hat{a} - \langle\hat{a}\rangle)]\rangle = 0 \ ,$$

or, carrying out the multiplications, if

$$\frac{1}{2}\,\langle\hat{a}\hat{b} + \hat{b}\hat{a}\rangle = \langle\hat{a}\rangle\langle\hat{b}\rangle \ . \tag{2.101b}$$

<u>Localization in three dimensions</u>. What if the wavepacket extends in three dimensions, i.e., $f = f(\underline{x})$? It is clear that it is then possible to ask more complicated questions about joint localization. For example, how small can the following joint spreads be?

$$\Delta x \cdot \Delta y, \qquad \Delta x \cdot \Delta k_y, \qquad \Delta z \cdot \Delta k_z, \qquad \Delta k_y \cdot \Delta k_z \ .$$

In each case Δq denotes the root mean square deviation of the variable q from its expectation value, e.g., in $\underline{x}$-space,

$$\Delta q = \left[\int_{\underline{x}} f^*(\underline{x})\,(\hat{q} - \langle\hat{q}\rangle)^2\, f(\underline{x})\, d^3\underline{x}\right]^{1/2} , \tag{2.102}$$

where $\hat{q}$ is the operator corresponding to the quantity q.

We can use the result of the preceding section to answer these questions very simply. In particular since

$$[\hat{x}_i, \hat{x}_j] = 0 \tag{2.103a}$$

$$[\hat{x}_i, \hat{k}_j] = i\,\delta_{ij} \tag{2.103b}$$

$$[\hat{k}_i, \hat{k}_j] = 0 \tag{2.103c}$$

we have,

$$(\Delta x_i) \cdot (\Delta x_j) \geq 0 \tag{2.104a}$$

$$(\Delta x_i) \cdot (\Delta k_j) \geq \frac{1}{2}\,\delta_{ij} \tag{2.104b}$$

$$(\Delta k_i) \cdot (\Delta k_j) \geq 0 \ , \tag{2.104c}$$

where δ_{ij} is the Kronecker delta. Consequently it is possible to localize a wave packet in two $\underline{x}$-directions, e.g., about $\langle x \rangle$ and $\langle y \rangle$, or about two $\underline{k}$-directions, e.g., about $\langle k_x \rangle$ and $\langle k_z \rangle$ and also about one $\underline{x}$-direction and one $\underline{k}$-direction, provided they are along different axes, i.e., $(\Delta x) \cdot (\Delta k_y)$ can be zero whereas $(\Delta y) \cdot (\Delta k_y)$ cannot.

Wave packet motion. Phase velocity and group velocity. Let us consider now a one-dimensional wave packet that can be expressed as a superposition of plane waves, i.e.,

$$f(x,t) = \frac{1}{2\pi} \int_k F_o(k)\, e^{i[kx - \omega(k)t]}\, dk \tag{2.105a}$$

where

$$F_o(k) = \int_x f(x,o)\, e^{-ikx}\, dx \quad . \tag{2.105b}$$

From equation 2.105a we can see that f(x, t) is a superposition of waves with a phase velocity

$$v_\varphi(k) = \frac{\omega(k)}{k} \quad . \tag{2.106}$$

If at $t = 0$, the wave packet is localized in a region

$$x_o - \Delta x_o \leqslant x \leqslant x_o + \Delta x_o \tag{2.107a}$$

and

$$k_o - \Delta k_o \leqslant k \leqslant k_o + \Delta k_o \quad , \tag{2.107b}$$

what can we say about the wave packet at a later time?

The "center of mass" $\langle x \rangle$ will move, since

$$\begin{aligned}\langle x(t) \rangle &= \int f^*(x,t)\; x\; f(x,t)\, dx \\ &= \frac{1}{2\pi} \int F_o^{\,*}(k)\, e^{i\omega(k)t} \left(i \frac{\partial}{\partial k} \right) F_o(k)\, e^{-i\omega(k)t}\, dk \quad ,\end{aligned}$$

which, if $\omega(k)$ is differentiable in the region covered by $F_o(k)$, yields

$$\begin{aligned}\langle x(t) \rangle &= \frac{1}{2\pi} \int F_o^{\,*}(k)\, [i\, F_o'(k) + F_o(k)\omega'(k)t]\, dk \\ &= \frac{i}{2\pi} \int F_o^{\,*}(k)\, F_o'(k)\, dk + \left[\frac{1}{2\pi} \int \omega'(k) \left| F_o(k) \right|^2 dk \right] t\end{aligned}$$

and therefore

$$\langle x(t) \rangle = \langle x(0) \rangle + \langle v \rangle\; t \quad , \tag{2.108}$$

where $\langle v \rangle$ is a constant (independent of time) with units of velocity and given by,

$$\langle v \rangle \equiv \frac{1}{2\pi} \int \omega'(k) \, |F_0(k)|^2 \, dk \quad . \tag{2.109}$$

The integral given by equation 2.109 is like the ones we use to compute the expectation value of a function of k, i.e.,

$$\langle v \rangle \equiv \langle \omega'(k) \rangle \equiv \frac{1}{2\pi} \int \omega'(k) \, |F_0(k)|^2 \, dk \quad . \tag{2.109'}$$

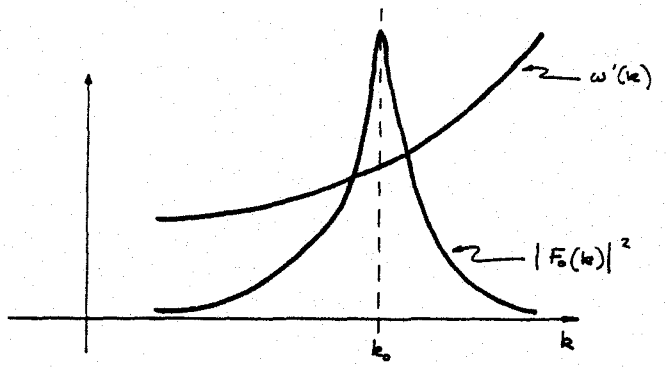
The function $\omega'(k)$ has units of velocity, which we call the group velocity

$$v_g(k) \equiv \frac{d\omega}{dk} \tag{2.110}$$

associated with the wavenumber k. The translational velocity of the center of mass is then given by

$$\langle v \rangle \equiv \frac{1}{2\pi} \int v_g(k) \, |F_0(k)|^2 \, dk \quad . \tag{2.111}$$

We can estimate the integral by noting that the integrand is the product of $\omega'(k)$ and $|F_0(k)|^2$, where $|F_0(k)|^2$ is localized about k_0.



If $\omega'(k)$ is a smooth function in the vicinity of k_0, we have

$$\omega'(k) = \omega'(k_0) + (k - k_0)\omega''(k_0) + \ldots \tag{2.112}$$

and, substituting 2.112 into 2.119, yields

$$\langle v \rangle = \omega'(k_0) \cancelto{1}{\int |F_0(k)|^2 \, d\left(\frac{k}{2\pi}\right)} + \omega''(k_0) \cancelto{0}{\int (k - k_0) \, |F_0(k)|^2 \, d\left(\frac{k}{2\pi}\right)} + \ldots$$

or

$$\langle v \rangle \simeq \omega'(k_0) = v_g(k_0) \ . \tag{2.113}$$

Therefore, the center of mass $\langle x(t) \rangle$ of the wave packet moves with constant velocity which is approximately given by the group velocity evaluated at k_0. These results are extendable in a straight forward fashion to three dimensions, where the phase velocity is given by

$$\underline{v}_\varphi(\underline{k}) \equiv \frac{\omega(\underline{k})}{|\underline{k}|}\, \hat{e}_{\underline{k}} \quad , \tag{2.114}$$

and where $\hat{e}_{\underline{k}}$ is a unit vector in the direction of $\underline{k}$. The group velocity is given by

$$\underline{v}_g(\underline{k}) \equiv \frac{\partial\omega(\underline{k})}{\partial\underline{k}} \quad , \tag{2.115}$$

where $\partial/\partial\underline{k}$ denotes the gradient in $\underline{k}$-space (see equation 2.36).

Note that if the wave packet can be described by an equation like 2.105, the expectation value of k is a constant, i.e.,

$$\langle k(t)\rangle = \frac{1}{2\pi}\int k \left| F_0(k)\, e^{-i\omega(k)t}\right|^2 dk$$
$$= \frac{1}{2\pi}\int k \; |F_0(k)|^2\, dk$$

and therefore

$$\langle k(t)\rangle = \langle k(o)\rangle = k_0 \quad . \tag{2.116}$$

Frequency and time. We can sit at a particular location $\underline{x} = \underline{x}_1$ and watch (if we had complex eyes!) one of these wave packets go by. This will give, of course, the time dependence of the wave amplitude at $\underline{x}_1$, i.e., $f(\underline{x}_1, t)$. Any function of time, however, can be decomposed into a Fourier superposition of frequencies ω, i.e.,

$$f(\underline{x}_1, t) \equiv \frac{1}{2\pi}\int \mathcal{F}(\underline{x}_1, \omega)\, e^{-i\omega t}\, d\omega \quad , \tag{2.117a}$$

where $\mathcal{F}(\underline{x}_1, \omega)$ is given by the inverse Fourier transform, i.e.,

$$\mathcal{F}(\underline{x}_1, \omega) = \int f(\underline{x}_1, t)\, e^{i\omega t}\, dt \quad . \tag{2.117b}$$

Note that as a result of Parseval's theorem (Appendix B, §14),

$$\int |f(\underline{x}_1, t)|^2\, dt = \frac{1}{2\pi}\int |\mathcal{F}(\underline{x}_1, \omega)|^2\, d\omega \tag{2.118}$$

permitting us, if either of the integrals exist, to normalize $f(\underline{x}, t)$ and $\mathcal{F}(\underline{x}, \omega)$ simultaneously, in t and ω respectively, as we did with $\underline{x}$ and $\underline{k}$.

Following the preceding formalism we can ask for the expected time of arrival $\langle t(\underline{x}_1)\rangle$ of the wave packet at $\underline{x} = \underline{x}_1$, given by

$$\langle t(\underline{x}_1)\rangle = \int f^*(\underline{x}_1, t)\; t\; f(\underline{x}_1, t)\, dt \quad , \tag{2.119}$$

and the mean frequency at $\underline{x}_1$

$$\langle \omega(\underline{x_1}) \rangle = \frac{1}{2\pi} \int \mathcal{F}^*(\underline{x_1}, \omega)\ \omega\ \mathcal{F}(\underline{x_1}, \omega)\ d\omega \ , \tag{2.120}$$

where we have assumed that f and $\mathcal{F}$ have been normalized. Equation 2.120 can be re-written in terms of a time integral if we note that

$$\omega\ \mathcal{F}(\underline{x_1}, \omega) = \int \left[i \frac{\partial}{\partial t}\ f(\underline{x_1}, t) \right] e^{i\omega t}\ dt \ ,$$

which we can substitute in equation .120 to obtain

$$\langle \omega(\underline{x_1}) \rangle = \int f^*(\underline{x_1}, t)\ \left(i \frac{\partial}{\partial t} \right) f(\underline{x_1}, t)\ dt \tag{2.121}$$

Similarly, the expectation value for t can be expressed in terms of an integral over all the frequencies, i.e.,

$$\langle t(\underline{x_1}) \rangle = \frac{1}{2\pi} \int \mathcal{F}^*(\underline{x_1}, \omega)\ \left(-i \frac{\partial}{\partial \omega} \right) \mathcal{F}(\underline{x_1}, \omega)\ d\omega \ . \tag{2.122}$$

The symmetry between t and ω reminds us of the properties of the (x, k) pair. In fact, from the above equations, we see that we can define a time and a frequency operator given by:

	t-space	ω-space
$\hat{t}$	$= t$	$= -i \frac{\partial}{\partial \omega}$
$\hat{\omega}$	$= i \frac{\partial}{\partial t}$	$= \omega$.

(2.123)

The similarity between 2.123 and 2.37 should not escape unnoticed.

We can derive the expected mean square deviation of t and ω from $\langle t \rangle$ and $\langle \omega \rangle$ respectively. In particular,

$$(\Delta t)^2 = \int f^*(\underline{x_1}, t)(t - \langle t \rangle)^2\ f(\underline{x_1}, t)\ dt \ , \tag{2.124}$$

and

$$(\Delta \omega)^2 = \frac{1}{2\pi} \int \mathcal{F}^*(\underline{x_1}, \omega)(\omega - \langle \omega \rangle)^2\ \mathcal{F}(\underline{x_1}, \omega)\ d\omega \ . \tag{2.125}$$

We now note that, for the same reasons as hold for the (x, k_x) pair, a wave packet cannot be localized in time and frequency simultaneously. Using the result from noncommuting operators, we then have

$$\Delta t \cdot \Delta \omega \geq \frac{1}{2}\ |[\hat{t}, \hat{\omega}]|$$

and since

$$\begin{aligned} [\hat{t}, \hat{\omega}]\ f(t) &\equiv i \left[t \frac{\partial f}{\partial t} - \frac{\partial}{\partial t}\ (t f) \right] \\ &= -i\ f(t) \ , \end{aligned}$$

we have that

$$[\hat{t}, \hat{\omega}] = -i \tag{2.126}$$

and therefore

$$\boxed{\Delta t \cdot \Delta\omega \geq 1/2} \quad . \tag{2.127}$$

This result is familiar to electrical engineers who know that they must use a detection system with a bandwidth $\Delta\omega$ to define the arrival time of a pulse to within a Δt of the order of $(\Delta\omega)^{-1}$. Equivalently, if we define the number of cycles N in the pulse by the ratio

$$N \simeq \langle\omega\rangle \, 2\Delta t/2\pi \quad ,$$

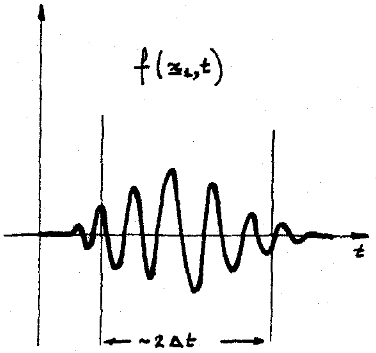
we have

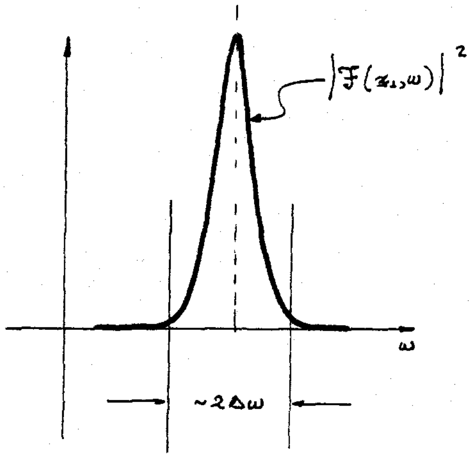
$$\Delta t \simeq \frac{\pi N}{\langle\omega\rangle} \quad ,$$

and therefore

$$\frac{\Delta\omega}{\langle\omega\rangle} \geqslant \frac{1}{2\pi N} \quad , \tag{2.128}$$

a result well-known in spectral analysis that says that the percentage frequency spread of an oscillatory wave, whose amplitude is modulated in time to include a certain number of cycles, is inversly proportional to the number of cycles.





2.3 Energy and Momentum. On December of 1900, Max Planck announced that he was able to account for the shape of the black body radiation spectrum by assuming that the energy in each mode of frequency ν, in a cavity filled with electromagnetic radiation, was discretized as an integer multiple of a quantity proportional to ν, i.e.,

$$E(\nu) = 0,\ h\nu,\ 2h\nu,\ 3h\nu,\ \ldots,\ nh\nu,\ \ldots., \tag{2.129}$$

where h, the constant of proportionality, was independent of frequency.

With this assumption, he computed the energy spectrum (we will derive this later)

$$S(\nu)\, d\nu = \left(\frac{8\pi\nu^2}{c^3}\right)\left(\frac{h\nu}{e^{h\nu/k_B T} - 1}\right) d\nu\ , \tag{2.130}$$

where c is the speed of light and k_B is the Boltzman constant ($k_B \simeq 1.38 \times 10^{-23}$ J/K). In this formula, h was an undetermined constant whose value was selected to provide a good fit to the experimental data. The data and the resulting fit are shown in figure 2.3. The resulting value of h, was very close to the presently accepted value of

$$h = 6.63 \times 10^{-34}\ \text{J}\cdot\text{s}\ .$$

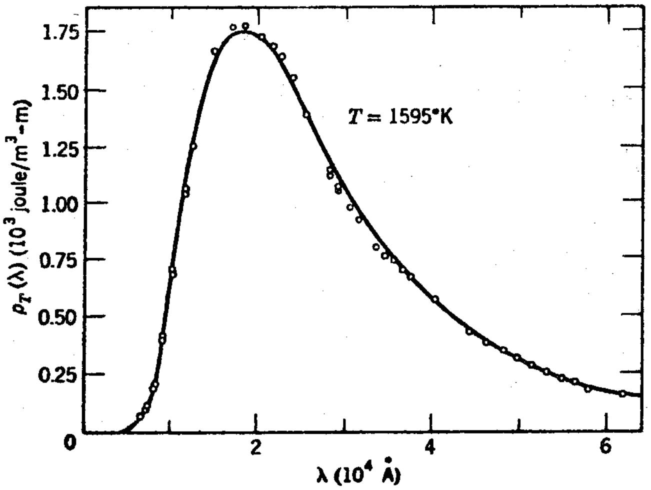


Figure 2.3. Planck's black body spectrum prediction (solid line) versus the experimental data. Note that $\lambda = c/\nu$. (Ref. 1, figure 1-11).

The idea in Planck's model was that the energy in each mode was the sum of the energies of little "bundles" (quanta) each of which had an energy

$$\varepsilon_\nu = h\nu = \hbar\omega \ , \qquad (2.131)$$

where

$$\hbar = h/2\pi \ .$$

The total energy was then equal to the number of "quanta" times the energy of each "quantum", i.e.,

$$E_n(\nu) = n\,\varepsilon_\nu = nh\nu = n\hbar\omega \ . \qquad (2.132)$$

Not long after (1905), Einstein proposed that an electron ejected from a metal by shining light of frequency ν, would have an energy given by

$$E = h\nu - W_0 \ , \qquad (2.133)$$

where W_0, called the work function, is the minimum energy that must be supplied to the electron for it to leave the metal. The prediction is that there exists a minimum frequency

$$\nu_0 = \frac{1}{h} W_0 \ ,$$

that is independent of the intensity of light, below which electrons cannot be ejected from the metal. This proposal was dramatically substantiated by Millikan when in 1914 he published his results of the voltage required to stop electrons ejected from a metal by light of frequency ν. See figure 2.4.

Einstein was awarded the 1921 Nobel prize in physics (which he received in 1922) for his work on the photoelectric effect, and should be credited for the formulation of the concept of an elementary excitation of electromagnetic radiation.[#] It is remarkable, at first sight, that the spectrum of thermal emission of light and the ejection of ~~photons~~ electrons from a metal under the influence of light, phenomena so different, could be explained in terms of the same idea. Given the electronic charge, the slope of the data from Millikan's measurements also gives, of course, the value of Planck's constant h.

[#] The term "photon" was coined by G. N. ~~Lentz~~ Lewis in 1926. See A. Pais, "Einstein and the Quantum Theory", Rev. of Mod. Phys., Oct. 1979, pp. 861-914, esp. p. 887 (footnote by M. Delbrück).

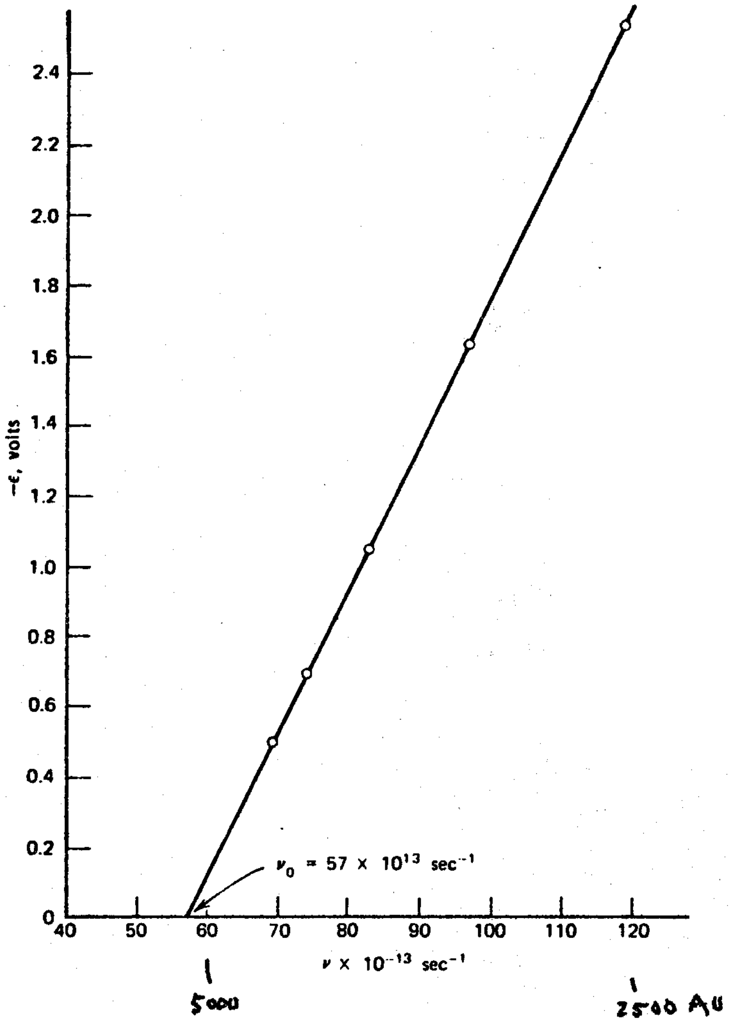


Figure 2.4. Photoelectric effect data of the stopping potential, as a function of light frequency, for electrons ejected from lithium (Ref. 2, figure 1-3).

As a result of a rather different line of research, Compton in 1923 published his experimental data of x-ray scattering from graphite, in which he found in addition to a strong maximum corresponding to light unshifted in wavelength, a second maximum of a longer wavelength shifted by an amount $\Delta\lambda$ which was a function of the scattering angle θ. See figure 2.5.

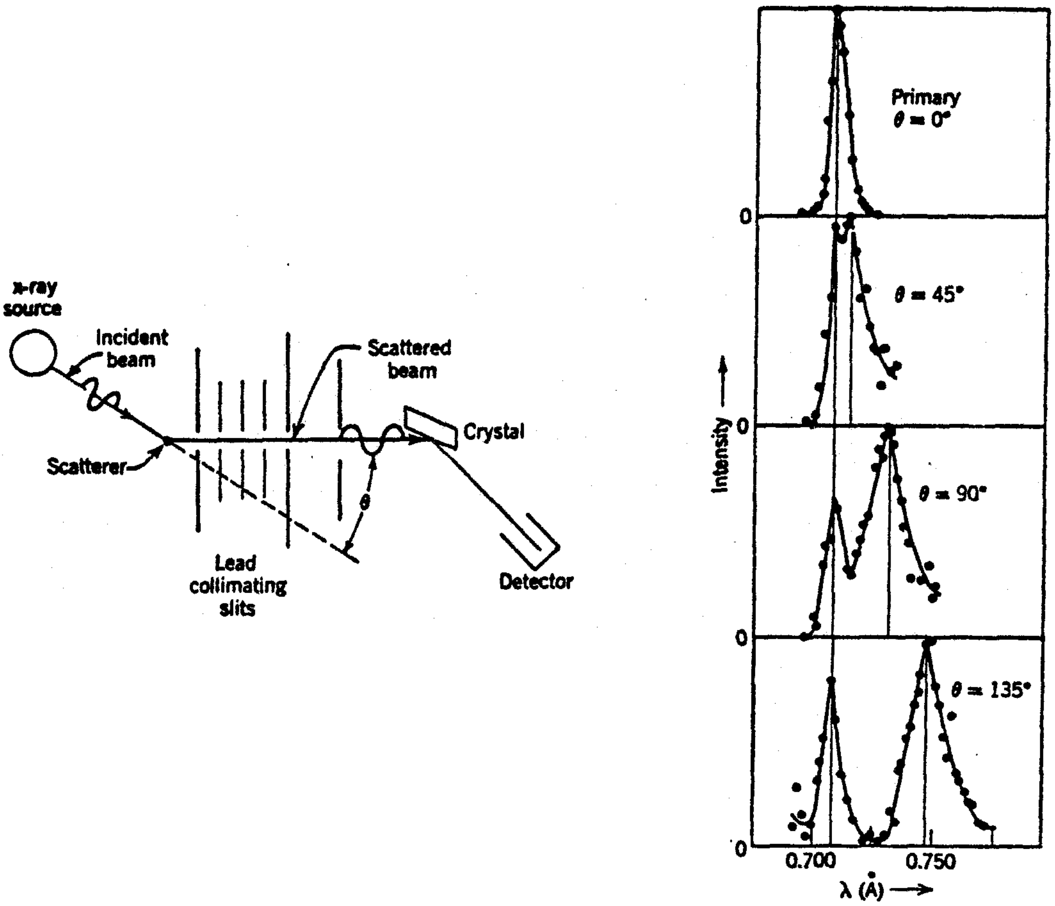


Figure 2.5. Compton's scattering experiment and results (taken from Ref. 1, figures 2-5 and 2-6).

Compton was able to explain his results by assuming that light whose wavelength is λ had to be associated with quanta (photons) whose momentum is given by

$$p = h/\lambda \tag{2.134}$$

or, in terms of the wavevector $\underline{k}$,

$$\underline{p} = \hbar \underline{k} \quad , \tag{2.135}$$

where h is the same constant that was derived from the blackbody theory of Planck and which had been used to explain the photoelectric effect. The shifted peak in intensity is to be understood as resulting from photons elastically scattered by electrons in the graphite. The momentum lost is then a function of the scattering angle and taken up by a recoiling electron. The intensity maximum at the unshifted light wavelength is still the result of electron scattering but, in this case, the momentum is transfered to the atom as a whole. Momentum is of course conserved here also. The atom, however, is so massive (compared to the electron) that the momentum lost to it is negligible. The scattering that results in the shifted peak, as a result of a single recoiling electron, is called Compton scattering. The scattering with the momentum transfered to a whole atom is called Thomson scattering.

The calculation of the momentum lost to the recoiling electron in Compton scattering must be carried out relativistically correctly and yields for the shift

$$\Delta\lambda(\theta) = \frac{h}{m_e c}(1 - \cos\theta) \quad , \tag{2.136}$$

where h is Planck's constant, m_e is the rest mass of the electron and c is the speed of light. The factor in front has units of length and is called the Compton wavelength,

$$\lambda_C \equiv \frac{h}{m_e c} = 2.43 \times 10^{-12} \mathrm{m} = 0.0243 \text{ Å} \; , \tag{2.137}$$

and can be seen to correspond to the wavelength shift $\Delta\lambda$ at a scattering angle of $\theta = 90^\circ$. Note that Compton's data (figure 2.5) are in good agreement with this prediction. Note that since, in general

$$E^2 = p^2c^2 + m_0^2c^4 \quad , \tag{2.138}$$

we must have

$$E = pc \tag{2.139}$$

for a photon, which has zero rest mass m_0, and therefore equations 2.131 and 2.135 are consistent with each other.

The preceding experimental data provide compelling evidence that light in addition to its obvious wave-like nature, can behave as if it is composed of "objects" whose energy is given by

$$\boxed{E = \hbar\omega} \quad , \qquad (2.140a)$$

corresponding to light of angular frequency $\omega = 2\pi\nu$, and momentum which is given by

$$\boxed{\underline{p} = \hbar\underline{k}} \quad , \qquad (2.140b)$$

corresponding to a wavevector $|\underline{k}| = 2\pi/\lambda$. The relationship between the angular frequency ω and the wavevector $|\underline{k}|$, i.e.,

$$\omega = c\,|\underline{k}| \quad , \qquad (2.141)$$

where c is the speed of light, can then be seen to be a re-statement of the fact that, for a particle of zero rest mass, we must have

$$E = c\,|\underline{p}| \quad . \qquad (2.142)$$

Impressed by the ideas of Planck, Einstein and Compton, Louis de Broglie proposed in 1924 in his Ph.D. thesis that the established dual wave-particle nature of light was also valid for material particles (non-zero rest mass). In particular, he proposed that a wave should be associated with the motion of a particle, with a wavenumber $\underline{k}$ given in terms of the particle momentum by

$$\underline{k} = \underline{p}/\hbar \quad , \qquad (2.143a)$$

and a frequency ω given in terms of the particle energy by

$$\omega = E/\hbar \quad . \qquad (2.143b)$$

The idea was very attractive having a unifying simplicity, as Einstein was quick to recognize, but there existed no experimental data to support or refute it.

Two years later, it was pointed out by Elsasser that if matter possessed wave-like properties it could be proven by a diffraction experiment from a crystal, as was done in the case of X-rays. This idea was confirmed by experiments performed by Davisson and Germer in the United States, and by Thomson in Scotland. See figure 2.6. These results established the relationship between wavelength of "matter waves" and the momentum of the associated particle that was predicted by de Broglie to a very high degree of accuracy. Louis de Broglie received his Nobel prize, as did Davisson and Thomson a few years later.

The results of section 2.3 on wave packets acquire a very important significance in the light of the energy-frequency and momentum-wavenumber association. In particular, the localization relations (equation 2.104 and 2.127)

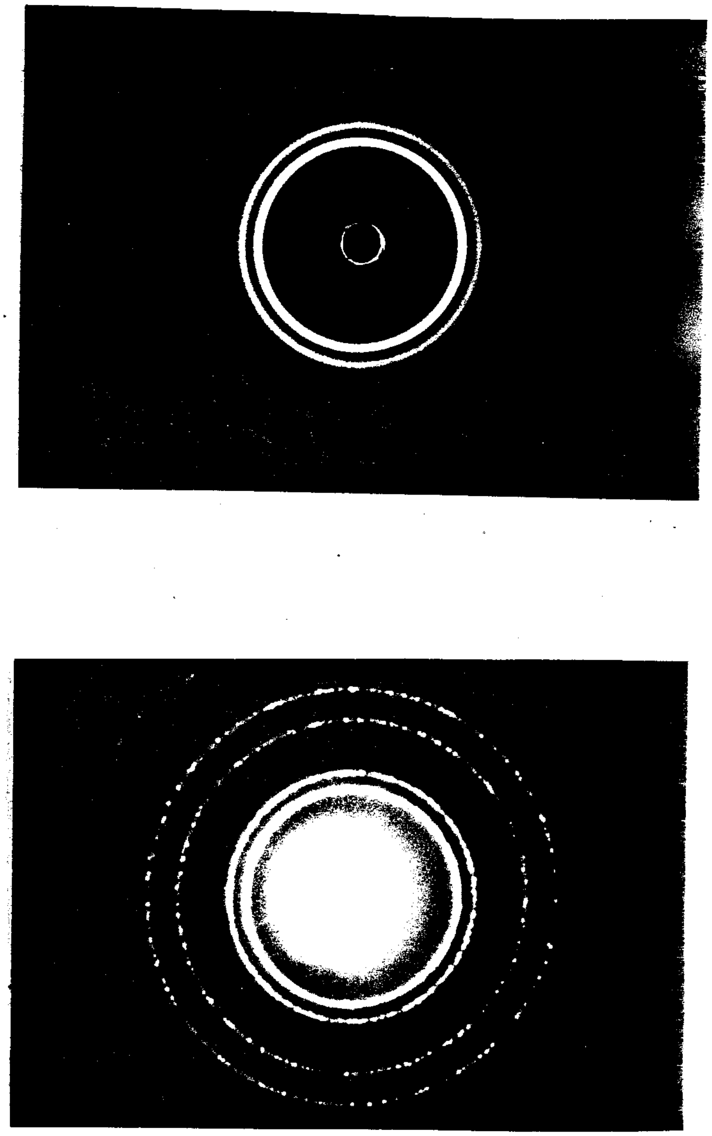


Figure 2.6. X-ray (top picture) and electron beam (bottom picture) diffraction from aluminum powder (reference 3, figures 9.2b and 9.6).

now become

$$(\Delta x_i)(\Delta x_j) \geq 0 \tag{2.144a}$$

$$(\Delta x_i)(\Delta p_j) \geq \frac{1}{2}\hbar\,\delta_{ij} \tag{2.144b}$$

$$(\Delta p_i)(\Delta p_j) \geq 0 \tag{2.144c}$$

$$(\Delta t)(\Delta E) \geq \frac{1}{2}\hbar \quad . \tag{2.144d}$$

Equation 2.144b, in particular, is the tombstone of classical mechanics stating that we cannot simultaneously specify the position and momentum of a particle along a particular direction to better than the joint spread given by 2.144b. This, of course violates the fundamental assumption of classical mechanics that both the position <u>and</u> velocity (momentum divided by mass) must be specified as initial conditions (equation 1.8) for the equation 1.19 of motion to be solved. In addition, the notion of a <u>path</u> must be discarded, since the results of joint localization suggest (when carried over to x, p - space) that a particle occupies <u>a region in phase space</u> (x, p - space), as it moves, given by

$$\langle x_i \rangle - \Delta x_i \leqslant x_i \leqslant \langle x_i \rangle + \Delta x_i$$

and

$$\langle p_j \rangle - \Delta p_j \leqslant p_j \leqslant \langle p_j \rangle + \Delta p_j \quad , \tag{2.145}$$

where, at any time t, the joint spread $\Delta x_i \cdot \Delta p_j$ must satisfy equation 2.144b. This is to be contrasted to classical motion given by the Hamilton equations

$$\dot{\underline{x}} = \frac{\partial H}{\partial \underline{p}} \tag{2.146a}$$

and

$$\dot{\underline{p}} = -\frac{\partial H}{\partial \underline{x}} \quad , \tag{2.146b}$$

where

$$H = H(\underline{p}, \underline{x}, t) \quad , \tag{2.146c}$$

is the <u>Hamiltonian</u> of the system. If H is not an explicit function of time then

$$\frac{dH}{dt} = \frac{\partial H}{\partial \underline{p}}\,\dot{\underline{p}} + \frac{\partial H}{\partial \underline{x}}\,\dot{\underline{x}} \tag{2.147}$$

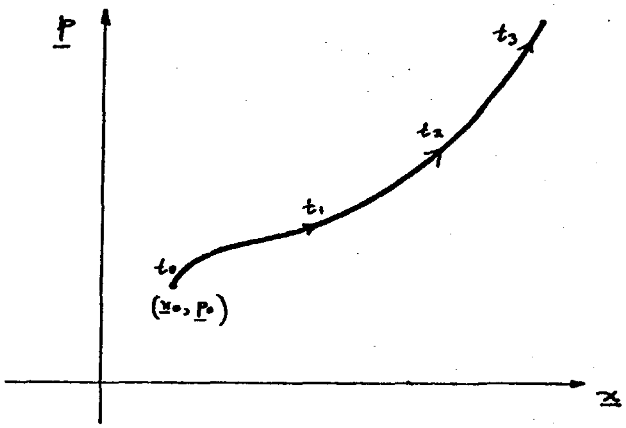
and, by virtue of equations 2.146,

$$\frac{dH}{dt} = 0 \quad .$$

Thus, <u>along the path described by equations 2.146</u>, the Hamiltonian is a constant, i.e.,

$$H = H(\underline{p}, \underline{x}) = E \quad . \tag{2.148}$$

The constant E is the <u>energy</u> of the system. The equations 2.146 of motion yield a path p(t) and x(t) in phase space which is fixed by specifying the



initial conditions

$$\underline{x}_0 = \underline{x}(0) \quad , \quad \underline{p}_0 = p(0) \quad . \tag{2.149}$$

See sections 1.3 and 1.5.

We now also understand the significance of the results on wave packet motion. There we found that for a wave packet composed of a superposition of plane waves, i.e., equation 2.105, the center of mass of the packet moves with constant velocity, i.e.,

$$\langle \underline{x}(t) \rangle = \langle \underline{x}(0) \rangle + \underline{v}\, t \tag{2.150}$$

where

$$\underline{v} \simeq \left. \frac{\partial \omega}{\partial \underline{k}} \right|_{k = \langle k \rangle} \tag{2.151}$$

The fact that the packet center of mass moves with constant velocity suggests that it represents a free particle for which

$$E = \frac{1}{2m} p^2 \quad . \tag{2.152}$$

Now since $E = \hbar\omega$ and $\underline{p} = \hbar \underline{k}$, i.e., both E and $\underline{p}$ are proportional to $\hbar$, we must have

$$v \simeq \frac{\partial \omega}{\partial \underline{k}} = \frac{\partial E}{\partial \underline{p}} = \frac{1}{m} \underline{p} \tag{2.153}$$

which is equal to the classical velocity of the particle! Note that the expectation value of x as a function of time that was derived, correctly predicts the velocity of the particle. Note also that the group velocity, as given by equation 2.151 is the proper velocity and not, for example, the phase velocity

$$\underline{v}_\varphi = \frac{\omega}{|\underline{k}|} \hat{e}_{\underline{k}} = \frac{E}{|\underline{p}|} \hat{e}_p = \frac{1}{2m} \underline{p}$$

of the plane waves

$$e^{i(\underline{p}\cdot\underline{x} - Et)/\hbar}$$

that are used in the superposition.

The fact that the expectation value of the wavenumber is constant in time can now be identified with the fact that the expectation of the momentum is constant, corresponding to a free particle, i.e.,

$$\langle \underline{p} \rangle = \hbar \langle \underline{k} \rangle = \text{const.}$$

or

$$\frac{d}{dt} \langle \underline{p} \rangle = 0 \ . \qquad (2.154)$$

There is an important parallel to be made between the equations of motion for a classical free particle and the results for a wave packet as given by equations 2.153 and 2.154. A particle that is free, i.e., not subject to forces, moves with constant momentum. From the classical equation of motion for $\dot{\underline{p}}$ (equation 2.146b) we see that this occurs if the Hamiltonian is not a function of position, i.e., $H = H(\underline{p})$. In that case we have

$$\dot{\underline{x}} = \underline{v} = \frac{\partial H}{\partial \underline{p}} \qquad \text{and} \qquad \dot{\underline{p}} = 0 \ . \qquad (2.155)$$

Compare this with the results given by equations 2.153 and the fact that the energy was a function of momentum only, behaving as the Hamiltonian of classical mechanics in describing the equations of motion of the expectation values.

Particular significance can now be attached to the operators $\hat{k}$ and $\hat{\omega}$ which after multiplication with $\hbar$ become the operators for momentum and energy, i.e., in (x, t) - space,

$$\boxed{\hat{\underline{p}} = \hbar \hat{\underline{k}} = -i\hbar \frac{\partial}{\partial \underline{x}}} \qquad (2.156)$$

and

$$\boxed{\hat{E} = \hbar \hat{\omega} = i\hbar \frac{\partial}{\partial t}} \ . \qquad (2.157)$$

Note that these operators are valid in general, having been derived by integration by parts of the corresponding integrals for the expectation values (see equations 2.24' and 2.120 and related discussion) without the assumption that f(x, t) was a superposition of plane waves $\exp[i(\underline{k}\cdot\underline{x} - \omega(\underline{k})t)]$. The only assumption is that $f(\underline{x}, t)$, at fixed t, possesses a Fourier transform with respect to $\underline{x}$, i.e.,

$$f(\underline{x}, t) = \left(\frac{1}{2\pi}\right)^3 \int_{\underline{k}} F(\underline{k}, t)\, e^{i\underline{k}\cdot\underline{x}}\, d^3\underline{k} \qquad (2.156a)$$

where

$$F(\underline{k},t) = \int_{\underline{x}} f(\underline{x},t)\, e^{-i\underline{k}\cdot\underline{x}}\, d^3\underline{x}\ , \tag{2.156b}$$

and also that f($\underline{x}$, t), at fixed $\underline{x}$ possesses a Fourier transform with respect to t, i. e.,

$$f(\underline{x},t) = \frac{1}{2\pi}\int_{\omega} \mathcal{G}(\underline{x},\omega)\, e^{-i\omega t}\, d\omega \tag{2.159a}$$

where

$$\mathcal{G}(\underline{x},\omega) = \int_{t} f(\underline{x},t)\, e^{i\omega t}\, dt\ . \tag{2.159b}$$

As a matter of interest, note that any function f($\underline{x}$, t) can be expressed as a superposition in wavenumber and frequency space, i. e.,

$$f(\underline{x},t) = \left(\frac{1}{2\pi}\right)^4 \int_{\underline{k},\omega} \mathfrak{F}(\underline{k},\omega)\, e^{i(\underline{k}\cdot\underline{x} - \omega t)}\, d^3\underline{k}\, d\omega \tag{2.160a}$$

where

$$\mathfrak{F}(\underline{k},\omega) = \int_{\underline{x},t} f(\underline{x},t)\, e^{-i(\underline{k}\cdot\underline{x} - \omega t)}\, d^3\underline{x}\, dt\ . \tag{2.160b}$$

The plane wave superposition, i. e., equation 2.105, is then the special case of

$$\mathfrak{F}(\underline{k},\omega) = 2\pi\ F(\underline{k},o)\ \delta[\omega - \boldsymbol{\omega}(\underline{k})]\ , \tag{2.161}$$

as can be shown by direct substitution into 2.160a. The function $\boldsymbol{\omega}(\underline{k})$ is the plane wave frequency associated with the wavevector $\underline{k}$ [denoted $\omega(\underline{k})$ previously], δ is the Dirac delta function. Consequently, a plane wave superposition is a special case of the most general superposition (equation 2.160) in that it associates a unique frequency $\boldsymbol{\omega}(\underline{k})$ with each wavevector $\underline{k}$.

2.4 Wave-particle duality and the uncertainty principle.

It is worth pausing, perhaps, to reflect on the significance of some of these results and conclusions. We have seen that light, apparently described quite successfully by Maxwell's equations as a wave, can behave like a collection of particles of definite momentum and/or energy. In particular, light of frequency ω can never interact with another object (or apparatus) to lose (or gain) energy that is a fraction of $\hbar\omega$. While this behavior is easily understood in terms of its particle-like behavior, it is difficult to reconcile with its wave-like behavior. Specifically, it is not clear, at this point, how the wave packet (that presumably represents it) can be distributed in $\underline{x}$, at some particular t, as described by f($\underline{x}$, t), so that the probability that it will interact and be detected as a whole is given by $|f(\underline{x},t)|^2$, which is obviously less than unity everywhere.

Moreover, the photon, whose amplitude is spread out according to $f(\underline{x}, t)$ prior to being detected, must be thought of as disappearing from everywhere else once it *has* been detected at some $\underline{x} = \underline{x}_1$. The square of the modulus of the wave function must be thought of as instantly "sucked in" at $\underline{x} = \underline{x}_1$, i.e.,

$$\begin{bmatrix}\text{before interaction}\end{bmatrix} = |f(\underline{x}, t)|^2$$

$$\begin{bmatrix}\text{at the moment of}\\ \text{interaction}\end{bmatrix} = \delta(\underline{x} - \underline{x}_1)\ .$$

These are difficult notions which have actually not been completely clarified as yet (to the best of my knowledge!). Fortunately, in the cases of light, the electron, and their interaction, the situation seems to be in good shape in as much as there exists a *formalism* that appears to answer most of the properly posed questions. Many difficulties arise, however, because the transition from the *formalism* to a *concept* is a difficult one. This duality of character (particle and wave) is a particularly elusive one, the two concepts being so different. In the words of one of the old masters (reference 4, pages 10-11):

> "The solution of the difficulty is that the two mental pictures which experiments lead us to form - the one of particles, the other of waves - are both incomplete and have only the validity of analogies which are accurate only in limiting cases. It is a trite saying that "analogies cannot be pushed too far," yet they may be justifiably used to describe things for which our language has no words. Light and matter are both single entities, and the apparent duality arises in the limitations of our language.
>
> It is not surprising that our language should be incapable of describing the processes occurring within the atoms, for, as has been remarked, it was invented to describe the experiences of daily life, and these consist only of processes involving exceedingly large numbers of atoms. Furthermore, it is very difficult to modify our language so that it will be able to describe these atomic processes, for words can only describe things of which we can form mental pictures, and this ability, too, is a result of daily experience. Fortunately, mathematics is not subject to this limitation, and it has been possible to invent a mathematical scheme - the quantum theory - which seems entirely adequate for the treatment of atomic processes; for visualization, however, we must content ourselves with two incomplete analogies - the wave picture and the corpuscular picture.",

and perhaps on a more elemental level,#

> "... our forms of perception are indeed not learned by the individual, ... but definitely products of evolution, like any other aspects of our body and mind."

#Max Delbrück, "On Einstein", Physics Coloquium lecture, California Institute of Technology, December 6, 1979.

One of the most dramatic manifestations of this duality is the fact that for objects like the electron, which we classically consider as particles, the same localization properties (or lack thereof) apply as do for the photon. Namely that the position and momentum (or velocity) cannot be specified jointly to better than $\hbar/2$ along the same direction (i.e., equation 2.144b). Similarly, the energy of the state and the instant in time that we must assign to that state can also not be specified jointly to better than $\hbar/2$ (i.e., equation 2.144d). Now admittedly $\hbar$ is a small number, i.e.,

$$\hbar = 1.05 \times 10^{-34}\ \mathrm{J}\cdot\mathrm{s} \quad , \tag{2.162}$$

and to the extent that it is, depending on the phenomenon, we can forget about it. It has no effect on the flight of a baseball, for example. Phenomena on the atomic scale, however, involve quantities (the corresponding value of the action to be more exact) that may not be large compared to $\hbar$. In that case the quantum description of such systems is imperative. Conceptually, however, we may have difficulties with the idea that a material particle cannot be localized any better than the limits we have quoted. As Heisenberg says (ref. 4 page 15)

> "Any use of the words 'position' and 'velocity' with an accuracy exceeding that given by equation 1[#] is just as meaningless as the use of words whose sense is not defined."

and he continues in a footnote,

> "In this connection one should particularly remember that the human language permits the construction of sentences which do not involve any consequences and which therefore have no content at all - in spite of the fact that these sentences produce some kind of picture in our imagination; e.g., the statement that besides our world there exists another world, with which any connection is impossible in principle, does not lead to any experimental consequence, but does produce a kind of picture in the mind. Obviously such a statement can neither be proved nor disproved. One should be especially careful in using the words 'reality,' 'actually,' etc., since these words very often lead to statements of the type just mentioned."

It should be clear, at this point, that the "uncertainty principle", as equations 2.144b and 2.144d are called, represents a property of superpositions, which can be expressed conveniently in this case in terms of the Fourier transform. The association of momentum with wavenumber, and energy with frequency comes after the fact, which is valid quite independently of that association. In particular, the "uncertainty principle" has nothing to do with

---

[#] our equation 2.144b.

measurement, even though it is, of course, possible to argue that you cannot violate it by means of a measurement, as is popularly done. The main trouble with such arguments, however, quite apart from drawing attention away from the actual basis of the "uncertainty principle", is that they are classical arguments and at the end one is left wondering, since the discrepancy can apparently be resolved in terms of classical physics, why one needs a quantum theory at all.

A second point that should be made concerns the probabilistic interpretation of $|f(\underline{x}, t)|^2$, i.e., that it represents the probability of detection (interaction with the apparatus). This has caused an awful lot of trouble in the past (and still causing it today!). The reason is that the word probability is usually associated with something that is random, like the throwing of a pair of dice. The difficulty has been compounded by the use of the word uncertainty to describe the joint localization properties of superpositions in phase space. There is nothing uncertain about the evolution in time of the wave function (wave packet). A single electron "knows" that it cannot land on a node of the interference pattern!

2.5 Summary and conclusions. The correspondence principle.

We have seen that the data on diffraction and interference for both light and matter (non-zero rest mass) suggest:

(i) The association of a complex amplitude with the motion of an object of interest that is, in general, a function of space and time, i.e., $f(\underline{x}, t)$.

(ii) This complex amplitude, in order for us to describe the interference data, must be expressible in terms of a superposition of other amplitudes, e.g.,

$$f(\underline{x}, t) = \sum_n C_n f_n(\underline{x}, t)$$

where the C_n's are complex coefficients, or, in the most general case, as a continuous superposition, e.g.,

$$f(\underline{x}, t) = \int_{\underline{k}} F(\underline{k}, t)\, e^{i\underline{k}\cdot\underline{x}}\, d^3\underline{k} .$$

(iii) The square of the modulus of $f(\underline{x}, t)$ must be identified with the <u>probability of detecting</u> the system at $(\underline{x}, t)$, i. e.,

$$\mathcal{P}(\underline{x}, t) = |f(\underline{x}, t)|^2 \quad .$$

(iv) If the superposition is going to be valid, the equations of motion must be <u>linear and homogeneous</u>, i. e., if $f_1(\underline{x}, t)$ and $f_2(\underline{x}, t)$ are solutions,

$$f(\underline{x}, t) = c_1 \; f_1(\underline{x}, t) + c_2 \; f_2(\underline{x}, t)$$

must also be a solution.

(v) If the equations of motion are expressible as differential equations, they must be of <u>first order in time</u>, i. e., $f(\underline{x}, 0)$ should be sufficient to specify $f(\underline{x}, t)$ for all t.

and finally

(vi) For large values of the <u>action</u>, i. e., if $\hbar$ is negligible, the solution of the equations of motion should coincide with the classical equation of motion, which we know, in that limit, are correct.

This latter constraint has been termed the <u>correspondence principle</u>, formulated by N. Bohr in 1923.

We have also seen that light and matter possess a dual character in that they <u>simultaneously display the properties of both particles and waves</u>.

(vii) An entity of energy E and momentum $\underline{p}$ is to be associated with a wave of angular frequency ω and wavevector $\underline{k}$ according to the relations

$$E = \hbar\omega$$

$$\underline{p} = \hbar\underline{k}$$

and vice versa, where $\hbar$ is Planck's constant.

(viii) When working in $(\underline{x}, t)$ - space, the angular frequency and wavevector can be represented in terms of the operators

$$\hat{\omega} = i \frac{\partial}{\partial t}$$

$$\underline{\hat{k}} = -i \frac{\partial}{\partial \underline{x}}$$

and corresponding to these, by virtue of (vii),

$$\hat{E} = i\hbar \frac{\partial}{\partial t}$$

$$\underline{\hat{p}} = -i\hbar \frac{\partial}{\partial \underline{x}} \ .$$

References

2.1 R. Eisberg and R. Resnick, Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles (John Wiley & Sons, 1974).

2.2 S. Gasiorowicz, Quantum Physics (John Wiley & Sons, 1974).

2.3 W. Blanpied, Modern Physics. An Introduction to its Mathematical Language (Holt, Rinehart and Winston, Inc.).

2.4 W. Heisenberg, The Physical Principles of the Quantum Theory (Dover 1949, first published by Chicago Press 1930).

Problems

2.1 In the light interference experiment (figure 2.1) assume that the frequency of the light source is a slowly varying function of time, i.e.,

$$\omega(t) = \omega_0 + \omega'(t)$$

where $\omega'(t)$ measures the excursions of $\omega(t)$ about a mean value ω_0. If $\omega'(t) << \omega(t)$ compute the resulting interference pattern on the plate P_2.

2.2. Instead of a single light source which illuminates an orifice plate to generate two spherical waves, we now use two lasers at $x = \pm \ell/2$ whose frequencies have been matched as well as possible. Nevertheless, the frequencies of the two lasers drift slightly so that

$$\omega_+(t) - \omega_-(t) = \omega'(t)$$

where $\omega'/\omega_+ << 1$ and $\omega'/\omega_- << 1$. Compute the resulting interference pattern assuming that the measurement always represents an average over a time T. What happens to the measured interference pattern as we reduce the laser intensity? What happens when the field between z_1 and z_2 contains the energy corresponding to one photon in a given time? Assume that $\omega'(t)$ is a <u>slowly</u> varying function of time.

2.3 A one-dimensional wave packet at $t = 0$ is given by $(a > 0)$,

$$f(x) = \begin{cases} A \cos[\pi(x - x_0)/a]\, e^{ik_0 x} & ; \ |x - x_0| < a/2 \\ 0 & ; \ |x - x_0| > a/2 \end{cases} .$$

(i) Find the distribution of wavenumbers contributing to $f(x)$.

(ii) Find $\langle x\rangle$, $\langle k\rangle$, $\Delta x = \langle\sqrt{\langle(x - \langle x\rangle)^2}\rangle$, $\Delta k = \langle\sqrt{\langle(k - \langle k\rangle)^2}\rangle$, and the joint spread $\Delta x \cdot \Delta k$.

2.4 Prove that $\langle(x - \langle x\rangle)^2\rangle = \langle x^2\rangle - \langle x\rangle^2$ in general.

2.5 Prove that the minimum of the joint spread

$$\Pi\{g(x)\} = \Delta x \cdot \Delta k ,$$

where the function $g(x)$ is complex and given by

$$g(x) = u(x)\, e^{i\varphi(x)} ,$$

is attained when $\varphi(x) = \text{const}$.

Problems, continued

2.6 From the inequality

$$\left|\frac{x}{2\Delta x}\, g(x) + \Delta x\, g'(x)\right|^2 \ge 0 ,$$

prove that

$$\Delta x \cdot \Delta k \ge \frac{1}{2} ,$$

and that the equality corresponds to a wave packet

$$g(x) = A\, e^{-x^2/4(\Delta x)^2} ,$$

where A may be a complex constant.

2.7 Prove the following properties of operator commutators

(i) $[\hat{a},\hat{b}] + [\hat{b},\hat{a}] = 0$

(ii) $[\hat{a},\hat{a}] = 0$

(iii) $[\hat{a},\hat{b}+\hat{c}] = [\hat{a},\hat{b}] + [\hat{a},\hat{c}]$

(iv) $[\hat{a}+\hat{b},\hat{c}] = [\hat{a},\hat{c}] + [\hat{b},\hat{c}]$

(v) $[\hat{a},\hat{b}\hat{c}] = [\hat{a},\hat{b}]\hat{c} + \hat{b}[\hat{a},\hat{c}]$

(vi) $[\hat{a}\hat{b},\hat{c}] = [\hat{a},\hat{c}]\hat{b} + \hat{a}[\hat{b},\hat{c}]$

(vii) $\big[\hat{a},[\hat{b},\hat{c}]\big] + \big[\hat{c},[\hat{a},\hat{b}]\big] + \big[\hat{b},[\hat{c},\hat{a}]\big] = 0 .$

2.8 If the operators $\hat{a}$ and $\hat{b}$ both commute with their commutators, show that

(i) $[\hat{a},\hat{b}^n] = n\,\hat{b}^{n-1}\,[\hat{a},\hat{b}]$

and

(ii) $[\hat{a}^n,\hat{b}] = n\,\hat{a}^{n-1}\,[\hat{a},\hat{b}]$.

2.9 Prove that

$$[Q(x),\hat{k}_x] = i\,\frac{dQ(x)}{dx} .$$

Compare this result with that of the previous problem if the function Q(x) can be expanded in a power series.

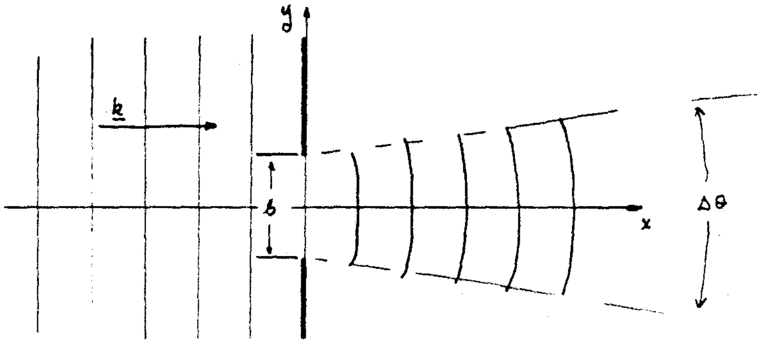
2.10 If $\hat{a}$ and $\hat{b}$ are Hermitian operators, prove that $\hat{c}$, where

$$[\hat{a},\hat{b}] = i\,\hat{c} ,$$

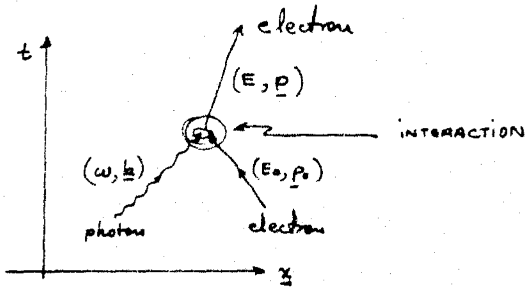
is also a Hermitian operator.

Problems (continued)

2.11 A monochromatic plane wave with a wavevector $\underline{k} = \hat{e}_x k_x$ is incident on a plate at $x = 0$ with a slit of width b. Estimate the spreading of the beam that emerges from the slit.



2.12 Discuss the momentum transferred to an isolated electron that absorbs a photon, i.e.



2.13 In computing the motion of the center of mass $\langle x(t)\rangle$ of a one-dimensional wave packet, we found

$$\langle x(t)\rangle = \langle x(0)\rangle + \langle v\rangle t \ , \qquad (2.108)$$

where v is approximately given by

$$\langle v\rangle = \frac{1}{2\pi}\int_k \frac{\partial\omega}{\partial k}\,|F_0(k)|^2\,dk \approx \left.\frac{\partial\omega}{\partial k}\right|_{k = k_0} \qquad (2.109) - (2.113)$$

where

$$k_0 = \langle k\rangle = \frac{1}{2\pi}\int_k k\,|F_0(k)|^2\,dk = \text{constant} \quad .$$

In comparing the velocity of a free particle with the classical result we found, however, that

$$\langle v\rangle = v = \frac{p}{m} = \frac{\partial E}{\partial p} = \frac{\partial\omega}{\partial k} \qquad (1.153)$$

exactly. Discuss the apparent discrepancy, in view of the fact that we require the expectation value of the center of mass of the wave packet to move classically.

2.14 A one-dimensional wave packet of light ($\omega = 2\pi c/\lambda$) at $t = 0$ is given by the superposition

$$f(x, o) \equiv f_0(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_0(k)\, e^{ikx}\, dk \quad ,$$

where ($k_1 > 0$)

$$F_0(k) = \begin{cases} \left(\dfrac{2\pi}{k_2 - k_1}\right)^{1/2} & \text{for } k_2 > k > k_1 \\ 0 & \text{for } k > k_2 \text{ or } k < k_1 \ . \end{cases}$$

Find $f(x, t)$ and describe the resulting evolution in time of $|f(x, t)|^2$.

2.15 We are given that the wave amplitude at $t = 0$, for a one-dimensional wave packet of light ($\omega = 2\pi c/\lambda$), is given by

$$f(x, o) = f_0(x) \ ,$$

where $f_0(x)$ is a real function. Find $f(x, t)$ and describe the resulting evolution in time of $|f(x, t)|^2$.

Hint: Careful!

2.16 Black body radiation. Compute the number of modes of electromagnetic radiation in a rectangular conducting cavity of volume V, with a frequency between ν and $\nu + \delta\nu$. Assuming that each mode is in thermal equilibrium with the cavity which is at a temperature T, compute the classical expectation value of the energy of the radiation in the cavity in the frequency interval $\nu < \nu' < \nu + \delta\nu$, if the probability that it has an energy E is given by the Boltzmann factor

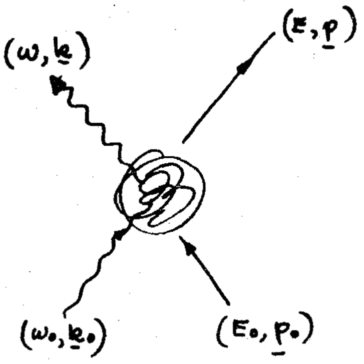
$$p(E)dE \propto e^{-E/kT}\, dE \ .$$

Compare your result with Planck's formula (2.130) in the limit of $h\nu/kT \ll 1$.

2.17 Photoelectric effect. Using the data in figure 2.4, compute h/e, the ratio of Planck's constant to the electronic charge. Compare with the accepted value of h using the known electronic charge.

Problems (continued)

2.18 Compton scattering. Derive the wavelength shift for Compton scattering (equation 2.136). Why is the relativistic calculation necessary?



3. THE SCHRÖDINGER EQUATION

3.1 Hermitian Operators. Eigenfunctions and eigenvalues.

We have seen from the preceding discussion that it is possible to associate with the physical quantities of coordinate, momentum, energy and time, etc., operators whose expectation values correspond to the classical values for these quantities.

Example 3.1. The expectation value of the momentum.

Given a wavepacket $\psi(\underline{x}, t)$, we can compute the expectation value of the momentum using the momentum operator $\hat{\underline{p}}$ for $\underline{x}$-space, i.e.

$$\begin{aligned}\langle \underline{p} \rangle &= \int \psi^*(\underline{x}, t)\ \hat{\underline{p}}\ \psi(\underline{x}, t) d^3\underline{x} \\ &= \hbar \int \psi^*(\underline{x}, t) \left(-i \frac{\partial}{\partial \underline{x}}\right) \psi(\underline{x}, t) d^3\underline{x} \\ &= \hbar \langle \underline{k} \rangle\end{aligned}$$

We would like to generalize this idea and assign an operator to every physical observable of interest. The expected value of the observable would then be computed for a system described by a wave $\psi(\underline{x}, t)$ by the integral

$$\langle Q \rangle = \int \psi^*(\underline{x}, t) [\hat{Q}\ \psi(\underline{x}, t)] d^3\underline{x} \ . \tag{3.1}$$

Now, if the expectation value $\langle Q \rangle$ corresponds to a physical quantity, it must be real, i.e.

$$\langle Q \rangle = \langle Q \rangle^* \ , \quad \text{or} \quad \int \psi^* (\hat{Q}\psi) d^3\underline{x} = \int \psi (\hat{Q}\psi)^* d^3\underline{x} \ , \tag{3.2}$$

for any $\psi(\underline{x}, t)$. The latter equation places important restrictions on the operator $\hat{Q}$ that satisfies the relation

$$\int \psi^* (\hat{Q}\varphi) d^3\underline{x} = \int (\hat{Q}\,\psi)^*\ \varphi\ d^3\underline{x} \ , \tag{3.3}$$

for any $\psi(\underline{x}, t)$ and $\varphi(\underline{x}, t)$ is called Hermitian. It therefore proves necessary to restrict the operators representing physical observables to Hermitian operators[#].

[#] Strictly speaking, we have shown that a Hermitian operator has real expectation values. Equation 3.3, however, appears more restrictive and whereas it is clearly sufficient that the operators be Hermitian, we have not shown that it is necessary.

Example 3.2. Show that the momentum operator is Hermitian.

We have

$$\int \psi^*(\hat{p}\psi)\, d^3\underline{x} = \hbar \int \psi^*\left(-i\frac{\partial}{\partial \underline{x}}\psi\right) d^3\underline{x}$$

$$= \cancelto{0}{-i\hbar\, \psi^*\psi \Big|_{-\infty}^{\infty}} + i\hbar \int\left(\frac{\partial}{\partial\underline{x}}\psi^*\right)\psi\, d^3\underline{x}$$

$$= \hbar\int\left(-i\frac{\partial}{\partial\underline{x}}\psi\right)^*\psi\, d^3\underline{x} \qquad \text{Q.E.D.}$$

Given an arbitrary operator $\hat{A}$, it is useful to define its adjoint operator $\hat{A}^\dagger$, by the equation

$$\int (A^\dagger\varphi)^*\,\psi\, d^3\underline{x} = \int \varphi^*(A\,\psi)\, d^3\underline{x}\ . \tag{3.4}$$

Evidently Hermitian operators are self-adjoint (their own adjoint).

Several important theorems can be proven:

Theorem 1. The expectation values of Hermitian operators are real. This follows from equation 3.3 and the previous discussion.

Theorem 2. The sum of two Hermitian operators is also Hermitian.

Theorem 3. If λ is a complex number, then λ^* (complex conjugate) is its adjoint operator.

Theorem 4. If A is not Hermitian then

$$A_+ = (A + A^\dagger) \tag{3.5a}$$

and

$$A_- = i(A - A^\dagger) \tag{3.5b}$$

are Hermitian. Therefore any operator can be written as a linear combination of two Hermitian operators, i.e.

$$A = \frac{1}{2}(A + A^\dagger) + \frac{1}{2i}\,[i(A - A^\dagger)] \tag{3.6}$$

Theorem 5. If C is the product of two arbitrary operators, i.e.

$$C = A\,B\ ,$$

update

then

$$C^\dagger = (A\,B)^\dagger = B^\dagger A^\dagger \ . \tag{3.7}$$

The action of an operator $\hat{Q}$ on a function $f(\underline{x})$, will in general yield some other function $g(\underline{x})$, i.e.

$$\hat{Q}\, f(\underline{x}) \;=\; g(\underline{x}) \ .$$

For an important class of functions, corresponding to an operator $\hat{Q}$, the action of the operator $\hat{Q}$ on the function results in a multiple of the same function, i.e.

$$\hat{Q}\, \psi_\lambda(\underline{x}) \;=\; \lambda\, \psi_\lambda(\underline{x}) \ , \tag{3.8}$$

where, in general, λ may be a complex number. Functions $\psi_\lambda(\underline{x})$, which for a given operator $\hat{Q}$ satisfy equation 3.8, are called eigenfunctions of $\hat{Q}$ ($\hat{Q}$'s own functions, in German), and the constants λ are called the corresponding eigenvalues. More often than not, additional constraints are placed on the solutions of equation 3.8, for example, we may require that $\psi(\underline{x})$ be *normalized*, i.e.

$$\int_V \psi^*(\underline{x})\, \psi(\underline{x})\, d^3\underline{x} \;=\; 1 \tag{3.9}$$

inside a region V, and we may also impose boundary conditions on some surface S, commonly the edges of V.

Example 3.3. Find the eigenvalues and corresponding eigenfunctions, which vanish outside a region

$$0 < x < a$$
$$0 < y < b$$
$$0 < z < c \ ,$$

of the kinetic energy operator.

We have,

$$\hat{K} = \frac{1}{2m}\,\hat{p}^2 = \frac{1}{2m}\,\hat{\underline{p}}\cdot\hat{\underline{p}} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \tag{3.10}$$

and we require

$$\hat{K}\, u_E(\underline{x}) = E\, u_E(\underline{x}) \tag{3.11}$$

or

$$\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) u_E(x,y,z) + E\, u_E(x,y,z) = 0 \ .$$

By separation of variables, let

$$u(x,y,z) = X(x)\cdot Y(y)\cdot Z(z) \tag{3.12}$$

where

$$X''(x) + k_x^2 X(x) = 0 \tag{3.13}$$

and similarly for Y and Z, and where

$$E = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2) . \tag{3.14}$$

The solution of equation 3.13 is given by

$$X(x) = A \sin(k_x x) + B \cos(k_x x) \quad . \tag{3.15}$$

From the boundary conditions at x=0, we have B=0, while from the boundary condition at x=a we have

$$X(L_x) = A \sin(k_x a) = 0$$

or

$$k_x = \frac{\pi}{a} n_x \quad ; \qquad n_x = 1, 2, \ldots\ldots \tag{3.16}$$

If we additionally impose the normalization condition, we have that

$$\int_0^{L_x} X^2(x)\, dx = A^2 \int_0^{L_x} \sin^2\left(\frac{\pi n_x x}{a}\right) dx = 1$$

or

$$A = (\frac{2}{a})^{1/2} \quad . \tag{3.17}$$

Note that $n_x = 0$ is disallowed by the normalization constraint. Collecting these results, we have the eigenfunctions

$$u_{n_x, n_y, n_z}(x, y, z) = (\frac{8}{V})^{1/2} \sin(n_x \frac{\pi x}{a}) \sin(n_y \frac{\pi y}{b}) \sin(n_z \frac{\pi z}{c}) , \tag{3.18a}$$

where V = abc, and the corresponding eigenvalues

$$E_{n_x, n_y, n_z} = \frac{\hbar^2}{2m} \left(\frac{\pi^2 n_x^2}{a^2} + \frac{\pi^2 n_y^2}{b^2} + \frac{\pi^2 n_z^2}{c^2} \right) , \tag{3.18b}$$

where $n_x, n_y, n_z = 1, 2, 3, \ldots\ldots$

<u>Example 3.4.</u> <u>Find the eigenvalues and corresponding eigenfunctions of the operator for the component of the angular momentum about the z-axis.</u>

Classically, we have

$$L = \underline{x} \times \underline{p}$$

and therefore[#], in $\underline{x}$-space

[#] This is not as casual as it may look. The cross product brings together components of $\underline{x}$ and $\underline{p}$ that commute as operators (see equation 2.144b) so that we don't have to worry about the order in which they operate.

$$\hat{L}_z = (\underline{x} \times \hat{\underline{p}})_z = x\hat{p}_y - y\hat{p}_x$$

or

$$\hat{L}_z = -i\hbar\,(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}) \quad , \tag{3.19}$$

and, transforming to cylindrical coordinates, we have

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\varphi} \quad . \tag{3.20}$$

The eigenvalue equation is then

$$\hat{L}_z\,\Phi_\mu(\varphi) = \mu\,\Phi_\mu(\varphi)$$

or

$$i\hbar\,\Phi'_\mu(\varphi) + \mu\,\Phi_\mu(\varphi) = 0 \quad . \tag{3.21}$$

Now the appropriate boundary condition, in this case is that the eigenfunction should be <u>periodic</u>, i.e.

$$\Phi(0) = \Phi(2\pi) \quad . \tag{3.22}$$

From equation 3.21 we then have

$$\Phi_\mu(\varphi) = e^{i\mu\varphi/\hbar} \quad ,$$

while from the periodic boundary condition we have that

$$1 = e^{i\mu 2\pi/\hbar}$$

or, μ must be an integer multiple of $\hbar$, i.e.

$$\mu = m\hbar \quad ; \qquad m = 0, \pm 1, \pm 2, \tag{3.23a}$$

and therefore

$$\Phi_m(\varphi) = e^{im\varphi} \quad , \tag{3.23b}$$

within a normalization factor.

Consider now two of the eigenvalue equations for $\hat{Q}$,

$$\hat{Q}\,u_n = \lambda_n u_n$$

$$\hat{Q}\,u_m = \lambda_m u_m \quad .$$

or, equivalently, taking the complex conjugate of the latter we have,

$$(\hat{Q}\,u_m)^* = \lambda_m^*\,u_m^* \quad ,$$

and therefore

$$\int_V u_m^*(\hat{Q}\,u_n)\,d^3\underline{x} - \int_V u_n(\hat{Q}\,u_m)^*\,d^3\underline{x} = (\lambda_n - \lambda_m^*)\int_V u_m^*\,u_n\,d^3\underline{x} \quad .$$

Now, if $\hat{Q}$ is Hermitian, the left hand side is zero and consequently

$$(\lambda_n - \lambda_m^*)\int_V u_m^*\,u_n\,d^3\underline{x} = 0 \quad . \tag{3.24}$$

Therefore we have, if m=n, that $\lambda_n = \lambda_n^*$ or,

Theorem 6. The eigenvalues of a Hermitian operator are real.

We also have, if $\lambda_n \neq \lambda_m$, that

$$\int_V u_m^* \, u_n \, d^3\underline{x} = 0 \quad ,$$

or,

Theorem 7. The eigenfunctions of a Hermitian operator corresponding to different eigenvalues are orthogonal.

We may invoke this latter property utilizing a normalized set of eigenfunctions to write

$$\int_V u_m^*(\underline{x}) \; u_n(\underline{x}) \; d^3\underline{x} = \delta_{mn} \quad , \tag{3.25a}$$

where δ_{mn} is the Kronecker delta,

$$\delta_{mn} \equiv \begin{cases} 1 & \text{if} \quad m=n \\ 0 & \text{if} \quad m \neq n \quad . \end{cases} \tag{3.26}$$

See appendix D. Such a set of functions is called orthonormal. The similarity of these properties and those of eigenvectors and eigenvalues of Hermitian matrices in linear algebra should not go unnoticed.

The set of eigenvalues of an operator $\hat{Q}$ is called the spectrum of $\hat{Q}$. The spectrum of eigenvalues may be discrete or continuous or both. In the case of the continuous spectrum, the orthonormality condition becomes

$$\int_V u_\mu^*(\underline{x}) \; u_\nu(\underline{x}) \; d^3\underline{x} = \delta(\mu - \nu) \quad , \tag{3.25b}$$

i.e. the (discrete) Kronecker delta is replaced by the (continuous) Dirac delta function. See appendix D.

Example 3.5. Find the eigenvalues and eigenfunctions of the momentum operator.

In 1-D, $\hat{p} = -i\hbar \frac{\partial}{\partial x}$. Eigenvalue equation

$$\hat{p}\, u = p\, u \quad , \qquad \text{or} \qquad i\hbar\, u'(x) + p\, u(x) = 0 \quad .$$

We have,

$$u(x) = (\text{const.})\, e^{ikx} \quad ; \qquad -\infty < k < \infty \quad ,$$

where

$$p = \hbar k \quad .$$

Note equation 3.25b and the fact that

$$\int_{-\infty}^{\infty} \left[(\frac{1}{2\pi})^{1/2} e^{ik'x}\right]^* \left[(\frac{1}{2\pi})^{1/2} e^{ikx}\right] dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(k-k')x} dx = \delta(k-k')$$

See appendix D, equation D.13.

Theorem 8. If the set of eigenvalues of a Hermitian operator $\hat{Q}$ possesses a minimum (maximum), then the minimum (maximum) expectation value of $\hat{Q}$ is given by the smallest (largest) eigenvalue.

Proof: Wish to minimize (maximize)

$$\langle Q \rangle = \int_V \psi^*(\hat{Q}\psi)\, d^3\underline{x} \quad , \tag{3.27a}$$

subject to the constraint that

$$\int_V \psi^* \psi \, d^3\underline{x} = 1 \quad . \tag{3.27b}$$

Consider instead the functional

$$J\{\psi\} = \int_V \psi^*(\hat{Q}\psi)\, d^3\underline{x} + \lambda[1 - \int_V \psi^*\psi\, d^3\underline{x}] \quad . \tag{3.28}$$

We require then that (see appendix A)

$$\delta J = J\{\psi+\eta\} - J\{\psi\} = 0(\eta^2) \quad .$$

This yields

$$\int_V [\eta^*(\hat{Q}\psi - \lambda\psi) + (\hat{Q}\psi - \lambda\psi)^*\eta]\, d^3\underline{x} = 0$$

where we have used the fact that $\hat{Q}$ is Hermitian to write the second term in the integrand. Therefore, if this is to hold for any $\eta(x)$, we must have

$$\hat{Q}\psi = \lambda\psi \quad , \tag{3.29}$$

or that ψ must be an eigenfunction of $\hat{Q}$.

Using this result, we then see that λ, which was introduced as a Lagrange multiplier, is in fact the expectation value under these conditions, i.e.

$$\langle Q \rangle = \lambda \quad .$$

Therefore

$$\begin{pmatrix}\min\\ \max\end{pmatrix} \{\langle Q \rangle\} = \begin{pmatrix}\min\\ \max\end{pmatrix} \{\lambda_n\} \quad , \tag{3.30}$$

and the minimum (maximum) is attained when ψ is the corresponding eigenfunction.

It should be emphasized that in the preceding discussion, integrals over the coordinate vector space $\underline{x}$ were used for the purposes of illustration only. All these results are independent of this assumption and are equally valid if the independent variables are wavenumbers, or time or frequency or anything else. For this reason it is advantageous to introduce a notation which is independent of the particular space in which the functions are expressed (sort of like vector notation which is independent of the coordinate system). We will denote a function $\psi(\underline{x}, t)$ as follows[#]

$$|\psi\rangle \rightleftarrows \psi(\underline{x}, t)\ ,\ \Psi(\underline{k}, t) \quad \text{etc.} \tag{3.31}$$

while its complex conjugate

$$\langle\psi| \rightleftarrows \psi^*(\underline{x}, t)\ ,\ \Psi^*(\underline{k}, t) \quad \text{etc.} \tag{3.32}$$

The action of an operator $\hat{Q}$ on this function will then be written as

$$\hat{Q}\psi \rightleftarrows \hat{Q}|\psi\rangle \quad , \tag{3.33}$$

whereas, if which to denote the result of this operation, i.e. $(\hat{Q}\psi) = \varphi$, we will write

$$\varphi = (\hat{Q}\psi) \rightleftarrows |\hat{Q}\psi\rangle \quad . \tag{3.34}$$

Equations 3.33 and 3.34 will be interchangeable for most purposes.

An integral over the space of the independent variable will be written as

$$\int_V \varphi^*(\hat{Q}\psi)\, d^3\underline{x} \rightleftarrows \langle\varphi|\hat{Q}\psi\rangle = \langle\varphi|\hat{Q}|\psi\rangle \quad . \tag{3.35}$$

Clearly, if $\hat{Q}$ is Hermitian

$$\langle\varphi|\hat{Q}|\psi\rangle = \langle\varphi|\hat{Q}\psi\rangle = \langle\hat{Q}\varphi|\psi\rangle \quad . \tag{3.36}$$

The expectation value of $\hat{Q}$ then becomes

$$\langle Q\rangle = \langle\psi|\hat{Q}|\psi\rangle \tag{3.37}$$

while the orthonormality of the eigenfunctions of $\hat{Q}$ can be written as (see equation 3.25a)

$$\langle u_m|u_n\rangle = \delta_{mn} \quad . \tag{3.38}$$

3.2 Superpositions of eigenfunctions.

Given a set of normalized eigenfunctions $\{u_n\}$ of a Hermitian operator Q, subject to some boundary conditions, we construct a linear superposition of some number of them, i.e.

$$|f\rangle = \sum_{n=0}^{N} |u_n\rangle c_n \quad . \tag{3.39}$$

[#] This notation is due to P.A.M. Dirac. See reference **3.2** .

update

If we take the scalar product of this equation with $\langle u_m|$, where $0 \leq m \leq N$, we would have

$$\langle u_m|f\rangle = \sum_{n=0}^{N} \langle u_m|u_n\rangle\, c_n{}^{\#} \quad . \tag{3.40}$$

or, using the orthogonality of the eigenfunctions (equation 3.38),

$$\langle u_m|f\rangle = \sum_{n=0}^{N} \delta_{mn}\, c_n = c_m \quad .$$

Therefore, we have that the relation of the coefficients, in a superposition of the type expressed by equation 3.39, and f, which represents the sum, is given by

$$c_n = \langle u_n|f\rangle \quad , \tag{3.41}$$

an expression which has the appealing geometrical interpretation of the <u>projection</u> of f on u_n. If we substitute equation 3.41 in 3.39, we have

$$|f\rangle = \sum_{n=0}^{N} |u_n\rangle\langle u_n|f\rangle \quad . \tag{3.42}$$

We can compute the expectation value of the operator $\hat{Q}$ in the superposition of equation 3.39, as follows, we have

$$\begin{aligned}\langle\hat{Q}\rangle = \langle f|\hat{Q}|f\rangle &= \sum_{m,n=0}^{N} c_m^* c_n \langle u_m|\hat{Q}|u_n\rangle \\ &= \sum_{m,n=0}^{N} \lambda_n c_m^* c_n \langle u_m|u_n\rangle \\ &= \sum_{m,n=0}^{N} \lambda_n c_m^* c_n \delta_{mn}\end{aligned}$$

or

$$\langle\hat{Q}\rangle = \sum_{n=0}^{N} \lambda_n |c_n|^2 = \sum_{n=0}^{N} \lambda_n |\langle u_n|f\rangle|^2 \quad . \tag{3.43}$$

Similarly, of course, for any integral power of $\hat{Q}$, i.e.

$$\langle\hat{Q}^m\rangle = \sum_{n=0}^{N} \lambda_n^m |c_n|^2 \tag{3.44}$$

and any function (resolvable in a power series), i.e.

[#] Note that if we were working in $\underline{x}$-space this equation would read,

$$\int_V u_m^*(\underline{x})\, f(\underline{x})\, d^3\underline{x} = \sum_{n=0}^{N} c_n \int_V u_m^*(\underline{x})\, u_n(\underline{x})\, d^3\underline{x}$$

$$\langle F(\hat{Q})\rangle = \sum_{n=0}^{N} F(\lambda_n)|c_n|^2 \quad . \tag{3.45}$$

It is therefore particularly convenient, in computing expectation values of operators, or functions of operators, if we are dealing with a superposition of eigenfunctions.

3.3 Eigenfunction expansions. Completeness.

The remarkable thing is that any function can be represented by an infinite superposition of eigenfunctions,[#] i.e., given any normalized $\psi(\underline{x})$, there exists a sequence of coefficients c_n, such that

$$\lim_{N\to\infty}\left\{\sum_{n=0}^{N} c_n u_n(\underline{x})\right\} = \psi(\underline{x}) \quad . \tag{3.46}$$

The equality in the limit is to be understood in a least squares sense, i.e.,

$$\lim_{N\to\infty}\left\{\int_V \left|\psi(\underline{x}) - \sum_{n=0}^{N} c_n u_n(\underline{x})\right|^2 d^3\underline{x}\right\} = 0 \quad . \tag{3.47}$$

To prove this property, we will order the $u_n(\underline{x})$ in a sequence of increasing absolute values of the corresponding eigenvalues, i.e.,

$$|\lambda_0| \le |\lambda_1| \le |\lambda_2| \le \ldots \le |\lambda_n| \le \ldots \quad . \tag{3.48}$$

Now, let $\psi_{N+1}(\underline{x})$ be the difference between the function $\psi(\underline{x})$ and a superposition of the first N+1 of the ordered eigenfunctions $0 \le n \le N$, i.e.,

$$\psi_{N+1}(\underline{x}) = \psi(\underline{x}) - \sum_{n=0}^{N} c_n u_n(\underline{x}) \ , \tag{3.49}$$

such that $\psi_{N+1}(\underline{x})$ is orthogonal to $\{u_0, u_1, \ldots, u_N\}$. This requirement dictates that the coefficients c_n be defined by (see problem 3.4)

$$c_n \equiv \int_V u_n^*(\underline{x})\,\psi(\underline{x})\,d^3\underline{x} \ . \tag{3.50}$$

Consider now the normalization integral,

[#] of a Hermitian operator

$$a_{N+1}^2 = \int_V \psi_{N+1}^*(\underline{x})\, \psi_{N+1}(\underline{x})\, d^3\underline{x} \tag{3.51}$$

$$= \int_V \left(\psi - \sum_{n=0}^{N} c_n u_n\right)^* \left(\psi - \sum_{n=0}^{N} c_n u_n\right) d^3\underline{x}$$

or

$$a_{N+1}^2 = \int_V \psi^* \psi\, d^2\underline{x} - \sum_{n=0}^{N} |c_n|^2 . \tag{3.52}$$

Note that if $a_{N+1}^2 \rightarrow 0$ as $N \rightarrow \infty$, we have proved the statement. See equations 3.51, 3.49 and 3.47. We can use the a_{N+1} to normalize the remainder function. We then have

$$\tilde{\psi}_{N+1}(\underline{x}) = \frac{1}{a_{N+1}} \psi_{N+1}(\underline{x}) \tag{3.53}$$

where

$$\int_V \tilde{\psi}_{N+1}^*(\underline{x})\, \tilde{\psi}_{N+1}(\underline{x})\, d^3\underline{x} = 1 , \quad \int_V u_n^*(\underline{x}) \tilde{\psi}_{N+1}(\underline{x})\, d^3\underline{x} = 0, \quad 0 \le n \le N \tag{3.54}$$

Consider now the expectation value of $\hat{Q}^2 - \lambda_0^2$ with the function $\tilde{\psi}_{N+1}(\underline{x})$, i.e.,

$$\langle(\hat{Q}^2 - \lambda_0^2)\rangle_{N+1} = \int \tilde{\psi}_{N+1}^*(\underline{x})\, (\hat{Q}^2 - \lambda_0^2)\, \tilde{\psi}_{N+1}(\underline{x})\, d^3\underline{x} \quad . \tag{3.55}$$

Since we know, however, that $\tilde{\psi}_{N+1}(\underline{x})$ is normalized and orthogonal to the first N+1 eigenfunctions, see equations 3.54, we must have

$$\int \tilde{\psi}_{N+1}^*(\underline{x})\, (\hat{Q}^2 - \lambda_0^2)\, \tilde{\psi}_{N+1}(\underline{x})\, d^3\underline{x} \ge \lambda_{N+1}^2 - \lambda_0^2 \tag{3.56}$$

(see problem 3.5). Using equations 3.49 and 3.53, we also have, however, that

$$\langle(\hat{Q}^2 - \lambda_0^2)\rangle_{N+1} = \frac{1}{a_{N+1}^2} \left[\int_V \left(\psi - \sum_{n=0}^{N} c_n u_n\right)^* (\hat{Q}^2 - \lambda_0^2) \left(\psi - \sum_{n=0}^{N} c_n u_n\right) d^3\underline{x} \right] ,$$

or

$$\langle(\hat{Q}^2 - \lambda_0^2)\rangle_{N+1} = \frac{1}{a_{N+1}^2} \left[1 - \sum_{n=0}^{N} (\lambda_n^2 - \lambda_0^2)\, |c_n|^2 \right] , \tag{3.57}$$

where we have used that $\psi(\underline{x})$ is normalized, i.e.,

$$\int_V \psi^*(\underline{x})\, \psi(\underline{x})\, d^3\underline{x} = 1 \quad . \tag{3.58}$$

Therefore, combining 3.56 and 3.57 we have that

$$\frac{1}{a_{N+1}^2}\left[1-\sum_{n=0}^{N}(\lambda_n^2-\lambda_0^2)\,|c_n|^2\right] \ge \lambda_{N+1}^2-\lambda_0^2$$

or, since $\lambda_n^2-\lambda_0^2 \ge 0$, we must have

$$(\lambda_{N+1}^2-\lambda_0^2)\,a_{N+1}^2 \le 1. \qquad (3.59)$$

Therefore, provided $|\lambda_{N+1}| \to \infty$ as $N \to \infty$, we must have that $a_{N+1} \to 0$ as $N \to \infty$, as was required.

A set of functions which can serve as a basis in terms of which we can expand any other function, in the sense of equation 3.47, is called <u>complete</u>. We have therefore proven that

<u>Theorem 9.</u> If the absolute values of the eigenvalues of a Hermitian operator form an unbounded set, the corresponding eigenfunctions form a complete set.

We therefore have, for any $\psi(\underline{x})$ that if $u_n(\underline{x})$ are the eigenfunctions of such an operator, then

$$\psi(\underline{x}) = \sum_{n=0}^{\infty} c_n u_n(\underline{x}) \qquad (3.60a)$$

where

$$c_n = \int u_n^*(\underline{x})\psi(\underline{x})\,d^3\underline{x} \quad , \qquad (3.60b)$$

or, in terms of our representation-independent notation,

$$|\psi\rangle = \sum_{n=0}^{\infty} |u_n\rangle c_n \qquad (3.61)$$

or

$$|\psi\rangle = \sum_{n=0}^{\infty} |u_n\rangle\langle u_n|\psi\rangle \quad . \qquad (3.61')$$

All the results of finite superpositions of eigenfunctions now apply (section 3.2), for example, if

$$\hat{\Omega}|u_n\rangle = \lambda_n|u_n\rangle$$

then

$$\langle\psi|\hat{\Omega}|\psi\rangle = \sum_n \lambda_n|c_n|^2 \quad , \qquad (3.62)$$

etc.

An interesting property of such expansions may be obtained as follows. From equation 3.60 we have that

$$\psi(\underline{x}) = \sum_{n=0}^{\infty} u_n(\underline{x}) \int_V u_n^*(\underline{x}') \, \psi(\underline{x}') \, d^3\underline{x}'$$

or

$$\psi(\underline{x}) = \int_V \left[\sum_{n=0}^{\infty} u_n^*(\underline{x}') \, u_n(\underline{x}) \right] \psi(\underline{x}') \, d^3\underline{x}' \quad .$$

Comparing the left and right hand side we see that we must have, for a complete set

$$\sum_{n=0}^{\infty} u_n^*(\underline{x}') \, u_n(\underline{x}) = \delta(\underline{x}' - \underline{x}) \quad . \tag{3.63}$$

This is called the closure relation. Compare this with equations 3.25a and 3.25b.

Note also that the normalization integral, i.e.

$$\langle \psi | \psi \rangle = 1$$

now becomes

$$\langle \psi | \psi \rangle = \sum_{n,m} c_m^* c_n \langle u_m | u_n \rangle = \sum_{n,m} c_m^* c_n \delta_{mn}$$

or

$$\sum_n |c_n|^2 = 1 \quad . \tag{3.64}$$

3.4 Hermitian operators and associated observables.

We have seen that the expected outcome of a set of measurements of an observable λ associated with an operator $\hat{Q}$ in a system described by a function $|\psi\rangle$ is given by

$$\langle \hat{Q} \rangle = \langle \psi | \hat{Q} | \psi \rangle \ , \tag{3.65}$$

with an expected mean square deviation from this expected mean, that would be given by

$$\Delta Q^2 = \langle (\hat{Q} - \langle \hat{Q} \rangle)^2 \rangle = \langle \psi | (\hat{Q} - \langle \hat{Q} \rangle)^2 | \psi \rangle \tag{3.66}$$

If we now consider $|\psi\rangle$ as an expansion in the eigenfunctions of $\hat{Q}$, i.e.

$$|\psi\rangle = \sum_{n=0}^{\infty} |u_n\rangle c_n \; ; \quad c_n = \langle u_n|\psi\rangle$$

then we have

$$\langle Q\rangle = \sum_{n=0}^{\infty} \lambda_n |c_n|^2 \tag{3.67}$$

and

$$\Delta Q^2 = \sum_{n=0}^{\infty} (\lambda_n - \langle\hat{Q}\rangle)^2 |c_n|^2 \quad , \tag{3.68}$$

where the λ_n are the eigenvalues corresponding to the $|u_n\rangle$. Let us now assume that we have prepared our system so that it is exactly described by one of the eigenfunctions of $\hat{Q}$, i.e. if

$$|\psi\rangle = |u_m\rangle \quad . \tag{3.69a}$$

We then have $c_m = 1$ and $c_n = 0$ for $n \neq m$. Correspondingly,

$$\langle\hat{Q}\rangle = \lambda_m \tag{3.69b}$$

and

$$\Delta Q^2 = 0 \quad . \tag{3.69c}$$

In other words, *if the system is described by an eigenfunction of the operator associated with the observable of interest, a measurement of that observable can only result in the corresponding eigenvalue*. Such a system will be described as being in an *eigenstate* of $\hat{Q}$. The converse is also true, namely if

$$\langle\hat{Q}\rangle = \lambda_m \quad \text{and} \quad \Delta Q^2 = 0$$

we must have $|\psi\rangle = |u_m\rangle$. See problem 3.6. Consistently with this result, we may now interpret the coefficient c_n as the *probability amplitude* that the system will be detected in the eigenstate $|u_n\rangle$, and therefore $|c_n|^2$ is the corresponding probability. In this light equation 3.66 can be understood as a classical computation of the expectation value of a variable λ_n with relative probability weights of $|c_n|^2$. The normalization

$$\langle\psi|\psi\rangle = \sum_n |c_n|^2 = 1$$

is then to be interpreted as a statement to the effect that the sum of the probabilities of detecting the system in the eigenstates $|u_n\rangle$ must equal unity, i.e. *there are no other possibilities*. Note, however, that we are not saying that the system *is* in *the* state $|u_n\rangle$ or $|u_m\rangle$ or whatever with probability $|c_n|^2$, $|c_m|^2$ etc. The system *is* in *all* these states [in accordance with the expansion,

$$|\psi\rangle = \sum_n |u_n\rangle c_n \; ; \quad c_n = \langle u_n|\psi\rangle$$

simultaneously.

3.5 The energy operator. Schrödinger's equation.

We have seen, from the discussion in chapter 2, that we can associate an energy $E = \hbar\omega$, with a system characterized by a frequency ω. We also saw that the appropriate corresponding operator to the frequency is given by equation 2.123, i.e.

$$\hat{\omega} = i\frac{\partial}{\partial t} \, ,$$

and that therefore the corresponding energy operator is given by equation 2.157, i.e.

$$\hat{E} = i\hbar\frac{\partial}{\partial t} \quad .$$

We therefore have, that the expectation value of the energy would be given by

$$\langle E\rangle = \langle\psi|\hat{E}|\psi\rangle = \langle\psi|i\hbar\frac{\partial}{\partial t}|\psi\rangle \quad , \tag{3.70}$$

or, for a system described in $(\underline{x}, t)$ space,

$$\langle E\rangle = i\hbar\int\psi^*(\underline{x}, t)\,\frac{\partial}{\partial t}\,\psi(\underline{x}, t)\,d^3\underline{x} \quad . \tag{3.70'}$$

We have also seen, however, that there exists a function of $\underline{x}$ and $\underline{p}$, which in classical physics when evaluated along the path of the evolution of the system is always equal to the energy, namely the Hamiltonian $H = H(\underline{x}, \underline{p})$. We would therefore expect, since by the correspondence principle (section 2.5 vi) the expectation value of $\hat{H} = H(\hat{\underline{x}}, \hat{\underline{p}})$ must behave classically, that we must also have that the energy must be given by

$$E = \langle\psi|\hat{H}|\psi\rangle \quad , \tag{3.71}$$

or, for a system in $\underline{x}, t$ space

$$E = \int\psi^*(\underline{x}, t)\;H(\underline{x}, -i\hbar\frac{\partial}{\partial\underline{x}})\;\psi(\underline{x}, t)\,d^3\underline{x} \quad . \tag{3.71'}$$

If we now define a new operator $\hat{E}'$, given by

$$\hat{E}' = \hat{H} - i\hbar\frac{\partial}{\partial t}$$

we must have, from the preceding discussion, that

$$\langle\hat{E}'\rangle = \langle\psi|(\hat{H} - i\hbar\frac{\partial}{\partial t})|\psi\rangle = 0 \quad , \tag{3.72}$$

for any admissible $|\psi\rangle$. Since we also require that we should never detect a violation of the conservation of energy, we must also have that

$$(\Delta E')^2 = \langle\psi|(\hat{H} - i\hbar \frac{\partial}{\partial t})^2 |\psi\rangle = 0 \ , \tag{3.73}$$

or, since $\hat{E}'$ is Hermitian

$$(\Delta E')^2 = \langle(\hat{H} - i\hbar \frac{\partial}{\partial t})\psi|(\hat{H} - i\hbar \frac{\partial}{\partial t})\psi\rangle = 0 \quad .$$

This is only possible if $|\psi\rangle$ is an eigenstate of the $\hat{E}'$ operator, corresponding to an eigenvalue of zero, i.e.

$$\hat{E}'|\psi\rangle = (\hat{H} - i\hbar \frac{\partial}{\partial t})|\psi\rangle = 0$$

or

$$\boxed{H(\underline{\hat{x}}, \hat{p})|\psi\rangle = i\hbar \frac{\partial}{\partial t}|\psi\rangle \quad ,} \tag{3.74}$$

or, in $(\underline{x}, t)$ space,

$$\boxed{H(\underline{x}, -i\hbar \frac{\partial}{\partial \underline{x}})\ \psi(\underline{x}, t) = i\hbar \frac{\partial}{\partial t}\ \psi(\underline{x}, t) \quad .} \tag{3.74'}$$

Equation 3.74, or 3.74', is the celebrated Schrödinger equation and serves as the basis of the non-relativistic quantum theory.

Example 3.6. Find the Schrödinger equation for the one-dimensional motion of a particle in a constant potential.

From the corresponding Lagrangian we obtain the classical Hamiltonian

$$H = \frac{1}{2m} p_x^2 + V_0 \tag{3.75}$$

and the corresponding operator

$$\hat{H} = \frac{1}{2m} p_x^2 + V_0$$

$$= \frac{1}{2m} (-i\hbar \frac{\partial}{\partial x})(-i\hbar \frac{\partial}{\partial x}) + V_0$$

or

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V_0$$

and therefore

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V_0 = i\hbar \frac{\partial \psi}{\partial t} \tag{3.76}$$

We can solve the Schrödinger equation, by separation of variables, by substituting

$$\psi(\underline{x}, t) = \mathcal{X}(\underline{x})\ T(t) \tag{3.77}$$

to obtain, for each admissible E,

$$\hat{H}\,\mathcal{X}_E(\underline{x}) = E\,\mathcal{X}_E(\underline{x}) \tag{3.78}$$

and

$$i\hbar\,\frac{\partial}{\partial t}\,T_E(t) = E\,T_E(t) \quad . \tag{3.79}$$

We then have ,

$$T_E(t) = c_E\,e^{-iEt/\hbar} \tag{3.80}$$

and therefore

$$\psi(\underline{x}, t) = \mathop{\mathcal{S}}_{E}\; c_E\,\mathcal{X}_E(\underline{x})\,e^{-iEt/\hbar} \quad , \tag{3.81}$$

where the symbol $\mathcal{S}$ denotes summation if E is discrete and integration if E is continuous. It will be recognized that equation 3.81 is an expansion of the time dependence of $\psi(\underline{x}, t)$ in the eigenfunctions of the operator $\hat{E} = i\hbar\partial/\partial t$. The coefficients c_E can be computed from the initial conditions, namely since

$$\psi(\underline{x}, 0) = \sum_E c_E\,\mathcal{X}_E(\underline{x}) \tag{3.82}$$

and the $\mathcal{X}_E(\underline{x})$ are orthonormal, we have

$$c_E = \int_V \mathcal{X}^*_E(\underline{x})\;\psi(\underline{x}, 0)d^3\underline{x} = \text{constants} \tag{3.83}$$

We can ask for the expectation value of the energy in a system described by equation 3.81. We have

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

$$= \sum_{E', E} c^*_{E'}\,c_E \int_V \mathcal{X}^*_{E'}(\underline{x})\;\hat{H}\;\;\mathcal{X}_E(\underline{x})d^3\underline{x}\;e^{-i(E-E')t/\hbar}$$

or

$$\langle E\rangle = \sum_E E\,|c_E|^2 \quad . \tag{3.84}$$

Several conclusions can be drawn from this result. One is that the energy is a constant. This can be seen from the fact that the possible values of E are the (fixed) eigenvalues of $\hat{H}$ and the fact that the coefficients c_E are independent of time (i. e. equation 3.83). This is not surprising, since we obtained the Schrödinger equation by requiring that the expectation value of $\hat{H}$ be equal to a constant along the path of evolution of the system. Conversly, we will not be able to describe dissipation properly, i. e. situations in which the energy of a system of interest is not conserved.

update

Is there a general method of discovering the constants of the motion, in the sense that we found energy to be a constant of the motion? Consider an operator $\hat{Q}$ with an expectation value

$$\langle\hat{Q}\rangle = \langle\psi|\hat{Q}|\psi\rangle = \int_V \psi^*(\underline{x},t)\,\hat{Q}\,\psi(\underline{x},t)\,d^3\underline{x} \quad , \tag{3.85}$$

which will, in general, be a function of time. The requirement that $\hat{Q}$ be a constant of the motion, in the sense that we found energy to be a constant of the motion, can be written as

$$\frac{d}{dt}\langle\hat{Q}\rangle = 0 \quad . \tag{3.86}$$

We see that we have to compute the time derivative of the expectation value of $\hat{Q}$. From equation 3.85, we have

$$\frac{d}{dt}\langle\hat{Q}\rangle = \langle\frac{\partial\psi}{\partial t}|\hat{Q}|\psi\rangle + \langle\psi|\frac{\partial\hat{Q}}{\partial t}|\psi\rangle + \langle\psi|\hat{Q}|\frac{\partial\psi}{\partial t}\rangle \quad , \tag{3.87}$$

i.e. the expectation value may vary in time for two reasons, either because the wavefunction $\psi(\underline{x},t)$ is a function of time, or because $\hat{Q}$ is an explicit function of time. We can now use the Schrödinger equation to substitute for $\partial\psi/\partial t$, in particular since

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

we have,

$$\frac{d}{dt}\langle\hat{Q}\rangle = \frac{i}{\hbar}\langle\hat{H}\psi|\hat{Q}|\psi\rangle - \frac{i}{\hbar}\langle\psi|\hat{Q}|\hat{H}\psi\rangle + \langle\psi|\frac{\partial\hat{Q}}{\partial t}|\psi\rangle \quad ,$$

or using the fact that $\hat{H}$ is Hermitian[#], we have

$$\frac{d}{dt}\langle\psi|\hat{Q}|\psi\rangle = \frac{i}{\hbar}\langle\psi|(\hat{H}\hat{Q} - \hat{Q}\hat{H})|\psi\rangle + \langle\psi|\frac{\partial\hat{Q}}{\partial t}|\psi\rangle$$

We therefore <u>define</u> an operator

$$\frac{d\hat{Q}}{dt} = \frac{i}{\hbar}[\hat{H},\hat{Q}] + \frac{\partial\hat{Q}}{\partial t} \quad , \tag{3.88}$$

where $[\hat{H},\hat{Q}]$ is the commutator of the Hamiltonian and the operator $\hat{Q}$, i.e.

$$[\hat{H},\hat{Q}] \equiv \hat{H}\hat{Q} - \hat{Q}\hat{H} \quad ,$$

whose expectation value, computed for systems which obey the Schrödinger equation, is always equal to the (total) time derivative of the expectation value of $\hat{Q}$. We therefore have that if an operator does not involve time explicitly, i.e., if $\partial\hat{Q}/\partial t = 0$, then a necessary and sufficient condition that it[†] be conserved, is that it commutes with the Hamiltonian. The similarity between equation 3.88 and the Poisson brackets equation 1.63 in classical mechanics should not go unnoticed.

[#] see also equations 3.33 and 3.34.

[†] its expectation values.

Let us use this result now to compute the time rate of change of the expectation value of the momentum $\underline{p}$ of a particle moving in a potential $V(\underline{x})$. The corresponding Hamiltonian is given by

$$H = \frac{1}{2m} \underline{p} \cdot \underline{p} + V(\underline{x}) \quad . \tag{3.89}$$

From equation 3.88 we then have,

$$\frac{d}{dt} \langle \hat{\underline{p}} \rangle = \frac{i}{\hbar} \langle [H, \hat{\underline{p}}] \rangle \quad ,$$

where

$$[H, \hat{p}] = [\frac{1}{2m} \hat{\underline{p}} \cdot \hat{\underline{p}} + V(\underline{x}), \hat{\underline{p}}] = [V(\underline{x}), \hat{\underline{p}}] = i\hbar \frac{\partial}{\partial \underline{x}} V(\underline{x}) \quad ,$$

see problems 2.7 (iv) and 2.9, and therefore

$$\frac{d}{dt} \langle \hat{\underline{p}} \rangle = \langle - \frac{\partial V}{\partial \underline{x}} \rangle \quad , \tag{3.90}$$

as we would expect from Newton's law and the requirement from the correspondence principle that the expectation values behave classically.

From the form of the Schrödinger equation and the fact that the total probability of detecting the system somewhere is conserved, i.e.

$$\int_V |\psi|^2 \, d^3\underline{x} = 1 \quad ,$$

we can derive a local conservation law for the probability density. As in every case where the integral of a density is conserved, we seek a conservation law of the form

$$\frac{\partial}{\partial t} (\text{density}) + \frac{\partial}{\partial \underline{x}} \cdot (\text{flux}) = 0 \quad , \tag{3.91}$$

i.e. the local time derivative of the field density, plus the divergence of the corresponding flux must equal zero. Equation 3.91 is nothing more than a statement of the fact that if the field is conserved, the local loss from a volume element must be accountable in terms of the flux leaving the volume element. In classical fluid flow this leads to the mass continuity equation,

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \underline{x}} \cdot \rho \underline{v} = 0 \quad ,$$

where ρ is the fluid density (gms/cm^3) and $\underline{v}(\underline{x}, t)$ is the local velocity, in electromagnetism we have the charge conservation equation

$$\frac{\partial \rho_c}{\partial t} + \frac{\partial}{\partial \underline{x}} \cdot \underline{j}_c = 0 \quad ,$$

where $\rho_c(\underline{x}, t)$ is the local charge density (Cb/m^3), and $\underline{j}_c(\underline{x}, t)$ is the current density (amperes/m^2). In quantum mechanics, we would like to identify the local density with the probability density, i.e.

$$\rho(\underline{x}, t) = |\psi(\underline{x}, t)|^2, \tag{3.92}$$

and the question is what is the corresponding flux $\underline{j}(\underline{x}, t)$, that solves the equation

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \underline{x}} \cdot \underline{j}(\underline{x}, t) = 0 \quad . \tag{3.93}$$

We wish to compute

$$\frac{\partial}{\partial t} \rho(\underline{x}, t) = \frac{\partial}{\partial t} [\psi^*(\underline{x}, t) \ \psi(\underline{x}, t)]$$

$$= \frac{\partial \psi^*}{\partial t} \psi + \psi^* \frac{\partial \psi}{\partial t} \quad ,$$

or, using the Schrödinger equation,

$$\frac{\partial \rho}{\partial t} = \frac{i}{\hbar} [(H\psi)^* \psi - \psi^*(H\psi)]$$

substituting for H, see equation 3.89, we have

$$\frac{\partial \rho}{\partial t} = \frac{i}{2m\hbar} [\psi(p^2\psi)^* - \psi^*(p^2\psi)]$$

$$= - \frac{i\hbar}{2m} [\psi(\nabla^2\psi)^* - \psi^*(\nabla^2\psi)]$$

or

$$\frac{\partial \rho}{\partial t} - \frac{i\hbar}{2m} [\psi^*(\nabla^2\psi) - \psi(\nabla^2\psi^*)] = 0 \quad .$$

We can transform this to conservation form, using a vector identity[#] to obtain

$$\frac{\partial}{\partial t} (\psi^*\psi) + \frac{\partial}{\partial \underline{x}} \cdot \left[- \frac{i\hbar}{2m} \left(\psi^* \frac{\partial \psi}{\partial \underline{x}} - \psi \frac{\partial \psi^*}{\partial \underline{x}}\right)\right] = 0 \tag{3.94}$$

and therefore we have for the flux vector

$$\underline{j}(\underline{x}, t) = - \frac{i\hbar}{2m} \left[\psi^*(\underline{x}, t) \frac{\partial}{\partial \underline{x}} \psi(\underline{x}, t) - \psi(\underline{x}, t) \frac{\partial}{\partial \underline{x}} \psi^*(\underline{x}, t)\right] , \tag{3.95}$$

sometimes called the <u>probability current</u>. Note that the definition of $\underline{j}(\underline{x}, t)$, in terms of equation 3.94 <u>is not unique</u>. Any $\underline{j}'(\underline{x}, t)$, given by

$$\underline{j}'(\underline{x}, t) = \underline{j}(\underline{x}, t) + \frac{\partial}{\partial \underline{x}} \times \underline{A}(\underline{x}, t) \tag{3.96}$$

will also satisfy the local conservation equation 3.93, since the divergence of the curl of any vector field is zero.

[#] If φ is a scalar field and $\underline{v}$ is a vector field, we have

$$\varphi \frac{\partial}{\partial \underline{x}} \cdot \underline{v} = \frac{\partial}{\partial \underline{x}} \cdot \varphi \underline{v} - \underline{v} \cdot \frac{\partial \varphi}{\partial \underline{x}}$$

Problems

3.1 Prove equation 3.7.

3.2 Derive equation 3.14.

3.3 Derive equation 3.20 from 3.19.

3.4 Given any function $|\psi\rangle$ and the normalized eigenfunctions $\{|u_0\rangle, |u_1\rangle, \ldots, |u_n\rangle, \ldots\}$ of some Hermitian operator $\hat{Q}$, show that if $c_n \equiv \langle u_n|\psi\rangle$

(i) $$|\psi_1\rangle = |\psi\rangle - |u_0\rangle c_0$$

is orthogonal to $|u_0\rangle$

(ii) $$|\psi_N\rangle = |\psi\rangle - \sum_{n=0}^{N-1} |u_n\rangle c_n$$

is orthogonal to $\{|u_n\rangle\}$; $n=1, 2, .., N-1$.

3.5 If $\lambda_0 < \lambda_1 < \lambda_2 \ldots < \lambda_n < \ldots$ are the eigenvalues of a Hermitian operator $\hat{Q}$, show that

(i) if $\langle\psi_1|\psi_1\rangle = 1$ and $\langle u_0|\psi_1\rangle = 0$, then

$$\min\{\langle\psi_1|\hat{Q}|\psi_1\rangle\} = \lambda_1$$

(ii) if $\langle\psi_N|\psi_N\rangle = 1$ and $\langle u_n|\psi_N\rangle = 0$ for $n=0, 1, \ldots, N-1$, then

$$\min\{\langle\psi_N|\hat{Q}|\psi_N\rangle\} = \lambda_N \; .$$

3.6 Show that if $|\psi\rangle = \sum_n |u_n\rangle c_n$

$$\langle\hat{Q}\rangle = \sum_n \lambda_n |c_n|^2 = \lambda_m$$

and

$$\Delta Q^2 = \langle(\hat{Q} - \langle\hat{Q}\rangle)^2\rangle$$

$$= \sum_n (\lambda_n - \langle\hat{Q}\rangle)^2 |c_n|^2 = 0$$

Then

$$|\psi\rangle = |u_m\rangle$$

Hint: $(\lambda_n - \lambda_m)^2 \geq 0$, $|c_n^2 \geq 0$.

3.7 Show that for a system obeying the Schrödinger equation,

$$\frac{d}{dt}\langle\underline{x}\rangle = \frac{1}{m}\langle\underline{p}\rangle \; .$$

3.8 Define a local phase field $\varphi(\underline{x}, t)$ by $[\rho(\underline{x}, t), \varphi(\underline{x}, t) = \text{real}]$

$$\psi(\underline{x}, t) = [\rho(\underline{x}, t)]^{1/2} e^{i\varphi(\underline{x}, t)} ,$$

where $\psi(\underline{x}, t)$ is a solution to the Schrödinger equation. Compute the probability current $\underline{j}(\underline{x}, t)$. Can you identify a velocity field from this? Compare with the classical fluid mass continuity equation.

3.9 (i) Using the Schrödinger equation, show that

$$\psi(\underline{x}, t+dt) \sim (1 - \frac{i}{\hbar} dt\, H)\, \psi(\underline{x}, t) ,$$

where H is the Hamiltonian for the system

(ii) If H is not a function of time, show that, for a finite time shift τ,

$$\psi(\underline{x}, t+\tau) = e^{-\frac{i}{\hbar}\tau H} \psi(\underline{x}, t) .$$

This allows us to define a time translation operator,

$$\hat{T}_t(\tau) \equiv e^{-\frac{i}{\hbar}\tau H} ,$$

such that

$$\hat{T}_t(\tau)\, \psi(\underline{x}, t) = \psi(\underline{x}, t+\tau) .$$

3.10 (i) Using the definition of the momentum operator, show that

$$\psi(\underline{x}+d\underline{x}) \sim (1 + \frac{i}{\hbar} d\underline{x}\cdot\hat{\underline{p}})\, \psi(\underline{x})$$

(ii) Show that for a finite translation in $\underline{x}$

$$\psi(\underline{x}+\underline{a}) = e^{\frac{i}{\hbar}\underline{a}\cdot\hat{\underline{p}}} \psi(\underline{x}) .$$

This allows us to define a space translation operator

$$\hat{T}_{\underline{x}}(\underline{a}) \equiv e^{\frac{i}{\hbar}\underline{a}\cdot\hat{\underline{p}}} ,$$

such that

$$\hat{T}_{\underline{x}}(\underline{a})\, \psi(\underline{x}) = \psi(\underline{x}+\underline{a}) .$$

4. QUANTUM BEHAVIOR IN ONE-DIMENSIONAL POTENTIALS.

In this section we will deal with one-particle systems described by a Hamiltonian

$$\hat{H} = \frac{1}{2m} p^2 + V(\underline{x}) \quad , \qquad (4.1)$$

where the potential $V(\underline{x})$ can be separated into the sum of potentials along each coordinate component, e.g.

$$V(\underline{x}) = V_x(x) + V_y(y) + V_z(z) \quad . \qquad (4.2)$$

In such cases the Hamiltonian itself is separable, i.e.

$$\hat{H} = \frac{1}{2m} p^2 + V(\underline{x}) = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + V_x(x) + V_y(y) + V_z(z) \quad ,$$

or

$$\hat{H} = \hat{H}_x + \hat{H}_y + \hat{H}_z \quad , \qquad (4.3a)$$

where

$$\hat{H}_i(x_i, p_i) = \frac{1}{2m} p_i^2 + V_i(x_i) \quad , \quad i=1,2,3 \qquad (4.3b)$$

This class of problems, provided the boundary conditions can also be separated along the corresponding coordinate components, can be solved by a further separation of the space variables[#].

Consider the Hamiltonian eigenvalue equation[†] (see equation 3.78)

$$\hat{H}\, \mathcal{X}_E(\underline{x}) = E\, \mathcal{X}(\underline{x}) \quad .$$

Now, if the Hamiltonian can be separated as a sum of the type of equation 4.3, we can consider solutions of the form

$$\mathcal{X}(\underline{x}) = X(x)\; Y(z)\; Z(z) \quad , \qquad (4.4)$$

where

$$\hat{H}_x\; X(x) = E_x X(x) \qquad (4.5a)$$

$$\hat{H}_y\; Y(y) = E_y Y(y) \qquad (4.5b)$$

$$\hat{H}_z\; Z(z) = E_z Z(z) \qquad (4.5c)$$

and where

$$E_x + E_y + E_z = E \quad , \qquad (4.5d)$$

the eigenvalue of the total Hamiltonian. See example 3.3.

In systems of this type, the behavior along each component of the coordinate is

---

[#] In addition to the separation of space - $\underline{x}$ and time - t, see equation 3.77 and related discussion.

[†] Sometimes also called the time-independent Schrödinger equation.

independent of the others, and the wavefunctions are products of functions of the separate coordinates. It should be noted that the essential ingredient for this behavior is the separability of the Hamiltonian, i.e. equation 4.3.

These results can be generalized in a straight-forward manner to cover systems of non-interacting subsystems (e.g. particles), i.e.

$$\hat{H}(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N, \underline{p}_1, \underline{p}_2, \ldots, \underline{p}_N) = \sum_{\alpha} \hat{H}^{(\alpha)}(\underline{x}_\alpha, \underline{p}_\alpha) \quad , \tag{4.6a}$$

where

$$\hat{H}^{(\alpha)}(\underline{x}_\alpha, \underline{p}_\alpha) = \frac{1}{2m_\alpha} p_\alpha^2 + V_\alpha(\underline{x}_\alpha) \quad . \tag{4.6b}$$

The total wave function for the system can then be written as a product

$$\mathcal{X}(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N) = \prod_{\alpha} \mathcal{X}_\alpha(\underline{x}_\alpha) \quad , \tag{4.7}$$

where

$$\hat{H}^{(\alpha)} \mathcal{X}_\alpha(\underline{x}_\alpha) = E_\alpha \mathcal{X}_\alpha(\underline{x}_\alpha) \tag{4.8a}$$

where

$$E_{\text{Total}} = \sum_{\alpha} E_\alpha \quad . \tag{4.8b}$$

Note again that the condition for this to be possible is that the Hamiltonian be separable into a sum of separate Hamiltonians for each subsystem (particle), which involve only the corresponding coordinate and momentum variables[#]. Note also the resulting total probability density

$$\mathcal{P}(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N) = |\mathcal{X}(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N)|^2 \tag{4.9}$$

is then expressible as a product

$$\mathcal{P}(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N) = \prod_{\alpha} \mathcal{P}_\alpha(\underline{x}_\alpha) \quad , \tag{4.10}$$

of probability densities, consistently with our notion of independence in classical probability theory.

As a result of these considerations, we are motivated to study the one-dimensional problem, for which the Hamiltonian is given by

$$\hat{H}(x, p) = \frac{1}{2m} p_x^2 + V(x) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \quad , \tag{4.11}$$

realizing that it is more broadly applicable than for motion of a single particle along one dimension.

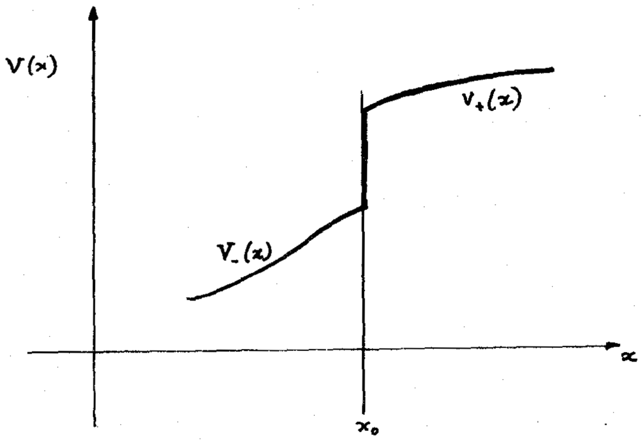
[#] actually, in this case, it is sufficient that the potential have this property, since the kinetic energy is already separable.

4.1 Matching conditions.

In many cases of interest, it is useful to divide the range of the coordinate into two or more regions that separate different analytical expressions for the potential. By way of example, let

$$V(x) = \begin{cases} V_-(x) & \text{for} \quad x < x_0 \\ V_+(x) & \text{for} \quad x > x_0 \end{cases} \tag{4.12}$$

where $V_-(x_0)$ need not be equal to $V_+(x_0)$, e.g.



We can then consider dividing the Schrödinger equation

$$[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)]\ u(x) = E\ u(x) \tag{4.13}$$

into two regions, i.e.

$$[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V_-(x)]\ u_-(x) = E\ u_-(x)\ , \qquad \text{for} \quad x < x_0 \tag{4.14a}$$

and

$$[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V_+(x)]\ u_+(x) = E\ u_+(x)\ , \qquad \text{for} \quad x > x_0 \tag{4.14b}$$

where

$$u(x) = \begin{cases} u_-(x) & \text{for} \quad x < x_0 \\ u_+(x) & \text{for} \quad x > x_0 \ . \end{cases} \tag{4.15}$$

The question is what happens to the wavefunction u(x) <u>at</u> $x = x_0$. From the Schrödinger equation 4.13, we have that

$$\frac{d^2}{dx^2}\,u(x) = -\frac{2m}{\hbar^2}\,[E - V(x)]\ u(x)\quad , \tag{4.15}$$

which we can integrate to obtain

$$\frac{du}{dx} = \int^{x} \frac{d^2u(x')}{dx'^2}\, dx' = -\frac{2m}{\hbar^2} \int^{x} [E - V(x')]u(x')\, dx' + c_1$$

and therefore the difference just after x_0 and just before is equal to

$$\left(\frac{du}{dx}\right)_{x_0+\epsilon} - \left(\frac{du}{dx}\right)_{x_0+\epsilon} = -\frac{2m}{\hbar^2} \int_{x_0-\epsilon}^{x_0+\epsilon} [E - V(x')]\, u(x')\, dx' \quad . \tag{4.16}$$

Now, provided $V(x)$ is finite and has no singularity at $x = x_0$ (it may have a discontinuity), we have

$$\lim_{\epsilon \to 0} \left\{\frac{2m}{\hbar^2} \int_{x_0-\epsilon}^{x_0+\epsilon} [E - V(x')]\, u(x')\, dx'\right\} = 0 \quad ,$$

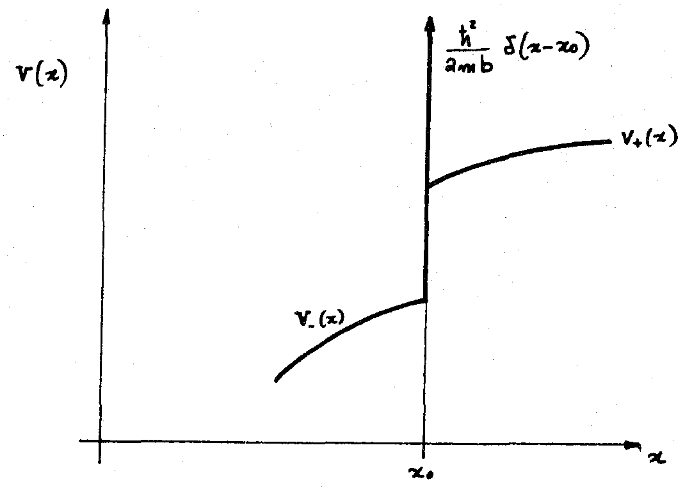
or

$$\left(\frac{du}{dx}\right)_{x_0^+} - \left(\frac{du}{dx}\right)_{x_0^-} = 0 \quad , \tag{4.17}$$

or, <u>the derivative of the wavefunction is continuous across a non-singular discontinuity of V(x)</u>.

What if $V(x)$ had a delta function singularity at x_0, i.e.

$$V(x) = \begin{cases} V_-(x) & \text{for} \quad x < x_0 \\ \dfrac{\hbar^2}{2mb}\, \delta(x - x_0) & \\ V_+(x) & \text{for} \quad x > x_0 \quad ? \end{cases} \tag{4.18}$$



Substituting 4.18 into 4.16 we have

$$\left(\frac{du}{dx}\right)_{x_0+\epsilon} - \left(\frac{du}{dx}\right)_{x_0-\epsilon} = \frac{1}{b} \int_{x_0-\epsilon}^{x_0+\epsilon} \delta(x' - x_0)\, u(x')\, dx'$$

$$- \frac{2m}{\hbar^2} \int_{x_0-\epsilon}^{x_0} [E - V_-(x')]\, u(x')\, dx' - \frac{2m}{\hbar^2} \int_{x_0}^{x_0+\epsilon} [E - V_+(x)]\, u(x')\, dx' \quad ,$$

or, for a delta function at $x = x_0$,

$$\left(\frac{du}{dx}\right)_{x_0^+} - \left(\frac{du}{dx}\right)_{x_0^-} = \frac{1}{b}\, u(x_0) \quad . \tag{4. 19}$$

Integrating 4. 19 we also have

$$u(x) = \int^{x} \frac{du(x')}{dx'}\, dx' + c_2$$

and therefore

$$u(x_0^{+}) - u(x_0^{-}) = 0$$

even if $V(x)$ has a delta function singularity at $x = x_0$.

4. 2 Motion in a constant potential. Free particle behavior.

The simplest problem of this type occurs if the potential is a constant, i. e.

$$V(x) = V_0 = \text{const.} \tag{4. 20}$$

We then have

$$\hat{H} = \frac{1}{2m}\, p_x^2 + V_0 = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V_0 \tag{4. 21}$$

and the Schrödinger equation becomes, see example 3. 6,

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t) + V_0 \psi(x, t) = i\hbar \frac{\partial \psi}{\partial t} \tag{4. 22}$$

separating x and t, we have

$$\psi(x, t) = u(x)\; T(t) \tag{4. 23}$$

where

$$-\frac{\hbar^2}{2m}\, u''(x) + V_0\, u(x) = E\, u(x) \tag{4. 24}$$

and

$$T(t) = T(0)\, e^{-iEt/\hbar} \quad . \tag{4. 25}$$

From equation 4. 24 we then have

$$u''(x) + \frac{2m}{\hbar^2}\,(E - V_0)\, u(x) = 0 \quad , \tag{4. 26}$$

whose solution is given by

$$u(x) = a_+ e^{ikx} + a_- e^{-ikx} \tag{4. 27a}$$

where

$$k = \sqrt{\frac{2m}{\hbar^2}\,(E - V_0)} \quad , \tag{4. 27b}$$

if $E > V_0$, and

$$u(x) = b_+ e^{\kappa x} + b_- e^{-\kappa x} \tag{4. 28a}$$

where

$$\kappa = \sqrt{\frac{2m}{\hbar^2}(V_0 - E)} \ , \tag{4.28b}$$

if $E < V_0$[#].

We see that if the energy is greater than the potential V_0, we have wave-like solutions, i.e.

$$\psi(x,t) = T(0)\,[a_+ e^{i(k_x - Et/\hbar)} + a_- e^{-i(k_x + Et/\hbar)}] \tag{4.29}$$

(superposition of right-going and left-going waves), whereas if $E < V_0$, we have exponentially growing or decaying amplitudes. The coefficients $a_\pm$ and $b_\pm$ would be determined by the boundary/initial conditions. Note that classically no solutions would exist for $E < V_0$.

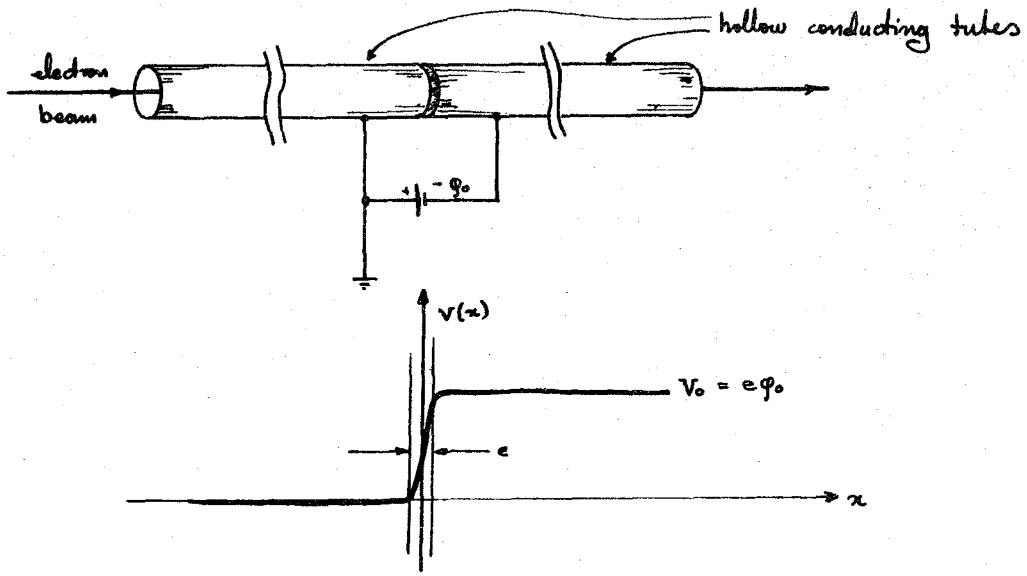
Solutions of different energies can be superimposed to solve the time dependent problems in each case. See equation 3.81 and problem 4.1.

4.2.1 Scattering by a potential step.

Consider a source of monoenergetic particles at $x = -\infty$, of strength J_0 (particles/cm^2)/sec, travelling from left to right on a step change in the potential $V(x)$, given by

$$V(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ V_0 & \text{for} \quad x > 0 \ . \end{cases} \tag{4.29}$$

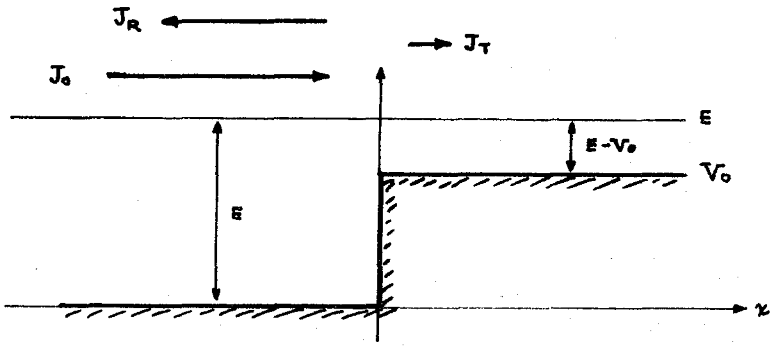
Such a situation might arise, for example, in an apparatus sketched below.



[#] The solutions of equation 4.20 are of course included in equation 4.19 if the radical is interpreted properly.

In the real experiment, the transition from $V = 0$ to $V = V_0$ will, of course, take place in a finite distance $\Delta x \sim \epsilon$. The approximation of a step change, however, we would expect to be valid as $\epsilon k_0 \to 0$, where k_0 is the wave number of the incident particles. See equation 4.31 below.

In the idealized situation, we then have a step change at $x = 0$ as depicted below.



The step change, we might expect, may reflect some particles back and transmit some particles forward. From the figure above and equation 4.27 we would expect that $(E > V_0)$

$$u(x) = \begin{cases} \left(c_0\, e^{ik_0 x} + R\, e^{-ik_0 x}\right) , & x < 0 \\ c_0\, T e^{ikx} , & x > 0 \end{cases} \tag{4.30}$$

where R and T are the complex reflection and transmission amplitudes, and k_0 and k are given by

$$k_0 = \sqrt{\frac{2mE}{\hbar}}^{\;1/2} , \qquad k = \sqrt{\frac{2m(E - V_0}{\hbar}} . \tag{4.31}$$

The coefficient c_0 can be computed by matching to the flux of the source. From the solution for $x < 0$, we have (see equation 3.95)

$$j(x) = -\frac{i\hbar}{2m}\left(u^* \frac{du}{dx} - u\,\frac{du^*}{dx}\right)$$

$$= -\frac{i\hbar}{2m}\left[|c_0|^2 \left(e^{-ik_0 x} + R^* e^{ik_0 x}\right)\left(ik_0\, e^{ik_0 x} - ik_0\, R\, e^{-ik_0 x}\right) - cc^{\#}\right]$$

therefore

$$j(x < 0) = \frac{\hbar k_0}{m}\, |c_0|^2 \,(1 - |R|^2) , \tag{4.32}$$

which must be equal to the flux of the source, minus the reflected flux, i.e.

$$j(x < 0) = J_0 - J_0|R|^2 = J_0(1 - |R|^2) .$$

[#] cc = complex conjugate.

Therefore

$$\frac{\hbar k_0}{m} \, |c_0|^2 = J_0$$

or

$$|c_0|^2 = \frac{m J_0}{\hbar k} \quad . \tag{4.33}$$

It should be noted that we have made an implicit assumption in assigning a single wavefunction to the total flux of the source, namely that we are dealing with a coherent source of particles. If the phases between the particles are uncorrelated, we should use instead a wavefunction for one particle at a time, whose extent (coherence) in x would be a function of the monochromaticity of the source[#], and at the end superimpose the probabilities for each outcome, as opposed to the amplitudes as we have done in writing down the wavefunction of equation 4.30. An analogous situation occurs in dealing with polarized light (coherent superposition of amplitudes), versus unpolarized light (incoherent superposition of amplitudes, superposition of probabilities). In this example, however, the assumption of a monoenergetic source, essentially makes the two cases result in the same outcome. Can you show this?

Going back to the problem, we have two unknowns, R and T, which we can determine from the matching conditions at $x = 0$, namely,

$$u(0^+) = u(0^-) \Rightarrow T = 1 + R \tag{4.34a}$$

$$u'(0^+) = u'(0^-) \Rightarrow ikT = ik_0(1 - R) \quad . \tag{4.34b}$$

Dividing the two equations, we have

$$\frac{1 - R}{1 + R} = \frac{k}{k_0}$$

or

$$R = \frac{1 - k/k_0}{1 + k/k_0} \tag{4.35a}$$

and

$$T = \frac{2}{1 + k/k_0} \quad . \tag{4.35b}$$

From these we can also compute the reflected and transmitted fluxes

$$J_R = |R|^2 J_0 = \left(\frac{1 - k/k_0}{1 + k/k_0}\right)^2 \tag{4.36a}$$

and

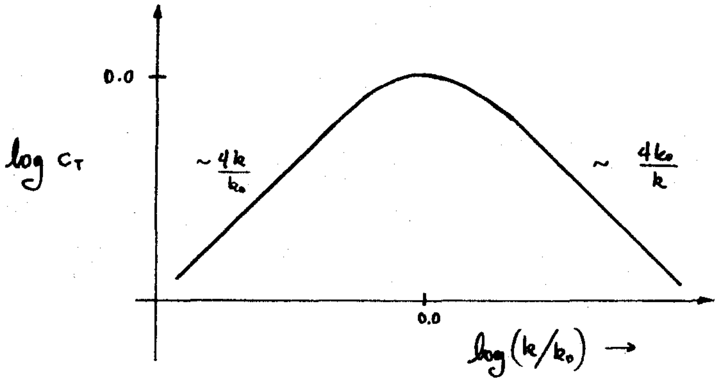
$$J_T = \frac{k}{k_0} \, |T|^2 \, J_0 = \frac{(4k k_0)}{(1 + k_0)^2} \, J_0 \quad . \tag{4.36b}$$

We can define the transmission coefficient, as

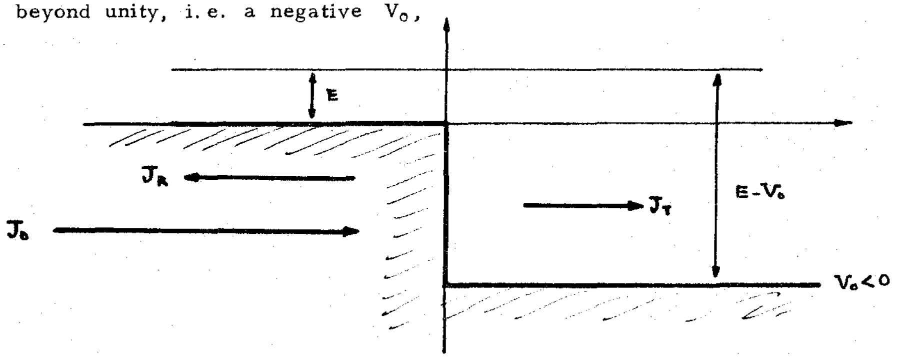
[#] i.e. a wavepacket for which $\Delta x \Delta k \sim 1$.

$$c_T = \frac{J_T}{J_0} = \frac{4(k/k_0)}{(1 + k/k_0)^2} \quad . \tag{4.37}$$

This is sketched below



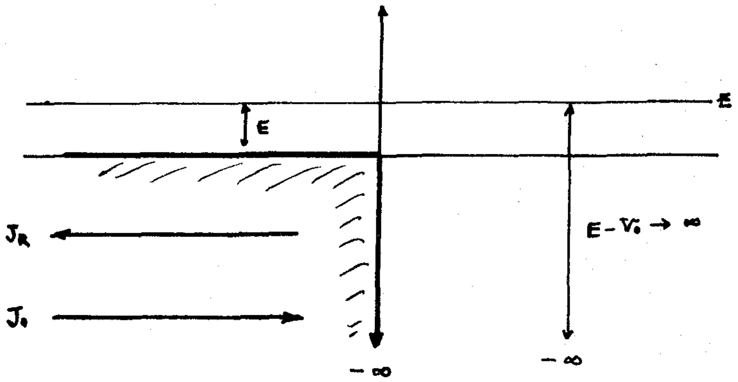
Note that classically there can be no reflected particles if $E > V_0$. The step change represents an impulse which will reflect the particles or not depending on its strength. The simultaneous possibility of reflection and transmission is a purely quantum mechanical result. Note also that for $k/k_0 = 1$ (no step) the transmitted flux is equal to the incident flux (no reflection), as one might expect. What one might not expect is that as k/k_0 increases beyond unity, i.e. a negative V_0,



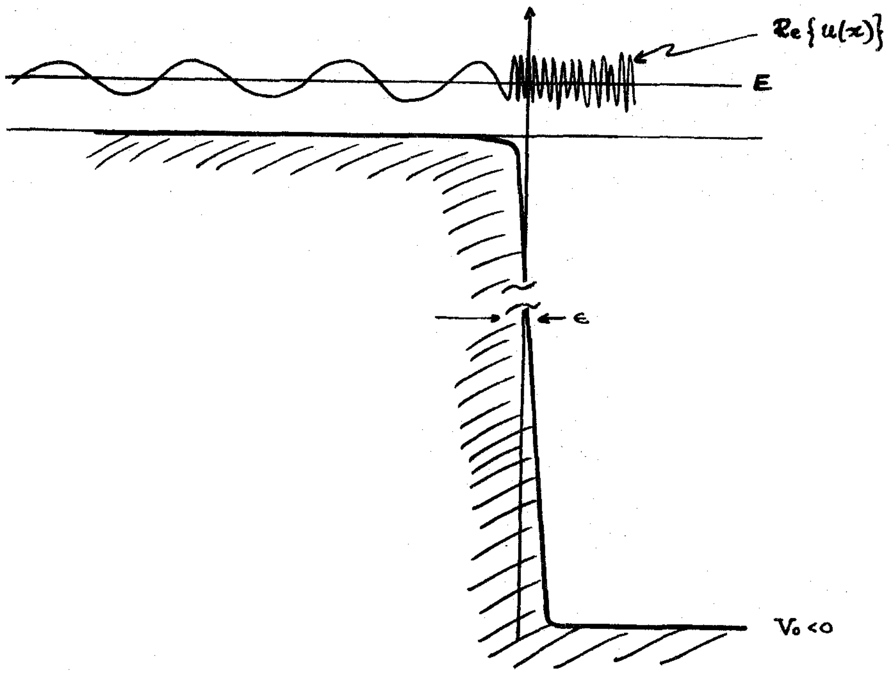
the transmission coefficient begins to fall again and for $k/k_0 \gg 1$ tends to

$$c_T \sim \frac{4}{k/k_0} \quad ,$$

which goes to zero, if the potential has an infinite drop.



It should be noted that these results depend, to some extent, on the assumption that the potential change occurs over a range ϵ in x which is small compared to the wavelength of the wavefunction (i. e. $\epsilon k << 1$). While this assumption might be valid for small potential step changes and low energies, it most likely will not be adequate in the limiting case of an infinite potential drop, since the wavenumber for $x > 0$ tends to infinity. To treat that problem more accurately one has to solve the Schrödinger equation in detail in the vicinity of the transition in the potential.



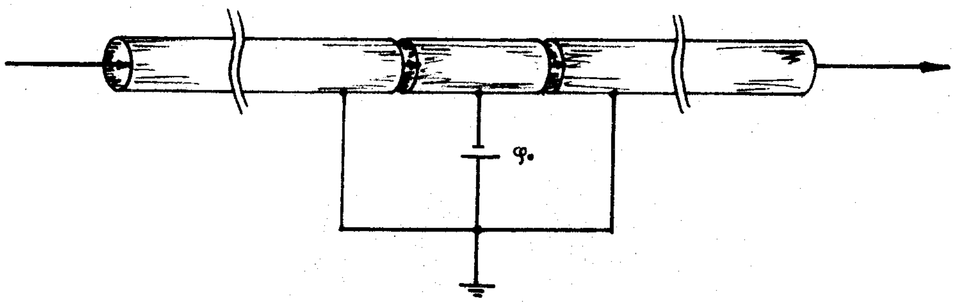
The solution to such a problem might best be handled numerically.

4. 2. 2 Scattering by a potential step of finite-length.

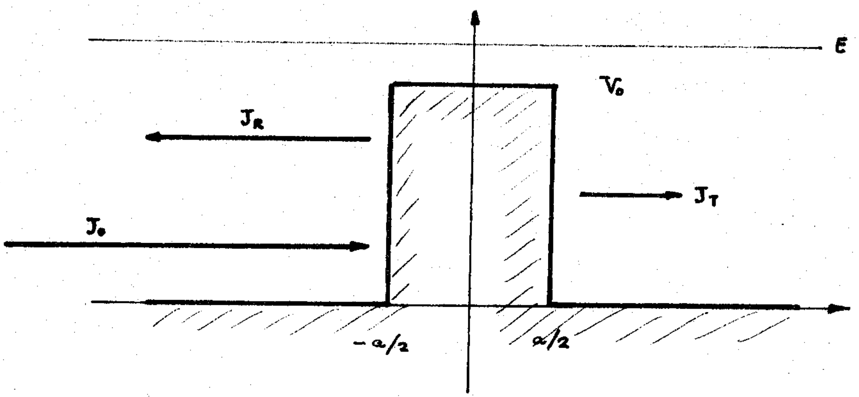
Consider the same source of monoenergetic particles incident a potential step of finite length, centered at $x = 0$, i. e.

$$V(x) = \begin{cases} 0 & \text{for} \quad |x| > a/2 \\ V_0 & \text{for} \quad |x| < a/2 \end{cases} \tag{4.38}$$

Such a situation might arise, in a similar apparatus to the one of section 4.2.1, as follows,



which we approximate by the rectangular potential of equation 4.38



We then have, for $E > V_0$,

$$u(x)^{\#} = \begin{cases} e^{ik_0x} + R\,e^{-ikx} & ; \quad x < -a/2 \\ A\,e^{ikx} + B\,e^{-ikx} & ; \quad -a/2 < x < a/2 \\ T\,e^{ik_0x} & ; \quad x > a/2 \;. \end{cases} \tag{4.39}$$

We now apply the matching conditions at $x = \pm\, a/2$ to obtain:

[#] normalized to unit incoming beam amplitude.

$$u\left(-\frac{a^-}{2}\right) = u\left(-\frac{a^+}{2}\right) \quad : \quad e^{-ik_0 a/2} + R\, e^{ik_0 a/2} = A\, e^{-ika/2} + B\, e^{ika/2} \tag{4.40a}$$

$$u\left(\frac{a^-}{2}\right) = u\left(\frac{a^+}{2}\right) \quad : \quad A\, e^{ika/2} + B\, e^{-ika/2} = T\, e^{ik_0 a/2} \tag{4.40b}$$

$$u'\left(-\frac{a^-}{2}\right) = u'\left(-\frac{a^+}{2}\right) : \quad k_0\left(e^{-ik_0 a/2} - R\, e^{ik_0 a/2}\right) = k\left(A\, e^{-ika/2} - B\, e^{ika/2}\right) \tag{4.40c}$$

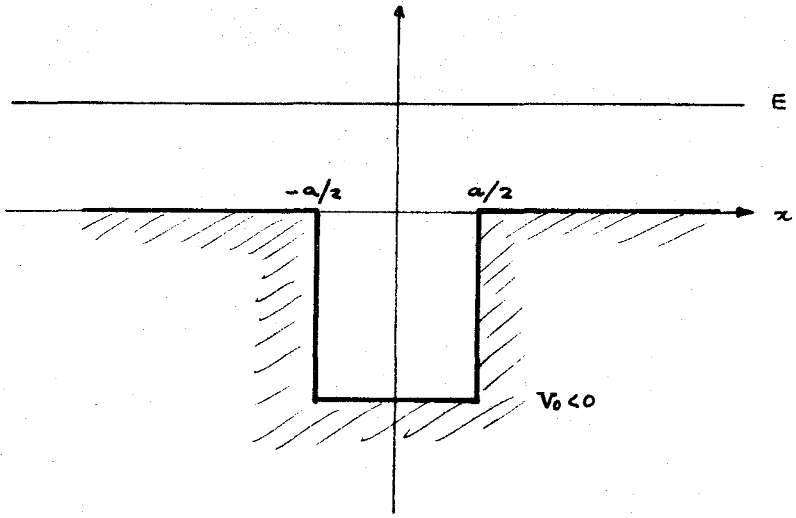
$$u'\left(\frac{a^-}{2}\right) = u'\left(\frac{a^+}{2}\right) \quad : \quad k\left(A\, e^{ika/2} - B\, e^{-ika/2}\right) = k_0\, T\, e^{ik_0 a/2} \quad . \tag{4.40d}$$

This is a system of four equations and four unknowns, which after a little algebra yields the reflection and transmission amplitude

$$R = \frac{-\left(\frac{k_0^2 - k^2}{2k\,k_0}\right) \sin ka \; e^{-ik_0 a}}{\cos ka - i\left(\frac{k_0^2 + k^2}{2k\,k_0}\right) \sin ka} \tag{4.41a}$$

$$T = \frac{e^{-ik_0 a}}{\cos ka - i\left(\frac{k_0^2 - k^2}{2k\,k_0}\right) \sin ka} \tag{4.41b}$$

It is noteworthy to observe that if $ka = n\pi$, the reflection amplitude is zero, and the beam experiences 100% transmission. This is a wave mechanical resonance phenomenon that occurs as a result of constructive interference between the transmitted and reflected waves at $x = \pm\, a/2$. It should also be pointed out that this solution is also valid if $E > 0 > V_0$, i. e. a potential well.



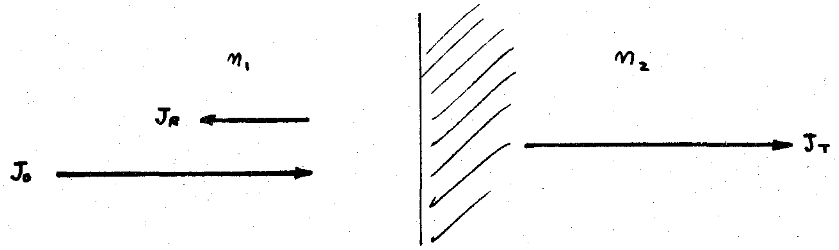
The results in sections 4.2.1 and 4.2.2 may be familiar in a different context. They represent wave behavior and have exact counterparts in optics and acoustics, for example. In section 4.2.1, the essential feature that results in the simultaneous reflection and transmission of the wave at the step is the abrupt change in the wavelength (wavenumber^{-1}), i.e. if $E > V_0$

$$\lambda(x < 0) \propto \frac{1}{\sqrt{E}} , \qquad \lambda(x > 0) \propto \frac{1}{\sqrt{E - V_0}} .$$

This also occurs when light (or sound) is incident from a medium of one index of refraction to one of another. The results are the same. From optics, for example, we have that the reflection coefficient for normal incidence is given by

$$c_R = |R|^2 = \left(\frac{1 - n_1/n_2}{1 + n_1/n_2}\right)^2 ,$$

where n_1 and n_2 are the indeces in the two media.

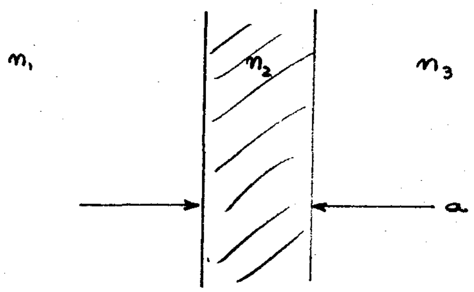


The ratio of the indeces of refraction is of course the reciprocal of the ratio of the wavenumbers and we recover the result of equation 4.35a. In optics, the approximation that the range of the transition in the properties of the medium ("potential"), times the wavenumber is much less than unity is a very good one, because the wavelength of visible light is of the order of several thousand Angstroms (4,000 < λ < 7,000 Å for visible light; 1Å = 10^{-8} cm), whereas the transition length for the index of refraction is of the order of the interatomic spacing in a solid (or liquid), which is of the order of a few Angstroms (range of interatomic potentials). Therefore

$$\epsilon k \sim \frac{\epsilon}{\lambda} \sim 10^{-3} ,$$

typically for visible light. Conversly, this is why it is not possible to construct a (conventional) reflector for x-rays ($\epsilon/\lambda \sim 1$).

The optical analogy to the potential step of finite length is the <u>etalon</u>. This usually consists of piece of glass of a certain thickness with plane parallel opposite faces (see refs. 4.1 and 4.2).



One then exploits the sharp transmission resonances at

$$ka = n\pi$$

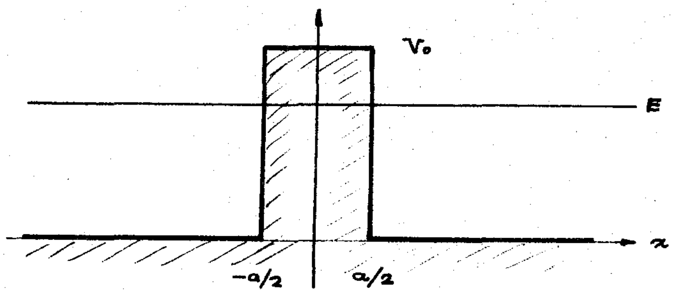
or

$$a = \frac{1}{2} n\lambda \quad ,$$

i.e. integral number of half wavelengths, for wavelength selection purposes. This device is commonly used to select a particular longitudinal mode in a laser cavity to force the laser to operate single frequency.

4.2.3 Tunneling.

The expressions for the reflection and transmission amplitudes for the finite potential step, i.e. equation 4.41 assume that $E - V(x) \geq$ everywhere. We also have the interesting situation, however, of $0 < E < V_o$, i.e.



Classically, of course, this represents a potential barrier, which will necessarily result in 100% reflection of the incoming beam. Quantum mechanically, however, it turns out that we again have both reflection and transmission. The wavefunctions in the three regions now become

$$u(x) = \begin{cases} e^{ik_o x} + R\,e^{-ik_o x} & ; \quad x < -a/2 \\ A\,e^{-\kappa x} + B\,e^{\kappa x} & ; \quad -a/2 < x < a/2 \\ T\,e^{ik_o x} & ; \quad a/2 < x \quad , \end{cases} \tag{4.42}$$

where, in this case

$$\kappa = \sqrt{\frac{2m}{\hbar^2}(V_0 - E)} \quad . \tag{4.43}$$

We may solve this system, the same way as before. If we observe, however, that the substitution

$$k = i\kappa \tag{4.44}$$

in equation 4.39 yields the new equations 4.42, we can obtain the results directly. In particular, the transmission amplitude becomes

$$T = \frac{e^{-ik_0 a}}{\cosh \kappa a - i\left(\frac{k_0^2 - \kappa^2}{2k_0\kappa}\right) \sinh \kappa a} \quad , \tag{4.45}$$

yielding a transmission coefficient,

$$c_T = |T|^2 = \left[1 + \left(\frac{k_0^2 + \kappa^2}{2k_0\kappa}\right)^2 \sinh^2 \kappa a\right]^{-1} \quad . \tag{4.46}$$

Now note that since

$$\sinh x \equiv \frac{e^x - e^{-x}}{2}$$

we have, for large κa

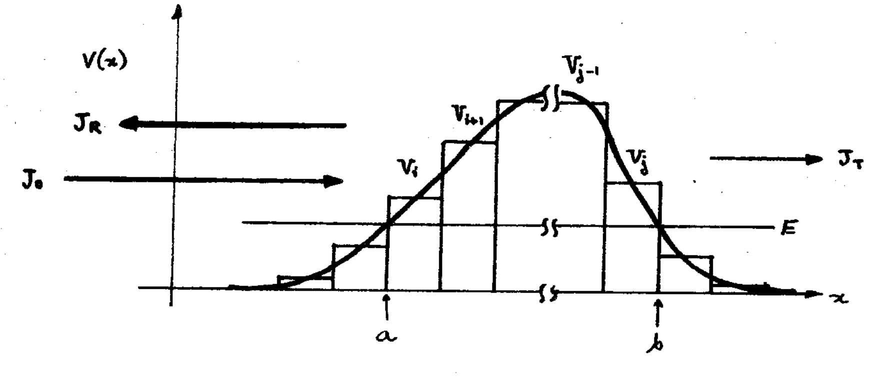
$$\sinh \kappa a \simeq \frac{1}{2} e^{\kappa a}$$

and therefore

$$c_T = |T|^2 \simeq \left(\frac{4k_0\kappa}{k_0^2 + \kappa^2}\right)^2 e^{-2\kappa a} \quad . \tag{4.47}$$

This is perhaps an unexpected result in that even though the incident particles do not have the energy to go over the potential hill classically, quantum mechanically they may make it to the other side, with some probability that decreases very rapidly, however, as the energy deficit increases. This phenomenon is called tunneling and is in fact the way particles are observed to behave.

In any practical situation, of course, potential barriers will not be square. It is possible, however, to approximate a smooth potential hill V(x) by a series of steps, i.e.



and consider the total transmission as the product of the individual transmissions through each of the consecutive barriers, i.e.

$$|T|^2 \simeq |T_i|^2 \cdot |T_{i+1}|^2 \ldots |T_{j-1}|^2 \cdot |T_j|^2 ,$$

where V_i is the first step for which $V(x) > E$ and V_j is the last one. If one ignores the much more slowly varying coefficient of the exponential in equation 4.47, we have

$$|T|^2 \simeq e^{-2\kappa_i\delta} \cdot e^{-2\kappa_{i+1}\delta} \ldots\ldots e^{-2\kappa_{j-1}\delta} \, e^{-2\kappa_j\delta}$$

where

$$\kappa_i = \sqrt{\frac{2m}{\hbar^2}(V_i - E)} ,$$

or

$$|T|^2 \simeq e^{-2\int_a^b \kappa(x)dx} \tag{4.48}$$

where $\kappa(x)$ is now a continuous variable, given by

$$\kappa(x) = \sqrt{\frac{2m}{\hbar^2}[V(x) - E]} . \tag{4.49}$$

This result is a good approximation and can be justified by a more rigorous treatment. It is valid when the potential changes smoothly and slowly, in particular if

$$\frac{1}{\kappa(x)} \frac{d\kappa(x)}{dx} << \kappa(x) . \tag{4.50}$$

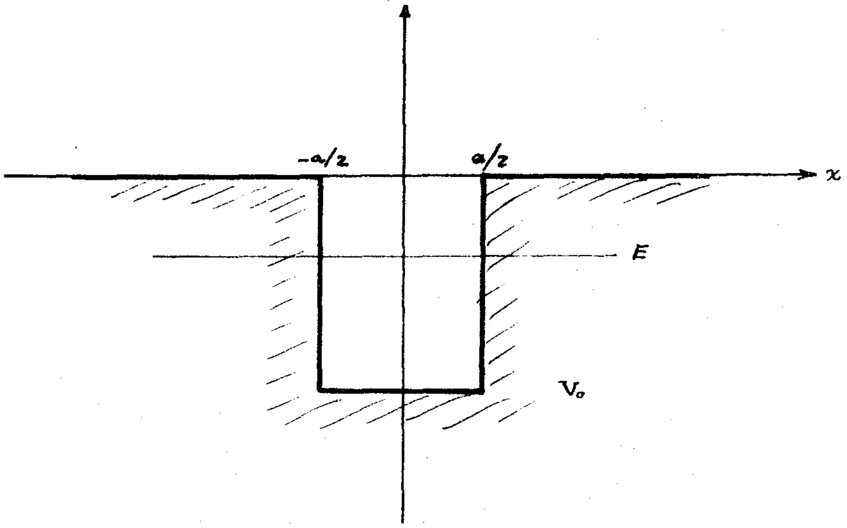
Even though this is not quite satisfied near the points $x = a$ and $x = b$, where $E \simeq V(x)$ and $\kappa(x)$ is small, the bulk of the contribution to the integral of equation 4.48 comes from the central portion and we may neglect the errors in those two regions.

4.2.4 Bound states of a rectangular potential well.

Consider now a potential

$$V(x) = \begin{cases} V_0 < 0 & \text{for } |x| < a/2 \\ 0 & \text{for } |x| > a/2 , \end{cases} \tag{4.51}$$

and a particle with an energy $V_0 < E < 0$.



Since $E - V(x) < 0$ for $x < -a/2$ and $x > a/2$, we will have real exponential solutions in those two regions, and oscillatory solutions in the region $-a/2 < x < a/2$, i.e.

$$u(x) = \begin{cases} A e^{+\kappa x} & ; \quad x < -a/2 \\ B e^{ikx} + C e^{-ikx} & ; \quad |x| < a/2 \\ D e^{-\kappa x} & ; \quad x > a/2 \end{cases} \tag{4.52}$$

where

$$\kappa = \sqrt{\frac{2m(-E)}{\hbar^2}} , \tag{4.53a}$$

and

$$k = \sqrt{\frac{2m(E - V_0)}{\hbar^2}} . \tag{4.53b}$$

Note that we have excluded the $e^{-\kappa x}$ solution for $x < -a/2$ and the $e^{+\kappa x}$ solution for $x > a/2$, since these two blow up at $x = \pm\infty$, and we know that since $E < 0$, the particle is not likely to be detected too far from the well.

Now, our expression for the wavefunction, as given by equation 4.52, has four unknowns and there are four matching conditions, namely two conditions from the requirement that the wavefunction be continuous at $x = \pm a/2$, and two conditions from the requirement that the derivative of the wavefunction be continuous at $x = \pm a/2$. These four conditions can, of course determine these constants[#]. It is useful, however, at this point to note that we are solving a problem with an important symmetry.

The Schrödinger equation, for this problem is given by,

$$\hat{H}(x,p)\, u(x) = [- \frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)]\, u(x) = E\, u(x)$$

[#] Actually they do not, because the determinant of the coefficients is zero. We have, however, one more condition by requiring that u(x) be normalized.

Now note that since, in this case, $V(x) = V(-x)$, i.e. the potential is symmetric, if $u(x)$ is a solution, $u(-x)$ must also be a solution. Therefore if $u(x)$ is a solution,

$$u_+(x) = \frac{1}{2}\,[u(x) + u(-x)]$$
$$u_-(x) = \frac{1}{2}\,[u(x) - u(-x)] \quad , \tag{4.54}$$

i.e. the even and odd linear combinations of $u(x)$ and $u(-x)$, must also be solutions. Evidently

$$u_+(x) = u_+(-x) \quad , \tag{4.55a}$$

whereas

$$u_-(x) = -\,u_-(-x) \quad . \tag{4.55b}$$

Consequently, it is sufficient, in such cases, to limit the search to functions of <u>even parity</u> and <u>odd parity</u>, as functions which satisfy equations 4.55a and 4.55b, respectively, are called. Using this information we now look for wavefunctions of the form

$$u_+(x) \propto \begin{cases} A_+\, e^{+\kappa x} & ; \quad x < -a/2 \\ \cos kx & ; \quad |x| < a/2 \\ A_+\, e^{-\kappa x} & ; \quad x > a/2 \quad , \end{cases} \tag{4.56a}$$

and

$$u_-(x) \propto \begin{cases} -\,A_-\, e^{+\kappa x} & ; \quad x < -a/2 \\ \sin kx & ; \quad |x| < a/2 \\ A_-\, e^{-\kappa x} & ; \quad x < a/2 \quad . \end{cases} \tag{4.56b}$$

Now the matching conditions at $x = \pm\, a/2$ are no longer independent, and it is sufficient to consider the pair at $x = a/2$ since our solutions now have a fixed even or odd behavior. Correspondingly, the number of unknowns is now two.

It is also useful to note, at this point that if both the derivative of a function <u>and</u> the function are continuous at $x = x_0$, we have

$$u'(x_0^-) = u'(x_0^+)$$
$$u(x_0^-) = u(x_0^+)$$

and therefore also

$$\left(\frac{u'}{u}\right)_{x_0^-} = \left(\frac{u'}{u}\right)_{x_0^+} \quad ,$$

or

$$\frac{d}{dx}\ln(u)\bigg|_{x_0^-} = \frac{d}{dx}\ln(u)\bigg|_{x_0^+} \quad ,$$

i.e. the <u>logarithmic derivative</u> must also be continuous. We then have, at $x = a/2$

$$\left(\frac{u_+'}{u_+}\right)_{a/2} = -\,k\tan ka/2 = -\,\kappa \quad ,$$

or

$$\cot \frac{ka}{2} = \frac{k}{\kappa} , \tag{4.57a}$$

for the even solutions, and

$$\left(\frac{u'_-}{u_-}\right)_{a/2} = k \cot \frac{ka}{2} = - \kappa$$

or

$$\cot \frac{ka}{2} = - \frac{\kappa}{k} . \tag{4.57b}$$

To solve equations 4.57a and 4.57b it is convenient to define dimensionless variables

$$\xi = \frac{ka}{2} \tag{4.58a}$$

$$\eta(\xi) = \cot \xi \tag{4.58b}$$

and

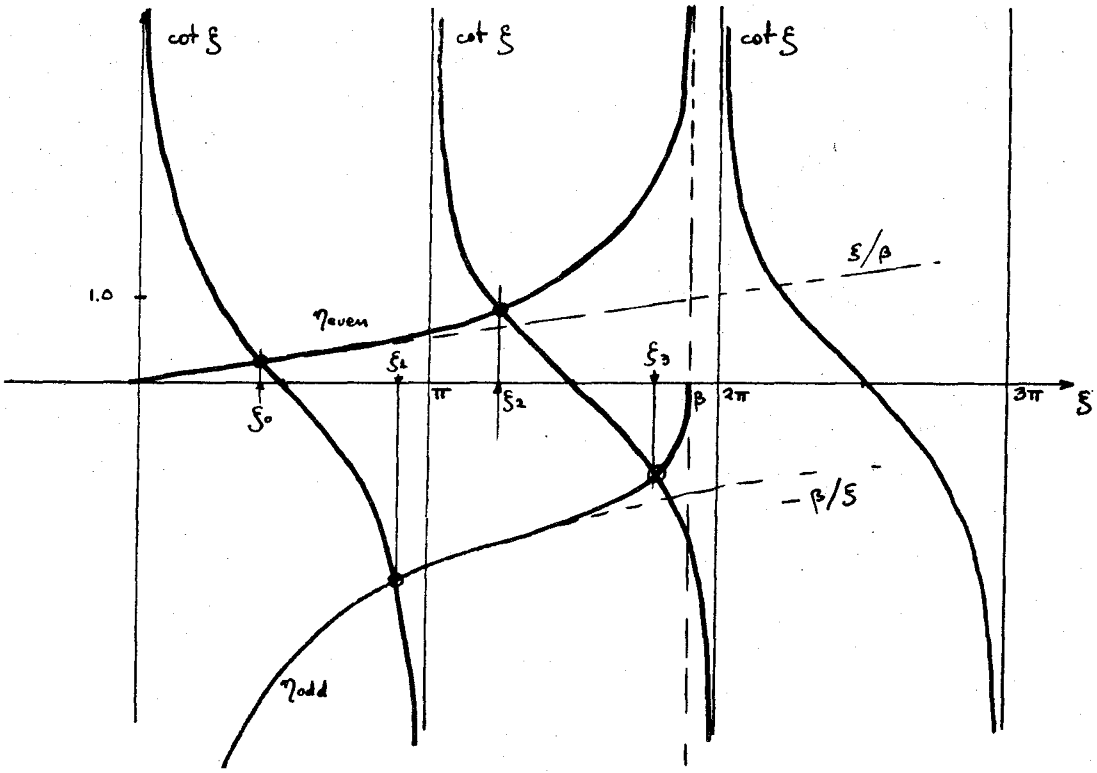
$$\beta = \frac{a}{2}\sqrt{\frac{2m(-V_0)}{\hbar^2}} \tag{4.58c}$$

note $V_0 < 0$. We then have, for the even solutions

$$\eta_{\text{even}}(\xi) = \frac{\xi}{\sqrt{\beta^2 - \xi^2}} \quad ; \quad \text{even solutions}$$

$$\eta_{\text{odd}}(\xi) = - \frac{\sqrt{\beta^2 - \xi^2}}{\xi} \quad ; \quad \text{odd solutions} . \tag{4.59}$$

It is useful to plot these equations, as a function of ξ. A value of $\beta^2 = 30$ has been chosen for this example.



Several important features of the solutions can be seen directly from this graphic solution:

(i) The problem depends on one parameter only, namely β,

(ii) Only specific values of ξ, corresponding to discrete negative energy eigenvalues of the bound states, can serve as solutions. The energies corresponding to the solutions ξ_n are then given by (see equations 4.58a and 4.53b),

$$E - V_0 = (\xi_n/\beta)^2 \, |V_0| \tag{4.60}$$

or,

$$E_n = - |V_0| + \frac{\hbar^2}{2ma^2} (2\xi_n)^2 < 0 \quad . \tag{4.61}$$

(iii) For a potential well of finite depth, there exist a finite number of solutions (bound states), which we can see from the graphic solution, is equal to the number of quadrants covered by the interval $0 < \xi < \beta$. In other words we have (n + 1) solutions, where

$$n \frac{\pi}{2} < \beta \leq (n + 1) \frac{\pi}{2} \quad . \tag{4.62}$$

We see that there exists at least one bound state, even for a vanishingly shallow well $(\beta \to 0)^{\#}$.

(iv) The eigenfunction corresponding to the lowest eigenvalue is even. The one corresponding to the next eigenvalue is odd. In general, for a symmetric potential, the eigenfunctions alternate between functions of even and odd parity*.

(v) As the well becomes deeper and deeper, the lower solutions tend to

$$\xi_n \to (n + 1) \frac{\pi}{2} \qquad \text{as } \beta \to \infty \tag{4.63}$$

for $\xi_n/\beta \ll 1$. Correspondingly, the energies tend to

$$(E_n - V_0) \to (n + 1)^2 \frac{\pi^2 \hbar^2}{2ma^2} \quad \text{as } V_0 \to -\infty \quad . \tag{4.63'}$$

Equations 4.59 may be solved by an asymptotic expansion of the arc cotangent, for the even solutions, and the arc tangent, for the odd solutions to obtain, for $\xi_n/\beta \ll 1$

[#] This result is only true for potentials in one dimension. A shallow potential in three dimensions need not have any bound states.

[*] For a general potential, the eigenfunction corresponding to the lowest eigenvalue has no nodes (zeros). The next highest has one, the following one has two, etc.

$$\frac{\xi_n}{\beta} \simeq \frac{\pi}{2}\frac{(n+1)}{(\beta+1)} \quad , \quad n = 0, 1, 2, 3, \ldots \tag{4.64}$$

Alternatively, for $\xi_n/\beta \leq 1$, they may be solved numerically. See appendix E. The numerical solutions, computed to four significant figures, are compared to the estimates of equation 4.62 in the table below for $\beta = \sqrt{30} = 5.4772$

n	$\frac{\pi}{2}\frac{(n+1)}{(\beta+1)}$	ξ_n/β#	$\frac{E_n}{\lvert V_0 \rvert}$
0	0.2425	0.2421	-0.941
1	0.4850	0.4818	-0.768
2	0.7275	0.7149	-0.489
3	0.9700	0.9294	-0.136

It can be seen that the asymptotic estimate for $\xi_n/\beta << 1$, i.e. equation 4.64, yields very good results for small ξ_n/β. In fact it is only 4% off at $\xi/\beta \simeq 0.93$!

The constants A_+ and A_- of equations 4.56 can now be evaluated for each eigenfunction, in particular, from the continuity of the wavefunctions at $x = a/2$, we have

$$A_+ = e^{\kappa a/2} \cos\frac{ka}{2} \tag{4.65a}$$

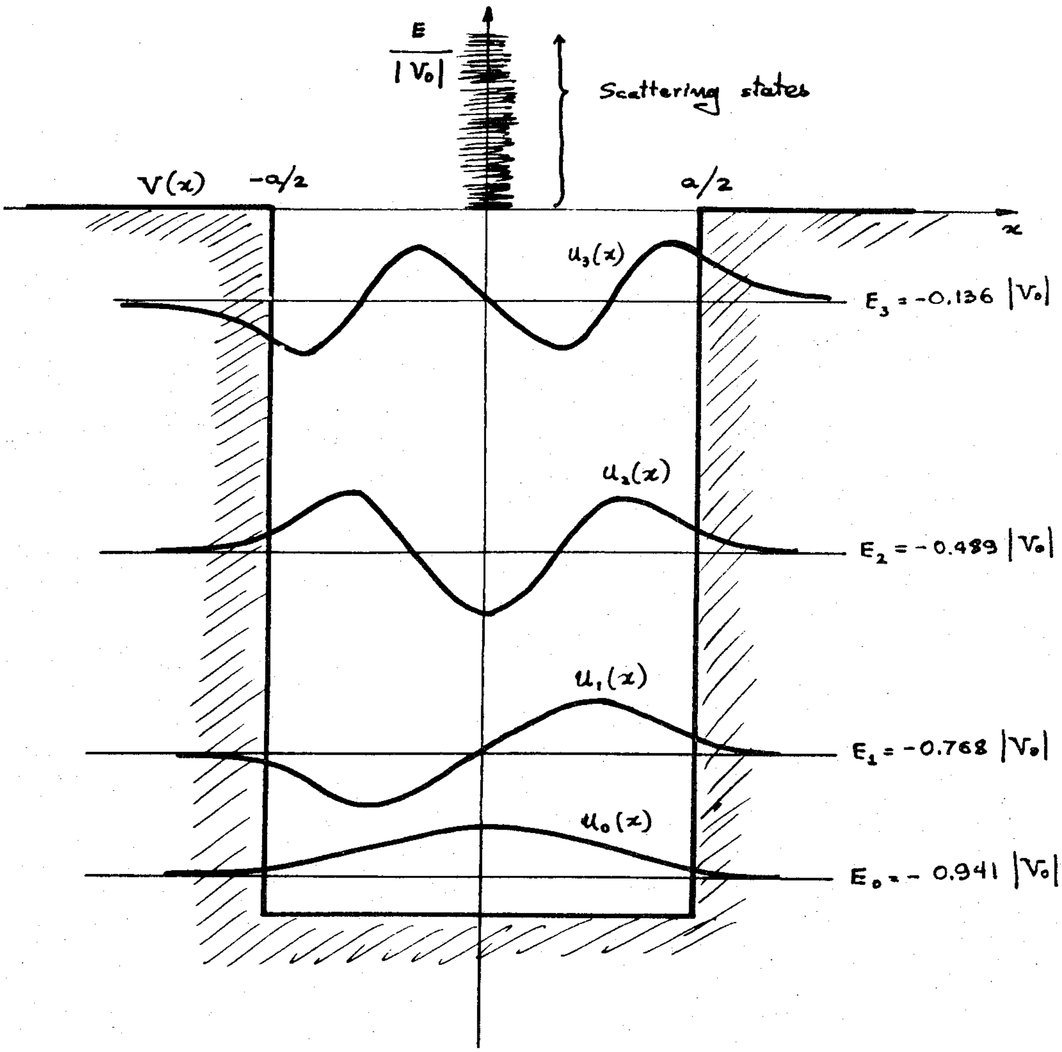
and

$$A_- = e^{\kappa a/2} \sin\frac{ka}{2} \quad . \tag{4.65b}$$

The constant of proportionality, in equations 4.56 can also be computed by requiring that the wavefunctions are normalized. The wavefunctions corresponding to the four bound states for the $\beta = \sqrt{30}$ potential are sketched below

---

#from the numerical computation in appendix E.



Eigenvalues and eigenfunctions for a rectangular potential well of a depth such that

$$\left(\frac{a}{2}\right)^2 \frac{2m\ (-\ V_0)}{\hbar^2} = 30$$

It is useful to note that, except for the normalization constant, we could have written the matched wavefunctions, from the beginning. In particular, if we require that $|u(x)| = 1$ at $x \pm a/2$, we have, from equations 4.56,

$$u(x) \propto \begin{cases} e^{\kappa(x + a/2)} & ; \quad x < -a/2 \\ \dfrac{\cos kx}{\cos(\frac{ka}{2})} & ; \quad |x| < a/2 \\ e^{\kappa(x - a/2)} & ; \quad x > a/2 \end{cases} \tag{4.66a}$$

and

$$u(x) \propto \begin{cases} -e^{\kappa(x + a/2)} & ; \quad x < -a/2 \\ \dfrac{\sin kx}{\sin(\frac{ka}{2})} & ; \quad |x| < a/2 \\ e^{\kappa(x - a/2)} & ; \quad x > a/2 \end{cases} \tag{4.66b}$$

which, for k and κ that satisfy equations 4.57, are matched by construction.

4.25 *The infinite potential well.*

Quite frequently, the fact that the depth of the well is finite has very little effect on the problem of interest. For example, if $|E| << |V_0|$, i.e. if we are dealing with low lying states for which (see equation 4.64), $n/\beta << 1$, and $\beta >> 1$, we go to the limits of equation 4.63. Note that, in that limit,

$$u_\pm(\pm a/2) \to 0$$

since $ka/2$ tends to an odd multiple of $\pi/2$ for the even (cosine) solutions and to a multiple of π for the odd (sine) solutions. We have therefore that in the limit of $\beta \to \infty$ and $n/\beta \to 0$, the wavefunctions are strictly confined to the region $|x| < a/2$, i.e. for n even

$$u_n(x) = \begin{cases} (\frac{2}{a})^{1/2} \cos k_n x & ; \quad |x| < a/2 \\ 0 & ; \quad |x| > a/2 \end{cases} , \tag{4.67a}$$

whereas for n odd,

$$u_n(x) = \begin{cases} (\frac{2}{a})^{1/2} \sin k_n x & ; \quad |x| < a/2 \\ 0 & ; \quad |x| > a/2 \end{cases} , \tag{4.67b}$$

where

$$\frac{k_n a}{2} = \frac{\pi}{2}(n + 1) , \tag{4.68}$$

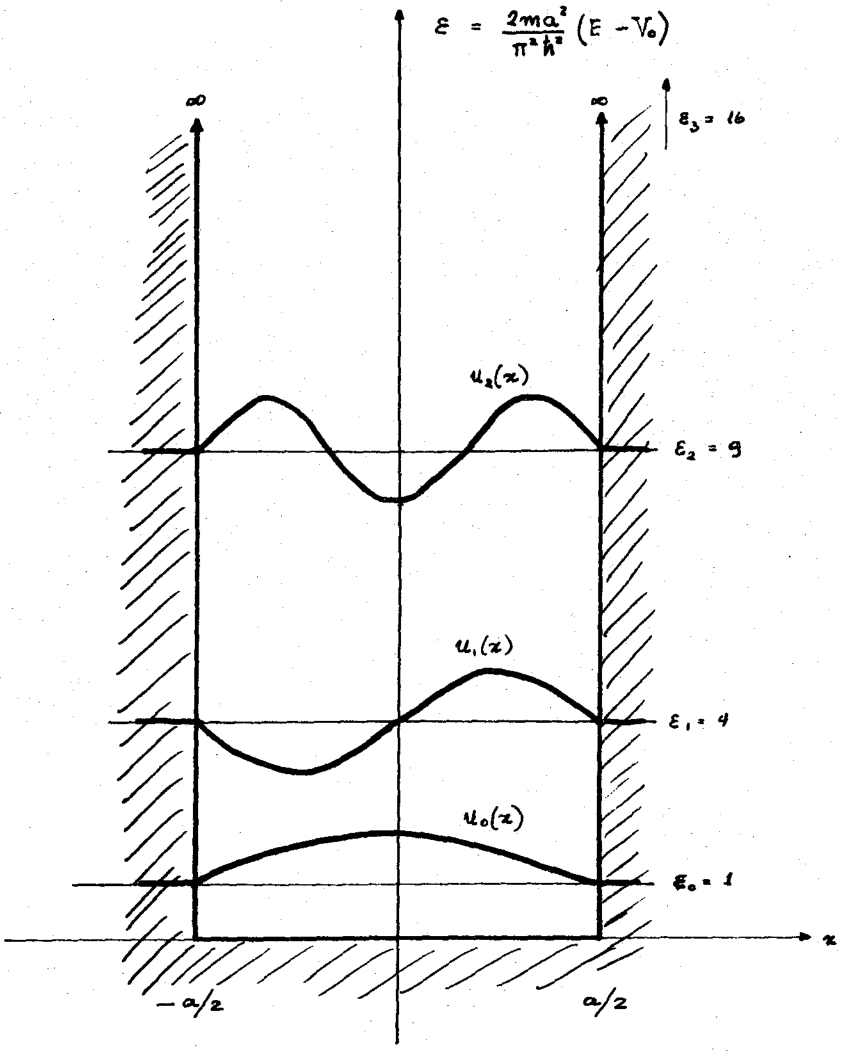
for all n. The expressions for the wavefunctions given by equations 4.67, include the normalization constant. See equation 3.17 in example 3.3, and problems 4.5 and 4.6.

We can now turn the problem around and recognize that these solutions are <u>the</u>

eigenfunctions of a free particle confined to a region of extent a, since the value of $V(x) = V_0 \to -\infty$ is irrelevant, if we measure the energies using V_0 as the origin. See equation 4.63'. These eigenfunctions are, of course, the eigenfunctions of the kinetic energy operator. See example 3.3. Note that the energy levels, scaled by $\frac{\pi^2 \hbar^2}{2ma^2}$, are the sequence of the index (plus one) squared, i.e.

$$\epsilon_n = \frac{2ma^2}{\pi^2 \hbar^2}(E_n - V_0) = (n+1)^2 \tag{4.69}$$

The index n is often referred to as the quantum number of the state. The first three wavefunctions, offset by the corresponding energy levels, are sketched below.



4.3 Density of states.

4.3.1 Motion in one dimension.

For a particle confined to a region

$$-\frac{L}{2} < x < \frac{L}{2} \tag{4.70}$$

we found the eigenvalues

$$E_n = \frac{\hbar^2}{2m} k_n^2 \tag{4.71}$$

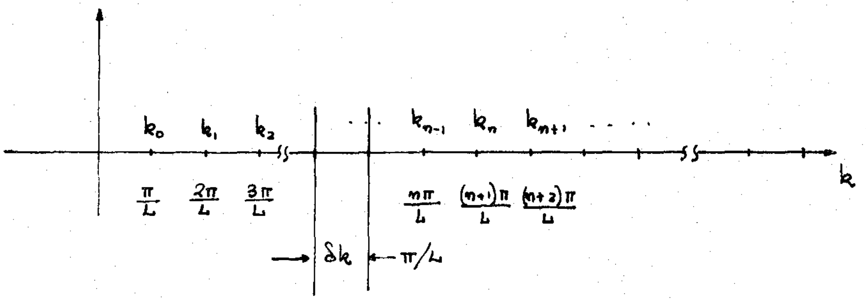
where

$$k_n L = (n + 1)\,\pi \quad , \tag{4.72}$$

corresponding to the eigenfunctions,

$$u_k(x) \propto \begin{cases} \cos k_n x & ; \quad |x| < L/2, \quad n = \text{even} \\ \sin k_n x & ; \quad |x| < L/2, \quad n = \text{odd} \\ 0 & ; \quad |x| > L/2 \quad . \end{cases} \tag{4.73}$$

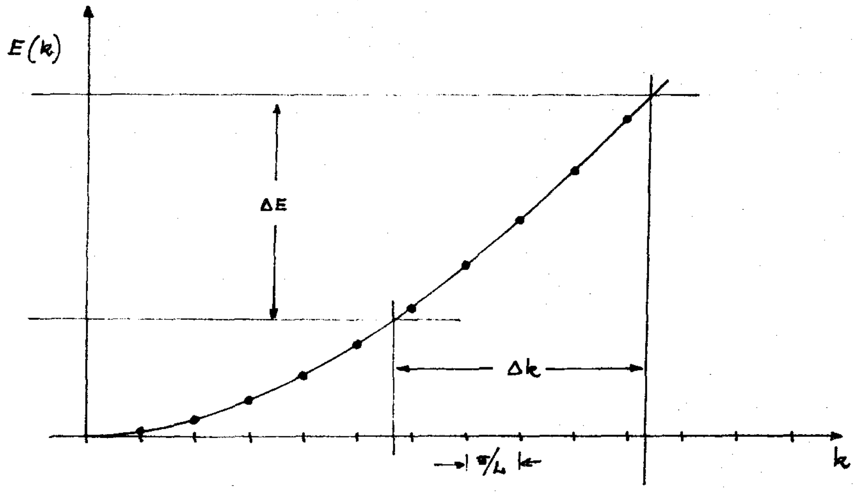
Note, from equation 4.72, that the wavenumbers corresponding to the solutions $u_k(x)$ are evenly spaced on the k-axis, i.e.



by an interval

$$\delta k = \frac{\pi}{L} \quad . \tag{4.74}$$

The corresponding energies are then discrete and fall on the parabola given by equation 4.71.



We now ask how many states there are in a particular energy interval ΔE at E. We can think of this problem as the difference

$$\Delta N = N(E + \Delta E) - N(E)$$

which we can express as

$$\Delta N \simeq \left(\frac{\Delta N}{\Delta E}\right) \Delta E$$

where

$$\frac{\Delta N}{\Delta E} = \frac{dk}{dE} \cdot \frac{\Delta N}{\Delta k} \qquad (4.74)$$

We can compute dE/dk from equation 4.71 in particular

$$k = \left(\frac{2mE}{\hbar^2}\right)^{1/2} \qquad (4.75)$$

and therefore

$$\frac{dk}{dE} = \frac{(2m)^{1/2}}{2\hbar} E^{-1/2} \qquad (4.76)$$

The second term $\Delta N/\Delta k$, can be seen to be equal to the number of states dN in an interval dk. Since, however, the states as indexed by k are uniformly spaced on the k-axis, we have, from equation 4.74,

$$\frac{\Delta N}{\Delta k} = \frac{1}{\delta k} = \frac{L}{\pi} \quad . \qquad (4.77)$$

Therefore, for one dimensional motion,

$$\Delta N \simeq \frac{L}{2\pi} \frac{(2m)^{1/2}}{\hbar} E^{-1/2} \Delta E \qquad (4.78)$$

We can think of the number of states N(E), in the limit of $L \to \infty$ as an integral over a density of states n(E), i.e.

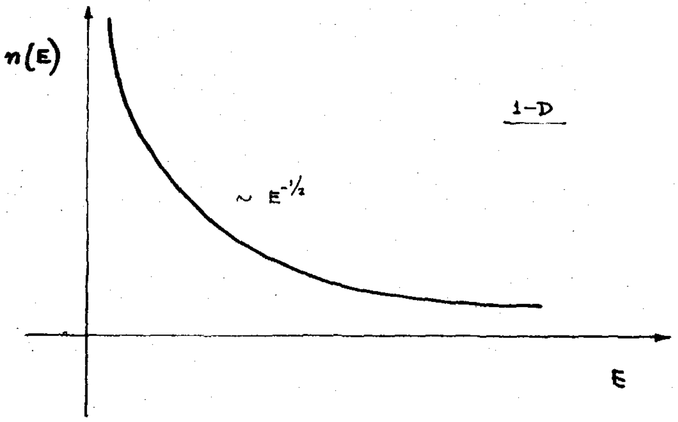
$$N(E) \simeq \int_0^E n(E)\,dE \quad ,$$

or

$$dN = n(E)\,dE \qquad (4.79)$$

where from 4.78 <u>for one-dimensional motion</u>

$$n(E) = \frac{L}{2\pi} \frac{(2m)^{1/2}}{\hbar} E^{-1/2} \qquad (4.80)$$



Note that the density of states $n(E)$ can be written as

$$n(E) = \frac{dE}{dk}^{-1} \cdot \frac{dN}{dk}$$

where, from chapter 2,

$$\frac{dE}{dk} = \hbar \frac{d\omega}{dk} = \hbar v_g(k)$$

where $v_g(k)$ is the group velocity at k. We therefore have

$$n(E) = \frac{1}{\hbar v_g(k)} \cdot \frac{dN}{dk} \tag{4.81}$$

In particular, note that the density of states goes to infinity wherever the group velocity $v_g(k)$ vanishes.

4.3.2 Periodic boundary conditions.

The set of solutions for a particle in a one dimensional box, as given by equations 4.73 can be seen to be standing waves $\cos kx$ and $\sin kx$ or, equivalently, superpositions of running waves e^{ikx} and e^{-ikx} but with equal amplitude. This restriction of the amplitude of e^{ikx} and e^{-ikx}, for a given k, arises from the details of the potential $V(x)$ at $x = \pm L/2$, which force the wavefunction to go to zero there. Often times, however, either that particular shape of the potential is an approximation, or the region of interest is localized in an interval away from the edges $x = \pm L/2$ and therefore the actual details of the boundary conditions there are unimportant. What is important, however, is that we preserve the proper accounting of the number of states per unit wavenumber interval and therefore also per unit energy interval.

An alternate solution set can be obtained in the region $|x| < L/2$ by replacing the infinite square well by the condition that the wavefunctions are <u>periodic</u>, i.e.

$$u(-L/2) = u(L/2) \tag{4.82}$$

Then, if we express the solutions in terms of the running waves e^{ikx}, we must have,

$$e^{ikL/2} = e^{-ikL/2} ,$$

or

$$e^{ikL} = 1 \tag{4.83}$$

where L is the extent of the region to which the particle is confined.

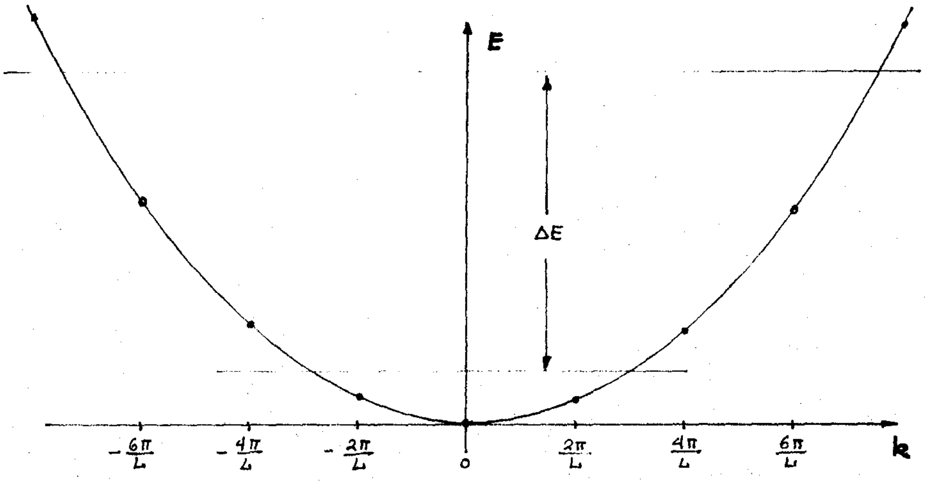
The solution to equation 4.83 is given by

$$kL = 2\pi\ell \qquad \ell = 0, \pm 1, \pm 2, \ldots$$

or

$$k = \frac{2\pi}{L}\ell \qquad \ell = 0, \pm 1, \pm 2 \tag{4.84}$$

Note that, in this representation, the values of k are both positive and negative, corresponding to the direction of the running waves, but are now spaced twice as far apart.



Therefore the density of states per unit energy interval remains the same as before.

4.3.3 Density of states in three dimensions.

Consider now a particle of mass m moving in a three dimensional

$$V(\underline{x}) = V_1(x) + V_2(y) + V_3(z) \qquad (4.85a)$$

where

$$V_i(x) = \begin{cases} \infty & \text{for} \quad |x| > L_i/2 \\ 0 & \text{for} \quad |x| > L_i/2 \end{cases} , \qquad (4.85b)$$

representing a rectangular box with sides L_1, L_2, L_3. From our previous discussion, we have that the time-independent Schrödinger equation

$$\left[-\frac{\hbar^2}{2m} \left(\frac{\partial}{\partial \underline{x}} \cdot \frac{\partial}{\partial \underline{x}} \right) + V(\underline{x}) \right] u(\underline{x}) = E\, u(\underline{x}) \quad ,$$

splits into three one-dimensional equations, i.e.

$$\left\{ \frac{\partial^2}{\partial x_i} + \frac{2m}{\hbar^2} [E_i - V_i(x_i)] \right\} u_i(x_i) = 0$$

where

$$E = E_1 + E_2 + E_3 \qquad (4.86)$$

and

$$u(\underline{x}) = u_1(x)\, u_2(y)\, u_3(z) \quad .$$

The energies are then given by

$$E_{i,n} = \frac{\hbar^2}{2m} k^2_{i,n_i} \quad , \qquad (4.87)$$

where

$$k_{i,n_i} L_i = (n_i + 1)\,\pi \quad , \tag{4.88}$$

and the eigenfunctions

$$u_{i,k_i}(x_i) \propto \begin{cases} \cos(k_{i,n_i} x_i) & ; \quad n_i = \text{even} \\ \sin(k_{i,n_i} x_i) & ; \quad n_i = \text{odd} \quad . \end{cases} \tag{4.89}$$

The total energy is then given, for each admissible $\underline{k} = (k_x, k_y, k_z)$, by

$$E_{\underline{k}} = \frac{\hbar^2}{2m} \left(k_x^2 + k_y^2 + k_z^2\right) \tag{4.90}$$

or

$$E_{\underline{n}} = \frac{\pi^2 \hbar^2}{2m} \left[\frac{(n_1 + 1)^2}{L_1} + \frac{(n_2 + 1)^2}{L_2} + \frac{(n_3 + 1)^2}{L_3} \right] , \tag{4.91}$$

for $n_1, n_2, n_3 = 0, 1, 2, \ldots$

We see, as before, that these solutions are <u>standing</u> waves confined to the extent of the box in each of the three dimensions. To admit <u>running</u> wave solutions, we consider instead the periodic boundary conditions,

$$u_i(-L_i/2) = u_i(L_i/2)$$

which yield the eigenfunctions

$$u_i(x_i) = e^{ik_i x_i} \quad , \tag{4.92}$$

where

$$k_i = \frac{2\pi}{L_i}\, \ell_i \quad ; \qquad \ell_i = 0, \pm 1, \pm 2, \ldots \tag{4.93}$$

with the corresponding eigenvalues

$$\begin{aligned} E &= \frac{\hbar^2}{2m} \left(k_1^2 + k_2^2 + k_3^2\right) \\ &= \frac{4\pi^2 \hbar^2}{2m} \left(\frac{\ell_1^2}{L_1^2} + \frac{\ell_2^2}{L_2^2} + \frac{\ell_3^2}{L_3^2} \right) \; ; \; \ell_1, \ell_2, \ell_3 = 0, \pm 1, \pm 2, \ldots \end{aligned} \tag{4.94}$$

Note that these wavefunctions, i. e.

$$u_{\underline{k}}(\underline{x}) = u_{1,k_1}(x)\, u_{2,k_2}(y)\, u_{3,k_3}(z)$$

are wavefunctions of definite momentum, i. e.

$$\underline{p} = \hbar \underline{k} = (\hbar k_1, \hbar k_2, \hbar k_3) \quad ,$$

which we can label unambiguously by the wavenumber $\underline{k}$, and denote in our representation independent notation as $|\underline{k}\rangle$. Any state of the system can then be represented as a superposition of these states, i. e.

$$|\psi\rangle = \sum_{\underline{k}} |\underline{k}\rangle\, c_{\underline{k}} \quad , \tag{4.95}$$

where the summation runs over the admissible values of $\underline{k} = (k_1, k_2, k_3)$, as given by equation

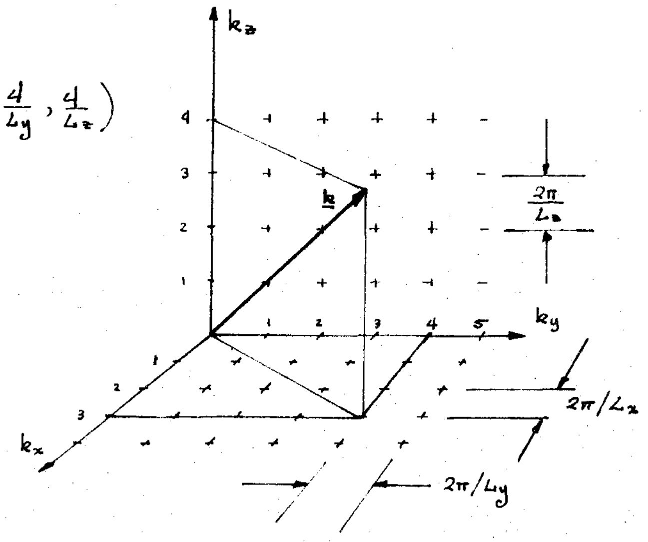
4.93. Note that

$$\hat{H}|\psi\rangle = \sum_{\underline{k}} \hat{H}|\underline{k}\rangle c_{\underline{k}} = \frac{\hbar^2}{2m} \sum_{\underline{k}} k^2 |\underline{k}\rangle c_{\underline{k}} \tag{4.96}$$

and that therefore, for example,

$$\langle E\rangle = \langle\psi|\hat{H}|\psi\rangle = \frac{\hbar^2}{2m} \sum_{\underline{k}} k^2 |c_{\underline{k}}|^2 = \sum_{\underline{k}} E_{\underline{k}} |c_{\underline{k}}|^2 \quad . \tag{4.97}$$

Note now that these states occupy a <u>rectangular lattice in wavenumber space</u>, i.e.



$\underline{k} = 2\pi\left(\frac{3}{L_x}, \frac{4}{L_y}, \frac{4}{L_z}\right)$ is indicated

of cell size

$$\delta^3 k = \left(\frac{2\pi}{L_1}\right)\left(\frac{2\pi}{L_2}\right)\left(\frac{2\pi}{L_3}\right) = \frac{(2\pi)^3}{\text{Volume}} , \tag{4.98}$$

alternatively there exists one state per $[(2\pi)^3/\text{Volume}]$ in $\underline{k}$-space.

We now wish to compute the number of states in an energy interval

$$E < E' < E + dE \quad ,$$

i.e.

$$dN(E) = N(E + dE) - N(E) \quad .$$

To compute $N(E)$ we note that the total number of states with energy

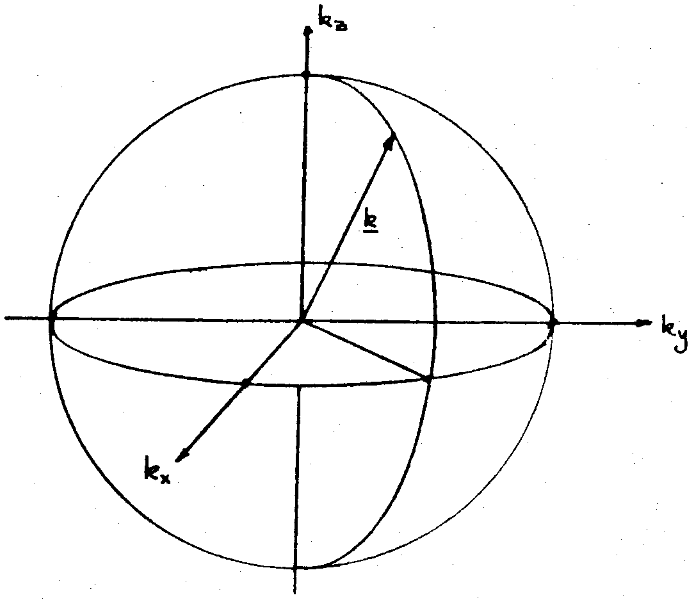
$$0 < E' < E$$

is given by

$$N(E' < E) = \frac{(\text{Volume in } \underline{k}\text{-space with } E' < E)}{(\text{Volume in } \underline{k}\text{-space per state})}$$

or

$$N(E' < E) = \frac{\frac{4}{3}\pi k^3(E)}{(2\pi)^3/\text{Vol.}} \quad , \tag{4.99}$$



where,

$$k = \left(\frac{2mE}{\hbar^2}\right)^{1/2} .$$

Therefore, after a little algebra

$$N(E' < E) = \frac{\text{Vol.}}{6\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} E^{3/2} , \qquad (4.100)$$

and therefore

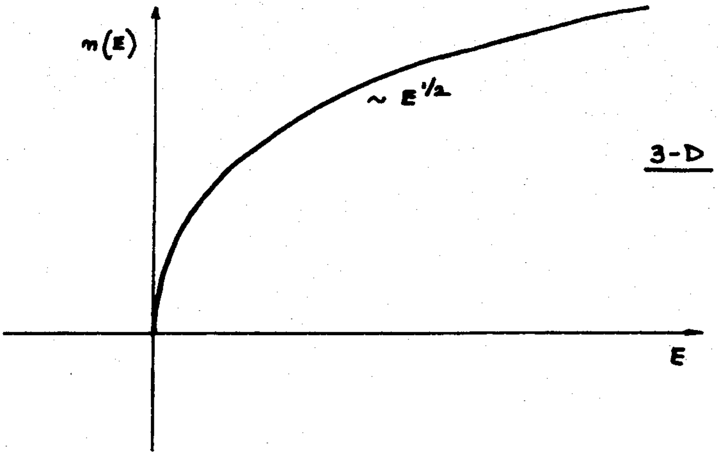
$$dN = \frac{dN}{dE}\,dE = \frac{\text{Vol.}}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} E^{1/2}\,dE .$$

Consequently, as in the one-dimensional case, we can define the <u>density of states</u> n(E), such that the total number of states N(E) between 0 and E is given by

$$N(E) = \int_0^E n(E')\,dE' , \qquad (4.101)$$

where

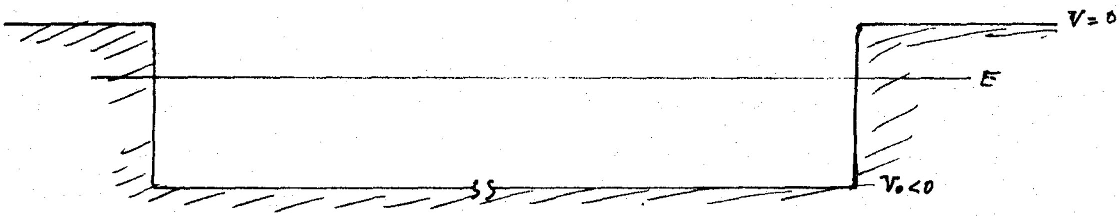
$$n(E) = \frac{dN}{dE} = \frac{\text{Vol.}}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} E^{1/2} . \qquad (4.102)$$



4.3.4 The Fermi energy of conduction electrons in a metal.

As we shall see later on, electrons possess intrinsic angular momentum called spin, whose projection along any axis is always found to be $+\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$, conventionally referred to as "spin up" or "spin down". We will also see that two electrons# that are in the same spin state (spin up or spin down) may not be described by the same spatial wavefunction. In other words, two electrons may not occupy the same state. This is the celebrated Pauli exclusion principle.

Consider now conduction electrons which are free to move around inside a metal. We can model this situation by a three-dimensional potential well whose extent is the piece of metal and whose depth is yet to be determined.



Now if the metal is an alkali (K, Na, etc.), which yields one electron per atom, we will have N_e electrons per unit volume,

$$N_e = \rho/m_a \tag{4.103}$$

where ρ is the density of the material and m_a is the mass per atom. An electron, however, moving inside the rectangular potential can occupy the wavenumber states $|\underline{k}\rangle = |k_x, k_y, k_z\rangle$ of a particle in a box.

If we had only one electron, then the lowest energy state of the system would find it in the ground state, corredponding to the lowest admissible values of k_x, k_y, k_z (see equation 4.93). If we now add a second electron, we could put it in the same (spatial) state provided it had a spin projection that was anti-parallel to that of the first. A third electron, however, cannot be accommodated in the same state and is forced to occupy the next highest admissible state. In this fashion, we can place two electrons in each (spatial) state. It is clear that the Pauli exclusion principle forces the electrons to occupy ever increasing energy states. How high up the energy ladder does this process go? Well, from the results of the preceding discussion, we have that the number of states per unit volume between 0 and E is given by

$$\frac{N(E)}{V} = \frac{1}{4\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2}\int_0^E E^{1/2}\,dE \tag{4.104}$$

or

---

or any particles with half-integral spin.

$$\frac{N(E)}{V} = \frac{1}{6\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2} E^{3/2} \quad . \tag{4.104'}$$

Since we can accommodate two electrons in each of these states, we must actually reach up to an energy such that

$$\frac{1}{6\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2} E^{3/2} = \frac{1}{2} N_e$$

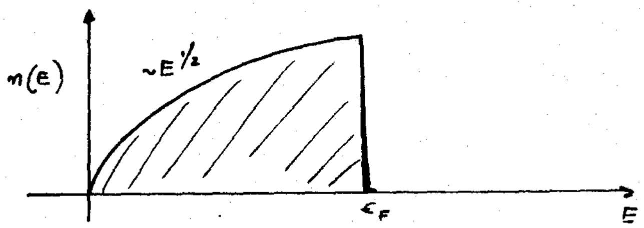
This energy is called the Fermi energy and is given by

$$\epsilon_F = \frac{\hbar^2}{2m_e}\left(3\pi^2 N_e\right)^{2/3} \quad . \tag{4.105}$$

Note that this has the correct dimensions and that the Fermi energy, corresponds to a wavenumber

$$k_F = \left(3\pi^2 N_e\right)^{2/3} \quad . \tag{4.106}$$

The electron energy spectrum is then given by equation 4.102 for $E < \epsilon_F$ and is zero for $E > \epsilon_F$.

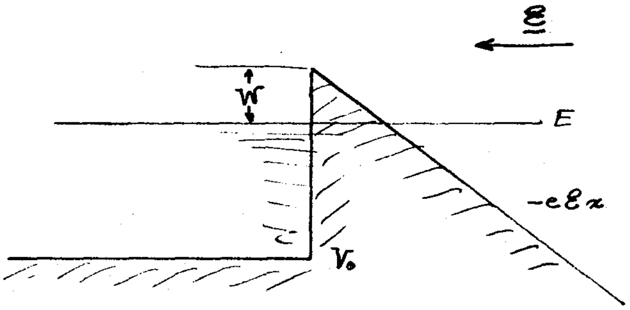


Strictly speaking, this is only true at zero temperature. Even at room temperature, however, this is a very good approximation since the smearing of the sharp cut-off at $E = \epsilon_F$ will be of the order of $\Delta \sim k_B T$ and, for typical values,

$$\frac{\epsilon_F}{k_B} > 10^4 \;{}^\circ\mathrm{K} \quad .$$

We can now determine the depth of the effective well since we know that the most energetic electrons (i.e. at ϵ_F) can be extracted photoelectrically if we deliver to them an energy equal to the work function which can readily be measured. This situation is depicted below. See also problem 4.11.

and sketched below.



As can be seen from this picture, the electrons can be extracted from the metal surface by <u>tunneling</u> through the potential barrier. The field emission current is then proportional to the tunneling transmission coefficient (equation 4.48).

$$j \propto e^{-2\int_0^a \kappa(x)dx}$$

where

$$\kappa(x) = \sqrt{\frac{2m}{\hbar^2}[V(x) - E]} \quad ; \quad 0 < x < a$$

or, for electrons at the top of the Fermi sea

$$\kappa(x) = \sqrt{\frac{2m}{\hbar^2}(W - e\mathcal{E}x)} \quad ; \quad 0 < x < a \quad , \tag{4.109}$$

where a is the location where the energy gain crosses the potential,

$$a = \frac{W}{e\mathcal{E}} \quad . \tag{4.110}$$

We then have,

$$j \propto e^{-\kappa_0 a \, 2\int_0^1 (1 - \xi)^{1/2} d\xi}$$

where the effective tunneling wavenumber is scaled by

$$\kappa_0 = \left(\frac{2mW}{\hbar^2}\right)^{1/2} \quad , \tag{4.111}$$

and ξ has been defined as x/a. Since

$$\int_0^1 (1 - \xi)^{1/2} d\xi = 2/3$$

we have

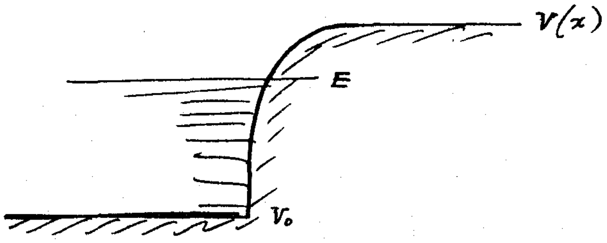
$$j \propto e^{-\frac{4}{3}\kappa_0 a} \tag{4.112}$$

or, substituting for κ_0 and a,

$$j \propto e^{-\frac{4}{3}\left(\frac{2mW}{\hbar^2}\right)^{1/2}\frac{W}{e\mathcal{E}}} \tag{4.113}$$

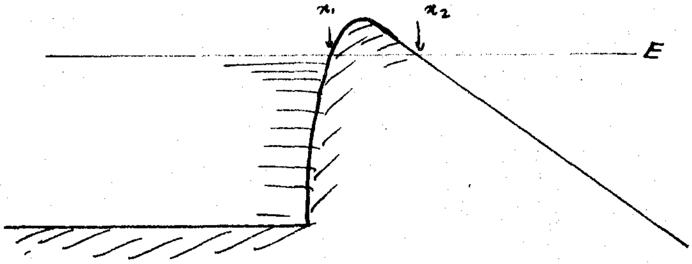
Equation 4.113 is known as the Fowler-Nordheim formula and provides a qualitatively correct description for field emission. To obtain a better expression, we have to take into account several effects we have ignored. The first is the electrostatic attraction an electron feels pulling it back to the metal, a consequence of the image charge induced in the metal. This leads to a potential which for $x > 0$ and zero electric field is given by

$$V(x) = -\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{4x} \quad . \tag{4.114}$$



When we then turn on the electric field, the potential for $x > 0$ becomes

$$V(x > 0) = -\left(\frac{e^2}{4\pi\varepsilon_0}\right)\frac{1}{4x} - e\mathcal{E}x \tag{4.115}$$



It can be seen that this results in an increase in the field emission current for two reasons:

(i) The range of integration is now slightly reduced, i.e.

$$x_2 - x_1 = a\left(1 - \frac{4b}{a}\right)^{1/2} \quad , \tag{4.116}$$

where b is defined by

$$W \equiv \frac{e^2}{16\pi\varepsilon_0 b} \quad . \tag{4.117}$$

and,

(ii) the maximum energy deficit is reduced, corresponding to a decrease of the largest $\kappa(x)$ in the integral ,

$$\kappa(x) = \kappa_0\sqrt{1 - \frac{x}{a} - \frac{b}{x}} \quad ; \quad x_1 < x < x_2 \quad ,$$

given by

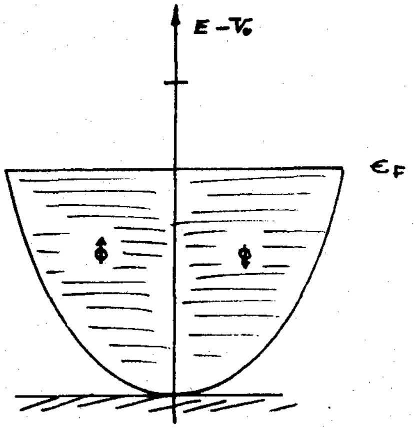
4.3.7. Paramagnetic behavior of metals.

It is known that classical paramagnetic behavior predicts a magnetization field, which for weak magnetic fields, is proportional to the applied field and inversly proportional to temperature, i.e.

$$(\mathcal{M}_z)_{\text{class}} \sim N_e \frac{\mu_0^2}{k_B T} \mathcal{B}_z \quad . \tag{4.120}$$

See appendix F, section 5. While this seems to describe the paramagnetic behavior of many salts, it has long been known that metals have a paramagnetic behavior, which is essentially independent of temperature.[‡] There is no way to account for this discrepancy classically, which is a consequence of the quantum behavior of the conduction electrons in a metal. This can be understood as outlined below.

We have seen that, in the absence of a magnetic field, the conduction electrons occupy the available free particle states in the box, defined by the extent of the metal, in pairs (spin up and spin down), up to the Fermi energy ε_F.



When we apply an external uniform magnetic field $\underline{\mathcal{B}} = \mathcal{B}_z \hat{e}_z$, we introduce an additional term to the Hamiltonian of each electron, representing the interaction of the electron spin magnetic moment $\underline{\mu}_S$ and the external field, i.e.

$$H = H_0 + H_S \quad , \tag{4.121}$$

where

$$H_S = -\underline{\mu}_S \cdot \underline{\mathcal{B}} \quad , \tag{4.122}$$

and

$$\underline{\mu}_S = -g_e\left(\frac{e}{2m_e}\right) \underline{S} \tag{4.123}$$

where $g_e \simeq 2.00$ and $\underline{S}$ is the spin angular momentum. See appendix F, section 5.

It can be seen that this new term does not operate on the space coordinates.[†] It has

---

[‡] for non ferromagnetic materials.

[†] There is actually an additional term we have not included, which couples to the orbital motion of the electrons. See problem F.11. This yields a diamagnetic contribution to the magnetization. We will discuss this later in the context of angular momentum.

only two values, depending on whether the electron spin is parallel or anti-parallel to the magnetic field. Correspondingly, the eigenfunctions are separable and can be written, as we have seen before, as a product of functions of the space variables (eigenfunctions of H_0) and functions of the spin variables (eigenfunctions of H_S), i.e.

$$\Psi(\underline{x}, \underline{S}) = \psi(\underline{x})\, \sigma(\underline{S}) \quad . \tag{4.124}$$

We know the eigenfunctions $\psi(\underline{x})$ corresponding to H_0. They are the free particle states in a box, corresponding to the admissible wavenumbers (subject to the periodic boundary condition restrictions) i.e.

$$|\psi\rangle = |\underline{k}\rangle = |k_x, k_y, k_z\rangle \quad . \tag{4.125}$$

We also know, in turn, the eigenfunctions of the spin variables since the electron spin (projection) can only be "up" or "down", i.e.

$$|\sigma\rangle = |m_s\rangle \quad , \tag{4.126}$$

where $m_s = \pm 1/2$, corresponding to the two possible states, and

$$S_z |m_s\rangle = \hbar m_s |m_s\rangle \tag{4.127}$$

corresponding to a z-projection of the spin angular momentum of ± $\hbar/2$.# Consequently the total eigenfunctions are labeled by four indeces, i.e.

$$|\Psi\rangle = |\underline{k}\rangle |m_s\rangle = |k_x, k_y, k_z, m_s\rangle = |\underline{k}, m_s\rangle \quad . \tag{4.128}$$

Note that

$$H_0 |\Psi\rangle = \frac{\hbar^2 k^2}{2m_e} |\Psi\rangle \quad , \tag{4.129a}$$

and

$$H_s |\Psi\rangle = g_e \left(\frac{\hbar e}{2m_e}\right) \mathcal{B}_z m_s \; |\Psi\rangle \tag{4.129b}$$

and that therefore the energy of these states

$$H|\Psi\rangle = E_{\underline{k}, m_s} |\Psi\rangle$$

is given by

$$E_{\underline{k}, m_s} = \frac{\hbar^2 k^2}{2m_e} + g_e \left(\frac{\hbar e}{2m_e}\right) \mathcal{B}_z m_s \quad . \tag{4.130}$$

The magnetic field, therefore results in an energy shift of the spin up electrons with respect to the spin down electrons, which is independent of $\underline{k}$, and given by (recall $m_s = \pm 1/2$)

$$\Delta E = 2\mu_0 \mathcal{B}_z = g_e \left(\frac{\hbar e}{2m_e}\right) \mathcal{B}_z \quad . \tag{4.131}$$

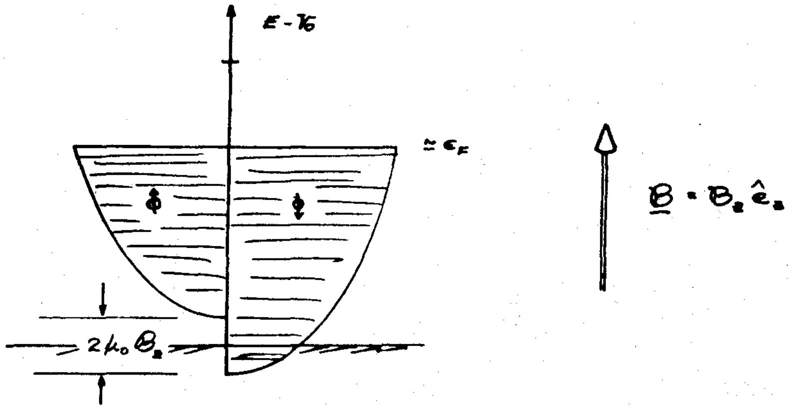
Consequently, if we now fill up the available states with the conduction electrons starting from the lowest energy up to the energy required to accommodate the number of electrons per unit volume, we have

---

#We will study these eigenfunctions later on in the context of angular momentum.

$$\underbrace{N_e}_{\text{electrons per unit volume}} = \underbrace{\int_0^{E-\mu_0 \mathcal{B}_z} n(\varepsilon)\,d\varepsilon}_{\text{spin-up electrons}^{\#}} + \underbrace{\int_0^{E+\mu_0 \mathcal{B}_z} n(\varepsilon)\,d\varepsilon}_{\text{spin-down electrons}^{\#}} , \qquad (4.\,132)$$

where $n(\varepsilon)$ is the density of kinetic energy states. This situation is sketched below.



Consequently, the number of spin-up electrons per unit volume is given by,

$$N_\uparrow \simeq \int_0^{E} n(\varepsilon)\,d\varepsilon - n(E)\,\mu_0 \mathcal{B}_z \qquad (4.\,133a)$$

while the number of spin-down electrons per unit volume is given by

$$N_\downarrow \simeq \int_0^{E} n(\varepsilon)\,d\varepsilon + n(E)\,\mu_0 \mathcal{B}_z \; . \qquad (4.\,133b)$$

and therefore,

$$N_e = N_\uparrow + N_\downarrow \simeq 2\int_0^{E} n(\varepsilon)\,d\varepsilon \; . \qquad (4.\,134)$$

We see that $E \simeq \varepsilon_F$, the unperturbed Fermi energy, to first order in $\mu_0 \mathcal{B}_z$. See also problem 4. 13.

Combining equations 4. 133 for the spin-up, spin-down electron number density, we have for the net magnetization[#]

$$\begin{aligned} \mathcal{M}_z &= (N_\downarrow - N_\uparrow)\,\mu_0 \\ &= 2n\,(\varepsilon_F)\,\mu_0^2\,\mathcal{B}_z \;, \end{aligned}$$

or, since

$$2n\,(\varepsilon_F) = \frac{3}{2}\left(\frac{N_e}{\varepsilon_F}\right) , \qquad (4.\,135)$$

we have, for the <u>paramagnetic</u> contribution to the magnetization,

$$(\mathcal{M}_z)_{para} = \frac{3}{2}\left(\frac{N_e \mu_0^2}{\varepsilon_F}\right)\mathcal{B}_z \;, \qquad (4.\,136)$$

[#] Note that the electron magnetic moment is <u>opposite</u> the direction of the spin.

first derived by Pauli in 1927.[#] The orbital electron motion (see footnote † on page 4.38) results in a diamagnetic contribution of

$$(\mathfrak{M}_z)_{dia} = -\frac{1}{2}\left(\frac{N_e\mu_0^2}{\epsilon_F}\right)\mathcal{B}_z \ ,$$

as we shall see later on, so that the total magnetization for a free electron gas (conduction electrons in a metal) is given by

$$\mathfrak{M}_z = \frac{N_e\mu_0^2}{\epsilon_F}\mathcal{B}_z \ . \tag{4.137}$$

As a matter of interest, note that this result would be obtained if we replaced the temperature T with the Fermi temperature T_F,

$$\epsilon_F \equiv k_B T_F \ , \tag{4.138}$$

in the classical formula (equation 4.120). Could you have argued for this without the benefit of the preceding derivation?

[#] W. Pauli, Z. Physik (1927), 41, 81

References

4.1. M. Born and E. Wolf, Principles of Optics (Pergammon Press, 1975, 5th Ed.), section 7.6.

4.2. M. Klein, Optics, (Pergammon Press, 1970), section 5.6.

Problems

4.1 A free particle is described at $t=0$ by the wavefunction (k_0 real and positive)

$$\psi(x,0) = (2\pi a^2)^{-1/4}\, e^{-x^2/4a^2}\, e^{ik_0 x} .$$

(i) Find $\langle x \rangle$, $\Delta x^2 = \langle (x - \langle x \rangle)^2 \rangle$ and $\langle p_x \rangle$ at $t=0$.

(ii) Find $\psi(x,t)$

(iii) Find $\langle x \rangle$, Δx^2 and $\langle p_x \rangle$ for all times.

Plot $\Delta x = \sqrt{\Delta x^2}$ as a function of time. Describe the motion of the free particle in your own words.

4.2 A particular measurement of the position and velocity of a (classical) free particle moving along a straight line at time $t=0$ yielded $x(0) = x_0$ and $\dot{x}(0) = v_0$, on the basis of which, the resulting trajectory was computed. In order to increase the confidence of the prediction, a multiple measurement experiment was undertaken as a result of which many estimates of x_0 and v_0 on the same particle were obtained. The histograms of the two sets of measurements (of x_0 and v_0) were found to approximate gaussian curves with standard deviations Δx_0 and Δv_0 for the x_0 and v_0 measurements respectively. You may assume, for the purposes of this problem, that the experimental accuracy for the determination of x_0 and v_0 was limited by small signal-to-noise ratios resulting in a product $m\Delta x_0 \Delta v_0$ several orders of magnitude larger than $\hbar$. On the basis of these data, give a quantitative description of the predicted motion of the particle. Discuss the difference between this result and that of problem 4.1.

4.3 Discuss the behavior of the particle beam in section 4.2.1 for $0 < E < V_0$.

4.4 Prove the asymptotic estimate of equation 4.64.

4.5 Compute the normalization constant, in closed form, for the wavefunctions in equation 4.56.

4.6 Compute the probability that the particle will be found <u>outside</u> the well for a rectangular well for which

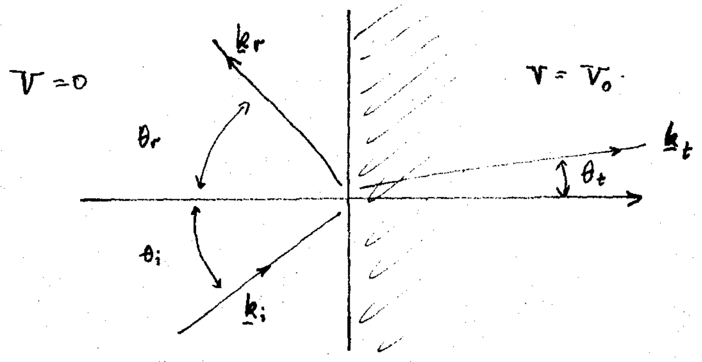
$$\frac{mV_0 a^2}{2\hbar^2} = -30 ,$$

for each of the bound states.

Problems (continued)

4.7 A two-dimensional potential is given by

$$V(x,y) = \begin{cases} V_0 & , \quad x > 0 \\ 0 & , \quad x < 0 \end{cases} .$$



A monoenergetic ($E > V_0$), non-relativistic beam of non interacting particles of mass m is incident on the $x = 0$ interface at an angle θ_i with respect to the perpendicular. See figure.

(i) Solve the Schrödinger equation. In particular, compute the angles and coefficients of reflection and transmission.

(ii) Solve the problem clasically.

4.8 An attractive one-dimensional potential for a particle of mass m, is approximated by

$$V(x) = -\frac{\hbar^2}{2ma}\,\delta(x) \quad ,$$

where $\delta(x)$ is the Dirac delta function.

a. Find the eigenfunctions and energies of all the bound states.

b. A beam of monoenergetic ($E > 0$) non-interacting particles of mass m travelling from left to right is incident on the potential described above. Derive the amplitude of the reflected and transmitted beams.

4.9 An electron is bound in a pair of attractive delta functions spaced by a distance a.

$$V(x) = -\frac{\hbar^2}{2mb}\,[\delta(x+a/2) + \delta(x-a/2)] \quad .$$

Plot the energies of the bound states as a function of the dimensionless separation $(a/2b)$.

Problems (continued)

4.10 Matching the bound state energy of an electron in a delta function potential (problem 4.8) to the ground state energy of the hydrogen atom (Rydberg)

$$-E_R = \frac{1}{2}\left(\frac{e^2}{4\pi\epsilon_0}\right)\frac{1}{a_B} = -13.6 \text{ eV} \quad ,$$

where a_B is the Bohr radius given by

$$a_B = \frac{\hbar^2}{m}\left(\frac{4\pi\epsilon_0}{e^2}\right) \approx 0.53 \times 10^{-8} \text{ cm} \quad ,$$

express the required delta function potential strength b in terms of a_B.

4.11 From handbook data, we find that the density of lithium (A.W. = 7) is given by

$$\rho_{Li} \simeq 0.534 \text{ grams/cm}^3 \quad .$$

(i) Find the conduction electron number density

Note: $m_p \sim 1.67 \times 10^{-24}$ grams.

(ii) Find the Fermi energy in electron volts

Note: $m_e \simeq 0.51$ MeV

$\hbar c \simeq 1.97 \times 10^{-5}$ eV-cm

(iii) Using the data in figure 2.4 (p. 2.31) compute the work function for lithium and the depth of the potential well for the conduction electrons.

(iv) Find the speed of an electron with an energy $E = \epsilon_F$. Compute the numerical value for lithium.

4.12 (i) Prove equation 4.114.

(ii) Prove equations 4.116 and 4.118.

(iii) Compute the electric field strength (Volts/meter) on the surface of the metal, required to decrease max $\{\kappa(x)\}$ by 10% for lithium.

4.13 Show that, there is a second order decrease of the Fermi energy with $\mathcal{B}_z$, given by

$$\frac{\epsilon_F(\mathcal{B}_z)}{\epsilon_F(0)} \sim 1 - \frac{1}{4}\left[\frac{\mu_0\mathcal{B}_z}{\epsilon_F(0)}\right]^2 \quad .$$

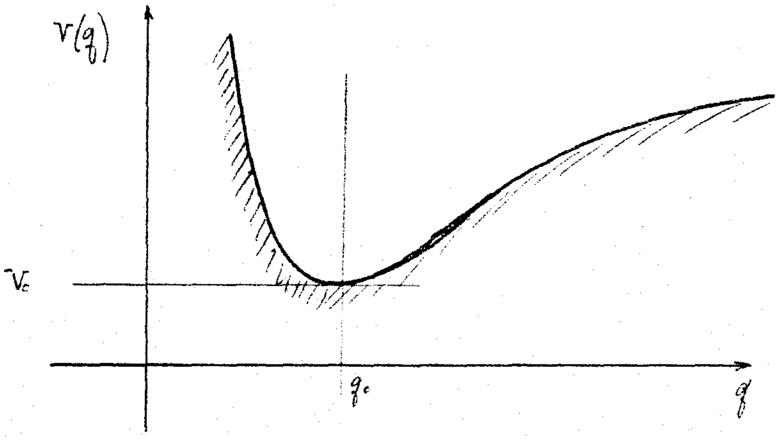
Problems (continued)

4.14 (i) Compute the energy of a free particle in a three-dimensional box in thermal equilibrium at a temperature T.

(ii) Compute the particle's heat capacity.

5. HARMONIC OSCILLATOR SYSTEMS

In many physical systems, the potential energy $V(q)$ will have a minimum at some value of the coordinate q, say q_0, which classically would be the rest position of the system.



In many cases of interest, where the excitation energy of the system E is small (measured from V_0), we can approximate the potential in the vicinity of q_0 by a Taylor expansion

$$V(q) \sim V_0 + \frac{1}{2} V''(q_0)(q - q_0)^2 + \ldots.. \tag{5.1}$$

The Lagrangian for the motion of a particle of mass m in this potential then becomes

$$L(\underline{x}, \dot{\underline{x}}, t) \sim \frac{1}{2} m|\dot{x}|^2 - \frac{1}{2} m\omega^2 x^2 \tag{5.2}$$

where x is measured from q_0, i. e.

$$x = q - q_0 \quad , \tag{5.3}$$

and where

$$\omega^2 \equiv \frac{1}{m} V''(q_0) \quad , \tag{5.4}$$

which we can identify as the classical frequency (squared) of oscillation. The constant V_0 is not included in the Lagrangian as not affecting the equations of motion, appearing only as a shift in the origin in the measurement of the energy.

From the Lagrangian of equation 5.2 we then define the Hamiltonian

$$H(x, p) \equiv p\dot{x} - L$$

or

$$H(x, p) = \frac{1}{2m} p^2 + \frac{1}{2} m\omega^2 x^2 \quad . \tag{5.5}$$

5.1. Eigenvalues and eigenfunctions

The transition to Quantum Mechanics may be made by replacing the conjugate momentum p to the coordinate x by

$$p = -i\hbar \frac{\partial}{\partial x} \quad , \tag{5.6}$$

leading to the (time independent) Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2} m\omega^2 x^2\right) u(x) = E\, u(x) \quad . \tag{5.7}$$

To solve this equation we introduce the natural dimensionless variables

$$\xi = \left(\frac{m\omega}{\hbar}\right)^{\frac{1}{2}} x \, , \qquad \beta = \frac{2E}{\hbar\omega} \, , \qquad v(\xi) = \left(\frac{\hbar}{m\omega}\right)^{\frac{1}{4}} u(x) \quad , \tag{5.8}$$

in terms of which the Schrödinger equation becomes

$$\left(\frac{d^2}{d\xi^2} + \beta - \xi^2\right) v(\xi) = 0 \quad . \tag{5.9}$$

We have encountered this equation before in the context of the minimum joint spread problem in chapter 2. Its analytical solution is discussed in Appendix C. There we found that the eigenvalues β are discrete and given by

$$\beta_n = 2n + 1$$

so that the energies E are given by

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \quad . \tag{5.10}$$

The wavefunction, corresponding to the n^{th} eigenvalue was also found to be given by

$$v_n(\xi) = A_n h_n(\xi)\, e^{-\xi^2/2} \quad , \tag{5.11}$$

where $h_n(\xi)$ is the n^{th} Hermite polynomial, defined by

$$h_n(\xi) = (-1)^n e^{\xi^2} \frac{d^n}{d\xi^n} e^{-\xi^2} \quad ,$$

and A_n is the normalization factor, given by

$$A_n = (\pi^{1/2} 2^n n!)^{-1/2} \quad . \tag{5.12}$$

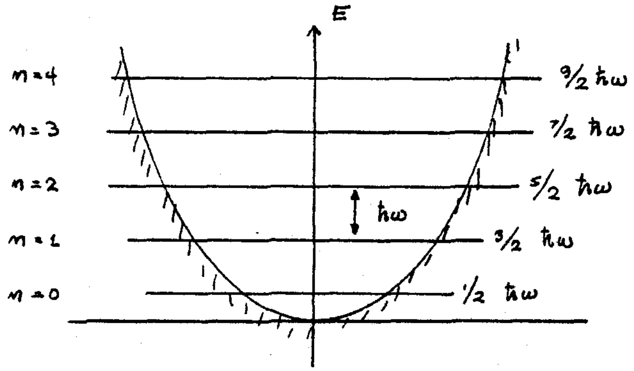
There are several noteworthy features of this solution:

(i) The lowest energy level (ground state) is characterized by a non-zero energy, i.e.

$$E_0 = \frac{1}{2}\hbar\omega \quad ,$$

as expected for a system localized to a region in space,

(ii) Subsequent excited levels are spaced uniformly by $\hbar\omega$ above E_0,



Recall[#] that for an infinite square well

$$E_n - E_0 \propto (n+1)^2 \ , \qquad \Delta E_n \equiv E_n - E_{n-1} \propto 2n+1 \quad .$$

(iii) The ground state wavefunction (a normalized Gaussian) satisfies the minimum joint spread condition as was shown in chapter 1. The reason for this is that the harmonic oscillator Hamiltonian, whose expectation value is minimized by the ground state wavefunction, is in fact the sum of the mean square deviation in p and x space respectively.

5.2. Ladder operators.

In discussing the behavior of systems described by a harmonic oscillator Hamiltonian, it is useful to define the Hermitian adjoint operators

$$a \equiv \frac{1}{\sqrt{2}} (\xi + i\kappa) \tag{5.13a}$$

$$a^\dagger \equiv \frac{1}{\sqrt{2}} (\xi - i\kappa) \ , \tag{5.13b}$$

where ξ is defined in equation 5.8 and κ is given by

$$\kappa = \frac{1}{(m\hbar\omega)^{1/2}} \, p \quad . \tag{5.14}$$

We then have, by direct substitution, that the commutator $[a, a^\dagger]$ is given by

$$[a, a^\dagger] = 1 \tag{5.15}$$

and the Hamiltonian is given by

[#] Section 4.2.5.

$$H = \frac{\hbar\omega}{2}\,(aa^\dagger + a^\dagger a) \quad ,$$

or, using the commutation relation 5.15

$$H = \hbar\omega\,(a^\dagger a + \tfrac{1}{2}) \quad . \tag{5.16}$$

We know, however, that H operating on a wavefunction $u_n(x)$ yields

$$Hu_n = \hbar\omega\,(n + \tfrac{1}{2})\,u_n$$

and can therefore identify the operator product $a^\dagger a$ as the number operator, i.e.

$$a^\dagger a = \hat{n} \quad . \tag{5.17}$$

What is the result of operating on an eigenvector with the operators a or $a^\dagger$ alone? Consider the quantity $(a^\dagger a)a|n\rangle$, where $|n\rangle$ denotes the n^{th} normalized eigenvector [eigenfunction $u_n(x)$]. Then we have[#]

$$\begin{aligned}(a^\dagger a)a|n\rangle &= (aa^\dagger - 1)a|n\rangle \\ &= a(a^\dagger a - 1)|n\rangle \\ &= a(n - 1)|n\rangle \\ &= (n - 1)a|n\rangle \quad ,\end{aligned}$$

or since

$$(a^\dagger a)|m\rangle = m|m\rangle \quad ,$$

the result of $a|n\rangle$ must be a simple multiple of the state $|n - 1\rangle$. We accordingly write

$$a|n\rangle = c_n|n - 1\rangle \quad , \tag{5.18}$$

and consider the scalar product

$$\langle n|a, a|n\rangle \equiv \int_x [a\,u_n(x)]^*\,[a\,u_n(x)]\,dx \quad .$$

Substituting equation 5.18, we have

$$\langle n|a, a|n\rangle = \langle n - 1|c_n^*, \; c_n|n - 1\rangle = |c_n|^2 \quad , \tag{5.19}$$

since $\langle n - 1|n - 1\rangle = 1$. We can also use the fact, however, that the operators a and $a^\dagger$ are Hermitian adjoints of each other to obtain

$$\langle n|a, a|n\rangle = \langle n|, a^\dagger a|n\rangle = \langle n|, n|n\rangle = n \quad , \tag{5.20}$$

where we again have used that the states $|n\rangle$ are normalized. Comparing equations 5.19 and 5.20 we must then have

$$|c_n|^2 = n \quad \Rightarrow \quad c_n = n^{1/2} \quad , \tag{5.21}$$

within a phase factor which without loss of generality may be set equal to unity. Consequently, we have

$$a|n\rangle = n^{1/2}|n - 1\rangle \quad . \tag{5.22}$$

[#] note from equation 5.15 that $aa^\dagger - a^\dagger a = 1$.

We see that the result of the operation of a on the state $|n\rangle$ is to produce the next lowest state or, equivalently, to annihilate a quantum of energy $\hbar\omega$. For this reason we call the operator a the <u>annihilation operator</u>. Combining equations 5.17 and 5.22 we also see that

$$a^\dagger|n\rangle = (n+1)^{1/2}|n+1\rangle \quad . \tag{5.23}$$

The operator $a^\dagger$ is correspondingly called the <u>creation operator</u>. The operators a and $a^\dagger$ are called <u>ladder</u> operators because they can be used to go up and down the ladder of eigenvalues. They are useful in problems possessing a uniformly spaced eigenvalue spectrum.

Using these operators, we can solve the Schrödinger equation rather effortlessly. In particular, the ground state, corresponding to $n=0$, must be the solution to the equation

$$a|0\rangle = 0 \quad , \tag{5.24}$$

which we may solve in our coordinate representation, if necessary, by substituting equation 5.13a for the annihilation operator, i.e.

$$(\xi + \frac{\partial}{\partial\xi})\ v_0(\xi) = 0 \quad , \tag{5.24'}$$

whose direct solution

$$v_0(\xi) = \text{const.}\ e^{-\xi^2/2} \quad ,$$

was obtained much more painfully by solving the second order differential Schrödinger equation in appendix C. Having the ground state, we may obtain all the higher excited states by repeated application of the creation operator, i.e.

$$\begin{aligned} |1\rangle &= a^\dagger|0\rangle \\ |2\rangle &= \frac{1}{\sqrt{2}}\ a^\dagger|1\rangle = \frac{1}{\sqrt{2}}\ (a^\dagger)^2\ |0\rangle \\ &\ \ \vdots \\ |n\rangle &= (n!)^{-1/2}\ (a^\dagger)^n\ |0\rangle \quad . \end{aligned} \tag{5.25}$$

If it is necessary to obtain the eigenfunctions in the coordinate representation, we substitute equation 5.13b to obtain, for example,

$$\begin{aligned} v_1(\xi) &= 2^{-1/2}(\xi - \frac{\partial}{\partial\xi})\ v_0(\xi) \\ &= 2^{-1/2}\ A_0\ (\xi - \frac{\partial}{\partial\xi})\ e^{-\xi^2/2} \\ v_1(\xi) &= 2^{-1/2}\ A_0\ (2\xi)\ e^{-\xi^2/2} \quad , \end{aligned}$$

which is correct including the proper normalization constant $A_1 = 2^{-1/2} A_0$.#

#Note that $h_1(\xi) = 2\xi$. See appendix C.

There is a one-to-one correspondence between the x space with the eigenfunctions $u_n(x)$ and the scalar product as defined in chapter 1, and the <u>occupation number</u> space with eigenvectors $|n\rangle$ and the scalar product that was implicitly defined in the preceding discussion, i.e.

$$\int_x u_m^*(x)\hat{Q}(\hat{x},\hat{p})\,u_n(x)\,dx = \langle m|\hat{Q}[\hat{x}(a^\dagger,a),\hat{p}(a^\dagger,a)]\,|n\rangle \tag{5.26}$$

where, by inverting the defining equations,

$$\hat{x} = \left(\frac{\hbar}{2m\omega}\right)^{1/2}(a^\dagger + a) \tag{5.27a}$$

and

$$\hat{p} = i\left(\frac{m\hbar\omega}{2}\right)^{1/2}(a^\dagger - a)\ . \tag{5.27b}$$

We can think of the sequence of eigenvectors $|n\rangle$ as forming the basis of an infinite dimensional vector space, i.e.

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \end{pmatrix},\qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \end{pmatrix},\qquad |n\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow n^{\text{th}}\ \text{slot} \tag{5.28}$$

Operators in this space are matrices. By way of example

$$a = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \sqrt{2} & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \sqrt{3} & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & & & & \\ \cdot & & & & & \cdot & & & \\ \cdot & & & & & & \cdot & & \\ \cdot & & & & & & & \cdot & \end{pmatrix} \tag{5.29a}$$

whereas

$$a^\dagger = \begin{pmatrix} 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \sqrt{1} & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \sqrt{2} & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \sqrt{3} & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & & & & & \\ \cdot & & & & & & & & \\ \cdot & & & & & & & & \end{pmatrix} \tag{5.29b}$$

Update

It can easily be verified by direct substitution that these basis vectors and matrix operators satisfy the fundamental relations as given by equations 5.22 and 5.23.

5.3. Harmonic oscillators in thermal equilibrium.

Any state a harmonic oscillator can assume can be represented by a superposition of the eigenfunctions $u_n(x)$, i.e.

$$\psi(x,t) = \sum_n c_n(t)\, u_n(x) \tag{5.30}$$

or, equivalently,

$$|\psi\rangle = \sum_n |n\rangle\, c_n(t) \tag{5.31a}$$

where

$$c_n(t) = c_n(0)\, e^{-iE_n t/\hbar} \tag{5.31b}$$

and $c_n(0)$ can be determined from the initial conditions, i.e.,

$$c_n(0) = \int u_n^*(x)\psi(x,0)\,dx = \langle n|\psi\rangle\big|_{t=0} \quad . \tag{5.31c}$$

The energy of the state is then given by

$$\langle E\rangle = \langle\psi|H|\psi\rangle$$

$$= \sum_n E_n |c_n|^2$$

$$\langle E\rangle = \hbar\omega \sum_n (n+\tfrac{1}{2})\,|c_n|^2 \quad . \tag{5.32}$$

Now for a system in thermal equilibrium with the surroundings at a temperature T, we have the Boltzmann relation,

$$|c_n|^2 = \frac{e^{-E_n/k_BT}}{\sum_m e^{-E_m/k_BT}} \quad , \tag{5.33}$$

and therefore

$$\langle E\rangle = \hbar\omega \frac{\sum_n (n+\frac{1}{2})\, e^{-x(n+\frac{1}{2})}}{\sum_n e^{-x(n+\frac{1}{2})}}$$

or

$$\langle E\rangle = -\hbar\omega \frac{\partial}{\partial x}\, \ell n \left[\sum_n e^{-x(n+\frac{1}{2})}\right] \quad , \tag{5.34}$$

where $x = \hbar\omega/k_BT$.

Now

$$\ell n\left[\sum_n e^{-x(n+\frac{1}{2})}\right] = \ell n\left[e^{-x/2}\sum_n\left(e^{-x}\right)^n\right]$$

$$= -\frac{x}{2} - \ell n\left(1 - e^{-x}\right)$$

and therefore

$$\langle E\rangle = \frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\hbar\omega/k_BT} - 1} \ . \tag{5.35}$$

Note that at high temperatures, i.e., $\hbar\omega/k_BT << 1$,

$$\langle E\rangle \simeq k_BT \ ,$$

in agreement with classical Maxwell-Boltzmann statistics and the correspondence principle.

5.4 Systems of uncoupled harmonic oscillators.

We can write the Hamiltonian of a system of N uncoupled harmonic oscillators as a sum

$$H(\underline{p},\underline{q}) = \sum_{j=1}^{N} H_j(p_j, q_j) \tag{5.36a}$$

where

$$H_j(p_j, q_j) = \frac{1}{2m_j} p_j^2 + \frac{1}{2} m_j\omega_j^2 q_j^2 \ , \tag{5.36b}$$

and $\underline{p}$ and $\underline{q}$ are the N-dimensional vectors

$$\underline{p} = (p_1, p_2, \ldots\ldots, p_N) \tag{5.37a}$$

$$\underline{q} = (q_1, q_2, \ldots\ldots, q_N) \ . \tag{5.37b}$$

Note that if the oscillators of the system were coupled, the Hamiltonian could not be separated into the sum of such terms, but would involve cross terms between the coordinates. By way of example, if the interaction between the various coordinates is a pair coupling, the Hamiltonian would be given by

$$H(\underline{p},\underline{q}) = \sum_j H_j(p_j, q_j) + \sum_{i\neq j} V_{ij}(p_i, p_j; q_i, q_j) \ . \tag{5.38}$$

See problem 5.8.

The Schrödinger equation for the system of the N uncoupled oscillators is given by

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\Psi(\underline{q}, t) &= H\,\Psi(\underline{q}, t) \\ &= \left(\sum_j H_j\right)\Psi(\underline{q}, t) \ . \end{aligned} \tag{5.39}$$

Equation 5.39 can be solved by separation of variables, and in particular, by assuming a wavefunction of the form

$$\Psi(\underline{q}, t) = \prod_j \psi_j(q_j, t) \, , \tag{5.40}$$

where

$$i\hbar \frac{\partial}{\partial t} \psi_j(q_j, t) = H_j \, \psi_j(q_j, t) \; ; \quad i = 1, N \, . \tag{5.41}$$

The overall system is in an eigenstate, if every one of the N oscillators is in a harmonic oscillator eigenstate, i.e.,

$$\Psi_{\underline{n}}(\underline{q}, t) = U_{\underline{n}}(\underline{q}) \, e^{-iE_{\underline{n}} t/\hbar} \tag{5.42}$$

where

$$U_{\underline{n}}(\underline{q}) = \prod_j u_{n_j}(q_j) \tag{5.43}$$

and

$$E_{\underline{n}} = \sum_j \varepsilon_{n_j} = \sum_j \hbar\omega_j (n_j + \tfrac{1}{2}) \, . \tag{5.44}$$

The vector $\underline{n}$ is an N dimensional index that labels the eigenstate of the total system, i.e.,

$$\underline{n} = (n_1, n_2, \ldots\ldots, n_N) \, .$$

We can now define, an alternate set of basis eigenvectors, by analogy to the single harmonic oscillator, in particular,

$$|\underline{n}\rangle \equiv |n_1, n_2, \ldots\ldots, n_N\rangle \tag{5.45}$$

and also annihilation and creation operators, corresponding to each pair of conjugate coordinates p_j, q_j, i.e.,

$$a_j = \left(\frac{1}{2\hbar}\right)^{1/2} \left[(m_j\omega_j)^{1/2} q_j + i(m_j\omega_j)^{-1/2} p_j \right] \tag{5.46a}$$

and

$$a_j^\dagger = \left(\frac{1}{2\hbar}\right)^{1/2} \left[(m_j\omega_j)^{1/2} q_j - i(m_j\omega_j)^{-1/2} p_j \right] \, , \tag{5.46b}$$

such that

$$a_j|\underline{n}\rangle = n_j^{1/2} |n_1, n_2, \ldots\ldots, n_j - 1, \ldots\ldots, n_N\rangle \tag{5.47a}$$

and

$$a_j^\dagger|\underline{n}\rangle = (n_j + 1)^{1/2} |n_1, n_2, \ldots\ldots, n_j + 1, \ldots\ldots, n_N\rangle \tag{5.47b}$$

The total Hamiltonian can then be expressed quite simply as

$$H = \sum_j \hbar\omega_j (a_j^\dagger a_j + \tfrac{1}{2}) \, . \tag{5.48}$$

Any state of the system can be expressed as a linear superposition of the $|\underline{n}\rangle$ eigenstates, i.e.,

$$|\Psi(t)\rangle = \sum_{n} |\underline{n}\rangle\, c_{\underline{n}}\, e^{-iE_{\underline{n}}t/\hbar} , \tag{5.49}$$

where

$$c_{\underline{n}} = \langle \underline{n}|\Psi(0)\rangle \tag{5.50}$$

or, equivalently

$$c_{\underline{n}} = \iint \cdots\cdots \int U_{\underline{n}}^{*}(\underline{q})\, \Psi(q,0)\, d^{N}\underline{q} \tag{5.51}$$

or, using equation 5. 40, we have for uncoupled oscillators,

$$c_{\underline{n}} = \prod_{j} \int_{q_j} u_{n_j}^{*}(q_j)\, \psi_j(q_j, 0)\, dq_j \ . \tag{5.52}$$

$$= \prod_{j} c_{n_j} \ .$$

The energy of such a state can then be computed by

$$\langle E\rangle = \langle \Psi|H|\Psi\rangle = \sum_{\underline{n}} E_{\underline{n}} |c_{\underline{n}}|^{2} \tag{5.53}$$

$$= \sum_{n_1} \sum_{n_2} \cdots\cdots \sum_{n_N} \left(\sum_{j} E_{n_j}\right) \prod_{j} |c_{n_{j'}}|^{2}$$

$$= \sum_{j} \sum_{n_j} E_{n_j} |c_{n_j}|^{2} \left(\sum_{n_1} \sum_{n_2} \cdots \sum_{n_{j-1}} \sum_{n_{j+1}} \cdots \sum_{n_N} \prod_{j'\neq j} |c_{n_{j'}}|^{2} \right)$$

$$= \sum_{j} \sum_{n_j} E_{n_j} |c_{n_j}|^{2} \prod_{j'\neq j} \left(\sum_{n_{j'}} |c_{n_{j'}}|^{2} \right)$$

or, since

$$\sum_{n_{j'}} |c_{n_{j'}}|^{2} = 1 \ ,$$

we have

$$\langle E\rangle = \sum_{j} \left(\sum_{n_j} |c_{n_j}|^{2} E_{n_j} \right) . \tag{5.54}$$

5. 5 Systems of uncoupled harmonic oscillators in thermal equilibrium.

If this system of oscillators is in thermal equilibrium at some temperature T, then the probability of finding the j^th oscillator at a level n_j would be given by a Boltzmann factor

$$|c_{n_j}|^2 = \frac{e^{-E_{n_j}/k_B T}}{\sum_n e^{-E_n/k_B T}}$$

which, substituting in equation 5. 54, yields

$$\langle E\rangle = \sum_j \left[\sum_{n_j} \left(\frac{E_{n_j} e^{-E_{n_j}/k_B T}}{\sum_n e^{-E_n/k_B T}} \right) \right] , \tag{5. 55}$$

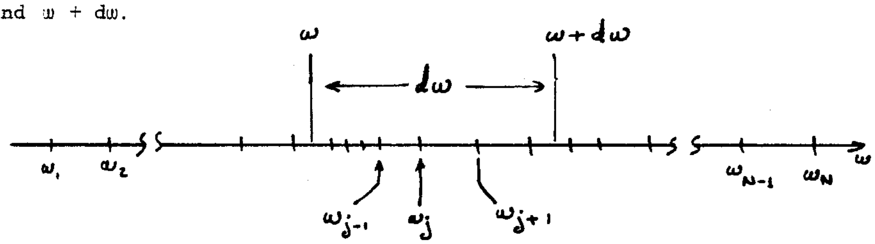
We note that each one of the terms in the summation over j, is equal to the energy at thermal equilibrium of an isolated harmonic oscillator and therefore, taking the result of equation 5.35, we have,

$$\langle E\rangle = \frac{1}{2} \sum_j \hbar\omega_j + \sum_j \frac{\hbar\omega_j}{e^{\hbar\omega_j/k_B T} - 1} . \tag{5. 56}$$

The first sum is the ground state energy E_G of the total system, which is fixed. The second sum, if the oscillators are many with closely spaced frequencies ω_j, we can approximate by an integral over ω, i. e.,

$$\langle E\rangle - E_G = \int_0^\infty \rho(\omega) \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} d\omega \tag{5. 57}$$

where $\rho(\omega) d\omega$ is the number of oscillators whose angular frequency is between ω and $\omega + d\omega$.



In any event, the summation can be recovered from the integral using

$$\rho(\omega) = \sum_j \delta(\omega - \omega_j) \tag{5. 58}$$

for the density of oscillators.

The integrand

$$S(\omega)d\omega = \rho(\omega) \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} d\omega \quad . \tag{5.59}$$

is called the spectrum of the energy.

Note that if we substitute for the density of oscillators $\rho(\omega)$, the density of modes of electromagnetic radiation in a box, i. e.

$$\rho(\omega)\, d\omega = \frac{\omega^2 d\omega}{\pi c^3} \quad , \tag{5.60}$$

see problem 2. 16, we obtain the black body spectrum of Planck, i. e. equation 2. 130. This result suggests that the light quanta of Planck behave like the excitation levels of harmonic oscillators whose fundamental frequency ω is that of the corresponding photon! Taken at face value it is a less than obvious model to associate with light, to be sure. At first sight, it would be difficult to assign a mass and a spring to a photon which is massless and not bound to any obvious equilibrium position by a quadratic potential! Nevertheless, we will show that there exist a set of coordinates for the photon field, in terms of which the Hamiltonian is given by equation 5. 36, or equivalently, equation 5. 48. To show that, of course, would be sufficient and we will then have derived Planck's formula through a rather unexpected route.

5.6 Quantum mechanics of a fluid. Phonons.

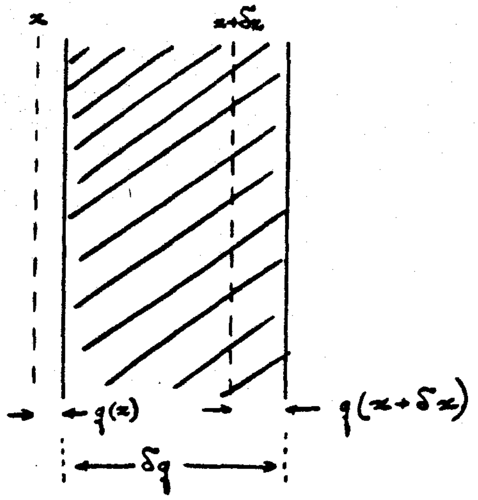
We can always treat a matter as a continuum when the smallest length scale of interest, say ℓ, is much larger than the (mean) spacing between the constituent particles. In other words, we can neglect the atomic nature of gases, liquids and solids if

$$\ell \left(\frac{\rho_m}{m}\right)^{1/3} \gg 1 \quad , \tag{5.61}$$

where ρ_m is the mass density and m is the mass of each particle. This approximation is usually associated with a classical description of matter. It should be noted, however, that the limit of equation 5.61 is quite different from the reasons that dictate a quantum versus a classical description of a phenomenon. The latter decision is based on whether the size of $\hbar$ (Planck's constant) is negligible or not. For many systems, especially at low temperatures where only the lowest energy levels are excited, a quantum mechanical treatment of the continuum behavior is in fact appropriate. This we will do by following the usual procedure of deriving the Lagrangian, the conjugate momenta and the Hamiltonian. The transition to Quantum Mechanics is then made by selecting a representation for the conjugate coordinate-momentum pairs such that

$$[q_j, p_j] = i\hbar \quad . \tag{5.62}$$

In describing the motion of the medium, consider the displacement q along the x-axis, of a medium element which has a rest position at x.



If we now compare $q(x)$, the displacement of the element from x, to $q(x+\delta x)$, the displacement of the element from $x+\delta x$, we have that the difference in the displacements δq is given by

$$\begin{aligned} \delta q &= \delta x + q(x+\delta x) - q(x) \\ &\simeq \delta x\left(1 + \frac{dq}{dx}\right) \quad . \end{aligned} \tag{5.63}$$

Now, if the density of the medium at rest is given by ρ_0, we must also have that

$$\rho_0 \, \delta x = \rho \, \delta q \quad , \tag{5.64}$$

where ρ is the resulting density of the medium between $x+q$ and $x+\delta x+q+\delta q$. Combining equations 5.63 and 5.64 we have

$$\frac{\rho_0}{\rho} = 1 + \frac{dq}{dx} \quad . \tag{5.65}$$

If we now consider a more general displacement in three dimensions, of a medium element originally at rest at $\underline{x}$, i. e.

$$\underline{q}(\underline{x}) = [q_x(\underline{x}), \; q_y(\underline{x}), \; q_z(\underline{x})] \quad ,$$

we can show by similar arguments that

$$\frac{\rho_0}{\rho(\underline{x}, t)} = 1 + \frac{\partial}{\partial \underline{x}} \cdot \underline{q}(\underline{x}, t) \quad , \tag{5.66}$$

where $\frac{\partial}{\partial \underline{x}} \cdot \underline{q}$ is the divergence of the displacement field.

Consider now the case of small departures of the density from its equilibrium value. We then have

$$\frac{\rho_0}{\rho} = \frac{\rho_0}{\rho_0 + (\rho - \rho_0)} = \frac{1}{1 + (\frac{\rho - \rho_0}{\rho_0})} \simeq 1 - (\frac{\rho - \rho_0}{\rho_0}) \tag{5.67}$$

or, using equation 5.66,

$$\frac{\rho(\underline{x}, t) - \rho_0}{\rho_0} \simeq - \frac{\partial}{\partial \underline{x}} \cdot \underline{q}(\underline{x}, t) \quad . \tag{5.68}$$

To define the Lagrangian, we note that the kinetic energy per unit volume of the medium is given by

$$\frac{1}{2} \rho(\underline{x}, t) \; |\dot{\underline{q}}(\underline{x}, t)|^2$$

and therefore the total kinetic energy of the medium is given by

$$K = \frac{1}{2} \int \rho(\underline{x}, t) \; |\dot{\underline{q}}(\underline{x}, t)|^2 \, d^3\underline{x} \quad , \tag{5.69}$$

or, to lowest order in the displacement field,

$$K \simeq \frac{\rho_0}{2} \int |\dot{\underline{q}}(\underline{x}, t)|^2 \, d^3\underline{x} \quad . \tag{5.70}$$

We will now restrict the discussion to matter that

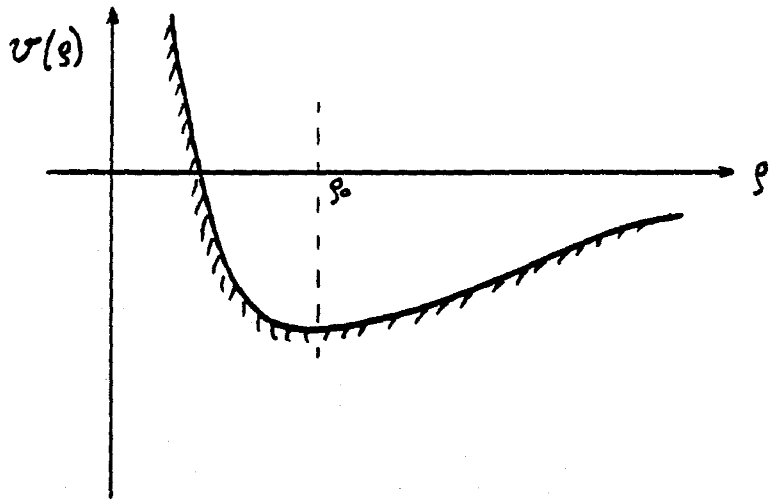
(i) possesses an equilibrium density, independent of container boundaries (i. e. not a gas),

(ii) is isotropic

and

(iii) cannot support static shear (i. e. a fluid).

These conditions are met by a liquid, for which we can consider the potential energy per unit volume as possessing a minimum at the equilibrium value $\rho = \rho_0$, as sketched below.



For small deviations from the equilibrium density we can then approximate the potential energy per unit volume by a Taylor expansion about its equilibrium value ρ_0, i.e.

$$\mathcal{V}(\rho) \approx \mathcal{V}(\rho_0) + \frac{1}{2}(\rho - \rho_0)^2 \, \mathcal{V}''(\rho_0) \tag{5.71}$$

and therefore

$$U \approx \frac{1}{2} \mathcal{V}''(\rho_0) \int_{\underline{x}} [\rho(\underline{x}, t) - \rho_0]^2 \, d^3\underline{x} \quad , \tag{5.72}$$

or, using equation 5.68,

$$U = \frac{1}{2} \rho_0^2 \mathcal{V}''(\rho_0) \int_{\underline{x}} \left[\frac{\partial}{\partial \underline{x}} \cdot \underline{q}(\underline{x}, t)\right]^2 d^3\underline{x} \quad . \tag{5.73}$$

Combining equations 5.70 and 5.73 we then have for the Lagrangian

$$L \approx \int \left\{ \frac{1}{2} \rho_0 \, |\dot{\underline{q}}(\underline{x}, t)|^2 - \frac{1}{2} \rho_0 \mathcal{V}''(\rho_0) \left[\frac{\partial}{\partial \underline{x}} \cdot \underline{q}(\underline{x}, t)\right]^2 \right\} d^3\underline{x} \quad . \tag{5.74}$$

Let us now assume that the displacement field satisfies periodic boundary conditions at the edge planes of a rectangular box of volume $V = L_x L_y L_z$, i.e.

$$\underline{q}(-L_x/2, y, z, t) = \underline{q}(L_x/2, y, z, t) \tag{5.75a}$$

$$\underline{q}(x, -L_y/2, z, t) = \underline{q}(x, L_y/2, z, t) \tag{5.75b}$$

and

$$\underline{q}(x, y, -L_z/2, t) = \underline{q}(x, y, L_z/2, t) \quad . \tag{5.75c}$$

If we now assume that the displacement field is irrotational, i.e. $\frac{\partial}{\partial \underline{x}} \times \underline{q} = 0$, we can expand the space dependence of $\underline{q}(\underline{x}, t)$ into a Fourier series of the form

$$\underline{q}(\underline{x}, t) = \sum_{\underline{k}} Q_{\underline{k}}(t) \, e^{i\underline{k} \cdot \underline{x}} \, \hat{e}_{\underline{k}} \tag{5.76}$$

where, from the periodic boundary conditions (equation 5.75), the permissible wavevectors are given by

$$\underline{k} = 2\pi \left(\frac{\ell}{L_x} \hat{e}_x + \frac{m}{L_y} \hat{e}_y + \frac{n}{L_z} \hat{e}_z \right) , \tag{5.77}$$

where

$$\ell, m, n = 0, \pm 1, \pm 2, \; \ldots . \tag{5.78}$$

Note that the orthogonality relation is given by

$$\int e^{i(\underline{k} - \underline{k}') \cdot x} \, d^3\underline{x} = V \, \delta_{\underline{k}, \underline{k}'} \tag{5.79}$$

where V is the volume, and

$$\delta_{\underline{k}, \underline{k}'} = \begin{cases} 1 & \text{if } \underline{k} = \underline{k}' \\ \\ 0 & \text{if } \underline{k} \neq \underline{k}' \end{cases} , \tag{5.80}$$

is the Kronecker delta, and therefore, the Fourier coefficients $Q_{\underline{k}}(t)$ are given by

$$Q_{\underline{k}}(t) = \frac{1}{V} \int [\hat{e}_{\underline{k}} \cdot \underline{q}(\underline{x}, t)] \, e^{-i\underline{k} \cdot \underline{x}} \, d^3\underline{x} \quad . \tag{5.81}$$

Note that the Fourier coefficients are complex, i.e.

$$Q_{\underline{k}} = \frac{1}{\sqrt{2}} (Q_{\underline{k},1} - i Q_{\underline{k},2}) \tag{5.82a}$$

where

$$Q_{\underline{k},1} = \frac{2^{\frac{1}{2}}}{V} \int [\hat{e}_{\underline{k}} \cdot \underline{q}(\underline{x}, t)] \cos(\underline{k} \cdot \underline{x}) \, d^3\underline{x} \tag{5.82b}$$

and

$$Q_{\underline{k},2} = \frac{2^{\frac{1}{2}}}{V} \int [\hat{e}_{\underline{k}} \cdot \underline{q}(\underline{x}, t)] \sin(\underline{k} \cdot \underline{x}) \, d^3\underline{x} \quad , \tag{5.82c}$$

Using the Fourier series expansion for the displacement field, we now have

$$\dot{\underline{q}}(\underline{x}, t) = \sum_{\underline{k}} \dot{Q}_{\underline{k}}(t) \, e^{i\underline{k} \cdot \underline{x}} \, \hat{e}_{\underline{k}} \tag{5.83}$$

and

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{q}(\underline{x}, t) = i \sum_{\underline{k}} k \, Q_{\underline{k}}(t) \, e^{i\underline{k} \cdot \underline{x}} \quad , \tag{5.84}$$

where $k = |\underline{k}|$. Substituting these in the expression for the Lagrangian (equation 5.74), we have

$$L = \sum_{\underline{k}} \frac{M}{2} \left\{ |\dot{Q}_{\underline{k}}|^2 - c_s^2 k^2 |Q_{\underline{k}}|^2 \right\} \tag{5.85}$$

where

$$M = \rho_0 V \tag{5.86}$$

is the total mass of the medium, and

$$c_s = [\rho_0 \mathcal{V}''(\rho_0)]^{\frac{1}{2}} \tag{5.87}$$

is a constant with units of velocity. Substituting the real and imaginary parts of $Q_{\underline{k}}$, we then have

$$L = \sum_{\underline{k},j} \frac{1}{4} M(\dot{Q}^2_{\underline{k},j} - c_s^2 k^2 Q^2_{\underline{k},j}) \qquad (5.88)$$

where $J = 1,2$.

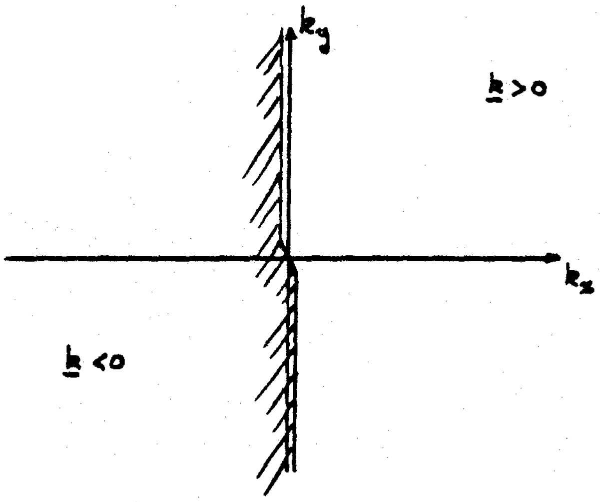
There is a slight complication that arises from the fact that the coordinates $Q_{\underline{k},j}$ and $Q_{-\underline{k},j}$ are not independent. In fact, from the defining equations 5.82a and 5.82b we see that

$$Q_{\underline{k},1} = Q_{-\underline{k},1} \qquad (5.89a)$$

and

$$Q_{\underline{k},2} = -Q_{-\underline{k},2} \quad , \qquad (5.88b)$$

consequently, even though the Lagrangian is correctly given by equation 5.88, not all the $Q_{\underline{k},j}$ can be used to form a set of orthogonal coordinates. In particular, if we use $Q_{\underline{k},j}$ as a coordinate, we must exclude $Q_{-\underline{k},j}$. The Lagrangian can be expressed, however, as a sum over some positive half space of $\underline{k}$, symbolically sketched below for two dimensions,



i.e.

$$L = \sum_{+\underline{k},j} \frac{1}{2} M(\dot{Q}^2_{\underline{k},j} - c_s^2 k^2 Q^2_{\underline{k},j}) \quad , \qquad (5.90)$$

where the summation over the positive $\underline{k}$-(half) space is denoted by the plus sign prefix of the $\underline{k}$ summation index.

Using equation 5.90, we can now define the conjugate momenta $P_{\underline{k},j}$ to the coordinates $Q_{\underline{k},j}$, i.e.

$$P_{\underline{k},j} = \frac{\partial L}{\partial \dot{Q}_{\underline{k},j}}$$

or, substituting for the Lagrangian,

$$P_{\underline{k},j} = M\dot{Q}_{\underline{k},j} \quad . \qquad (5.91)$$

Using these momenta, we can now express the Hamiltonian of the system, i.e.

$$H \equiv \sum_{+\underline{k},j} \dot{Q}_{\underline{k},j} \frac{\partial L}{\partial \dot{Q}_{\underline{k},j}} - L$$

or

$$H = \sum_{+\underline{k},j} \left[\frac{1}{2M} P^2_{\underline{k},j} + \frac{M}{2} \omega^2_{\underline{k}} Q^2_{\underline{k},j}\right] \tag{5.92}$$

where

$$\omega^2_{\underline{k}} = c_s^2 k^2 \tag{5.93}$$

We see that the fluid, in this continuum approximation, behaves like a system of uncoupled harmonic oscillators of mass M and frequency $\omega_{\underline{k}} = c_s k$, where c_s can be identified with the speed of sound. See problem 5.14. In other words

$$H = \sum_{+\underline{k},j} H_{\underline{k},j} \tag{5.94a}$$

where

$$H_{\underline{k},j} = \frac{1}{2M} P^2_{\underline{k},j} + \frac{M}{2} \omega^2_{\underline{k}} Q^2_{\underline{k},j} \quad , \tag{5.94b}$$

We can now compute the energy of the liquid in thermal equilibrium at a temperature T using the results of the preceding section. In particular, using equation 5.57 we have

$$\langle E \rangle - E_G = \hbar \int_0^\infty \frac{\omega \eta(\omega)}{e^{\hbar\omega/k_B T} - 1} d\omega \tag{5.95}$$

where $\eta(\omega)$ is the density of states per angular frequency.

To compute the density of states, we note that there are two modes for each wave-number $\underline{k}$, corresponding to $j = 1$ and $j = 2$, the real and imaginary part of the complex Fourier coefficient $Q_{\underline{k}}$ of the displacement field. Recall however, that we must include only a half space of $\underline{k}$ in counting the independent modes. Therefore the number of modes with frequency $\omega' < \omega$ is given by

$$N(\omega' < \omega) = 2 \frac{\frac{2}{3}\pi k^3(\omega)}{(2\pi)^3/V} \quad ,$$

or, since $k = \omega/c_s$

$$N(\omega' < \omega) = \frac{(4\pi/3)\omega^3 V}{(2\pi)^3 c_s^3}$$

and therefore

$$\eta(\omega) = \frac{dN}{d\omega} \quad \frac{V}{2\pi^2 c_s^3} \omega^2 \tag{5.96}$$

and therefore

$$\langle E \rangle - E_G = \frac{V\hbar}{2\pi^2 c_s^3} \int_0^\infty \frac{\omega^3 d\omega}{e^{\hbar\omega/k_B T} - 1}$$

or, changing to a dimensionless integration variable $x = \hbar\omega/k_B T$,

$$\langle E \rangle - E_G = \frac{V(k_B T)^4}{2\pi^2 c_s^3 \hbar^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} \quad . \tag{5.97}$$

The definite integral is a pure number and given by (see for example F. Reif, Fundamentals of Statistical and Thermal Physics, Appendix 11)

$$\int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15} \quad . \tag{5.98}$$

The energy is then given by

$$\langle E \rangle - E_G = \frac{\pi^2 V (k_B T)^4}{30 (c_s \hbar)^3} \quad , \tag{5.99}$$

and we find the heat capacity per unit mass,

$$c_V = \frac{2\pi^2 k_B T^3}{\rho_0 15 (\hbar c_s)^3} \quad , \tag{5.100}$$

which is seen to be proportional to the cube of the temperature.

Substituting for the Boltzmann and Planck's constants,

$$k_B = 1.3800 \times 10^{-23} \quad J/^\circ K$$
$$\hbar = 1.055 \times 10^{-34} \quad J \sec$$

and, for liquid ^{4}He at low temperatures ($T \leqslant 0.6^\circ K$)

$$\rho_0 = 0.1450 \ g/cm^3$$
$$c_s = 239 \ m/sec$$

we have the theoretical prediction for the heat capacity of liquid ^{4}He at low temperatures

$$c_s = (0.0205 \pm \ 0.0003)(T/^\circ K)^3, \ (J/g)^\circ K^{-1}$$

where the uncertainty in the theoretical prediction results from the finite accuracy with which the speed of sound is known. The measured heat capacity of liquid ^{4}He below $0.6^\circ K$ is found to be [J. Wiebes, C.G. Niels-Hakkenberg and H.C. Kramers (1957), Physica 23, 625-632],

$$c_V = (0.0204 \pm 0.0004) \ (T/^\circ K)^3 \ \ (J/g) \ ^\circ K^{-1} \quad .$$

This expression for the heat capacity fails at temperatures higher than $0.6^\circ K$ for an interesting reason: our continuum assumption fails! As the temperature increases higher frequency states become occupied and since

$$E_{\underline{k}} = \hbar \omega_{\underline{k}} = \hbar c_s k \quad ,$$

states with higher and higher wave numbers become occupied. You will recall, however, that our original assumption of a continuum, as given by equation 5.61, requires that

$$k << (\rho/m)^{1/3}$$

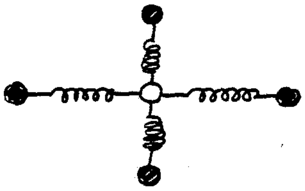
since, for a mode of excitation of wavenumber k, the appropriate length scale (wavelength) goes like 1/k. The interesting thing is that the model of a system of uncoupled harmonic oscillators continues to be valid even beyond that limit. The effect of the discrete particles (helium atoms) comprising the liquid, however, is to produce a slightly more complicated dispersion relation [$\omega = \omega(k)$]. The Hamiltonian, as given by equation 5.92, is still correct but the dependence of ω on $|\underline{k}|$ is as sketched below. See problem 5.18.

Problems

5.1 Derive the eigenfunctions and energy levels of a two-dimensional harmonic oscillator whose potential is given by

$$V(x, y) = \frac{1}{2} k \left[x^2 \cos^2 \alpha + y^2 \sin^2 \alpha\right]$$

This is a good model for the vibrations of the central atom of a planar molecule of the type sketched below



5.2 Show that the operator for κ, as defined by 5.14, is given by

$$\hat{\kappa} = -i \frac{\partial}{\partial \xi} \quad .$$

5.3 Show that the operator $a^\dagger$, as defined by equation 5.13b, is the Hermitean adjoint of the operator a.

5.4 Show that

$$[a, a^\dagger] = 1 \quad ,$$

where the a and $a^\dagger$ are defined by equation 5.13.

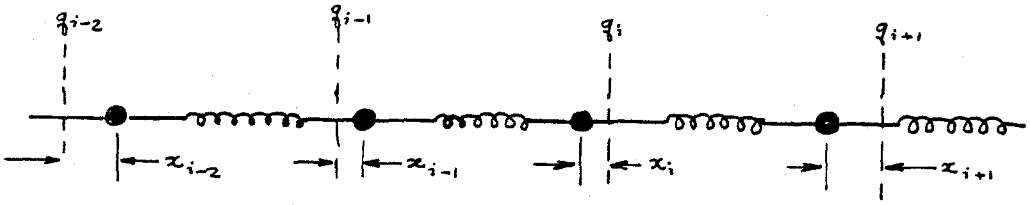
5.5 Prove that the harmonic oscillator Hamiltonian can be written as

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

5.6 Using the operators $\hat{x}$ and $\hat{p}$, expressed in terms of the ladder operators (equations 5.27), compute the matrix elements $\langle m|\hat{x}|n\rangle$ and $\langle m|\hat{p}|n\rangle$. Compare with the corresponding procedure as outlined in problems C.4 and C.5 of Appendix C.

5.7 Write down the matrix representation corresponding to the basis defined by equations 5.28 for $\hat{x}$ and $\hat{p}$, the coordinate and momentum operators.

5.8 Consider a line of identical masses connected to each other by identical springs.



Let q_i be the equilibrium position of the i^{th} mass, and x_i be the departure of the i^{th} mass from its equilibrium position. Derive the Lagrangian and Hamiltonian for this system.

5.9 Express the Hamiltonian of the preceding problem using the annihilation and creation operators defined by equation 5.46.

5.10 Find the value of the commutators $[a_i, a_j]$, $[a_i, a_j^\dagger]$, $[a_i^\dagger, a_j^\dagger]$, where a_i and $a_i^\dagger$ are the annihilation and creation operators defined by equation 5.46.

5.11 Compute and plot the heat capacity of a single harmonic oscillator as a function of $\hbar\omega / k_B T$.

5.12 Compute the heat capacity of a system of uncoupled oscillators, if the density of oscillators per unit frequency interval is given by

$$n(\omega)d\omega = n_0 \cdot \left(\frac{\omega}{\omega_0}\right)^s d\omega \quad .$$

What are the units of n_0?

5.13 Prove equation 5.66

5.14 Equation 5.73 can be written as

$$L = \int \mathcal{L} d^3\underline{x} \tag{i}$$

where

$$\mathcal{L} = \frac{1}{2}\rho_0 |\underline{\dot{q}}(\underline{x}, t)|^2 - \frac{1}{2}\rho_0^2 \mathcal{V}'(\rho_0)\left[\frac{\partial}{\partial \underline{x}} \cdot \underline{q}(\underline{x}, t)\right]^2 \tag{ii}$$

is called the Lagrangian density. Show that the principle of stationary action, i.e.

$$\delta \int_t L dt = \delta \int_t \int_{\underline{x}} \mathcal{L} d^3\underline{x}\, dt = 0 \quad , \tag{iii}$$

leads to the classical equations of motion

$$\frac{\partial \mathcal{L}}{\partial q_i} - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) - \sum_j \frac{\partial}{\partial x_j}\left(\frac{\partial \mathcal{L}}{\partial q_{i,j}}\right) = 0 \tag{iv}$$

where

$$q_{i,j} = \frac{\partial}{\partial x_j} q_i(\underline{x}, t) \quad . \tag{v}$$

Derive the classical equation of motion from the Lagrangian density given by equation (ii) and show that $[\rho_0 \mathcal{V}'(\rho_0)]^{\frac{1}{2}}$ should be identified with the speed of sound.

5.15 A point mass m is executing harmonic oscillations described by

$$x(t) = x_0 \sin(\omega t + \phi)$$

(i) Compute the classical momentum $p(t)$.

(ii) Compute the classical average values of x, x^2, p, p^2 and the energy E. Is the energy a function of time?

(iii) Compare with the quantum mechanical behavior as $E/\hbar\omega \to \infty$.

Problems (continued)

where $\Phi(|\underline{x} - \underline{x}'|)$ can be written as

$$\Phi(|\underline{x} - \underline{x}'|) = \frac{c_s^2}{\rho_0}\,\delta(\underline{x} - \underline{x}') - g(|\underline{x} - \underline{x}'|) \quad ,$$

where c_s is the speed of sound, ρ_0 is the equilibrium density, $\delta(\underline{x} - \underline{x}')$ is the three dimensional delta function, and $g(|\underline{x} - \underline{x}'|)$ is a function that possesses a peak at the mean distance between atoms. Note that if $g(|\underline{x} - \underline{x}'|) = 0$, we recover the continuum model for the liquid as given by the expression for the potential of equation 5.72

(i) Derive the Hamiltonian for the system, if the kinetic energy is given by equation 5.70 and the potential energy is given by the expression above.

(ii) Derive the dispersion relation $\omega(\underline{k})$ for this system.

(iii) Invert this relation to express $g(r)$ as a function of $\omega(k)$, where $r = |\underline{x} - \underline{x}'|$. Can you sketch $g(r)$? What must its behavior be as $r \to 0$, $r \to \infty$?

6. ANGULAR MOMENTUM

We shall begin the discussion of angular momentum in quantum mechanics by considering a few examples, also of interest in their own right, to illustrate the points.

6.1 The two-dimensional harmonic oscillator.

The Hamiltonian of an isotropic harmonic oscillator in two dimensions is given by

$$H = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{m\omega^2}{2}(x^2 + y^2) \quad . \tag{6.1}$$

Classically, the equations of motion are given by

$$\dot{p}_j = \frac{\partial H}{\partial x_j} = - m\omega^2 x_j \; ; \quad j=1,2 \tag{6.2a}$$

and

$$\dot{x}_j = \frac{\partial H}{\partial p_j} = \frac{1}{m} p_j \; ; \quad j=1,2 \quad . \tag{6.2b}$$

These can be integrated to yield

$$x(t) = x_0 \cos(\omega t - \delta_x) \tag{6.3a}$$

$$p_x(t) = - m\omega x_0 \sin(\omega t - \delta_x) \tag{6.3b}$$

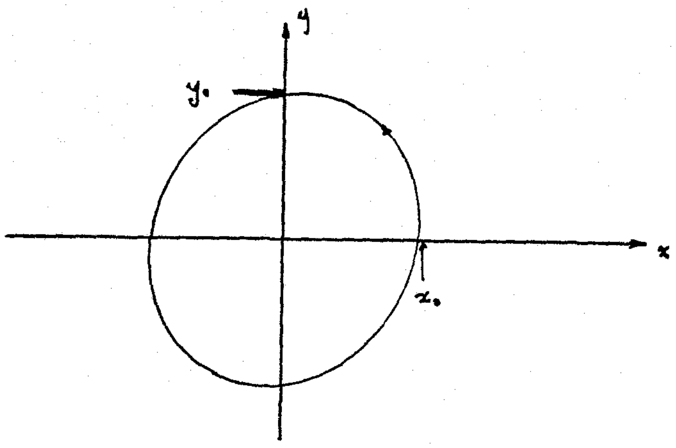
and

$$y(t) = y_0 \cos(\omega t - \delta_y) \tag{6.3c}$$

$$p_y(t) = - m\omega y_0 \sin(\omega t - \delta_y) \quad , \tag{6.3d}$$

where x_0, δ_x, y_0, δ_y are constants that depend on the initial conditions.

The motion described by equation 6.3 is an ellipse in the (x, y)-plane.



Note that when $-\pi < (\delta_y - \delta_x) < 0$, the particle moves clockwise along the ellipse ("*left-handed*" motion), while when $0 < (\delta_y - \delta_x) < \pi$, the particle moves counter-clockwise along the ellipse ("right-handed" motion). Note that the energy along the x-axis is conserved, i.e.

$$E_x = \frac{1}{2m} p_x^2 + \frac{m}{2} \omega^2 x^2 = \frac{1}{2} m\omega^2 x_0^2 \qquad (6.4a)$$

as well as along the y-axis, i.e.

$$E_y = \frac{1}{2m} p_y^2 + \frac{m}{2} \omega^2 y^2 = \frac{1}{2} m\omega^2 y_0^2 \qquad (6.4b)$$

and therefore the total energy

$$E = \frac{1}{2} m\omega^2 (x_0^2 + y_0^2) \qquad (6.5)$$

is also conserved. Note also that since the corresponding Lagrangian is not a function of the azimuthal angle φ, the conjugate momentum p_ϕ, equal to the z-component of the angular momentum

$$p_\phi = L_z = (\underline{r} \times \underline{p})_z = xp_y - yp_x \quad , \qquad (6.6)$$

is also conserved and given by

$$L_z = [mx_0 y_0 \sin(\delta_y - \delta_x)]\omega \qquad (6.7)$$

It can be seen that L_z is positive for right-handed motion (counter-clockwise) and negative for left-handed motion (clockwise).

<u>Quantum mechanically</u>, we would solve the problem by defining annihilation and creation operators

$$a_j = \left(\frac{1}{2\hbar}\right)^{1/2} [(m\omega)^{1/2} x_j + i(m\omega)^{-1/2} p_j] \qquad (6.8a)$$

$$a_j^\dagger = \left(\frac{1}{2\hbar}\right)^{1/2} [(m\omega)^{1/2} x_j - i(m\omega)^{-1/2} p_j] \qquad (6.8b)$$

in terms of which we have

$$H = \hbar\omega(a_x^\dagger a_x + a_y^\dagger a_y + 1) \qquad (6.9)$$

with eigenstates

$$|\underline{n}\rangle = |n_x, n_y\rangle \qquad (6.10)$$

and corresponding energies

$$E_n = \hbar\omega(n + 1) \quad , \qquad (6.11)$$

where $n = n_x + n_y$.

Note that the energy E_n is (n+1)-fold degenerate, i.e., for a given n there exist $n+1$ orthogonal eigenstates $|n_x, n_y\rangle$ which have the same energy, namely

$$|0, n\rangle \;,\; |1, n-1\rangle \;.....\; |n-1, 1\rangle \;,\; |n, 0\rangle \;.$$

By analogy to the classical expression for the z-component of the angular momentum, we now define an operator

$$\hat{L}_z \equiv \hat{x}\hat{p} - \hat{y}\hat{p}_x \quad , \qquad (6.12)$$

which we can express in terms of the annihilation and creation operators, i.e.

$$\hat{L}_z = i\hbar\,(a_x a_y^\dagger - a_x^\dagger a_y) \quad . \qquad (6.13)$$

Using equation 6.9 for the Hamiltonian, we can verify by direct substitution that L_z commutes with the Hamiltonian, i.e.

$$[H, L_z] = 0 \quad , \qquad (6.14)$$

and therefore L_z is a constant of the motion, in agreement with the classical result. See Equation 3.88 and related discussion.

The operators H and L_z are Hermitian and they commute. There exists, therefore[#], a set of basis states (vectors) that simultaneously diagonalize H and L_z. In other words it should be possible to find a common set of eigenstates for H and L_z. In particular, consider the operators

$$a_R = \frac{1}{\sqrt{2}}(a_x - ia_y) \tag{6.14a}$$

$$a_R^\dagger = \frac{1}{\sqrt{2}}(a_x^\dagger + ia_y^\dagger) \tag{6.14b}$$

and

$$a_L = \frac{1}{\sqrt{2}}(a_x + ia_y) \tag{6.14c}$$

$$a_L^\dagger = \frac{1}{\sqrt{2}}(a_x^\dagger - ia_y^\dagger) \quad . \tag{6.14d}$$

It can be readily verified that ($\sigma = R, L$)

$$\left.\begin{aligned} [a_\sigma, a_{\sigma'}] &= 0 \\ [a_\sigma, a_{\sigma'}^\dagger] &= \delta_{\sigma,\sigma'} \\ [a_\sigma^\dagger, a_{\sigma'}^\dagger] &= 0 \end{aligned}\right\} \quad , \tag{6.15}$$

and that the Hamiltonian, expressed in terms of the new operators, is given by

$$H = \hbar\omega(a_R^\dagger a_R + a_L^\dagger a_L + 1) \quad , \tag{6.16}$$

while L_z is given by

$$\hat{L}_z = \hbar(a_R^\dagger a_R - a_L^\dagger a_L) \quad . \tag{6.17}$$

From equations 6.15 and the form of the Hamiltonian, we see that the operators a_σ, $a_\sigma^\dagger$ are appropriate annihilation and creation operators (see discussion in section 5.2), that operate on a set of basis vectors $|n_R, n_L\rangle$. It is then clear that both H and $\hat{L}_z$ are diagonal in this basis, i.e.

$$H|n_R, n_L\rangle = \hbar\omega(n+1)|n_R, n_L\rangle \tag{6.18}$$

and

$$\hat{L}_z|n_R, n_L\rangle = m\hbar|n_R, n_L\rangle \quad , \tag{6.19}$$

where

$$n = n_R + n_L \tag{6.20a}$$

$$m = n_R - n_L \quad . \tag{6.20b}$$

see also problem 6.4.

These results motivate the identification of $a_R, a_R^\dagger$ as the annihilation and creation operators of "right-circular quanta", and $a_L, a_L^\dagger$ as the annihilation and creation operators for "left-circular quanta". The total energy of the system (above the ground state energy), in a particular eigenstate, is then equal to $\hbar\omega$ times the total right and left quanta, i.e.

[#] We can always simultaneously diagonalize two Hermitian matrices that commute.

$n_R + n_L$, whereas the angular momentum L_z is equal to $\hbar$ times the difference $(n_R - n_L)$ of the number of right and left circular quanta. It is then convenient to re-label the $|n_R, n_L\rangle$ eigenstates such that

$$|n, m\rangle \equiv |n_R + n_L, n_R - n_L\rangle \quad , \tag{6.21}$$

where n is the energy quantum number and m is the z-angular momentum quantum number.

Note that if n=0, m=0 necessarily, if n=1, we have m=1, -1, if n=2, we have m=2, 0, -2. In other words, if n is odd, m is a positive or negative odd integer, while if n is even, m is a positive or negative even integer. In both cases

$$-n \le m \le n \quad , \tag{6.22}$$

while, for the same n, two different m's must differ by a multiple of 2. The (n+1)-fold degeneracy of an energy eigenstate $|n, m\rangle$ can now be identified with the (n+1) possible values of the angular momentum quantum number m, corresponding to (n+1) orthogonal angular momentum states of the same energy.

What about the wavefunctions corresponding to the various representations? Well, the $|n_x, n_y\rangle$ basis states are products of the familiar one-dimensional wave functions (see equation 5.11 and related discussion), i.e.

$$\Psi_{n_x, n_y}(x, y) = \alpha \frac{h_{n_x}(\alpha x)\, h_{n_y}(\alpha y)}{(\pi 2^{n_x + n_y} n_x! n_y!)^{1/2}} e^{-\alpha^2 (x^2 + y^2)/2} \tag{6.23a}$$

where (see equation 5.8),

$$\alpha = \left(\frac{m\omega}{\hbar}\right)^{1/2} \quad , \tag{6.23b}$$

and $h_n(\xi)$ is the n^{th} Hermite polynomial.

The ground state of the system is $\Psi_{00}(x, y)$ and given by equation 6.23,

$$|0, 0\rangle = \Psi_{00}(x, y) = \frac{\alpha}{\pi^{1/2}} e^{-\alpha^2 (x^2 + y^2)/2} \quad , \tag{6.24}$$

i.e., a two-dimensional Gaussian.

To obtain the wavefunctions corresponding to the $|n_R, n_L\rangle$ basis vectors we use the annihilation and creation operators $a_R, a_R^\dagger$ and $a_L, a_L^\dagger$. In particular, since

$$a_R^\dagger = \frac{1}{\sqrt{2}} (a_x^\dagger + i a_y^\dagger)$$

we have

$$|1_R, 0_L\rangle = \frac{1}{\sqrt{2}} (|1_x, 0_y\rangle + i|0_x, 1_y\rangle) \tag{6.25}$$

and therefore

$$\Psi_{1_R, 0_L}(x, y) = \frac{\alpha}{2\pi^{1/2}} [h_1(\alpha x) + i h_1(\alpha y)]\, e^{-\alpha^2 (x^2 + y^2)/2}$$

or, since $h_1(\xi) = 2\xi$ (see Appendix C),

$$\Psi_{1_R,0_L}(x,y) = \frac{\alpha^2}{\pi^{1/2}}\,(x+iy)\,e^{-\alpha^2(x^2+y^2)/2} \quad . \tag{6.26}$$

This function suggests the introduction of polar coordinates, i.e.

$$x = r_\perp \cos\varphi$$

$$y = r_\perp \sin\varphi$$

in terms of which

$$\Psi_{0,0}(r_\perp,\varphi) = \frac{\alpha}{\pi^{1/2}}\,e^{-\alpha^2 r_\perp^2/2} \tag{6.27}$$

and

$$\Psi_{1_R,0_L}(r_\perp,\varphi) = \frac{\alpha^2}{\pi^{1/2}}\,r_\perp e^{-\alpha^2 r_\perp^2/2}\,e^{i\varphi} \tag{6.28a}$$

Similarly we then also have

$$\Psi_{0_R,1_L}(r_\perp,\varphi) = \frac{\alpha^2}{\pi^{1/2}}\,r_\perp\,e^{-\alpha^2 r_\perp^2/2}\,e^{-i\varphi} \tag{6.28b}$$

The identification of $a_R^\dagger$ and $a_L^\dagger$ as the operators that create a right and left circular (about the z-axis) quantum is now, perhaps, clearer. The wavefunctions of equation 6.28 can be obtained directly from the ground state wavefunction $\Psi_{0,0}(r_\perp,\varphi)$ by expressing the annihilation and creation operators in polar coordinates. In particular, from equation 6.28 we have

$$a_x = \frac{1}{\sqrt{2}}\,(\alpha x + \frac{1}{\alpha}\frac{\partial}{\partial x}) \tag{6.29a}$$

$$a_y = \frac{1}{\sqrt{2}}\,(\alpha y + \frac{1}{\alpha}\frac{\partial}{\partial y}) \tag{6.29b}$$

and therefore

$$a_R = \frac{1}{2}\,[\alpha(x-iy) + \frac{1}{\alpha}\,(\frac{1}{\partial x} - i\,\frac{\partial}{\partial y})]$$

or, in polar coordinates

$$a_R = \frac{e^{-i\varphi}}{2}\,(\alpha r_\perp + \frac{1}{\alpha}\frac{\partial}{\partial r_\perp} - \frac{i}{\alpha r_\perp}\frac{\partial}{\partial\varphi}) \tag{6.30a}$$

Similarly,

$$a_R^\dagger = \frac{e^{i\varphi}}{2}\,(\alpha r_\perp - \frac{1}{\alpha}\frac{\partial}{\partial r_\perp} - \frac{i}{\alpha r_\perp}\frac{\partial}{\partial\varphi}) \quad , \tag{6.30b}$$

and

$$a_L = \frac{e^{i\varphi}}{2}\,(\alpha r_\perp + \frac{1}{\alpha}\frac{\partial}{\partial r_\perp} + \frac{i}{\alpha r_\perp}\frac{\partial}{\partial\varphi}) \quad , \tag{6.30c}$$

$$a_L^\dagger = \frac{e^{-i\varphi}}{2}\,(\alpha r_\perp - \frac{1}{\alpha}\frac{\partial}{\partial r_\perp} + \frac{i}{\alpha r_\perp}\frac{\partial}{\partial\varphi}) \quad . \tag{6.30d}$$

We see that the result of the operation of $a_R^\dagger$ on a function $e^{im\varphi}\,F(r_\perp)$ is to produce a function

update

$$a_R^\dagger\,[e^{im\varphi}\,F(r_\perp)] = \frac{e^{i(m+1)\varphi}}{2}\,[(\alpha r_\perp + \frac{m}{\alpha r_\perp})\,F - \frac{1}{\alpha}\,F'] \ ,$$

similarly

$$a_L^\dagger\,[e^{im\varphi}\,F(r_\perp)] = \frac{e^{i(m-1)\varphi}}{2}\,[(\alpha r_\perp - \frac{m}{\alpha r_\perp})\,F - \frac{1}{\alpha}\,F'] \ ,$$

where $F' = dF/dr_\perp$. Consequently, the φ-dependence of an eigenfunction $\Psi_{n_R, n_L}(r_\perp, \varphi)$ is given by

$$\Psi_{n_R, n_L}(r_\perp, \varphi) \propto e^{i(n_R - n_L)\varphi}$$

or in terms of the energy and angular momentum quantum numbers n and m, we have

$$\Phi_{n,m}(r_\perp, \varphi) \propto e^{im\varphi} \quad . \tag{6.31}$$

Recall now, that

$$\hat{L}_z \Phi_{n,m}(r_\perp, \varphi) = m\hbar\, \Phi_{n,m}(r_\perp, \varphi)$$

and therefore using 6.31 we must have

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial\varphi} , \tag{6.32}$$

which, of course, we could have shown directly.

Example 6.1 Motion of a charged particle in a uniform magnetic field

Consider a uniform magnetic field along the z-axis

$$\underline{\mathcal{B}} = \hat{e}_z \mathcal{B} \quad . \tag{6.33}$$

We then have for the corresponding Hamiltonian

$$H = \frac{1}{2m} p^2 - \frac{q_c \mathcal{B}}{2m} L_z + \frac{q_c^2 \mathcal{B}^2}{8m} (x^2 + y^2) , \tag{6.34}$$

where L_z is the z-component of the angular momentum and q_c is the charge of the particle. See appendix F, section 5 and problem F.11. This Hamiltonian can be separated into a free particle Hamiltonian along the z-axis and a Hamiltonian corresponding to the motion in the (x, y)-plane, i.e.

$$H = H_z + H_\perp \tag{6.35a}$$

where

$$H_z = \frac{1}{2m} p_z^2 \tag{6.35b}$$

and

$$H_\perp = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{q_c^2 \mathcal{B}^2}{8m}(x^2 + y^2) - \left(\frac{q_c \mathcal{B}}{2m}\right) L_z \quad . \tag{6.35c}$$

Correspondingly, the eigenfunctions are separable and given by

$$\Psi(\underline{x}) = \psi(x, y)\; Z(z) \tag{6.36}$$

or, for a free particle along the z-axis,

$$\Psi(\underline{x}) = \psi(x, y)\; e^{i p_z z/\hbar} \quad . \tag{6.36'}$$

We now note that the first two terms of $H_\perp$ correspond to a <u>two-dimensional harmonic oscillator</u> in the (x, y)-plane, whose eigenstates would have been the $|n_x, n_y\rangle$ states of equation 6.10. These states, however, are not eigenstates of L_z and therefore cannot serve as the eigenstates of $H_\perp$. Recall, however, that the states $|n_R, n_L\rangle$ were constructed to diagonalize the two dimensional harmonic oscillator Hamiltonian and the z-component of the angular momentum simultaneously, i.e.

$$H_\perp = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{1}{2} m\omega_c^2(x^2 + y^2) - \omega_c L_z \tag{6.37}$$

where

$$\omega_c = \frac{q_c \mathcal{B}}{2m} , \tag{6.38}$$

is the classical cyclotron frequency, and

$$H_\perp |n_R, n_L\rangle = [\hbar|\omega_c|(n_R + n_L + 1) - \hbar\omega_c(n_R - n_L)]\,|n_R, n_L\rangle \tag{6.39}$$

Note that if $q_c > 0$, we have $\omega_c > 0$ and

$$E_\perp = \hbar|\omega_c|[n_R + n_L + 1 - (n_R - n_L)] , \tag{6.40}$$

where n_R and n_L are the right and left circularly polarized quanta of the previous section, so that

$$E_\perp = \hbar|\omega_c|\;(2n_L + 1) \;,\quad q_c > 0 \tag{6.41a}$$

and similarly,

$$E_\perp = \hbar|\omega_c|\;(2n_R + 1) \;,\quad q_c < 0 . \tag{6.41b}$$

The total wavefunctions are then given by

$$|\Psi\rangle = |p_z, n_R, n_L\rangle , \tag{6.42}$$

with energies

$$E = \begin{cases} \dfrac{1}{2m}p_z^2 + \hbar|\omega_c|\;(2n_L + 1) \;, & \text{for} \quad q_c > 0 \\ \\ \dfrac{1}{2m}p_z^2 + \hbar|\omega_c|\;(2n_R + 1) \;, & \text{for} \quad q_c < 0 . \end{cases} \tag{6.43}$$

We can combine these two expressions using the $n = n_R + n_L$ and $m = n_R - n_L$ quantum numbers. In particular,

$$E = \frac{1}{2m}p_z^2 + \hbar|\omega_c|\;[n - m\,\mathrm{sign}(q_c) + 1] , \tag{6.44}$$

where

$$\mathrm{sign}(x) \equiv \begin{cases} +1 \;, & \text{for} \quad x > 0 \\ \\ -1 \;, & \text{for} \quad x < 0 . \end{cases} \tag{6.45}$$

6.2 Angular momentum in three dimensions.

In the preceding section we studied the angular momentum along the z-axis, corresponding to motion restricted to two-dimensions [(x, y)-plane]. The z-component of $\underline{L}$, the angular momentum was then computed to be

$$L_z = (\underline{r} \times \underline{p})_z = xp_y - yp_x \ . \tag{6.46a}$$

For motion in three dimensions, however, the other two components of angular momentum, namely

$$L_x = (\underline{r} \times \underline{p})_x = yp_z - zp_y \tag{6.46b}$$

and

$$L_y = (\underline{r} \times \underline{p})_y = zp_x - xp_z \ , \tag{6.46c}$$

must also be considered. Since quantum mechanically, the quantities in all the products in equation 6.46 commute, we may replace the linear momenta p_x, p_y, p_z by the corresponding components of the momentum operator to obtain the operator expressions for the component of the angular momentum, i.e.

$$L_x = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y}\right) \ , \tag{6.47a}$$

$$L_y = -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z}\right) \ , \tag{6.47b}$$

and

$$L_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}\right) \ . \tag{6.47c}$$

The three components of the angular momentum do not commute with each other. In fact, we can show by direct substitution of the commutation relations between the components of $\underline{x}$ and the components of $\underline{p}$, that

$$[L_x, L_y] = i\hbar L_z \tag{6.48a}$$

$$[L_y, L_z] = i\hbar L_x \tag{6.48b}$$

and

$$[L_z, L_x] = i\hbar L_y \ . \tag{6.48c}$$

The vector angular momentum $\underline{L}$ is given by

$$\underline{L} = L_x \hat{e}_x + L_y \hat{e}_y + L_z \hat{e}_z \tag{6.49}$$

and has a magnitude squared, given by

$$L^2 = \underline{L} \cdot \underline{L} = L_x^2 + L_y^2 + L_z^2 \ . \tag{6.50}$$

It is easy to show that it commutes with the three components L_x, L_y, L_z. In particular,

$$[L^2, L_x] = [L_x^2, L_x] + [L_y^2, L_x] + [L_z^2, L_x] \ .$$

The first commutator is zero since L_x commutes with itself and therefore with its square (see problem 2.7 v). The other two can be computed as follows,

$$[L_y^2, L_x] \equiv [L_y, L_x]L_y + L_y[L_y, L_x] ,$$

or, using equation 6.48a

$$[L_y^2, L_x] = -i\hbar(L_zL_y + L_yL_z) .$$

Similarly

$$[L_z^2, L_x] = i\hbar(L_zL_y + L_yL_z) ,$$

and therefore

$$[L^2, L_x] = [L^2, L_y] = [L^2, L_z] = 0 . \qquad (6.51)$$

In the study of the behavior of angular momentum it proves convenient to define two new operators

$$L_+ = L_x + iL_y \qquad (6.52a)$$

and

$$L_- = L_x - iL_y . \qquad (6.52b)$$

The three operators L_+, L_- and L_z satisfy the following commutation relations,

$$[L_z, L_+] = \hbar L_+ \qquad (6.53a)$$

$$[L_z, L_-] = -\hbar L_- \qquad (6.53b)$$

$$[L_+, L_-] = 2\hbar L_z , \qquad (6.53c)$$

as can readily be shown by direct substitution. The operator L^2 also commutes with L_+ and L_-, since they are linear combinations of L_x and L_y which, in turn, commute with L^2. Inverting equations 6.52 and substituting for L_x^2 and L_y^2 in equation 6.50 we also have

$$L^2 = \frac{1}{2}(L_+ L_- + L_- L_+) + L_z^2 \qquad (6.54)$$

from which we can also show that

$$L_- L_+ = L^2 - L_z(L_z + \hbar) \qquad (6.55a)$$

and

$$L_+ L_- = L^2 - L_z(L_z - \hbar) . \qquad (6.55b)$$

From the commutation relations for the three cartesian components of the angular momentum (equations 6.48) and the fact that the necessary and sufficient condition for two Hermitean matrices to be diagonalizable by the same transformation is that they commute, we see that it is not possible to find a set of the angular momentum, say L_z, which are also eigenstates of any one of the other two (much less of both). Using the fact, however, that the modulus squared of the angular momentum, i.e., L^2, commutes with any one component, a transformation exists that simultaneously diagonalizes L^2 and one component of $\underline{L}$. Consequently there exists a set of simultaneous eigenstates of L^2 and one component of the angular momentum which, without loss of generality, we can choose to be L_z. Equivalently there exist a set of eigenfunctions $\psi_{\lambda,\mu}$ such that

$$L^2 \psi_{\lambda,\mu} = \lambda \psi_{\lambda,\mu} \qquad (6.56a)$$

and

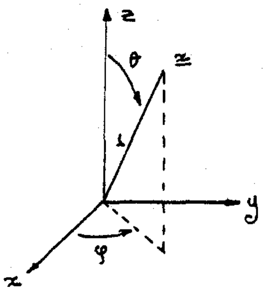
$$L_z \psi_{\lambda,\mu} = \mu \psi_{\lambda,\mu} . \qquad (6.56b)$$

Equations 6.56 are the eigenvalue equations for the eigenfunctions $\psi_{\lambda,\mu}$. These we can find by expressing L^2 and L_z as differential operators, and solving the resulting differential equations. If we transform the cartesian x, y, z coordinates in equations 6.47 into spherical coordinates, i.e.

$$x = r \sin\theta \cos\varphi \qquad (6.57a)$$

$$y = r \sin\theta \sin\varphi \qquad (6.57b)$$

$$z = r \cos\theta , \qquad (6.57c)$$



we have, after a little algebra,

$$L_x = i\hbar(\sin\varphi \frac{\partial}{\partial\theta} + \cot\theta \cos\varphi \frac{\partial}{\partial\varphi}) \qquad (6.58a)$$

$$L_y = i\hbar(- \cos\varphi \frac{\partial}{\partial\theta} + \cot\theta \sin\varphi \frac{\partial}{\partial\varphi}) \qquad (6.58b)$$

and

$$L_z = -i\hbar \frac{\partial}{\partial\varphi} , \tag{6.58c}$$

the latter of which we have already encountered in the discussion of angular momentum in the two-dimensional harmonic oscillator. It can be seen that angular momentum operates on the angular coordinates θ and φ, and does not operate on the radial coordinate r.

Substituting equations 6.58 into the expression for L^2 and $L_\pm$, we also have

$$L^2 = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} (\sin\theta \frac{\partial}{\partial\theta}) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right] \tag{6.59}$$

and

$$L_\pm = \hbar e^{\pm i\varphi} (\pm \frac{\partial}{\partial\theta} + i \cot\theta \frac{\partial}{\partial\varphi}) . \tag{6.60}$$

From equations 6.59 and 6.58c for L^2 and L_z respectively, we see that the eigenfunctions $\psi_{\lambda,\mu}$ must be given by

$$\psi_{\lambda,\mu} = \psi_{\lambda,\mu}(\theta,\varphi) = \Theta_{\lambda,\mu}(\theta)\, \Phi_\mu(\varphi) , \tag{6.61}$$

where, since L_z operates on φ only (equation 6.58c), we have that

$$L_z \Phi_\mu(\varphi) = \mu\, \Phi_\mu(\varphi) . \tag{6.62}$$

Substituting for L_z we have

$$i\hbar\, \Phi_\mu'(\varphi) + \mu\, \Phi_\mu(\varphi) = 0 , \tag{6.63a}$$

with the boundary condition

$$\Phi_\mu(0) = \Phi_\mu(2\pi) . \tag{6.63b}$$

The solution to equation 6.63a is evidently

$$\Phi_\mu(\varphi) = e^{i\mu\varphi/\hbar} \tag{6.64a}$$

where, from the boundary condition, we have that

$$\mu = m\hbar , \tag{6.64b}$$

where m is any positive or negative integer or zero.

To study the spectrum of L^2, we note that L^2 is a sum of three positive definite Hermitean operators, i.e., for any state $|\psi\rangle$, we have from equation 6.50,

$$\langle\psi|L^2|\psi\rangle = \langle\psi|L_x^2|\psi\rangle + \langle\psi|L_y^2|\psi\rangle + \langle\psi|L_z^2|\psi\rangle .$$

But, since L_x is a Hermitean operator

$$\langle\psi|L_x^2|\psi\rangle = \langle\psi|L_x, L_x|\psi\rangle = \| L_x|\psi\rangle \|^2 \geq 0 .$$

where $\| |\chi\rangle \|^2$ denotes the modulus squared of the state vector $|\chi\rangle$. By similar arguments we have that

$$\langle\psi|L_y^2|\psi\rangle \geq 0$$

and

$$\langle\psi|L_z^2|\psi\rangle \geq 0$$

for any state $|\psi\rangle$ *and therefore also*

$$\langle\psi|L^2|\psi\rangle \geq 0 \tag{6.65}$$

for any $|\psi\rangle$. Consequently the eigenvalues of L^2 must be non-negative, i.e.

$$\lambda \geq 0 \; . \tag{6.66}$$

It will subsequently prove convenient to define a non-negative real number ℓ, such that

$$\lambda = \hbar^2 \ell(\ell+1) \; , \tag{6.67}$$

where $\hbar^2$ is chosen for dimensional reasons and ℓ is the non-negative root of 6.67. It is clear that this does not result in any loss of generality since for any λ there exists a unique non-negative root of equation 6.67. We will henceforth label the common eigenstates of L^2 and L_z with the new indeces ℓ and m, in particular, we have

$$L^2|\ell,m\rangle = \hbar^2\ell(\ell+1)|\ell,m\rangle \tag{6.68a}$$

and

$$L_z|\ell,m\rangle = \hbar m|\ell,m\rangle \; , \tag{6.68b}$$

where m has been determined to be a positive or negative integer or zero, and ℓ is a non-negative real number whose behavior is yet to be determined.

We now note, by virtue of equations 6.55, that

$$L_-L_+|\ell,m\rangle = [L^2 - L_z(L_z+\hbar)]|\ell,m\rangle = \hbar^2[\ell(\ell+1) - m(m+1)]|\ell,m\rangle$$

and

$$L_+L_-|\ell,m\rangle = [L^2 - L_z(L_z-\hbar)]|\ell,m\rangle = \hbar^2[\ell(\ell+1) - m(m-1)]|\ell,m\rangle \; .$$

We have, however, that

$$\langle\ell,m|,L_-L_+|\ell,m\rangle = \langle\ell,m|L_+,L_+|\ell,m\rangle = \||L_+|\ell,m\rangle\|^2 \geq 0$$

and similarly

$$\langle\ell,m|,L_+L_-|\ell,m\rangle = \||L_-|\ell,m\rangle\|^2 \geq 0 \; ,$$

since L_+ and L_-, like the annihilation and creation operators of the Harmonic oscillator, are the adjoint of each other. Therefore, we must have

$$\ell(\ell+1) - m(m+1) \equiv (\ell-m)(\ell+m+1) \geq 0$$

and

$$\ell(\ell+1) - m(m-1) \equiv (\ell+m)(\ell-m+1) \geq 0 \; .$$

The first inequality requires that

$$-(\ell+1) \le m \le \ell$$

while the second requires that

$$-\ell \le m \le (\ell+1) .$$

These latter two inequalities can evidently be simultaneously satisfied if and only if

$$-\ell \le m \le \ell . \tag{6.69}$$

To proceed further, it proves convenient to study the behavior of the operators L_+ and L_- separately. In particular, we note from the preceding discussion that

$$\| L_+ | \ell, m\rangle \|^2 = \langle \ell, m | L_- L_+ | \ell, m\rangle = (\ell - m)(\ell + m + 1) ,$$

and therefore

$$L_+ | \ell, m\rangle = 0 , \qquad \text{iff} \quad m = \ell . \tag{6.70a}$$

Similarly, we also have

$$L_- | \ell, m\rangle = 0 , \qquad \text{iff} \quad m = -\ell . \tag{6.70b}$$

If, however, $m \neq \ell$ we have, using the fact that L_+ commutes with L^2, that

$$[L^2, L_+] | \ell, m\rangle = 0$$

or, equivalently

$$\begin{aligned} L^2 L_+ | \ell, m\rangle &= L_+ L^2 | \ell, m\rangle \\ &= \hbar^2 \ell(\ell+1) L_+ | \ell, m\rangle . \end{aligned}$$

This latter result shows that, if $m \neq \ell$, the state $L_+ | \ell, m\rangle$ is also an eigenstate of L^2 with the same eigenvalue as $| \ell, m\rangle$, namely $\hbar^2 \ell(\ell+1)$.

We now operate on $| \ell, m\rangle$ with the commutator of equation 6.53a, again if $m \neq \ell$,

$$[L_z, L_+] | \ell, m\rangle = \hbar L_+ | \ell, m\rangle ,$$

and since

$$[L_z, L_+] | \ell, m\rangle = L_z L_+ | \ell, m\rangle - L_+ L_z | \ell, m\rangle ,$$

we have that, if $m \neq \ell$,

$$\begin{aligned} L_z L_+ | \ell, m\rangle &= L_+ L_z | \ell, m\rangle + \hbar L_+ | \ell, m\rangle \\ &= \hbar(m+1) L_+ | \ell, m . \end{aligned}$$

Consequently, we see that $L_+ | \ell, m\rangle$ is an eigenstate of L_z with eigenvalue

$\hbar(m+1)$. Similarly, if $m \neq -\ell$,

$$L_z L_- |\ell, m\rangle = \hbar(m-1) L_- |\ell, m\rangle$$

and therefore $L_- |\ell, m\rangle$ must be an eigenstate of L_z with eigenvalue $\hbar(m-1)$. Combining these results, we see that if $m \neq \ell$

$$L_+ |\ell, m\rangle = C_{\ell, m} |\ell, m+1\rangle$$

where $C_{\ell, m}$ is a complex constant. If the states $|\ell, m\rangle$ are normalized, we then have

$$|C_{\ell, m}|^2 = \langle \ell, m | L_+^\dagger L_+ | \ell, m\rangle = \| L_+ | \ell, m\rangle \|^2$$

and using the results of the preceding discussion,

$$|C_{\ell, m}|^2 = \hbar^2 [\ell(\ell+1) - m(m+1)] \ .$$

Therefore, if we choose the relative phase of $|\ell, m\rangle$ and $|\ell, m+1\rangle$, such that $C_{\ell, m}$ is real and positive, we have

$$L_+ |\ell, m\rangle = \hbar [\ell(\ell+1) - m(m+1)]^{1/2} |\ell, m+1\rangle \ , \qquad (6.71a)$$

and by similar arguments

$$L_- |\ell, m\rangle = \hbar [\ell(\ell+1) - m(m-1)]^{1/2} |\ell, m-1\rangle \ . \qquad (6.71b)$$

Note that equations 6.70 are included in equations 6.71.

Consider now a state $|\ell, m\rangle$. The result of operating on this state with L_+ is to produce a state proportional to $|\ell, m+1\rangle$, i.e.,

$$L_+ |\ell, m\rangle \propto |\ell, m+1\rangle \ .$$

If we operate again with L_+, we have

$$L_+^2 |\ell, m\rangle = L_+(L_+ |\ell, m\rangle) \propto L_+ |\ell, m+1\rangle \propto |\ell, m+2\rangle \ .$$

Similarly, we have after m' applications of this procedure,

$$(L_+)^{m'} |\ell, m\rangle \propto |\ell, m+m'\rangle \ .$$

Clearly, unless this process terminates, we can generate an infinity of states $|\ell, m+m'\rangle$, corresponding to any integer or zero value for m'. This, however, is disallowed by equation 6.69 which requires that

$$m + m' \leq \ell \ .$$

<u>This can only be achieved if ℓ is an integer</u>, in which case

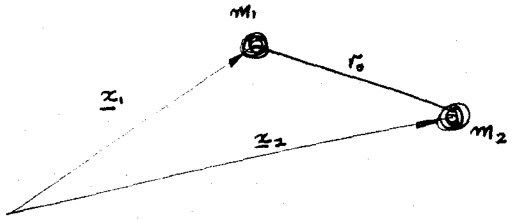
$$L_+ |\ell, m=\ell\rangle = 0 \ .$$

Note that therefore, for a given ℓ, there exist $(2\ell+1)$ simultaneous (orthogonal) eigenstates $|\ell,m\rangle$ corresponding to the $(2\ell+1)$ possible values of m permitted by equation 5.69, i.e.,

$$|\ell,-\ell\rangle,\ |\ell,-\ell+1\rangle,\ \ldots.\ ,\ |\ell,0\rangle,\ \ldots.\ ,\ |\ell,\ell-1\rangle,\ |\ell,\ell\rangle\ .$$

Example 6.2 Rotational spectra of ~~diatomic~~ diatomic molecules at low energies.

Consider two atoms of mass m_1 and m_2 respectively, at an equilibrium distance r_0 along the line of centers.



If we define the position of the center of mass, i.e.,

$$\underline{X} = \frac{m_1\underline{x}_1 + m_2\underline{x}_2}{m_1 + m_2}\ , \tag{6.72a}$$

and the relative position, i.e.,

$$\underline{x} = \underline{x}_2 - \underline{x}_1\ , \tag{6.72b}$$

we can write the total kinetic energy of the system as the kinetic energy of the center of mass plus the kinetic energy about the center of mass (see, for example, problem 1.4). The resulting Hamiltonian (in this approximation of fixed interatomic distance) in the absence of external forces is given by

$$H \simeq \frac{1}{2M}\underline{P}^2 + \frac{1}{2m}p^2 = -\frac{\hbar^2}{2M}\left|\frac{\partial}{\partial\underline{X}}\right|^2 - \frac{\hbar^2}{2m}\left|\frac{\partial}{\partial\underline{x}}\right|^2\ ,$$

where

$$M = m_1 + m_2 \tag{6.73a}$$

is the total mass,

$$\frac{1}{m} = \frac{1}{m_1} + \frac{1}{m_2} \tag{6.73b}$$

is the reduced mass, and $\underline{P}$ and $\underline{p}$ are the conjugate momenta to $\underline{X}$ and $\underline{x}$ respectively. The total wavefunction $\Psi(\underline{X},\underline{x})$ is then separable into the product

$$\Psi(\underline{X},\underline{x}) = \Psi(\underline{X})\ \psi(\underline{x})\ .$$

The wavefunction $\Psi(\underline{X})$ corresponds to a free particle wavefunction

$$\Psi(\underline{X}) \propto e^{i\underline{P}\cdot\underline{X}/\hbar}$$

where

$$P = \left(\frac{2M\,E_{CM}}{\hbar^2}\right)^{1/2}$$

and E_{CM} is the kinetic energy of the center of mass of the molecule.

The wavefunction $\psi(\underline{x})$, can be expressed as a superposition of the eigenfunctions of the Hamiltonian for the <u>relative</u> motion, i.e.

$$H_{rel}\,\psi_\varepsilon(\underline{x}) = \varepsilon\,\psi_\varepsilon(\underline{x})$$

or

$$\left[-\frac{\hbar^2}{2m}\left|\frac{\partial}{\partial\underline{x}}\right|^2 - \varepsilon\right]\psi_\varepsilon(\underline{x}) = 0 \quad .$$

We can now express the Laplacian

$$\left|\frac{\partial}{\partial\underline{x}}\right|^2 = \frac{\partial}{\partial\underline{x}}\cdot\frac{\partial}{\partial\underline{x}} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

in spherical coordinates,[#] i.e.

$$\left|\frac{\partial}{\partial\underline{x}}\right|^2 \equiv \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] \tag{6.74}$$

and use the fact that, in this approximation, $r \simeq r_0$ is held constant and therefore we have

$$\left\{-\frac{\hbar^2}{2mr_0^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right] - \varepsilon\right\}\psi_\varepsilon(\theta,\varphi) = 0 \quad .$$

At this point, however, we observe that <u>the differential operator in the brackets (times $\hbar^2$) is in fact L^2</u>. See equation 6.59. Therefore

$$(H_{rel} - \varepsilon)\psi_\varepsilon(\theta,\varphi) = \left(\frac{1}{2mr_0^2}\,L^2 - \varepsilon\right)\psi_\varepsilon(\theta,\varphi) = 0 \quad .$$

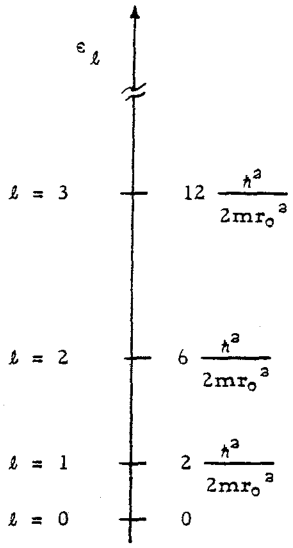
Therefore the eigenstates, are the states $|\ell, m\rangle$ with energies

$$\varepsilon_\ell = \frac{\hbar^2}{2mr_0^2}\,\ell(\ell+1) \quad . \tag{6.75}$$

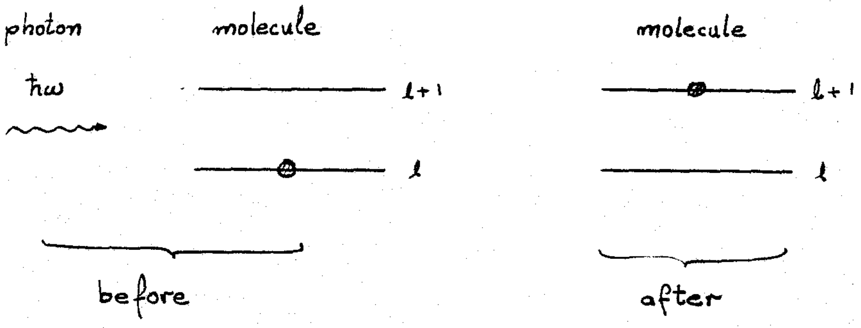
Note that successive energy levels are spaced by

$$\Delta\varepsilon_\ell = \varepsilon_\ell - \varepsilon_{\ell-1} = \frac{\hbar^2}{2mr_0^2}\,(2\ell) \quad . \tag{6.76}$$

[#] See equations 6.57 and appendix G.



In an absorption (or emission) experiment, these are in fact the observable quantities, since



where, by conservation of energy,

$$[\hbar\omega]_m = \Delta\epsilon_\ell + \Delta E_{CM} . \qquad (6.77)$$

The subscripted brackets indicate that the photon energy (frequency) has to be computed in the frame in which the absorption (or emission) takes place (Doppler shifted). See problems 6.15 and 6.16.

We will show later that for absorption or emission of light by a heteronuclear (unlike atoms) diatomic molecule (HCℓ, CO, etc.) we must have a molecular transition between two states whose angular momentum quantum numbers differ by ±1, i.e., $\Delta\ell = \pm 1$. We will also show that homonuclear diatomic molecules (O_2, N_2, etc.) cannot occupy energy levels with odd ℓ and therefore, necessarily, do not exhibit radiative transitions with $\Delta\ell = \pm 1$.

6.3 The Spherical Harmonics

What are the eigenfunctions corresponding to the states $|\ell, m\rangle$? From equations 6.61, 6.64 and the related discussion, we see that

$$\psi_{\lambda,\mu}(\theta,\varphi) \equiv Y_{\ell,m}(\theta,\varphi) = F_{\ell,m}(\theta)\, e^{im\varphi} \tag{6.78}$$

where ℓ and m are integers and related to the eigenvalues λ and μ of L^2 and L_z (see equations 6.56) by equations 6.67 and 6.64 respectively.

We can find the function $F_{\ell,\ell}(\theta)$ by a trick that also proved useful in deriving the harmonic oscillator wavefunctions using ladder operators (see equation 5.24 and related discussion). In the case of the angular momentum wavefunctions, we note (from equation 6.70a) that

$$L_+ Y_{\ell,\ell}(\theta,\varphi) = 0 \quad . \tag{6.79}$$

Consequently, substituting for the differential operator for L_+ from equation 6.60, we have

$$\left(\frac{\partial}{\partial\theta} + i \cot\theta \frac{\partial}{\partial\varphi}\right) F_{\ell,\ell}(\theta)\, e^{i\ell\varphi} = 0$$

or

$$\left(\frac{d}{d\theta} - \ell \cot\theta\right) F_{\ell,\ell}(\theta) = 0 \quad . \tag{6.80}$$

This is a first order differential equation, which can be integrated directly, note that

$$\cot\theta\, d\theta = \frac{d \sin\theta}{\sin\theta} \quad ,$$

to give

$$F_{\ell,\ell}(\theta) = c_\ell (\sin\theta)^\ell \quad , \tag{6.81}$$

where c_ℓ is some normalization constant.

We can determine c_ℓ by requiring that $Y_{\ell,\ell}(\theta,\varphi)$ be normalized on the surface of the unit sphere, i.e.

$$\int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi} |Y_{\ell,\ell}(\theta,\varphi)|^2 \sin\theta\, d\theta\, d\varphi = 1 \quad .$$

Substituting for $Y_{\ell,\ell}(\theta,\varphi)$, we have

$$1 = 2\pi\, |c_\ell|^2 \int_{\theta=0}^{\pi} (\sin\theta)^{2\ell} \sin\theta\, d\theta$$

or, substituting $x = \cos\theta$

$$2\pi\, |c_\ell|^2 \int_{-1}^{1} (1 - x^2)^\ell\, dx = 1$$

We can compute the integral by noting that

$$I_\ell = \int_{-1}^{1} (1 - x^2)^\ell\, dx = \int_{-1}^{1} (1 - x^2)^{\ell-1} (1 - x^2)\, dx$$

$$= I_{\ell-1} - \frac{1}{2\ell} I_\ell \quad ,$$

where the second term is obtained by integration by parts. We can now solve I_ℓ in terms of $I_{\ell-1}$ to obtain

$$I_\ell = \frac{2\ell}{(2\ell+1)} I_{\ell-1} \quad ,$$

which we can solve recursively to obtain

$$I_\ell = \frac{2\ell}{2\ell+1} \cdot \frac{2(\ell-1)}{2(\ell-1)+1} I_{\ell-2}$$

$$\vdots$$

$$I_\ell = \frac{2\ell}{2\ell+1} \cdot \frac{2(\ell-1)}{2(\ell-1)+1} \cdot \frac{2(\ell-2)}{2(\ell-2)+1} \cdots\cdots I_0$$

or

$$I_\ell = \frac{(2\ell)!!}{(2\ell+1)!!} I_0 \quad , \tag{6.82}$$

where the double factorial is defined as

$$n!! \equiv n(n-2)(n-4) \ldots 6\cdot 4\cdot 2 \quad ,$$

and I_0 is equal to

$$\int_{-1}^{1} dx = 2 \quad .$$

Noting that

$$(2\ell+1)!! = \frac{(2\ell+1)!}{(2\ell)!!}$$

we have, from equation 6.82,

$$I_\ell = \frac{2^{2\ell+1}\ (\ell!)^2}{(2\ell+1)!} \tag{6.83}$$

and therefore

$$|c_\ell| = \frac{1}{2^\ell\ \ell!} \left[\frac{(2\ell+1)!}{4\pi}\right]^{1/2} \quad .$$

This determines c_ℓ within a phase which we can choose such that

$$c_\ell = \frac{(-1)^\ell}{2^\ell\ \ell!} \left[\frac{(2\ell+1)!}{4\pi}\right]^{1/2} \quad ,$$

by convention. Therefore, substituting these results, we have

$$Y_{\ell,\ell}(\theta,\varphi) = \frac{(-1)^\ell}{2^\ell\ \ell!} \left[\frac{(2\ell+1)}{4\pi}\right]^{1/2} \sin^\ell\theta\ e^{i\ell\varphi} \quad . \tag{6.85}$$

We can now generate the remainder of the functions $Y_{\ell,\ell-1}(\theta,\varphi)$ by repeated application of equation 6.71b, e.g.

$$L_-\ Y_{\ell,\ell}(\theta,\varphi) = \hbar[\ell(\ell+1) - \ell(\ell-1)]^{1/2}\ Y_{\ell,\ell-1}(\theta,\varphi) \quad ,$$

or, substituting for L_- from equation 6.60, we have

$$Y_{\ell,\ell-1}(\theta,\varphi) = \frac{e^{-i\varphi}}{(2\ell)^{-1/2}} \left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right) c_\ell(\sin\theta)^\ell\ e^{i\ell\varphi}$$

$$= -\frac{c_\ell}{(2\ell)^{1/2}}\ e^{i(\ell-1)\varphi} \left(\frac{\partial}{\partial\theta} + \ell\cot\theta\right)(\sin\theta)^\ell$$

or

$$Y_{\ell,\ell-1}(\theta,\varphi) = -\ c_\ell(2\ell)^{1/2}\ (\sin\theta)^{\ell-1}\cos\theta\ e^{i(\ell-1)\varphi} \quad . \tag{6.86}$$

Similarly, we can generate $Y_{\ell,\ell-2}(\theta,\varphi)$ and so on to $Y_{\ell,-\ell}(\theta,\varphi)$. We can actually derive an explicit expression for $Y_{\ell,m}(\theta,\varphi)$ in this manner, i.e. (the algebra is ... uninteresting)

$$Y_{\ell,m}(\theta,\varphi) = c_\ell \left[\frac{(\ell+m)!}{(\ell-m)!}\right]^{1/2} (\sin\theta)^{-m}\, e^{im\varphi} \frac{d^{\ell-m}}{d(\cos\theta)^{\ell-m}} (1-\cos^2\theta)^{\ell} \quad , \tag{6.87}$$

where c_ℓ is given by equation 6.84. It should be noted that equation 6.87 generates the normalized wavefunctions $Y_{\ell,m}(\theta,\varphi)$. The first few of these are:

$$\begin{aligned}
Y_{0,0}(\theta,\varphi) &= \left(\frac{1}{4\pi}\right)^{1/2} \\
Y_{1,\pm1}(\theta,\varphi) &= \mp \left(\frac{3}{8\pi}\right)^{1/2} \sin\theta\, e^{\pm i\varphi} \\
Y_{1,0}(\theta,\varphi) &= \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta \\
Y_{2,\pm2}(\theta,\varphi) &= \left(\frac{15}{32\pi}\right)^{1/2} \sin^2\theta\, e^{\pm 2i\varphi} \\
Y_{2,\pm1}(\theta,\varphi) &= \mp \left(\frac{15}{8\pi}\right)^{1/2} \sin\theta\cos\theta\, e^{\pm i\varphi} \\
Y_{2,0}(\theta,\varphi) &= \left(\frac{5}{16\pi}\right)^{1/2} (3\cos^2\theta - 1) \quad .
\end{aligned} \tag{6.88}$$

The functions $Y_{\ell,m}(\theta,\varphi)$ are called spherical harmonics and are generally useful in solving spherically symmetric problems. They are often expressed in terms of the associated Legendre functions, defined by

$$P_{\ell,m}(x) \equiv (1-x^2)^{m/2} \frac{d^m}{dx^m} P_\ell(x) \quad , \tag{6.89a}$$

where the $P_\ell(x)$ are the Legendre polynomials, defined in turn by

$$P_\ell(x) \equiv \frac{1}{2^\ell \ell!} \frac{d^\ell}{dx^\ell} (x^2-1)^\ell \quad . \tag{6.89b}$$

In particular, for $m \geq 0$

$$Y_{\ell,m}(\theta,\varphi) = (-1)^m \left(\frac{2\ell+1}{4\pi}\right)^{\frac{1}{2}} \left[\frac{(\ell-m)!}{(\ell+m)!}\right]^{\frac{1}{2}} P_{\ell,m}(\cos\theta)\, e^{im\varphi} \quad , \tag{6.90}$$

while the functions for $m < 0$ can be obtained from the complex conjugate relationship

$$Y_{\ell,m}(\theta,\varphi) = (-1)^{-m}\, Y^*_{\ell,-m}(\theta,\varphi) \quad . \tag{6.91}$$

update

The Legendre polynomials can be shown to possess a generating function[#], given by (for $\rho < 1$)

[#] See G. F. Carrier, M. Krook and C. E. Pearson, Functions of a Complex Variable (McGraw-Hill, 1968) pp. 101-103.

$$(1 - 2\rho x + \rho^2)^{-1/2} = \sum_{\ell=0}^{\infty} \rho^{\ell} P_{\ell}(x) \quad , \tag{6.92}$$

which is useful in computing the reciprocal of the distance between two points $\underline{r}$ and $\underline{R}$, where $|\underline{r}|/|\underline{R}| < 1$. See problem 6.19. The first few of these are given by,

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{1}{2}(3x^2 - 1) \\ P_3(x) &= \frac{1}{2}(5x^3 - 3x) \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3) \quad . \end{aligned} \tag{6.93}$$

The ℓ^{th} Legendre polynomial of the cosine of a subtended angle θ between two directions $\hat{e}_1 = (\theta_1, \varphi_1)$ and $\hat{e}_2 = (\theta_2, \varphi_2)$, i.e.

$$\hat{e}_1 \cdot \hat{e}_2 = \cos\theta \quad , \tag{6.94}$$

can also be expanded as a summation of the products of the spherical harmonics $Y_{\ell,m}(\theta_1, \varphi_1)$ and $Y_{\ell,m}(\theta_2, \varphi_2)$.

In particular[#],

$$P_{\ell}(\cos\theta) = \frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{\ell} Y^*_{\ell,m}(\theta_1, \varphi_1)\, Y_{\ell,m}(\theta_2, \varphi_2) \quad , \tag{6.95}$$

known as the addition theorem for the Legendre polynomials. See also problem 6.20.

Before leaving this section, it should be pointed out that care should be exercised in consulting references on the spherical harmonics because the various conventions of signs, phases, etc., are, unfortunately, not universal. The choices here were made so as to have $Y_{\ell,0}(\theta) > 0$ at $\theta = 0$ and are consistent, for example, with A. Messiah, Quantum Mechanics, v. I, appendix B.

[#] See, for example, C. Cohen-Tannoudji, B. Diu and F. Laloë, Quantum Mechanics (John Wiley, 1977), pp. 688-689.

Problems

6.1 Prove equation 6.13.

6.2 Give a reference and/or outline a proof of the statement that two commuting Hermitian matrices can be simultaneously diagonalized.

6.3 Prove equations 6.15.

6.4 Prove equations 6.25 and 6.28b.

6.5 Prove equations 6.30.

6.6 Prove equation 6.32 directly from 6.12.

6.7 Show that the $n = 2$ eigenfunctions $\Phi_{n,m}(r_\perp, \varphi)$ of the two dimensional harmonic oscillator, are given by

$$\Phi_{2,\pm 2}(r_\perp, \varphi) = \frac{\alpha}{(2\pi)^{1/2}} (\alpha r_\perp)^2 \, e^{-(\alpha r_\perp)^2/2} \, e^{\pm 2i\varphi}$$

$$\Phi_{2,0}(r_\perp, \varphi) = \frac{\alpha}{\pi^{1/2}} [(\alpha r_\perp)^2 - 1] \, e^{-(\alpha r_\perp)^2/2} \quad .$$

6.8 Express the Hamiltonian of equation 6.1 as a differential operator in $r_\perp$ and φ, where

$$x = r_\perp \cos \varphi$$

$$y = r_\perp \sin \varphi$$

and show directly that the corresponding time-independent Schrödinger equation possesses solutions of the form

$$\Psi_{n,m}(r_\perp, \varphi) = F_{n,m}(r_\perp) \, e^{im\varphi} \quad .$$

6.9 Compute the matrix elements

$$\langle n', m' | r_\perp^2 | n, m \rangle \quad ,$$

between the eigenstates of a two-dimensional Harmonic oscillator.

Caution: There are several hard ways to do this!

6.10 (i) Derive the density of eigenstates per unit volume of a charged particle of mass m, charge q_c that is bound by a two-dimensional potential

$$V = \frac{1}{2} m\omega_o^2 (x^2 + y^2) \quad ,$$

and free to move along the z-axis.

(ii) Derive the density of eigenstates for the preceding problem if, in addition to the two-dimensional harmonic potential, the particle moves in the presence of a uniform magnetic field

$$\underline{\mathcal{B}} = \mathcal{B} \, \hat{e}_z$$

(iii) Discuss the behavior of the density of states in the limit of

$$\frac{\omega_o}{|\omega_c|} \to 0 \quad .$$

6.11 Prove equations 6.48.

6.12 Prove equations 6.53.

6.13 Prove equations 6.54 and 6.55.

<u>Problems</u> (continued)

6.14 Substituting the transformation into spherical coordinates (equation 6.57) into the expression for the three components of angular momentum as given by equations 6.47, prove equations 6.58.

6.15 In the absorption spectrum of $HC\ell$ vapor in the far infra-red, the following lines are observed (wavelenths per centimeter):

λ^{-1} (cm^{-1})

.

.

.

83.03

104.1

124.30

145.03

165.51

185.86

206.38

226.50

.

.

(this is not an exhaustive list). Compute the distances r_0 between the two atoms.

<u>Data</u>: G. Herzberg, <u>Spectra of Diatomic Molecules</u> (1950, 2nd Ed.), table 10.

<u>Note</u>: Strictly speaking, the interatomic distance is a function of ℓ, since at higher ℓ's centrifugal forces in fact slightly stretch the molecule. This effect is less important at the lower ℓ's, which should therefore be used in estimating the interatomic distance.

6.16 a. If the photons absorbed (or emitted) by the molecules have an energy $\hbar\omega$ in a fixed laboratory frame, compute the spectrum of photon energies encountered (or emitted) by a moving molecule of total mass M in thermal equilibrium at a temperature T. This phenomenon is called <u>Doppler broadening.</u>

b. Using the data of the preceding problem for $HC\ell$, compute the expected line widths $\Delta\omega_D$ as a result of this phenomenon.

c. Compute the recoil ΔE_{CM} for the emission (or absorption) of the photon by a molecule of mass M.

6.17 Compute the matrix elements

$$\langle \ell, m' | L_x | \ell, m \rangle$$

and

$$\langle \ell, m' | L_y | \ell, m \rangle \quad .$$

6.18 (i) Using equation 6.89 show that the associated Legendre functions are given by the product of $(1 - x^2)^{m/2}$ with an even or odd polynomial depending on whether $(\ell - m)$ is even or odd respectively, and consequently

$$P_{\ell, m}(-x) = (-1)^{\ell - m} P_{\ell, m}(x)$$

Problems (continued)

(ii) Show that inversion, i.e.

$$\underline{x} \rightarrow -\underline{x}$$

is equivalent to the transformation in spherical coordinates.

$$\theta \rightarrow \pi - \theta$$

$$\varphi \rightarrow \pi + \varphi \quad .$$

(iii) Show that

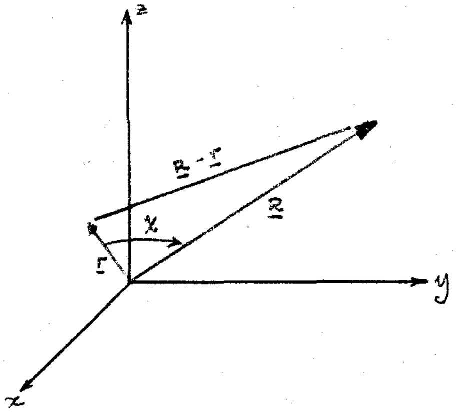
$$Y_{\ell, m}(\pi - \theta, \pi + \varphi) = (-1)^{\ell} \, Y_{\ell, m}(\theta, \varphi) \quad .$$

In other words, eigenstates with odd ℓ have odd parity (anti-symmetric), where as eigenstates with even ℓ have even parity (symmetric under inversion).

6.19 Using the generating function for the Legendre polynomials, show that

$$\frac{1}{|\underline{R} - \underline{r}|} = \frac{1}{R} \sum_{\ell=0}^{\infty} \left(\frac{r}{R}\right)^{\ell} P_{\ell}(\cos \chi) \quad ,$$

where $r = |\underline{r}|$, $R = |\underline{R}|$, $r/R < 1$, P_{ℓ} is the ℓ^{th} Legendre polynomial, and χ is the angle subtended between $\underline{r}$ and $\underline{R}$.



6.20 Using the results of the preceding problem, show also that

$$\frac{1}{|\underline{R} - \underline{r}|} = \frac{4\pi}{R} \sum_{\ell=0}^{\infty} \frac{(r/R)^{\ell}}{2\ell + 1} \sum_{m=-\ell}^{\ell} Y^{*}_{\ell, m}(\Theta, \Phi) \, Y_{\ell, m}(\theta, \varphi) \quad ,$$

where $\hat{e}_{\underline{r}} = (\theta, \varphi)$ and $\hat{e}_{\underline{R}} = (\Theta, \Phi)$.

update

7. MOTION IN A CENTRAL POTENTIAL

We have seen that the motion of a pair of particles subject to a mutual potential that is a function of their separation only can be reduced to the free motion of the center of mass and the motion about the center of mass. In particular, if the potential is given by the function U(r), where $r = |\underline{x}_2 - \underline{x}_1|$ is the separation of the two particles located at $\underline{x}_1$ and $\underline{x}_2$ respectively, the total Hamiltonian of the system would be given by

$$H = \frac{1}{2M} P^2 + \frac{1}{2m_r} p^2 + U(r) \ , \tag{7.1}$$

where $\underline{P}$ is the conjugate momentum to the coordinate of the center of mass (equation 6.72a), $\underline{p}$ is the conjugate momentum to the relative coordinate (equation 6.72b), M is the total mass (equation 6.73a) and m_r is the reduced mass (equation 6.73b). As in example 6.2, the total eigenfunctions $\Phi(\underline{X}, \underline{x})$ are separable into the product

$$\Phi(\underline{X}, \underline{x}) = \Psi(\underline{X}) \ \psi(\underline{x}) \tag{7.2}$$

while the total energy E_T is given by the sum of the kinetic energy of the center of mass E_{CM} and the energy of the relative motion about the center of mass E, i.e.

$$E_T = E_{CM} + E \ . \tag{7.3}$$

Substituting 7.2 into the Schrödinger equation we have

$$- \frac{\hbar^2}{2M} \left(\frac{\partial}{\partial \underline{X}} \cdot \frac{\partial}{\partial \underline{X}}\right) \Psi(\underline{X}) = E_{CM} \Psi(\underline{X}) \tag{7.4a}$$

and

$$\left[- \frac{\hbar^2}{2m} \left(\frac{\partial}{\partial \underline{x}} \cdot \frac{\partial}{\partial \underline{x}}\right) + U(|\underline{x}|)\right] \psi(\underline{x}) = E\psi(\underline{x}) \ . \tag{7.4b}$$

The wavefunction $\Psi(\underline{X})$ corresponds to free particle solutions (see example 6.2). Equation 7.4 can be expressed in spherical coordinates to yield (see equation 6.74 and 6.59)

$$\left[- \frac{\hbar^2}{2m_r r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{2m_r r^2} L^2 + U(r) - E\right] \psi(r, \theta, \varphi) = 0 \ , \tag{7.5}$$

where L^2 is the angular momentum (squared) operator.

Equation 7.5 is separable, in particular if

$$\psi(r, \theta, \varphi) = R_{E, \ell}(r) \ Y_{\ell, m}(\theta, \varphi) \ , \tag{7.6}$$

then, since

$$L^2 \ Y_{\ell, m}(\theta, \varphi) = \hbar^2 \ell(\ell+1) \ Y_{\ell, m}(\theta, \varphi) \ , \tag{7.7}$$

we have for the radial function $R_{E, \ell}(r)$,

$$\left\{\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) - \frac{\ell(\ell+1)}{r^2} + \frac{2m_r}{\hbar^2}\left[E - U(r)\right]\right\} R_{E,\ell}(r) = 0 \ . \tag{7.8}$$

We now note that if we make the substitution

$$R_{E,\ell}(r) = \frac{1}{r}\,\chi_{E,\ell}(r) \ , \tag{7.9}$$

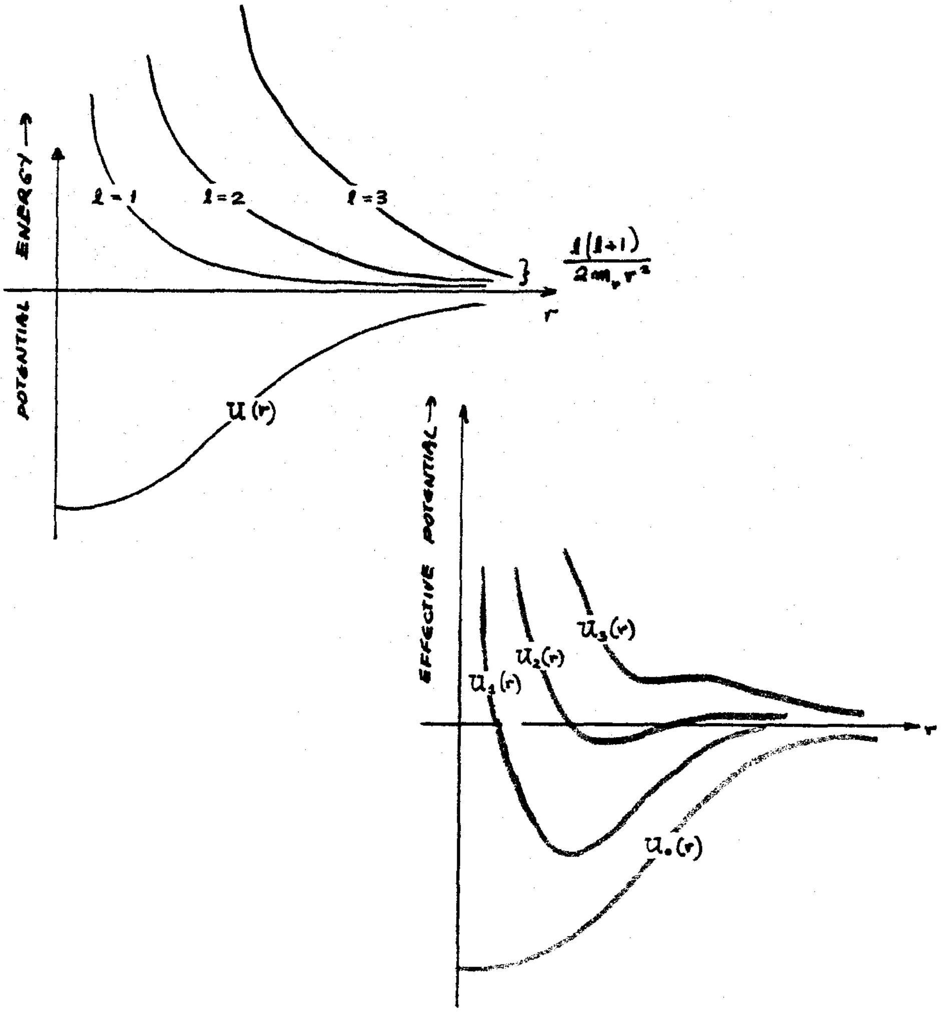
the equation simplifies to

$$\frac{d^2}{dr^2}\chi_{E,\ell}(r) + \left\{\frac{2m_r}{\hbar^2}\left[E - U(r)\right] - \frac{\ell(\ell+1)}{r^2}\right\} \chi_{E,\ell}(r) = 0 \ . \tag{7.10}$$

This latter equation is equivalent to <u>one dimensional motion in an effective potential</u> (that includes the centrifugal forces)

$$U_\ell(r) = U(r) + \left(\frac{\hbar^2}{2m_r}\right)\frac{\ell(\ell+1)}{r^2} \ , \tag{7.11}$$

with the restriction, however, that only positive values of r are allowed.



If the wavefunction is to remain finite and square integrable we must have

$$\chi(0) = \chi(\infty) = 0 \ . \tag{7.12}$$

Equations 7.10 and 7.11 define the eigenvalue problem. Note that the radial function R(r) is a function of both E and ℓ, in as much as both of these quantities enter as parameters in the (reduced) radial equation 7.10.

One can study the radial behavior of the eigenfunctions near the origin as a function of the angular momentum quantum number ℓ. In particular, if the potential is such that

$$\lim_{r\to 0} \{r^2 U(r)\} = 0 \ , \tag{7.13}$$

then if

$$R_{E,\ell}(r) \to r^s \quad \text{as} \quad r \to 0 \ ,$$

we have from equation 7.8

$$\frac{d}{dr}\left(r^2 \frac{dr^s}{dr}\right) - \ell(\ell+1)\, r^s = 0 \ ,$$

or

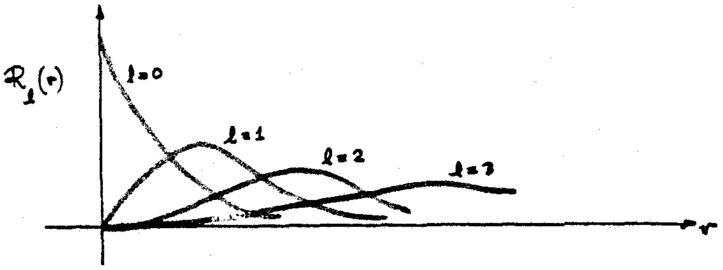
$$s(s+1) = \ell(\ell+1)$$

and therefore

$$s = \ell \quad \text{or} \quad s = -(\ell+1) \ .$$

The latter solution is unacceptable because it is infinite at $r = 0$, and therefore

$$R_{\varepsilon,\ell}(r) \to r^\ell \quad \text{as} \quad r \to 0 \ . \tag{7.14}$$



It can be seen that the wavefunctions are pushed further and further away from the origin as the angular momentum quantum number ℓ increases, consistently with our notion of an increasing centrifugal tendency with increasing angular momentum.

If the potential is localized, i.e., if

$$\lim_{r\to\infty} \{U(r)\} = 0 \ , \tag{7.15}$$

we can also study the behavior of these solutions away from the origin. In particular, for a bound state ($E < 0$) we would have

$$\chi''_{E,\ell}(r) + \frac{2m_r E}{\hbar^2}\,\chi_{E,\ell}(r) \to 0 \quad \text{as} \quad r \to \infty\,, \tag{7.16}$$

and therefore

$$\chi_{E,\ell}(r) \to e^{-\kappa r} \quad \text{as} \quad r \to \infty\,, \tag{7.17}$$

where

$$\kappa = \left(-\frac{2m_r E}{\hbar^2}\right)^{1/2}\,, \tag{7.18}$$

representing a spatially confined solution around the origin.

If the radial potential $U(r)$ satisfies both conditions, i.e., equation 7.13 at $r \to 0$ and equation 7.15 at $r \to \infty$, it is often convenient to remove the behavior at $r \to 0, \infty$ from the solutions and to define a new function $w(\rho)$ such that

$$\chi_{E,\ell}(r) = \rho^{\ell+1}\, e^{-\rho}\, w(\rho)\,, \tag{7.19}$$

where ρ is the dimensionless radius

$$\rho = \kappa r\,. \tag{7.20}$$

See equations 7.9, 7.14 and 7.17. The equation for $w(\rho)$ then becomes

$$w''(\rho) + 2\left(\frac{\ell+1}{\rho} - 1\right) w'(\rho) + \left[\frac{U(\rho/\kappa)}{E} - \frac{2(\ell+1)}{\rho}\right] w(\rho) = 0\,. \tag{7.21}$$

7.1 Motion in a constant potential. Free particle motion in spherical coordinates.

The $U(r) = 0$ (or a constant) is, of course, a special case of a central potential. Consequently, it is possible to seek wavefunction solutions that are seperable in spherical coordinates. We have already discussed free particle behavior in three dimensions (sections 4.2, 4.3.3), where we found that imposing periodic boundary conditions on the wavefunctions results in eigenfunctions of definite linear momentum (see equation 4.92 and related discussion). The discussion of free particle motion in the present context of spherical coordinates is important because it yields eigenfunctions of definite angular momentum. It goes without saying that the two representations are entirely equivalent, both formulations yielding a complete set of eigenstates in terms of which any particular solution can be expressed. The latter set of solutions are generally usefull in studying scattering problems far from the immediate vicinity of a localized scattering potential, where the particles are essentially free.

Setting $U(r) = 0$, we have for the Schrödinger equation in spherical coordinates, from equation 7.5,

$$\left[-\frac{\hbar^2}{2m_r r^2}\frac{\partial}{\partial r}(r^2\frac{\partial}{\partial r}) + \frac{1}{2m_r r^2}L^2 - E\right]\psi(r,\theta,\varphi) = 0 \tag{7.22}$$

which is separable, as in equation 7.6, to yield a radial equation

$$\left[\frac{1}{r^2}\frac{d}{dr}(r^2\frac{d}{dr}) + k^2 - \frac{\ell(\ell+1)}{r^2}\right]\mathcal{R}_{k,\ell}(r) = 0 \quad , \tag{7.23}$$

where

$$k^2 = \frac{2m_r E}{\hbar^2} \quad . \tag{7.24}$$

The solution of equation 7.23 is given in terms of the <u>spherical Bessel functions</u>, $j_\ell(kr)$ and $n_\ell(kr)$, which have an asymptotic behavior for small kr,

$$j_\ell(kr) \sim \frac{1}{(2\ell+1)!!}(kr)^\ell \quad , \quad kr \to 0 \tag{7.25a}$$

and

$$n_\ell(kr) \sim \frac{(2\ell+1)!!}{(2\ell+1)}(kr)^{-(\ell+1)} \quad , \quad kr \to 0 \tag{7.25b}$$

or, alternatively, the <u>spherical Hankel functions</u>, defined by

$$h_\ell^{(\pm)}(kr) = n_\ell(kr) \pm j_\ell(kr) \quad , \tag{7.26}$$

which have an asymptotic behavior for large kr, given by

$$h_\ell^{(\pm)}(kr) \sim \frac{1}{kr}e^{\pm ikr} \quad , \quad kr \to \infty \quad . \tag{7.27}$$

See appendix H. Consequently, for solutions in regions including the origin, we must have

$$\mathcal{R}_{k,\ell}(r) = j_\ell(kr) \quad , \tag{7.28}$$

within a normalization factor, since the other solution is infinite at $r = 0$.

The transition from plane travelling waves of definite linear momentum to a superposition of radial functions can be realized using the important identity

$$e^{ikz} = \sum_{\ell=0}^{\infty} i^\ell(2\ell+1)\, j_\ell(kr)\, P_\ell(\cos\theta) \quad , \tag{7.29}$$

where $P_\ell(\cos\theta)$ is the ℓ^{th} <u>Legendre polymonial</u>, equal to the corresponding m=0 associated Legendre function (see pages 6.20 - 21)

$$P_\ell(x) = P_{\ell,o}(x) \equiv \frac{1}{2^\ell \ell!}\frac{d^\ell}{dx^\ell}(x^2 - 1)^2 \quad , \tag{7.30}$$

or, in terms of the m=0 spherical harmonics,

$$P_\ell(\cos\theta) = (\frac{2\pi}{2\ell+1})^{1/2}\; Y_{\ell,o}(\theta,\varphi) \quad . \tag{7.31}$$

update

This is useful in partial wave analysis of scattering problems. See for example reference 7.1, section 16-3, and reference 7.2, section 11.5 and 11.6.

coordinates, from equation 7.5,

$$\left[-\frac{\hbar^2}{2m_r r^2}\frac{\partial}{\partial r}(r^2\frac{\partial}{\partial r}) + \frac{1}{2m_r r^2} L^2 - E\right]\psi(r,\theta,\varphi) = 0 \qquad (7.22)$$

which is separable, as in equation 7.6, to yield a radial equation

$$\left[\frac{1}{r^2}\frac{d}{dr}(r^2\frac{d}{dr}) + k^2 - \frac{\ell(\ell+1)}{r^2}\right]\mathcal{R}_{k,\ell}(r) = 0 \quad , \qquad (7.23)$$

where

$$k^2 = \frac{2m_r E}{\hbar^2} \quad . \qquad (7.24)$$

The solution of equation 7.23 is given in terms of the spherical Bessel functions, $j_\ell(kr)$ and $n_\ell(kr)$, which have an asymptotic behavior for small kr,

$$j_\ell(kr) \sim \frac{1}{(2\ell+1)!!}(kr)^\ell \quad , \quad kr \to 0 \qquad (7.25a)$$

and

$$n_\ell(kr) \sim \frac{(2\ell+1)!!}{(2\ell+1)}(kr)^{-(\ell+1)} \quad , \quad kr \to 0 \qquad (7.25b)$$

or, alternatively, the spherical Hankel functions, defined by

$$h_\ell^{(\pm)}(kr) = n_\ell(kr) \pm j_\ell(kr) \quad , \qquad (7.26)$$

which have an asymptotic behavior for large kr, given by

$$h_\ell^{(\pm)}(kr) \sim \frac{1}{kr} e^{\pm ikr} \quad , \quad kr \to \infty \quad . \qquad (7.27)$$

See appendix H. Consequently, for solutions in regions including the origin, we must have

$$\mathcal{R}_{k,\ell}(r) = j_\ell(kr) \quad , \qquad (7.28)$$

within a normalization factor, since the other solution is infinite at $r = 0$.

The transition from plane travelling waves of definite linear momentum to a superposition of radial functions can be realized using the important identity

$$e^{ikz} = \sum_{\ell=0}^{\infty} i^\ell(2\ell+1)\, j_\ell(kr)\, P_\ell(\cos\theta) \quad , \qquad (7.29)$$

where $P_\ell(\cos\theta)$ is the ℓ^{th} Legendre polymonial, equal to the corresponding $m=0$ associated Legendre function (equation 6.89)

$$P_\ell(x) = P_{\ell,0}(x) \equiv \frac{1}{2^\ell \ell!}\frac{d^\ell}{dx^\ell}(x^2 - 1)^2 \quad , \qquad (7.30)$$

or, in terms of the $m=0$ spherical harmonics,

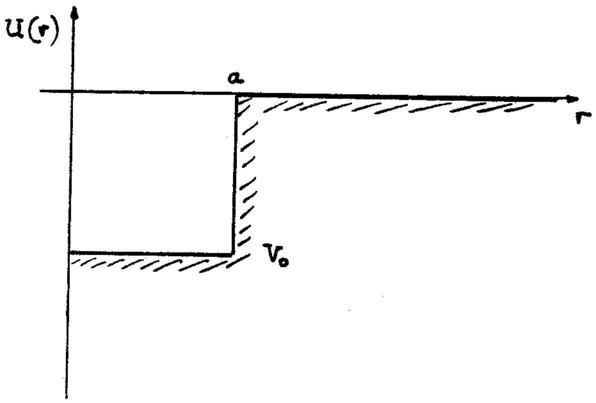
$$P_\ell(\cos\theta) = (\frac{2\pi}{2\ell+1})^{1/2}\, Y_{\ell,0}(\theta,\varphi) \quad . \qquad (7.31)$$

This is useful in partial wave analysis of scattering problems. See for example reference 7.1, section 16-3, and reference 7.2, section 11.5 and 11.6.

7.2 Bound states of a spherical potential well.

Consider the radial potential

$$U(r) = \begin{cases} V_0 < 0 , & \text{for } r < a \\ 0 , & \text{for } r > a . \end{cases} \qquad (7.32)$$



From the general results for motion in a radial potential, we have that the wavefunctions are given by

$$\psi(r,\theta,\varphi) = \mathcal{R}_{E,\ell}(r)\, Y_{\ell,m}(\theta,\varphi) \qquad (7.33)$$

where

$$\left[\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) + k^2 - \frac{\ell(\ell+1)}{r^2}\right]\mathcal{R}_{E,\ell}(r) = 0 \quad , \quad \text{for } r < a \qquad (7.34)$$

$$\left[\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) - \kappa^2 - \frac{\ell(\ell+1)}{r^2}\right]\mathcal{R}_{E,\ell}(r) = 0 \quad , \quad \text{for } r > a ,$$

and where (recall $V_0 < E < 0$)

$$k^2 = \frac{2m_r}{\hbar^2}(E - V_0) \quad , \qquad \kappa^2 = \frac{2m_r}{\hbar^2}(-E) \qquad (7.35)$$

The solutions to these equations can be expressed in terms of the free particle solutions of the preceding section. In particular, within a normalization constant,

$$\mathcal{R}_{k,\ell}(r) = \begin{cases} \dfrac{j_\ell(kr)}{j_\ell(ka)} , & r < a \\[2ex] \dfrac{h_\ell^{(+)}(i\kappa r)}{h_\ell^{(+)}(i\kappa a)} , & r > a \end{cases} \qquad (7.36)$$

where the choice for $r < a$ is dictated by matching the behavior at $r \to 0$ (see equations 7.14 and 7.25a), and for $r > a$ by matching the behavior at $r \to \infty$ (see equations 7.17 and 7.27). The energy eigenvalues are then obtained by matching the derivative of $\mathcal{R}_{k,\ell}(r)$ at $r=a$ $[R_{k,\ell}(r=a^-) = R_{k,\ell}(r=a^+)$ by construction].

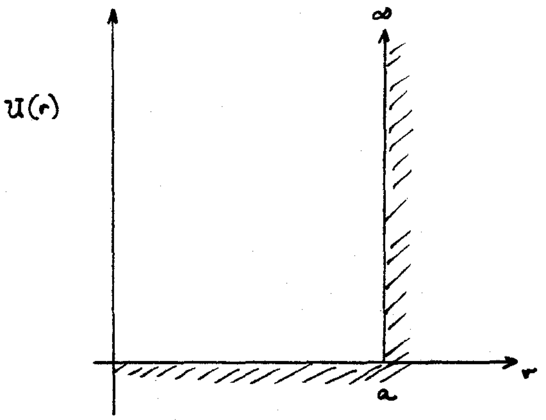
It should be noted that the zero angular momentum solutions ($\ell=0$) of the corresponding one dimensional equation of

$$\chi_{E,0}(r) = r\,\mathcal{R}_{E,0}(r) \tag{7.37}$$

are the same as the <u>odd</u> solutions of the one dimensional rectangular potential well discussed in section 4.2.4 (why odd?).

To examine the behavior of solutions with angular momentum, we will consider instead the mathematically simpler infinite spherical well

$$U(r) = \begin{cases} 0 \ , & \text{for} \quad r < a \\ \infty \ , & \text{for} \quad r > a \ , \end{cases} \tag{7.38}$$



for which $\mathcal{R}(r) = 0$ for $r \geq a$. We then have

$$\mathcal{R}_{E,\ell}(r) = j_\ell(kr) \tag{7.39}$$

where, if E is measured from the bottom of the well,

$$k^2 = \frac{2m_r E}{\hbar^2} \ , \tag{7.40}$$

and such that

$$j_\ell(ka) = 0 \ . \tag{7.41}$$

Equation 7.41 is the eigenvalue equation for this problem in that the admissible values of k correspond to the roots of the ℓ^{th} spherical Bessel function. If we denote the n^{th} root of the ℓ^{th} Bessel function by $\rho_{n,\ell}$, i.e.

$$j_\ell(\rho_{n,\ell}) = 0 \quad ; \quad n = 1,2,\dots. \tag{7.42}$$

we have

$$k_{n,\ell} = \frac{1}{a}\,\rho_{n,\ell} \ . \tag{7.43}$$

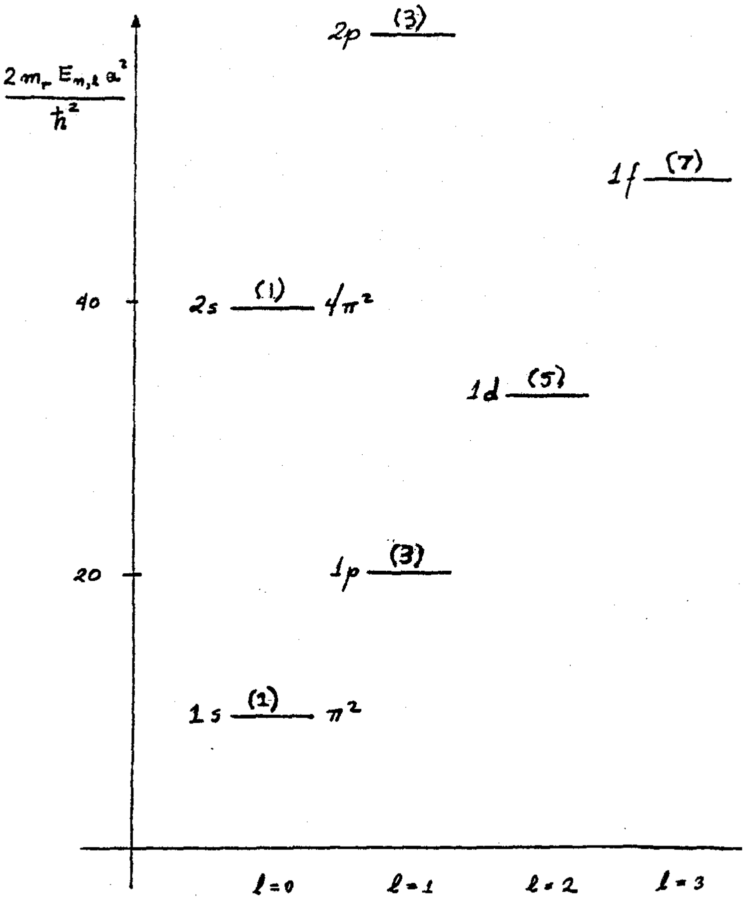
The first few of these, in increasing order, are tabulated below.

n	ℓ	$\rho_{n,\ell}$	$\rho^2_{n,\ell}$
1	0	$3.142 = \pi$	9.872
1	1	4.493	20.187
1	2	5.763	33.212
2	0	$6.283 = 2\pi$	39.476
1	3	6.988	48.832
2	1	7.725	59.676

The energy eigenvalues are then given by

$$E_{n,\ell} = \frac{\hbar^2}{2m_r a^2}\,\rho^2_{n,\ell} \quad , \tag{7.44}$$

and are sketched below.



Note that the $n=1$ states, corresponding to the first root of $j_\ell(\rho)$ at $\rho = ka$, have no nodes (in the interior of the well). The $n=2$ states have one node, the $n=3$ have two nodes, etc. In general, for a given angular momentum quantum number ℓ, the energy is monotonic with the number of radial nodes. A similar result was found for the eigenvalues of the one-dimensional potential well [page 4.20, article (iv) and second footnote]. This is a general property of the differential Schrödinger equation, which is a Sturm-Louisville problem. See for example reference 3, section 6.3, pages 719-725.

It should also be noted that the energy eigenvalues $E_{n,\ell}$ are $(2\ell+1)$-fold degenerate, corresponding to the $2\ell+1$ orthogonal eigenstates

$$\psi_{n,\ell,m}(r,\theta,\varphi) = \mathcal{R}_{n,\ell}(r)\, Y_{\ell,m}(\theta,\varphi) \rightleftharpoons |n,\ell,m\rangle \tag{7.45}$$

with different m's, the z-angular momentum quantum numbers. The degeneracy of each (n,ℓ) eigenvalue is indicated in parentheses in the sketch above.

7.3 Motion in an attractive Coulomb potential. Hydrogenic wavefunctions.

An important special case of motion in a central potential is the motion of two oppositely charged particles about each other subject to their mutual electrostatic forces. In that case, we have

$$U(r) = -\gamma/r\ . \tag{7.46}$$

Note that for the motion of a single electron around a nucleus of atomic number Z we have

$$\gamma = \frac{Ze^2}{4\pi\epsilon_0}\ , \tag{7.47}$$

where e is the charge of an electron

$$e = 1.6 \times 10^{-19}\ \text{Cb}\ . \tag{7.48}$$

Since this potential satisfies equation 7.13 at the origin and equation 7.15 at infinity, we can make the substitution of equation 7.19 and solve for the function $w(\rho)$. Note that in this case

$$\frac{U(\rho/\kappa)}{E} = \frac{\gamma\kappa}{(-E)\rho} = \frac{\rho_0}{\rho} \tag{7.49}$$

where

$$\rho_0 = \frac{\gamma}{\hbar}\left(\frac{2m_r}{-E}\right)^{1/2}\ . \tag{7.50}$$

Therefore equation 7.21 becomes

$$\rho w''(\rho) + 2(\ell+1-\rho)\,w'(\rho) + [\rho_0 - 2(\ell+1)]\,w(\rho) = 0\ . \tag{7.51}$$

Substituting a power series

$$w(\rho) = \sum_{k=0}^{\infty} C_k \rho^k , \tag{7.52}$$

we find the recursion relation

$$C_{k+1} = \frac{2(k+\ell+1) - \rho_0}{(k+1)(k+2\ell+2)} C_k . \tag{7.53}$$

Note that

$$\frac{C_{k+1}}{C_k} \rightarrow \frac{2}{k} \qquad \text{as} \qquad k \rightarrow \infty$$

so that, unless the series terminates,

$$w(\rho) \rightarrow e^{2\rho} .$$

Consequently, we must have that for some k, say $k = K$,

$$\rho_0 = 2(K+\ell+1) \tag{7.54}$$

where $\ell=0, 1, 2, \ldots$ and $K=0, 1, 2, \ldots$. Alternatively, if n is an integer, such that

$$n = K+\ell+1 = 1, 2, 3, \ldots \tag{7.55}$$

we must have

$$\rho_0 = 2n . \tag{7.56}$$

Substituting for ρ_0, we recover for the energies of the bound states, i.e.

$$E_n = - \frac{m_r \gamma^2}{2\hbar^2 n^2} . \tag{7.57}$$

For the motion of an electron around a nucleus of charge Ze, γ is given by equation 7.47 and the energy levels are given by

$$E_n = - \left(\frac{Ze^2}{4\pi\varepsilon_0} \right)^2 \frac{m_r}{2\hbar^2 n^2} . \tag{7.58}$$

Substituting in the expression for κ (equation 7.18) we have

$$\kappa = Z/na_0 \tag{7.59}$$

where

$$a_0 = \left(\frac{4\pi\varepsilon_0}{e^2} \right) \frac{\hbar^2}{m_r} = 0.529 \times 10^{-8} \text{ cm} \tag{7.60}$$

is the <u>Bohr radius</u> of the hydrogen atom of the old quantum theory. In terms of a_0 the energy can then be expressed as

$$E_n = - \frac{1}{2n^2} \left(\frac{Z^2 e^2}{4\pi\varepsilon_0 a_0} \right) \tag{7.61}$$

Alternatively, we can multiply and divide equation 7.58 by the square of the speed of light to obtain

$$E_n = -\frac{Z^2}{2}\left(\frac{\alpha}{n}\right)^2 m_r c^2 \qquad (7.62)$$

where α is a dimensionless constant

$$\alpha \equiv \frac{e^2}{4\pi\varepsilon_0 \hbar c} \simeq \frac{1}{137} \ , \qquad (7.63)$$

called the *fine structure constant*.

Setting $Z = 1$ and $n = 1$, we obtain *the ground state (binding) energy of the hydrogen atom*

$$E_1 = -\frac{1}{2}\left(\frac{e^2}{4\pi\varepsilon_0 a_0}\right) = -13.6\,\text{eV} \ . \qquad (7.64)$$

This quantity is called the *Rydberg*. Successive energy levels are given by the sequence

$$E_n = \frac{E_1}{n^2} \ , \quad n = 1, 2, \ldots . \qquad (7.65)$$

and represent the *Balmer series* of hydrogen.

Note that the integer n, called the *principal quantum number*, is given by equation 7.55, where K is the degree of the polynomial representing the function $w(\rho)$. Note also that since

$$\ell = n - K - 1 \ ,$$

for a given n, ℓ can take the values

$$\ell = 0, 1, \ldots , \quad n - 1 \ . \qquad (7.66)$$

For a given n, however, the corresponding eigenfunctions

$$\psi_{n,\ell,m}(r,\theta,\varphi) = R_{n,\ell}(r)\, Y_{\ell,m}(\theta,\varphi)$$

to each value of ℓ share the same energy E_n. On the other hand, for a given n *and* ℓ we can have $(2\ell+1)$ orthogonal eigenstates corresponding to the $(2\ell+1)$ possible values of m, the eigenvalues of $L_z/\hbar$. Consequently, the degeneracy of a particular energy level E_n is given by

$$\sum_{\ell=0}^{n-1} (2\ell+1) = n(n-1) + n = n^2 \ . \qquad (7.67)$$

update

It should be pointed out that the $(2\ell+1)$ degeneracy for a given value of ℓ is a consequence of the isotropy (spherical symmetry) of the potential (and therefore also the Hamiltonian). The remaining degeneracy, however, is a coincidence of the eigenvalues, i.e.

the energy is a function of the radial quantum number n, not of n and ℓ, as would be natural for this type of problem. This is peculiar to the Coulomb potential (and the spherically symmetric harmonic oscillator potential) and is not true for any other potential. Recall, for example, the spherical potential well (section 7.2), whose energy eigenvalues are a function of both n and ℓ.

In spectroscopic vernacular, an $\ell = 0$ state is called an "s" state, an $\ell = 1$ state is called a "p" state, the sequence is given by

ℓ	state	
0	s	
1	p	
2	d	
3	f	
4	g	alphabetical .
5	h	
.	.	
.	.	
.	.	
.	.	

A state (n, ℓ) is then given a number corresponding to n and a letter corresponding to ℓ, e.g., 3p is a $n = 3$ and $\ell = 1$ state.

It is illustrative to make an energy diagram of the various states. The spectroscopic designation is indicated to the left of every state and the degeneracy is indicated in parentheses on top. In fact the degeneracy is twice that if the spin, which represents another degree of freedom, is taken into account as we shall see later.

The functions $w_{n,\ell}(\rho)$ are proportional to the associated Laguerre polynomials. In particular, within a normalization factor,

$$w_{n,\ell}(\rho) = L_{n-\ell-1,\,2\ell+1}(2\rho) \quad , \tag{7.68}$$

where

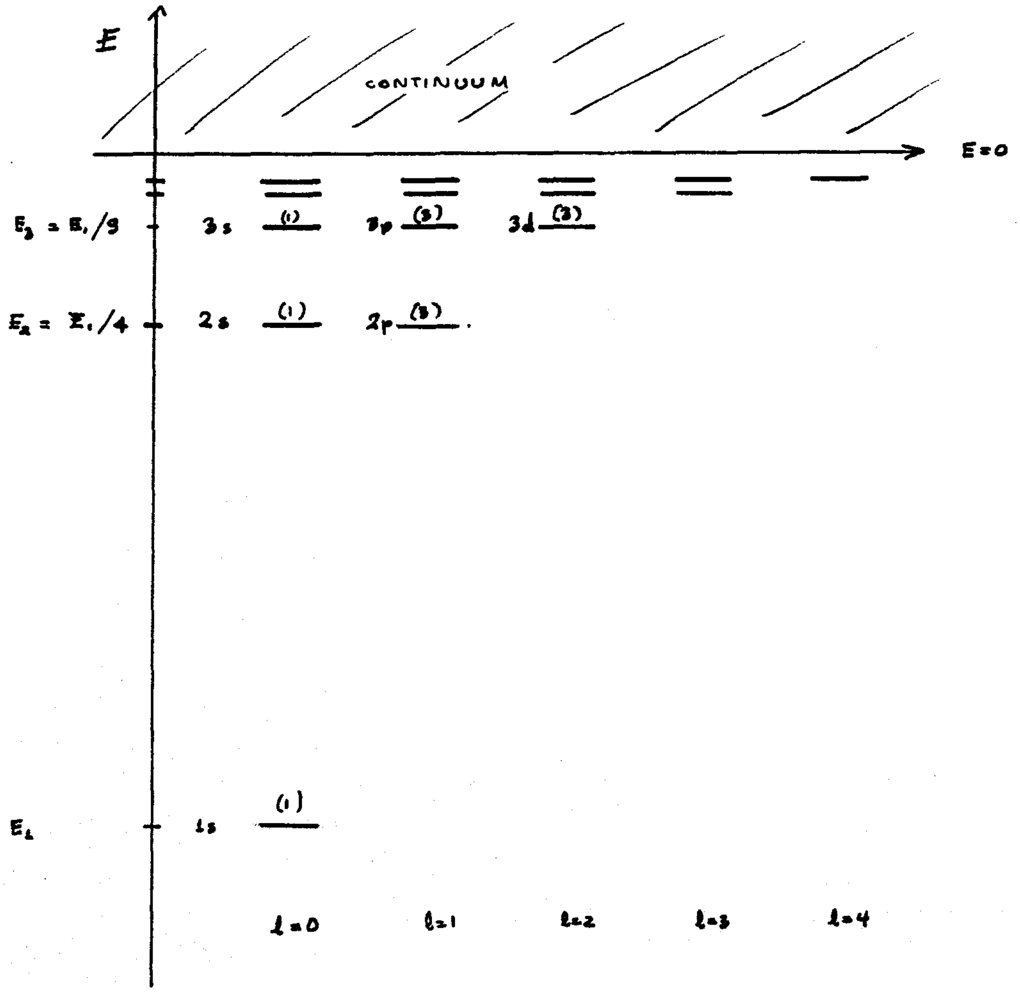
$$L_{n,m}(x) \equiv (-1)^m \frac{d^m}{dx^m} L_{n+m}(x) \quad , \tag{7.69}$$

and $L_n(x)$, the Laguerre polynomial, is defined by,

$$L_n(x) \equiv e^x \frac{d^n}{dx^n} (x^n e^{-x}) \quad . \tag{7.70}$$

Since the spherical harmonics are already normalized, it is sufficient to normalize the radial function, i.e.

$$\int_0^\infty R^2_{n,\ell}(r)\, r^2\, dr = \int_0^\infty \chi^2_{n,\ell}(r)\, dr = 1 \quad . \tag{7.71}$$



Coulomb Potential Bound State
Energy Levels

Using the definite integral

$$\int_0^\infty e^{-x}\, x^{2(\ell+1)}\, [L_{n-\ell-1,\,2\ell+1}(x)]^2\, dx = \frac{2n[(n+\ell)!]^3}{(n-\ell-1)!} \tag{7.72}$$

we can compute the normalization constants and write the normalized wavefunctions as

$$\psi_{n,\ell,m}(r,\theta,\varphi) = \left[\frac{(n-\ell-1)!}{2n}\right]^{1/2} \left[\frac{2\kappa}{(n+\ell)!}\right]^{3/2} (2\kappa r)^{\ell}\, e^{-\kappa r}\, L_{n-\ell-1,\,2\ell+1}(2\kappa r)\, Y_{\ell,m}(\theta,\varphi) \tag{7.73}$$

The radial functions $R_{n,\ell}(r)$ of the first few of these are listed below

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

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

$$R_{2,1}(r) = \left(\frac{Z}{2a_0}\right)^{3/2} \frac{2}{\sqrt{3}}\left(\frac{Zr}{2a_0}\right)e^{-\left(\frac{Zr}{2a_0}\right)}$$

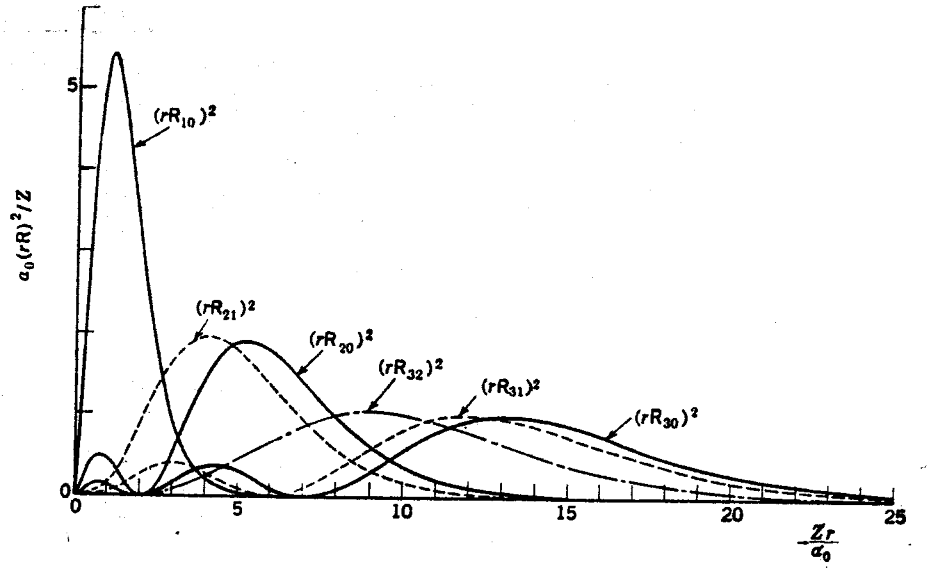
$$R_{3,0}(r) = \left(\frac{Z}{3a_0}\right)^{3/2} 2\left[1 - 2\left(\frac{Zr}{3a_0}\right) + \frac{2}{3}\left(\frac{Zr}{3a_0}\right)^2\right] e^{-\left(\frac{Zr}{3a_0}\right)} \tag{7.74}$$

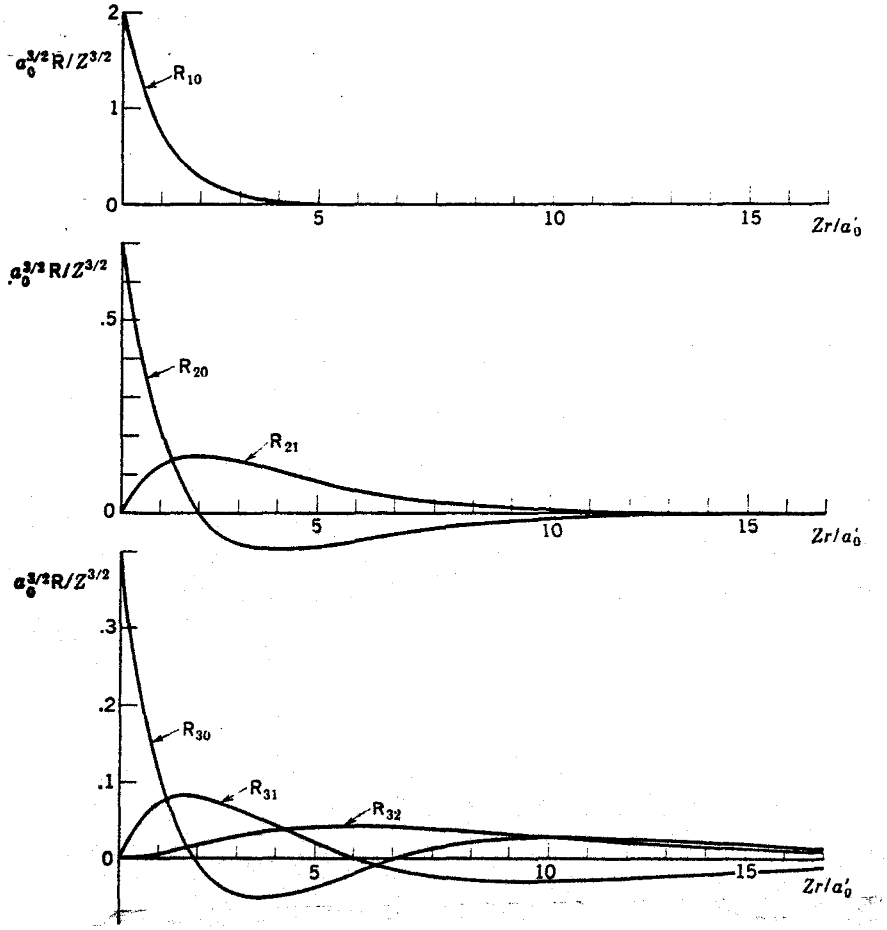
$$R_{3,1}(r) = \left(\frac{Z}{3a_0}\right)^{3/2} 4\cdot 2^{1/2} \left(\frac{Zr}{3a_0}\right)\left[1 - \frac{1}{2}\left(\frac{Zr}{3a_0}\right)\right] e^{-\left(\frac{Zr}{3a_0}\right)}$$

and

$$R_{3,2}(r) = \left(\frac{Z}{3a_0}\right)^{3/2} \frac{2}{3}\left(\frac{2}{5}\right)^{1/2} \left(\frac{Zr}{3a_0}\right)^2 e^{-\left(\frac{Zr}{3a_0}\right)},$$

where the substitution $\kappa = Z/(na_0)$ has been made. These functions are plotted below. The figures are from R. Leighton, Principles of Modern Physics.





References

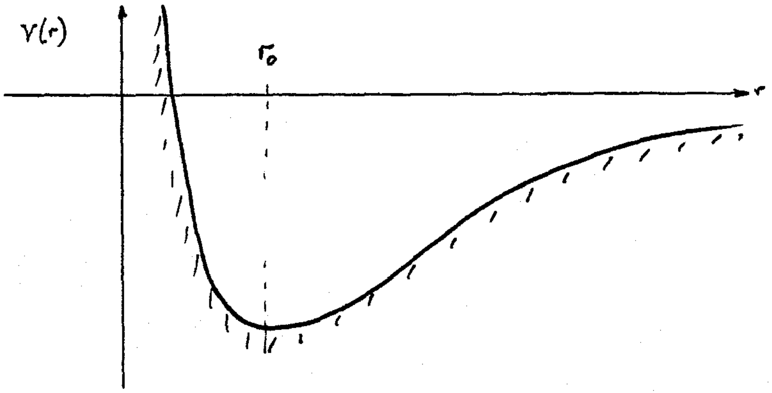
7.1 R.H. Dicke and J.P. Wittke, Introduction to Quantum Mechanics (Addison-Wesley, 1960).

7.2 E. Merzbacher, Quantum Mechanics (2nd edition, John Wiley, 1970).

7.3 P.M. Morse and H. Feshbach, Methods of Theoretical Physics (McGraw-Hill, 1953).

7.4 Vibrational-Rotational spectra of diatomic molecules.

The general form of the diatomic molecule potential is a repulsive region at close distances followed by an attractive region at larger distances. Such a potential will have a minimum at some value of $r = r_0$, which in the low energy approximation in example 6.2 of a rigid rotator we took to be equal to the interatomic separation. A typical diatomic molecule potential is sketched below.



We can approximate the potential in the vicinity of r_0 by a Taylor series $(V_0 < 0)$

$$V(r) \simeq V_0 + \frac{1}{2}(r - r_0)^2 \; V''(r_0) + \frac{1}{6}(r - r_0)^3 \; V'''(r_0) + \ldots \quad . \tag{7.75}$$

Keeping second order terms, we can then use this expansion as the radial potential $U(r)$ in equations 7.5 and 7.10 to yield the effective potential

$$U_\ell(r) \simeq V_0 + \frac{1}{2} m_r \omega_0^2 (r - r_0)^2 + \frac{\hbar^2}{2m_r} \cdot \frac{\ell(\ell+1)}{r^2} \quad , \tag{7.76}$$

where the substitution

$$m_r \omega_0^2 = V''(r_0) \quad , \tag{7.77}$$

has been made.

The molecule will rotate and execute vibrations about the minimum r_ℓ of the effective potential $U_\ell(r)$. To find this minimum, we compute the root of the first derivative of $U_\ell(r)$, i.e.

$$U'_\ell(r_\ell) = m_r \omega_0^2 (r_\ell - r_0) - 2 B_0 \ell(\ell+1) \frac{r_0^2}{r_\ell^3} = 0 \quad , \tag{7.78}$$

where

$$B_0 = \frac{\hbar^2}{2m_r r_0^2} \quad , \tag{7.79}$$

and therefore, we compute the stretching of the molecule in a rotational level ℓ,

$$r_\ell - r_0 \simeq r_0 \left(\frac{2B_0}{\hbar\omega_0}\right)^2 \ell(\ell+1) \quad , \tag{7.80}$$

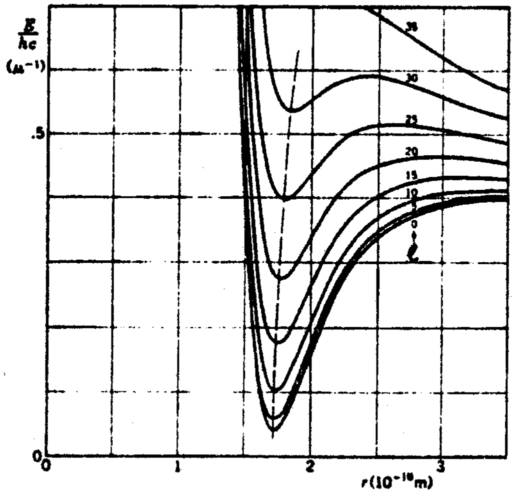


FIG. 9-7. The effective-potential curves for the diatomic molecule HgH. [After Villars and Condon, ***Phys. Rev.***, **35**, 1028 (1930).]

(The figure above is taken from R. Leighton, Principles of Modern Physics.)

where the assumption (almost always valid) has been made that

$$\left(\frac{2B_0}{\hbar\omega_0}\right)^2 \ll 1 \ . \tag{7.81}$$

We can now expand the effective potential $U_\ell(r)$ about its minimum at $r = r_\ell$, i.e.

$$U_\ell(r) \simeq U_\ell(r_\ell) + \frac{1}{2}(r - r_\ell)^2 U_\ell''(r_\ell) \tag{7.82}$$

or

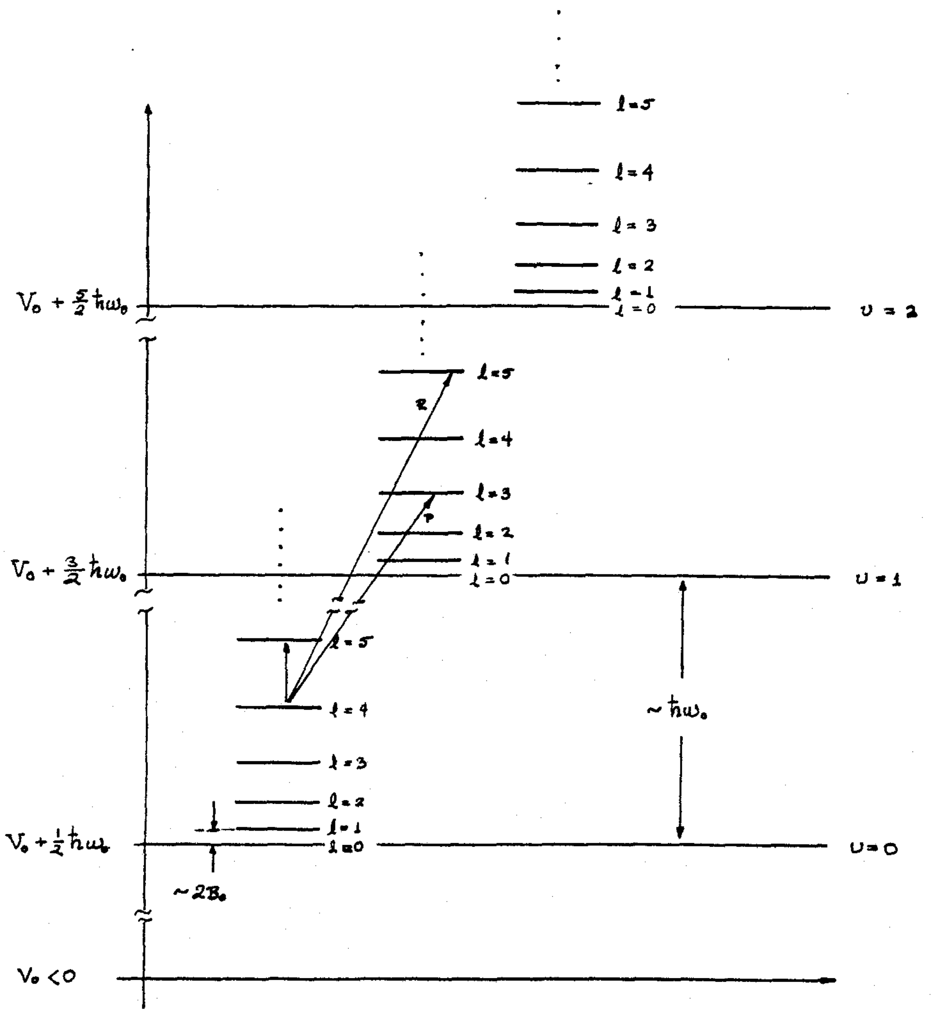
$$U_\ell(r) \simeq V_\ell + \frac{1}{2} m_r \omega_\ell^2 (r - r_\ell)^2 \ , \tag{7.83a}$$

where

$$V_\ell \simeq V_0 + B_0 \left[1 - \left(\frac{2B_0}{\hbar\omega_0}\right)^2 \ell(\ell+1)\right] \ell(\ell+1) \ , \tag{7.83b}$$

and

$$\omega_\ell \simeq \omega_0 \left[1 + \frac{3}{2}\left(\frac{2B_0}{\hbar\omega_0}\right)^2 \ell(\ell+1)\right] \ , \tag{7.83c}$$



Vibrational-rotational absorption transitions of a diatomic molecule.

The solutions of the resulting time-independent Schrödinger equation are then of the form

$$\psi_{v,\ell,m}(r,\theta,\varphi) = \frac{1}{r}\chi_{v,\ell}(r-r_\ell)\,Y_{\ell,m}(\theta,\varphi)\ , \qquad (7.84)$$

where $(\xi = r - r_\ell)$,

$$\chi''_{v,\ell}(\xi) + \frac{2m_r}{\hbar^2}\left[E_{v,\ell} - V_\ell - \frac{1}{2}m_r\omega_\ell^2\xi^2\right]\chi_{v,\ell}(\xi) = 0\ . \qquad (7.85)$$

Equation 7.85 is the Schrödinger equation for a one-dimensional harmonic oscillator, whose solution yields (see section 5.1)

$$\chi_{v,\ell}(r-r_\ell) = \left(\frac{m_r\omega_\ell}{\hbar}\right)^{1/4} A_v h_v\left[\left(\frac{m_r\omega_\ell}{\hbar}\right)^{1/2}(r-r_\ell)\right] e^{-\frac{1}{2}\left(\frac{m_r\omega_\ell}{\hbar}\right)(r-r_\ell)^2} \qquad (7.86)$$

and

$$E_{v,\ell} - V_o = \hbar\omega_o(v+\frac{1}{2}) + + B_o\,[1 + 3\,(\frac{2B_o}{\hbar\omega_o})(v+\frac{1}{2}) - (\frac{2B_o}{\hbar\omega_o})^2\,\ell(\ell+1)]\,\ell(\ell+1)\ . \qquad (7.87)$$

where $v = 0, 1, 2, 3, \ldots$ is called the <u>vibrational quantum number.</u>

Several important conclusions can be drawn from these results. In this approximation:

(i) The molecule executes harmonic vibrations about an equilibrium interatomic separation which increases with the orbital quantum number in accordance with equation 7.80.

(ii) For a given vibrational level v, the molecule rotates with a moment of inertia that <u>increases</u> with the rotational quantum number (reciprocal of quantity in the brackets of equation 7.87.

(iii) For a given rotational level ℓ, the molecule occupies energy levels that are equally spaced by an amount $\hbar\omega_\ell$, where ω_ℓ is given by equation 7.83c. This last conclusion is an artifact of the <u>quadratic</u> expansion of the effective potential $U_\ell(r)$ about $r \simeq r_\ell$. A better treatment will be described below.

Radiative transitions allow these levels to be observed directly through absorption and/or emission of light. One then observes that the transitions can only occur between states which differ from each other by one unit in the rotational quantum number ℓ.# In particular, if

$$\Delta v = v_{final} - v_{initial} = 0\ ,$$

we have

$$\Delta\ell = \ell_{final} - \ell_{initial} = \begin{cases} +1 & \text{(absorption)} \\ -1 & \text{(emission)} \end{cases} \qquad (7.88a)$$

and we obtain the <u>rotational</u> spectrum of the molecule (see for example problem 5.15). If the transition involves different vibrational levels, we have

---

#The reason for this, as we shall see later, is that a photon carries one unit of angular momentum.

$$\Delta v = v_{\text{final}} - v_{\text{initial}} = \begin{cases} +1 & \text{(absorption)} \\ -1 & \text{(emission)} \end{cases}$$

and (7.88b)

$$\Delta \ell = \ell_{\text{final}} - \ell_{\text{initial}} = \begin{cases} +1 & \text{(R-branch)} \\ -1 & \text{(P-branch)} \end{cases}$$

which yield the vibrational-rotational spectrum of the molecule. Note that for the P-branch ($\Delta\ell = -1$) $\ell_{\text{lower}} = 0$ is not possible.

Within the quadratic approximation of the effective potential, transitions between states which differ by more than one vibrational quantum number are not allowed. This conclusion, however, is also an artifact of this approximation and in practice, even though transitions between states with a $|\Delta v| > 1$ are found to be weak (occur with a small transition probability), they are observed nevertheless.

It is possible to give a better description of the diatomic molecule that would remove these discrepancies by including more terms in the Taylor expansion of the effective potential $U_\ell(r)$ about its minimum. The convergence of such an expansion, however, is slow and the analysis cumbersome. Alternatively, a better analytical functional representation of the interatomic potential $V(r)$ could be used that possesses the proper features. Such an analytical expression was proposed in 1929 by P. M. Morse (Phys. Rev. 34, 57), and is given by

$$V(r) = V_0 + \frac{1}{2} m_r \omega_0^2 \sigma^2 \left[1 - e^{-(r-r_0)/\sigma}\right]^2 , \tag{7.89}$$

where ω_0, σ and r_0 are selected to fit the particular diatomic molecule.

The Morse potential gives very accurate results for nearly all molecules and yields energy levels given by

$$\begin{aligned} E_{v,\ell} = {} & \hbar\omega_0 \left(v + \frac{1}{2}\right) - B_0 \left(\frac{r_0}{\sigma}\right)^2 \left(v + \frac{1}{2}\right)^2 + \\ & B_0 \left[1 - \alpha_0 \left(v + \frac{1}{2}\right) - \beta_0\, \ell(\ell+1)\right] \ell(\ell+1) , \end{aligned} \tag{7.90}$$

where α_0 and β_0 are (small) positive constants that can be computed from the parameters of the potential (see, for example, L. Pauling and E. B. Wilson, Introduction to Quantum Mechanics, §35d).

Equation 7.90 for the vibrational-rotational spectrum predicts correctly the decrease of the spacing between successive vibrational levels, and the increase in the moment of inertia of the molecule (reciprocal of B_0 and the square brackets) with increasing vibrational and rotational quantum number. This can be understood from the asymmetry of the potential about its minimum (increase with v) and the stretching of the molecule with ℓ.

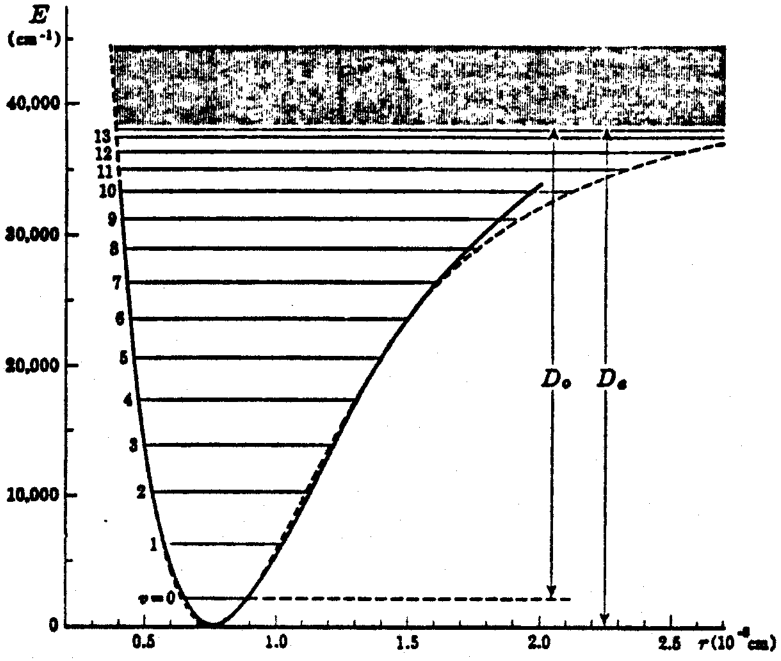


Fig. 50. Potential Curve of the H_2 Ground State with Vibrational Levels and Continuous Term Spectrum. The full curve is drawn according to Rydberg's data (610). The broken curve is a Morse curve. The continuous term spectrum, above $v = 14$, is indicated by vertical hatching. The vibrational levels are drawn up to the potential curve, that is, their end points correspond to the classical turning points of the vibration. It must be remembered that in quantum theory these sharp turning points are replaced by broad maxima of the probability amplitude ψ.

(Figure taken from G. Herzberg, *Spectra of Diatomic Molecules.*)

By fitting spectroscopic data of the rotational-vibrational spectrum all the parameters of the Morse potential can be deduced (see problems 6.15, 7.14 7.15). A short table of r_0, the minimum of $V(r)$, of $\hbar\omega_0$, the vibrational quantum $[V''(r_0)/m_r]$ and B_0 the rotational quantum is given below.

A more extensive listing can be found in G. Herzberg, *Spectra of Diatomic Molecules*, table 39. To use that table note that the present discussion concerns itself with the electronic ground state of the molecule, i.e. when the two atoms comprising the molecule are in their respective (electronic) ground state (this will be recognized in the table by the state prefix X). Note also that the units in table 39 (and traditionally in spectroscopy) are given in terms of reciprocal wavelengths (in units of reciprocal centimeters) and that

$$E_{\text{photon}} = \frac{1.24}{(\lambda/\mu m)} \text{ eV} \tag{7.91}$$

or

$$E_{\text{photon}} = 1.24 \times 10^{-4} \left(\frac{\text{cm}}{\lambda}\right) \text{ eV} . \tag{7.92}$$

Molecule	r_0/A°	$\hbar\omega_0/\text{eV}$	B_0/eV
H_2	0.74	5.45×10^{-1}	7.56×10^{-3}
HD	0.74	4.73×10^{-1}	5.69×10^{-3}
D_2	0.74	3.87×10^{-1}	3.79×10^{-3}
Li_2	2.67	4.35×10^{-2}	8.39×10^{-5}
N_2	1.09	1.03×10^{-2}	2.48×10^{-4}
0_2	1.21	1.96×10^{-1}	1.78×10^{-4}
LiH	1.60	1.74×10^{-1}	9.27×10^{-4}
$HC\ell^{35}$	1.27	3.71×10^{-1}	1.32×10^{-3}
$NaC\ell^{35}$	2.51	4.71×10^{-2}	2.36×10^{-5}
$KC\ell^{35}$	2.79	3.47×10^{-2}	1.43×10^{-5}
KBr^{79}	2.94	2.86×10^{-2}	9.1×10^{-6}
HBr^{79}	1.41	3.29×10^{-1}	1.06×10^{-3}

References

7.4 G. Herzberg, Molecular Spectra and Molecular Structure. I. Spectra of Diatomic Molecules (Van Nostrand, 1950, 2nd Edition).

7.5 R. B. Leighton, Principles of Modern Physics (McGraw-Hill, 1959), chapter 9.

7.6 L. Pauling and E. B. Wilson, Introduction to Quantum Mechanics (McGraw-Hill, 1935), chapter 10.

7.7 L. Landau and E. M. Lifshitz, Quantum Mechanics (Pergamon, 1977, 3rd edition), §82, §89.

7.4.1. Diatomic molecules in thermal equilibrium.

Any state of a diatomic molecule (not electronically excited) can be described as a superposition of the eigenstates of the Hamiltonian, i.e.

$$|\psi\rangle = \sum_{\underline{P}_{CM}, v, \ell, m} c_{\underline{P}_{CM}, v, \ell, m} \, |\underline{P}_{CM}, v, \ell, m\rangle \; , \tag{7.93}$$

where $\underline{P}_{CM}$ is the (linear) momentum of the center of mass, v is the vibrational quantum number, ℓ is the orbital angular momentum quantum number and m is the eigenvalue of $L_Z/\hbar$. If the diatomic molecule is in thermal equilibrium with an environment at a temperature T, we have that the modulus squared of the probability amplitude is given by the Boltzmann factor, i.e.

$$|c_{\underline{P}_{CM}, v, \ell, m}|^2 = \frac{1}{Z_T} e^{-(E_{CM}+E)/k_B T} \; , \tag{7.94}$$

where E is the energy of the relative motion about the center of mass, and where, Z_T is called the (total) partition function and given by

$$Z_T \equiv \sum_{\underline{P}_{CM}, v, \ell, m} e^{-(E_{CM}+E)/k_B T} \; , \tag{7.95}$$

since, of necessity,

$$\sum_{\underline{P}_{CM}, v, \ell, m} |c_{\underline{P}_{CM}, v, \ell, m}|^2 = 1 \; . \tag{7.96}$$

Using these results we could compute the total energy of the diatomic molecule in thermal equilibrium at a temperature T, i.e.,

$$\langle E_T \rangle = \frac{1}{Z_T} \sum_{\underline{P}_{CM}, v, \ell, m} (E_{CM} + E) \, e^{-(E_{CM}+E)/k_B T} \tag{7.97}$$

or

$$\langle E_T \rangle = \langle E_{CM} \rangle + \langle E \rangle \tag{7.98a}$$

where

$$\langle E_{CM} \rangle = \frac{1}{Z_{CM}} \sum_{\underline{P}_{CM}} E_{CM} \, e^{-E_{CM}/k_B T} \tag{7.98b}$$

and

$$\langle E \rangle = \frac{1}{Z} \sum_{v, \ell, m} E \, e^{-E/k_B T} \; , \tag{7.98c}$$

where

$$Z_T = Z_{CM} \cdot Z \tag{7.99a}$$

and

$$Z_{CM} = \sum_{\underline{P}_{CM}} e^{-E_{CM}/k_B T} \tag{7.99b}$$

$$Z = \sum_{v,\ell,m} e^{-E/k_BT} . \tag{7.99c}$$

From the results of a single particle in a box in thermal equilibrium at a temperature T, we have (see problem 4.13)

$$\langle E_{CM} \rangle = \frac{3}{2} k_BT . \tag{7.100}$$

To compute $\langle E \rangle$ we note that the energy of the relative motion about the center of mass is not a function of m and therefore we can perform the summation over m to obtain

$$\langle E \rangle = \frac{1}{Z} \sum_{v,\ell} (2\ell+1)\, E_{v,\ell}\, e^{-E_{v,\ell}/k_BT} , \tag{7.101}$$

where

$$Z = \sum_{v,\ell} (2\ell+1)\, e^{-E_{v,\ell}/k_BT} . \tag{7.102}$$

Neglecting, higher order corrections, we have from equation 7.90,

$$\begin{aligned} Z &\simeq \sum_{v,\ell} (2\ell+1)\, e^{-[\hbar\omega_0 (v+\frac{1}{2}) + B_0 \ell(\ell+1)]/k_BT} \\ &= \left[\sum_{v} e^{-(v+\frac{1}{2})\theta_v/T}\right] \left[\sum_{\ell} (2\ell+1)\, e^{-\ell(\ell+1)\theta_r/T}\right] \\ &= Z_{vib} \cdot Z_{rot} \end{aligned} \tag{7.103}$$

and

$$\begin{aligned} \langle E \rangle &\simeq \frac{\hbar\omega_0}{Z_{vib}} \sum_{v} (v+\tfrac{1}{2})\, e^{-(v+\frac{1}{2})\theta_v/T} \\ &\quad + \frac{B_0}{Z_{rot}} \sum_{\ell} (2\ell+1)\, \ell(\ell+1)\, e^{-\ell(\ell+1)\theta_r/T} \\ &\simeq \langle E_{vib} \rangle + \langle E_{rot} \rangle , \end{aligned} \tag{7.104}$$

where θ_v is the <u>characteristic vibrational temperature</u>,

$$\theta_v \equiv \hbar\omega_0 / k_B \tag{7.105a}$$

and θ_r is the <u>characteristic rotational temperature</u>,

$$\theta_r \equiv B_0 / k_B \tag{7.105b}$$

From the data on page 7.22 we see that θ_v/T is usually small at room temperature ($k_B \cdot 300^\circ K \simeq \frac{1}{40}$ eV) or lower. Therefore, for $T \leqslant 300^\circ K$

$$\langle E \rangle \simeq \langle E \rangle_{\text{rot}} = \frac{B_0}{Z_{\text{rot}}} \sum_{\ell} (2\ell+1)\ell(\ell+1)\, e^{-\ell(\ell+1)\theta_r/T} \quad . \tag{7.106}$$

At low temperatures, i.e., if $(\theta_r/T) >> 1$, we need only consider the $\ell = 1$ term, i.e., for heteronuclear molecules

$$\langle E \rangle_{\text{rot}} \sim 6\, B_0\, e^{-2 B_0/k_B T} \quad . \tag{7.107}$$

At high temperatures, we can approximate the sums by integrals to obtain

$$\langle E_{\text{rot}} \rangle \sim B_0 \frac{\int_0^{\infty} (2\ell+1)\,\ell(\ell+1)\, e^{-\ell(\ell+1)\theta_r/T}\, d\ell}{\int_0^{\infty} (2\ell+1)\, e^{-\ell(\ell+1)\theta_r/T}\, d\ell}$$

$$= B_0 \frac{\int_0^{\infty} x\, e^{-x\theta_r/T}\, dx}{\int_0^{\infty} e^{-x\theta_r/T}\, dx} ,$$

or

$$\langle E_{\text{rot}} \rangle \simeq k_B T \quad . \tag{7.108}$$

Most molecules are in this regime at room temperature and consequently

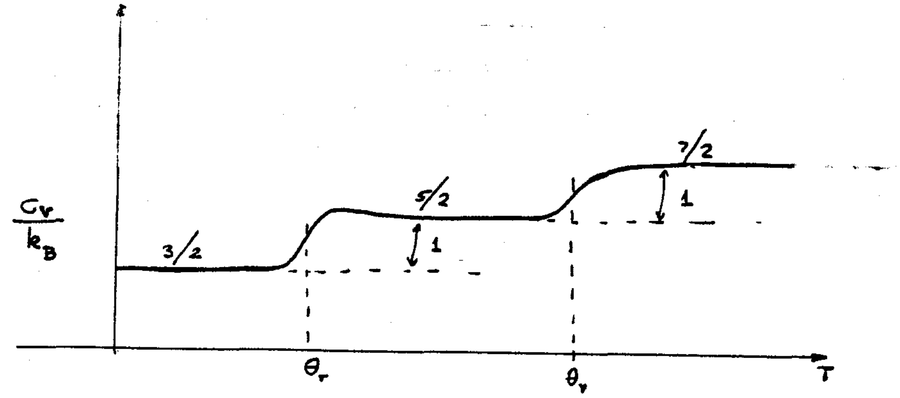
$$C_v \simeq \frac{\partial}{\partial T} \langle E_{\text{CM}} \rangle + \frac{\partial}{\partial T} \langle E_{\text{rot}} \rangle$$

$$= \frac{3}{2} k_B + k_B ,$$

or,

$$C_v = \frac{5}{2} k_B , \tag{7.109}$$

as is well known, derived here from first principles. For intermediate temperatures, $\langle E_{\text{rot}} \rangle$ can be computed numerically as a function of θ_r/T.

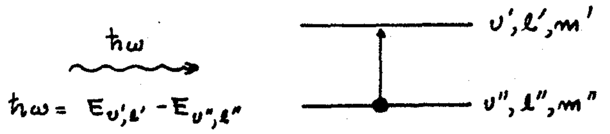


General form for the heat capacity of diatomic molecules as a function of temperature.

At higher temperatures, of course, the vibrational modes can also be excited and will contribute to the heat capacity. This can be computed from $\langle E_{vib} \rangle$, which is the energy at thermal equilibrium of a one-dimensional harmonic oscillator at a frequency $\omega = \omega_0$ (see equation 5.35), by differentiation with respect to temperature.

7.4.2 Radiative transitions of diatomic molecules in thermal equilibrium.

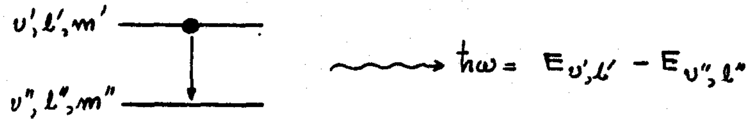
Using the intensity distribution of the absorption or emission spectrum of a diatomic molecule in thermal equilibrium at a temperature T, it is possible to measure the temperature of the molecule directly. In particular, the intensity of an absorption or emission line will be proportional to the sum of the transition rates for all possible initial and final states at the same frequency, times the probability that the initial state is occupied. For absorption



we then have (at line center)

$$\sigma_{abs} \propto \sum_{m'',m'} |\langle v',\ell',m'|\hat{V}_T|v'',\ell'',m''\rangle|^2 \, e^{-E_{v'',\ell''}/k_BT} , \qquad (7.110)$$

while for emission



$$\sigma_{em} \propto \sum_{m'',m'} |\langle v'',\ell'',m''|\hat{V}_T|v',\ell',m'\rangle|^2 \, e^{-E_{v',\ell'}/k_BT} , \qquad (7.111)$$

The quantities σ_{abs} and σ_{em} are the absorption and emission cross-sections, respectively. These are proportional to the sum over all possible initial and values of L_z of the molecule, of the matrix element for the transition (we will derive this later) times the probability of finding the molecule in the initial state, given by the corresponding Boltzmann factor.

Strictly speaking these results are for induced emission and absorption. The square of the matrix elements of the operator $\hat{V}_T$ that induces the transition, is in fact the Einstein coefficient $B_{\underline{n}'',\underline{n}'}$ (see for example R. P. Feynman, R. B. Leighton and M. Sands, The Feynman Lectures on Physics, v. I, ch. 42-5). Spontaneous emission has been ignored in equation 7.111. We will show later on that

$$\sum_{m'',m'} |\langle v',\ell',m'|\hat{V}_T|v'',\ell'',m''\rangle|^2 \propto (\ell'+\ell''+1) \qquad (7.112)$$

and therefore

$$\sigma_{abs}[(v'',\ell'') \rightarrow (v''+1,\ell')] \propto (\ell'+\ell''+1)\, e^{-\ell''(\ell''+1)\,B_0/k_B T} \qquad (7.113a)$$

and

$$\sigma_{em}[(v',\ell') \rightarrow (v'-1,\ell'')] \propto (\ell'+\ell''+1)\, e^{-\ell'(\ell'+1)\,B_0/k_B T} \qquad (7.113b)$$

Note that for an R-transition ($\Delta\ell=\ell'-\ell''=1$)

$$(\sigma_{abs})_R \propto (2\ell''+2)\, e^{-\ell''(\ell''+1)\theta_r/T}, \qquad (7.114a)$$

whereas for a P-transition ($\Delta\ell=\ell'-\ell''=-1$)

$$(\sigma_{abs})_P \propto 2\ell''\, e^{-\ell''(\ell''+1)\theta_r/T}. \qquad (7.114b)$$

For absorption, the R-transitions are always more intense than the P-transitions.

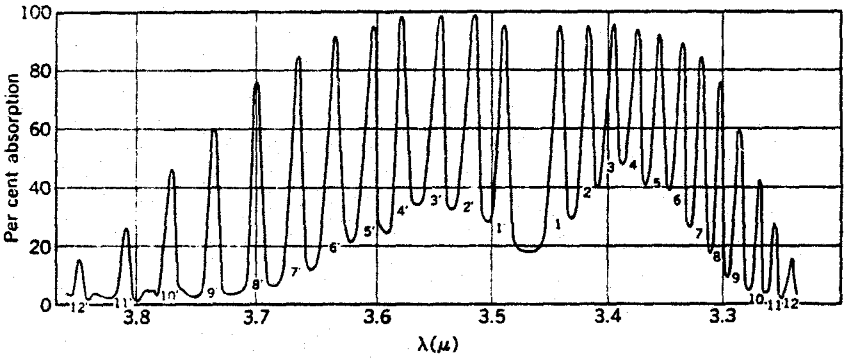


FIG. 9-13. The absorption spectrum of HCl vapor in the near infrared. [After Imes, *Astrophys. J.*, **50**, 251 (1919).]

[Figure taken from R. Leighton, Principles of Modern Physics, (v=0) → (v=1) absorption spectrum.]

For emission, however,

$$\sigma_{em} \propto \begin{cases} 2\ell'\, e^{-\ell'(\ell'+1)\theta_r/T} & \text{(R-transition)} \\ (2\ell'+2)\, e^{-\ell'(\ell'+1)\theta_r/T} & \text{(P-transition)}. \end{cases} \qquad (7.115)$$

This is why a diatomic molecule gas laser operating on a vibrational-rotational line will first lase on a P-transition near or at threshold, and, at pumping rates above threshold, the P-lines are stronger.

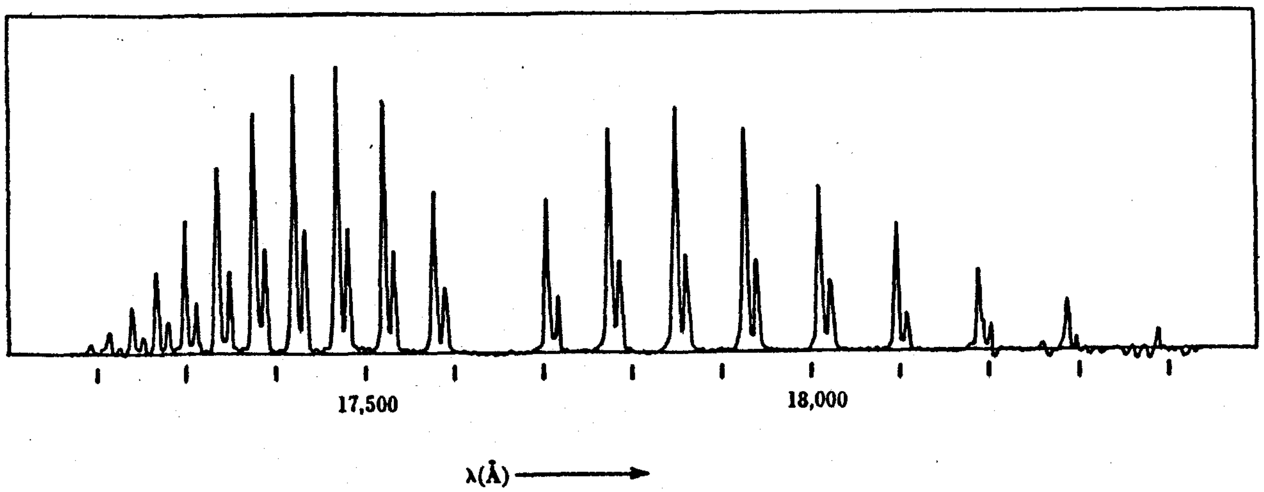


FIGURE 5-4 *The first overtone band of* HCl *taken on a Cary model 14 spectrometer. The strong set of lines is from* HCl^{35}; *the weak set is from* HCl^{37}. *The unequal spacings between lines show the effects of vibration-rotation coupling. Each division on the wavelength scale equals* 100 Å.

Hanna, Melvin W. (1969) Quantum Mechanics In Chemistry, Second Edition, W.A. Benjamin, Inc., Publishers.

(v=0) → (v=2) Transition Spectrum.

Problems

7.1 Derive equation 7.21 for the function $w(\rho)$ of equation 7.19.

7.2 Explain why only the odd solutions of the rectangular potential well are acceptable $\ell = 0$ solutions of the spherical potential well of equations 7.32.

7.3 (i) Sketch the effective potential, as given by equation 7.11, for $\ell = 0, 1, 2$, corresponding to the spherical well potential of equation 7.36.

(ii) Estimate the minimum depth of a spherical well possessing one bound state. See footnote # on page 4.20. Express your answer in dimensionless form.

7.4 Derive equation 7.21.

7.5 Compute the bound state energies of an electron in an infinitely deep spherical potential well in the presence of a weak, uniform magnetic field. Ignore the effects of electron spin.

7.6 Derive the recursion relation, given by equation 7.53, for the coefficients of the power series expansion of equation 7.52 for the function $w(\rho)$.

7.7 Compute the energies of the bound states of a hydrogen atom in a weak uniform magnetic field, ignoring the spin of the electron.

7.8 The radius of the proton is of the order of

$$r_p \sim 1.2 \times 10^{-13} \text{ cm} \quad .$$

Compute the fraction of the time that the hydrogen atom electron in the ground state will be found inside the proton.

7.9 The solutions for a Coulomb potential were obtained assuming a point particle of charge Ze. A nucleus, however, is found to have a finite radius (approximately given by $r_n \simeq r_0 A^{1/3} \cdot 10^{-13}$ cm, where $1.2 \leqslant r_0 \leqslant 1.5$ and A is the atomic mass number). Compute the potential that an electron would move in, assuming that the nucleus is spherical and

(i) all the charge resides on the surface,

(ii) the charge is uniformly distributed throughout the nuclear volume.

7.10 Estimate the correction to the ground state binding energy of the hydrogen atom for the two cases examined in problem 7.9.

7.11 Prove equations 7.83.

7.12 Show that the absorption spectrum corresponding to the vibrational transition $v \rightarrow v+1$, can be represented by a parabola

$$\Delta E_{v,k} = c_{v,v+1} + D_{v,v+1}k + e_{v,v+1}k^2$$

where $k = 1,2,3, \ldots$ for an R-branch transition and $k = -1,-2,-3, \ldots$ for P-branch transition. Compute $c_{v,v+1}$, $d_{v,v+1}$ and $e_{v,v+1}$ in terms of the Morse energy level formula (equation 7.90).

7.13 Show that the coefficient $c_{v,v+1}$ in the spectrum parabola (see previous problem) can be represented by another parabola

$$c_{v,v+1} = av - bv^2$$

where $v = 1,2,3, \ldots$. Compute a and b in terms of the Morse energy level formula.

7.14 The following absorption lines for HCℓ^{35} at 3.46 μ are found to correspond to the (v=0) → (v=1) transitions.

λ^{-1} (cm^{-1})

3085.62
72.76
59.07
44.88
29.96
14.29
2997.78
80.90
63.24
44.89
25.78
06.25

2865.09
43.56
21.49
2798.78
75.79
52.03
27.75
03.06
2677.73
51.97
25.74
2599.00

Compute the spectrum parabola coefficients (see problem 7.12).

Hint: Generate a finite difference table, or, if you have access to a calculator or computer do a least squares parabolic fit. Note that k=0 is missing.

Problems (continued)

7.15 The coefficients $c_{o,v}$ are found experimentally for $HC\ell^{35}$ to be equal to

v	$c_{o,v}(cm^{-1})$
1	2885.9
2	5668.1
3	8347.0
4	10923.1
5	13396.6

Using the results of the last problem and the data above, compute the relevant parameters of the corresponding Morse potential.

7.16 In table 39 of G. Hertzberg, Spectra of Diatomic Molecules, the following spectroscopic data are given for I_2^{127}.

$$\hbar\omega_0 = 214.57 \text{ cm}^{-1}$$

$$B_0 = 0.03735 \text{ cm}^{-1}$$

Compute the heat capacity of I_2 at room temperature.

Hint: ... Careful!

7.17 Using the data in the figure on page 7.28, estimate the temperature in Dr. Ime's laboratory.

7.18 Using the plane wave expansion of equation 7.29, show that

$$e^{i\underline{k}\cdot\underline{x}} = 4\pi \sum_{\ell=0}^{\infty} i^{\ell}\, j_{\ell}(kr) \sum_{m=-\ell}^{\ell} Y^{*}_{\ell,m}(\theta_{\underline{k}},\varphi_{\underline{k}})\, Y_{\ell,m}(\theta,\varphi)$$

in general, where

$$\hat{e}_{\underline{k}} = (\theta_{\underline{k}},\varphi_{\underline{k}}) \quad , \qquad \hat{e}_{x} = (\theta,\varphi) \quad .$$

update

8. METHODS OF APPROXIMATION

As with every field of physics, so in Quantum Mechanics, the number of problems which can be solved exactly is very small indeed. In fact, it is probably fair to say that there is no problem in nature which can be solved exactly. The art and the science of the physicist concerns itself in selecting, on the one hand, a model which in some way offers a fair description of the phenomenon he wishes to investigate and on the other hand finding a way of solving this model. In this sense, the whole of Quantum Mechanics, as is discussed in these lectures, is also in turn a model in which we make certain approximations so as to be able to solve the particular problems of interest. By way of example, in the preceding section we discussed the Schrödinger equation solution corresponding to the problem of two point charges moving about each other subject to a mutual electrostatic attraction. That is in many ways a far cry from the solution of, say, the hydrogen atom, a problem considerably more complicated than the one discussed. In particular, the following effects, of varying importance, were ignored in that discussion (this is by no means an exhaustive list!):

(i) The electron, as we shall see, possesses intrinsic angular momentum (spin), so that in addition to the Coulomb attractive force with the proton, we also have an interaction between the electron magnetic moment and the magnetic field the electron generates as it is orbiting around the nucleus. This is actually a relativistic effect that is called <u>spin-orbit</u> coupling.

(ii) The proton also possesses intrinsic angular momentum and has an associated magnetic moment which interacts with the magnetic field generated by the electron's orbital motion, on the one hand, as well as the electron's magnetic moment on the other.

(iii) The proton is not a point charge, but possesses a finite radius, so that the resulting Coulomb field is not proportional to $1/r$ all the way to $r = 0$. See problems 7.8, 7.9 and 7.10.

(iv) The electron in an excited state can interact with the electromagnetic field to emit a photon. That means that <u>the excited states are not in fact stationary states</u>, which we have assumed to be when we identified them with the eigenstates of the Coulomb potential hamiltonian.

(v) We have ignored relativistic effects.

Nevertheless, it is still true that the effect of the electrostatic Coulomb forces between the proton and the electron is the most important one and it is possible to use the preceding discussion as a starting point and compute these other effects as corrections to that solution.# There exists a formalism for computing these corrections, which we will now discuss, called perturbation theory. To be precise, we will assume that the Hamiltonian of interest can be expressed as a sum

$$H = H_0 + H' , \tag{8.1}$$

where H_0 leads to a Schrödinger equation

$$H_0 |n, m\rangle_0 = E_n^{(o)} |n, m\rangle_0 , \tag{8.2}$$

where the energy eigenvalues $E_n^{(o)}$ and the associated eigenstates $|n, m\rangle_0$ are known. The second index (m) in the eigenstates is included here to allow for labeling of orthogonal eigenstates of energy eigenvalues which may be degenerate. The perturbation Hamiltonian H' is assumed small in some sense, which we will quantify later.

We can imagine that the perturbation H' is proportional to a dimensionless perturbation parameter $\lambda_1 \ll 1$, such that

$$H' \equiv \lambda_1 V . \tag{8.3}$$

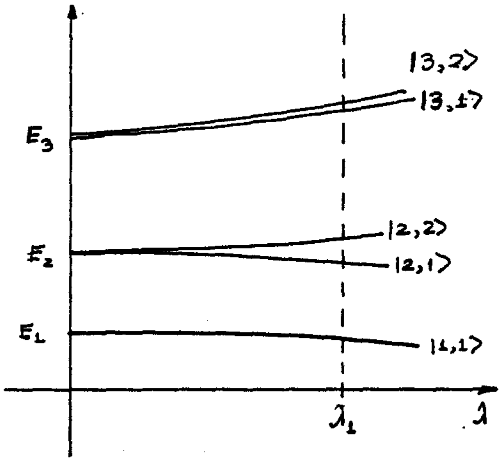
The total Hamiltonian can then be thought as a function of this perturbation parameter, i.e.,

$$H_1(\lambda) = H_0 + \lambda V , \tag{8.4}$$

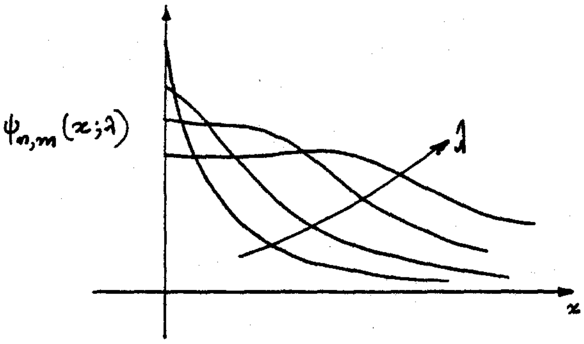
which we can imagine as an independent variable which gauges the strength of the perturbation. For each value of λ we have a new Hamiltonian with its eigenvalues and associated eigenfunctions, while at $\lambda = \lambda_1$ we obtain the original Hamiltonian of interest, i.e.

$$H = H_1(\lambda_1) .$$

[#] Better yet, we can start with an equation which is relativistically correct, i.e., the Dirac equation, so that we have to make fewer corrections.



Eigenvalue variation with λ.



Eigenfunction variation with λ.

Example 8.1 Charged particle in a harmonic potential exposed to a uniform electric field.

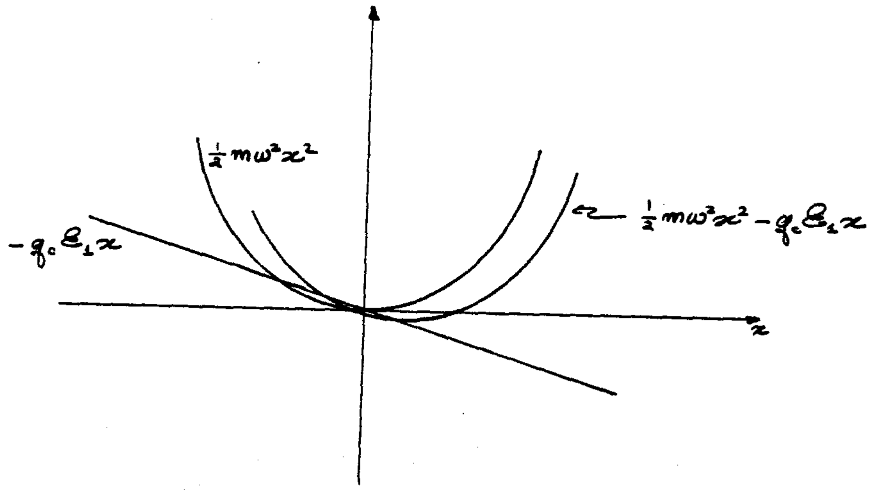
The total Hamiltonian is given by

$$H = \frac{1}{2m} p^2 + \frac{1}{2} m\omega^2 x^2 - q_c \mathcal{E}_1 x ,$$

or

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} m\omega^2 x^2 - q_c \mathcal{E}_1 x ,$$

where q_c is the particle charge and $\mathcal{E}_1$ is the electric field.



Assuming that the electric field is not too strong, we can consider the kinetic energy plus the harmonic potential as comprising H_0, i.e.

$$H_0 = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 ,$$

and the electrostatic energy as the perturbation, i.e.

$$H' = -q_c\mathcal{E}_1 x .$$

In the discussion of the Harmonic oscillator (see section 5.1), the quantity

$$a = \left(\frac{\hbar}{m\omega}\right)^{1/2} ,$$

emerged as a characteristic length scale (see equation 5.8). This we can use to construct the perturbation parameter λ_1, i.e.

$$H' = -q_c\mathcal{E}_1 x = -\left(\frac{q_c\mathcal{E}_1 a}{m\omega^2 a^2}\right)\cdot m\omega^2(ax)$$

or

$$H' = \lambda_1 V$$

where

$$\lambda_1 = \frac{q_c\mathcal{E}_1}{m\omega^2 a} = \frac{q_c\mathcal{E}_1}{(\hbar m\omega^3)^{1/2}}$$

and

$$V = -m\omega^2(ax) = -(\hbar m\omega^3)^{1/2}x$$

This problem can, of course, be solved exactly by completing the square.

In the subsequent discussions we shall consider the solution to the perturbation problem for various types of H_0 and H'.

8.1 Time-independent perturbation of non-degenerate states.

In this case the perturbation Hamiltonian is not a function of time,

$$H' \neq H'(t)$$

while the eigenfunctions of the unperturbed Hamiltonian are non-degenerate, i.e.

$$H_0 |n\rangle_0 = E_n^{(o)} |n\rangle_0 \; . \tag{8.5}$$

Consider now the eigenvalues and eigenstates of the parametrized Hamiltonian $H_1(\lambda)$, i.e.

$$H_1(\lambda)|k;\lambda\rangle = E_k(\lambda)|k;\lambda\rangle \; , \tag{8.6}$$

and note that since $H_1(0) = H_0$,

$$|k;0\rangle = |k\rangle_0 \; . \tag{8.7}$$

Note also that if E_k and $|k\rangle$ are the eigenstates of H, i.e.

$$H|k\rangle = E_k|k\rangle \tag{8.8}$$

we must have, since $H = H_1(\lambda_1)$

$$E_k = E_k(\lambda_1) \quad \text{and} \quad |k\rangle = |k;\lambda_1\rangle \; . \tag{8.9}$$

We can now expand the eigenfunctions $|k;\lambda\rangle$ of the parametrized Hamiltonian $H_1(\lambda)$ in the complete set $\{|n\rangle_0\}$ of basis states of H_0, i.e.

$$|k;\lambda\rangle = \sum_n |n\rangle_0 C_{nk}(\lambda) \; , \tag{8.10a}$$

where

$$C_{nk}(\lambda) = {}_0\langle n|k;\lambda\rangle \; . \tag{8.10b}$$

Substituting this expansion in the eigenvalue equation 8.6 we have

$$(H_0 + \lambda V)[\sum_n |n\rangle_0 \; C_{nk}(\lambda)] = E_k(\lambda)[\sum_n |n\rangle_0 \; C_{nk}(\lambda)] \; ,$$

or

$$\sum_n |n\rangle_0 [E_k(\lambda) - E_n^{(o)}] \; C_{nk}(\lambda) = \lambda \sum_n V|n\rangle_0 \; C_{nk}(\lambda) \; .$$

If we now multiply both sides of this equation with ${}_0\langle m|$, we obtain a set of equations

$$[E_k(\lambda) - E_m^{(o)}] \; C_{mk}(\lambda) = \lambda \sum_n V_{mn} C_{nk}(\lambda) \tag{8.11a}$$

where V_{mn} is the matrix element.

$$V_{mn} \equiv {}_0\langle m|V|n\rangle_0 \; . \tag{8.11b}$$

Equations 8.11 are the eigenvalue equations that would arise in the problem of the diagonalization of the original Hamiltonian

$$[H_1(\lambda)]_{mn} = E_m^{(o)}\delta_{mn} + \lambda V_{mn} ,$$

expressed in the basis vectors of H_o, a problem, of course, no easier to solve. Here, however, we can make use of the fact that λ is a small quantity and expand the eigenvalues $E_k(\lambda)$ and the coefficients $C_{nk}(\lambda)$ in a power series of λ, i.e.

$$E_k(\lambda) = E_k^{(o)} + \lambda E_k^{(1)} + \lambda^2 E_k^{(2)} + \ldots. \tag{8.12a}$$

and

$$C_{nk}(\lambda) = C_{nk}^{(o)} + \lambda C_{nk}^{(1)} + \lambda^2 C_{nk}^{(2)} + \ldots. \tag{8.12b}$$

Note that we have, in fact, also expanded the state $|k;\lambda\rangle$, which depends parametrically on λ, in a power series in λ through the expansion of $C_{nk}(\lambda)$, i.e.

$$\begin{aligned} |k;\lambda\rangle &= \sum_n |n\rangle_o \, C_{nk}(\lambda) \\ &= \sum_n |n\rangle_o \sum_q \lambda^q C_{nk}^{(q)} \\ &= \sum_q \lambda^q \left[\sum_n |n\rangle_o \, C_{nk}^{(q)} \right] \end{aligned}$$

or

$$|k;\lambda\rangle = \sum_q \lambda^q |k\rangle_q , \tag{8.13}$$

where

$$|k\rangle_q \equiv \sum_n |n\rangle_o \, C_{nk}^{(q)} . \tag{8.14}$$

From equation 8.13 and 8.14 we see that

$$|k;0\rangle = |k\rangle_o = \sum_n |n\rangle_o \, C_{nk}^{(o)} ,$$

and therefore

$$C_{nk}^{(o)} = \delta_{nk} . \tag{8.15}$$

We now substitute the power series expansion of the eigenvalues $E_k(\lambda)$ and the coefficients $C_{nk}(\lambda)$ in equation 8.11. This yields a matrix equation in k and m,

$$[E_k^{(o)} - E_m^{(o)} + \lambda E_k^{(1)} + \lambda^2 E_k^{(2)} + \ldots][\delta_{mk} + \lambda C_{mk}^{(1)} + \lambda^2 C_{mk}^{(2)} + \ldots] =$$

$$= \lambda \sum_n V_{mn}[\delta_{nk} + \lambda C_{nk}^{(1)} + \lambda^2 C_{nk}^{(2)} + \ldots] .$$

This equality must hold as the value of the perturbation parameter λ changes from $\lambda = 0$ to $\lambda = \lambda_1$ to give us the Hamiltonian of interest. Consequently, we must have, collecting terms of equal powers of λ, for the diagonal elements $k = m$

$$\lambda: \qquad E_k^{(1)} = V_{kk} \tag{8.16a}$$

$$\lambda^2: \qquad E_k^{(2)} + E_k^{(1)} C_{kk}^{(1)} = \sum_n V_{kn} C_{nk}^{(1)} \tag{8.16b}$$

$$\vdots$$

and for the off-diagonal elements $k \neq m$,

$$\lambda: \qquad (E_k^{(0)} - E_m^{(0)}) C_{mk}^{(1)} = V_{mk} \tag{8.17a}$$

$$\lambda^2: \quad (E_k^{(0)} - E_m^{(0)}) C_{mk}^{(2)} + E_k^{(1)} C_{mk}^{(1)} = \sum_n V_{mn} C_{nk}^{(1)} \tag{8.17b}$$

$$\vdots$$

Equations 8.16 and 8.17 contain the desired results. In particular, to first order in λ, a particular eigenvalue $E_k(\lambda)$ is given by

$$E_k(\lambda) = E_k^{(0)} + \lambda E_k^{(1)} + 0(\lambda^2)$$

or, substituting from equation 8.16a,

$$E_k(\lambda) \sim E_k^{(0)} + \lambda V_{kk}$$

and, at $\lambda = \lambda_1$, $E_k = E_k(\lambda_1)$ and therefore

$$E_k \sim E_k^{(0)} + H'_{kk} \, , \tag{8.18a}$$

where $H'_{kk} = \lambda V_{kk}$, or

$$H'_{kk} \equiv {}_0\langle k | H' | k \rangle_0 \; . \tag{8.18b}$$

Similarly, to first order in λ, the expansion coefficients $C_{nk}(\lambda)$ are given by

$$C_{nk}(\lambda) = \delta_{nk} + \lambda \, C_{nk}^{(1)} + 0(\lambda^2) \; ,$$

where, substituting from equation 8.17a,

$$C_{nk}^{(1)} = C_{kk}^{(1)} \, \delta_{nk} - \frac{V_{nk}}{E_n^{(0)} - E_k^{(0)}} (1 - \delta_{nk}) \, , \tag{8.17}$$

where $C_{kk}^{(1)}$ is yet to be determined.

From equation 8.11, we have, to first order in λ,

$$|k;\lambda\rangle = |k\rangle_0 + \lambda |k\rangle_1 + 0(\lambda^2) \tag{8.18}$$

or, using equation 8.14,

$$|k;\lambda\rangle \sim |k\rangle_o + \lambda \sum_n |n\rangle_o C^{(1)}_{nk}$$

or, substituting for $C^{(1)}_{nk}$ from 8.19,

$$|k;\lambda\rangle \sim |k\rangle_o [1 + \lambda C^{(1)}_{kk}] - \lambda \sum_{n \neq k} |n\rangle_o \frac{V_{nk}}{E^{(o)}_n - E^{(o)}_k} \quad . \tag{8.20}$$

To compute $C^{(1)}_{kk}$, we now use the fact that the states $|k;\lambda\rangle$ must be normalized. Consequently, the scalar products $\langle k;\lambda|k;\lambda\rangle$, using the states $|k;\lambda\rangle$ correct to first order, must be equal to unity plus an error which may be second order or higher. This leads to the condition that

$$C^{(1)*}_{kk} + C^{(1)}_{kk} = 0 \ , \tag{8.21}$$

or, in other words, that $C^{(1)}_{kk}$ must be purely imaginary, say

$$C^{(1)}_{kk} = i\gamma_k \ ,$$

where γ_k is real. Note now that

$$1 + \lambda C^{(1)}_{kk} = 1 + i\lambda\gamma_k = e^{i\lambda\gamma_k} + O(\lambda^2) \ ,$$

so that a non-zero $C^{(1)}_{kk}$ results only in a phase-shift of the perturbed wave-functions $|k;\lambda\rangle$, correct to first order. Since, however, we may arbitrarily select the phase of each eigenstate, we may set

$$C^{(1)}_{kk} = 0 \ . \tag{8.22}$$

Consequently

$$|k;\lambda\rangle \sim |k\rangle_o - \lambda \sum_{n \neq k} |n\rangle_o \frac{V_{nk}}{E^{(o)}_n - E^{(o)}_k} \ , \tag{8.23}$$

or, since $|k\rangle = |k;\lambda_1\rangle$, we have to first order in the perturbation H',

$$\boxed{|k\rangle \sim |k\rangle_o - \sum_{n \neq k} |n\rangle_o \frac{H'_{nk}}{E^{(o)}_n - E^{(o)}_k}} \ , \tag{8.24a}$$

where

$$H'_{nk} \equiv {}_o\langle n|H'|k\rangle_o \ . \tag{8.24b}$$

Equation 8.24 allows us to quantify the size of the perturbation Hamiltonian, for the perturbation procedure in terms of a power series in λ to be valid. In particular, we require that the perturbation expansion is an <u>asymptotic expansion</u>, i.e. the magnitude of the $(n+1)^{st}$ term in the expansion should be of the same order as the error after n terms.# This will be the case if the correction to the wavefunctions is small, or if

$$|H'_{nk}| << |E_n^{(o)} - E_k^{(o)}| \, , \tag{8.25}$$

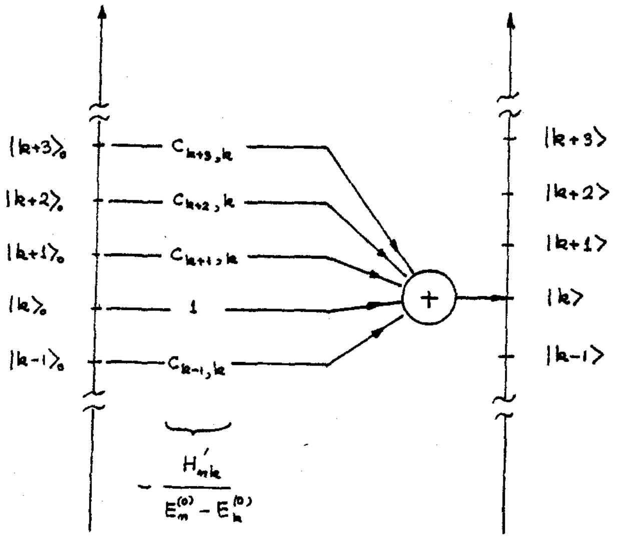
i.e. the matrix elements of the perturbation must be small compared to the energy difference of the corresponding unperturbed energy levels. Equation 8.25 also illustrates where the assumption of non-degenerate states comes in.

Let us pause to see what the results of the first order corrections tell us. For the energy (equation 8.18), we see that to first order the eigenvalues are computed as the expectation values (diagonal matrix elements) of the perturbed Hamiltonian, using the unperturbed wavefunctions, i.e.

$$E_k \sim E_k^{(o)} + \lambda_1 V_{kk} = {}_0\langle k|(H_0 + \lambda_1 V)|k\rangle_0 \, . \tag{8.26}$$

The reason for this is that the expectation values that could be computed using the corrected wavefunctions to first order would contain second order terms (cross terms between the perturbation $\lambda\hat{V}$ and the first order corrections $\lambda|k\rangle_1$ to the unperturbed wavefunctions).

The effect of the perturbation on the wavefunctions, however, is more complicated. We see from equations 8.23 or 8.24 that unless the perturbation is diagonal, i.e. if $V_{nk} \neq V_{kk}\delta_{nk}$ (in which case we are done, since the eigenstates $|k\rangle_0$ of H_0 would also be eigenstates of $H = H_0 + \lambda_1 V$), a perturbed eigenstate is a mixture of the original eigenstates,



#Note that such an expansion does not necessarily converge and that the best approximation may be realized after a finite number of terms (see for example E.T. Copson, Asymptotic Expansions, Section 7).

i. e.

$$|k\rangle = \sum_n |n\rangle_0 \, C_{nk} \tag{8.27}$$

where, to first order, the matrix of coefficients is given by (see equations 8. 15, 8. 19 and 8. 22),

$$C_{nk} \sim \delta_{nk} - (1 - \delta_{nk}) \, \frac{H_{nk}}{E_n^{(o)} - E_k^{(o)}} \; . \tag{8.28}$$

In other words, a state that originally was an eigenstate of the unperturbed Hamiltonian H_0, is no longer an eigenstate of the system under the action of the perturbation H', unless H' is diagonal in the original basis.

Quite frequently, the first order corrections to the eigenvalues are not adequate. This usually occurs if the matrix elements H'_{kk} happen to be zero identically. In that case we must go to the second order corrections. From equation 8. 14b we have for the second order correction (recall that $C_{kk}^{(1)} = 0$),

$$E_k^{(2)} = \sum_{n \neq k} V_{kn} C_{nk}^{(1)}$$

or, substituting for $C_{nk}^{(1)}$, we have

$$E_k^{(2)} = - \sum_{n \neq k} \frac{|V_{nk}|^2}{E_n^{(o)} - E_k^{(o)}} \quad , \tag{8.29}$$

where we have used that $V_{kn}^* = V_{nk}$. Consequently, <u>to second order in the perturbation</u> (equation 8. 12a),

$$\boxed{E_k = E_k^{(o)} + H'_{kk} - \sum_{n \neq k} \frac{|H'_{nk}|^2}{E_n^{(o)} - E_k^{(o)}}} \; . \tag{8.30}$$

The second order corrections to the wavefunctions are very seldom used. The algebra to obtain them is straightforward but tedious. The second order corrections to the coefficients are derived from equation 8 . 17b and, with the normalization condition, yield the eigenstates correct to second order,

$$\begin{aligned} |k\rangle = {} & |k\rangle_0 \left[1 - \frac{1}{2} \sum_{n \neq k} \frac{|H'_{nk}|^2}{(E_n^{(o)} - E_k^{(o)})^2} \right] \\ & - \sum_{n \neq k} \frac{|n\rangle_0}{E_n^{(o)} - E_k^{(o)}} \left[\left(1 + \frac{H'_{kk}}{E_n^{(o)} - E_k^{(o)}} \right) H'_{nk} - \sum_{m \neq k} \frac{H'_{nm} H'_{mk}}{E_m^{(o)} - E_k^{(o)}} \right] . \end{aligned} \tag{8.31}$$

In practice, one usually computes the eigenstates to first order (equation 8.24) and the eigenvalues to second order (equation 8.30). If the diagonal elements of H' are zero, these are the lowest order perturbation corrections to the eigenstates respectively.

These results, which were derived for discrete eigenvalues, can be generalized to continuous eigenvalues. By way of example, if part of the eigenvalue spectrum is discrete and the rest is continuous, the first order correction to an eigenstate in the discrete part of the spectrum would be given by,

$$|k\rangle_1 = - \sum_{n \neq k} |n\rangle_0 \frac{V_{nk}}{E_n^{(0)} - E_k^{(0)}} - \int_\nu |\nu\rangle_0 \frac{V_{\nu k}}{E^{(0)}(\nu) - E_k^{(0)}} \, d\nu \, . \tag{8.32}$$

Example 8.2 Polarizability of a harmonically bound charge.

Consider a particle of mass m and charge q_c bound by an isotropic harmonic potential, i.e.

$$H_0 = \frac{1}{2m} p^2 + \frac{1}{2} m\omega_0^2 r^2 \, , \tag{8.33}$$

where $r^2 = x^2 + y^2 + z^2$. In the presence of a weak electric field

$$\underline{\mathcal{E}} = \mathcal{E}_z \hat{e}_z \, ,$$

the Hamiltonian becomes

$$H = \frac{1}{2m} p^2 + \frac{1}{2} m\omega_0^2 r^2 - q_c \mathcal{E}_z z \, , \tag{8.34}$$

with a perturbation Hamiltonian

$$H' = - q_c \mathcal{E}_z z$$

(see example 8.1).

The unperturbed eigenstates are then given by, in the occupation number representation,

$$|n\rangle_0 = |n_x, n_y, n_z\rangle_0 \tag{8.35}$$

with energy levels

$$E = (n_x + n_y + n_z + \tfrac{3}{2}) \hbar\omega_0 \, . \tag{8.36}$$

The Hamiltonian is separable to the sum of three Hamiltonians corresponding to motion along each of the three coordinates, i.e.

$$H = H_x + H_y + H_z \tag{8.37a}$$

where

$$H_x = \frac{1}{2m} \hat{p}_x^{\,2} + \frac{1}{2} m\omega_0^2 x^2 \tag{8.37b}$$

$$H_y = \frac{1}{2m}\,\hat{p}_y^{\,2} + \frac{1}{2}\,m\omega_0^2 y^2 \tag{8.37c}$$

$$H_z = \frac{1}{2m}\,\hat{p}_z^{\,2} + \frac{1}{2}\,m\omega_0^2 z^2 - q_c\mathcal{E}_z z\ , \tag{8.37d}$$

so that the unperturbed eigenstates are separable into the product

$$|\underline{n}\rangle_0 = |n_x\rangle_0\ |n_y\rangle_0\ |n_z\rangle_0 \tag{8.38}$$

(see section 5.4). It can be seen that the perturbation does not affect the eigenstates for the motion along the x and y axes so that the perturbed eigenstates will be expressible as products of the form

$$|\underline{n}\rangle = |n_x\rangle_0\ |n_y\rangle_0\ |n_z\rangle\ . \tag{8.39}$$

This, of course, can be immediately deduced from the form of the Hamiltonian as given by equation 8.37. It is therefore sufficient to study the solutions to the Hamiltonian corresponding to the motion along the z-axis.

We then have, to first order in the perturbation, the correction to the energies,

$$E_{n_z} = E_{n_z}^{(o)} + {}_0\langle n_z|H'|n_z\rangle_0$$

where

$${}_0\langle n_z|H'|n_z\rangle_0 = -q_c\mathcal{E}_z\ {}_0\langle n_z|z|n_z\rangle_0\ . \tag{8.40}$$

Using the annihilation and creation operators (equation 5.27a and related discussion), we have

$$z = \left(\frac{\hbar}{2m\omega_0}\right)^{1/2} (a_z^\dagger + a_z) \tag{8.41}$$

and therefore

$${}_0\langle n_z|z|n_z\rangle_0 = \left(\frac{\hbar}{2m\omega_0}\right)^{1/2} {}_0\langle n_z|(a_z^\dagger + a_z)|n_z\rangle_0\ .$$

But

$$(a_z^\dagger + a_z)|n_z\rangle_0 = |n_z + 1\rangle_0\ (n_z+1)^{1/2} + |n_z - 1\rangle_0\ n_z^{\,1/2} \tag{8.42}$$

and therefore

$${}_0\langle n_z|H'|n_z\rangle_0 = -q_c\mathcal{E}_z\ {}_0\langle n_z|z|n_z\rangle_0 = 0\ ,$$

so that the correction to the energy levels to first order is zero.

For the eigenstates we then have

$$|n_z\rangle \sim |n_z\rangle_0 - \sum_{m_z \neq n_z} |m_z\rangle_0\ \frac{{}_0\langle m_z|H'|n_z\rangle_0}{E_{m_z}^{(o)} - E_{n_z}^{(o)}}$$

where

$$\begin{aligned} H'|n_z\rangle_0 &= -\ |n_z\rangle_0\ q_c\mathcal{E}_z\hat{z} \\ &= -\ (a_z^\dagger + a_z)|n_z\rangle_0\ q_c\mathcal{E}_z \left(\frac{\hbar}{2m\omega_0}\right)^{1/2} \\ &= -q_c\mathcal{E}_z\left(\frac{\hbar}{2m\omega_0}\right)^{1/2} \left[\,|n_z+1\rangle_0\ (n_z+1)^{1/2} + |n_z-1\rangle_0\ n_z^{\,1/2}\right] \end{aligned}$$

and therefore

$$_o\langle m_z|H'|n_z\rangle_o = -q_c\mathcal{E}_z\left(\frac{\hbar}{2m\omega_o}\right)^{1/2}\left[(n_z+1)^{1/2}\delta_{m_z,n_z+1} + n_z^{1/2}\delta_{m_z,n_z-1}\right] \quad (8.43)$$

so that, to first order in the eigenstates we have for the corrected ground state

$$|0_z\rangle \sim |0_z\rangle_o + |1_z\rangle_o \frac{q_c\mathcal{E}_z}{\hbar\omega_o}\left(\frac{\hbar}{2m\omega_o}\right)^{1/2} ,$$

or

$$|0_z\rangle \sim |0_z\rangle_o + \frac{1}{\sqrt{2}}\ |1_z\rangle_o \frac{q_c\mathcal{E}_z}{(m\hbar\omega_o^3)^{1/2}} \quad (8.44a)$$

while for $n_z \geq 1$

$$|n_z\rangle \sim |n_z\rangle_o + \frac{1}{\sqrt{2}}\frac{q_c\mathcal{E}_z}{(m\hbar\omega_o^3)^{1/2}}\left[|n_z+1\rangle_o(n_z+1)^{1/2} - |n_z-1\rangle_o\, n_z^{1/2}\right] \quad (8.44b)$$

Note that the coefficient in front of the brackets is in fact the perturbation parameter for the problem as derived in example 8.1.

The dipole moment of a charge distribution $\rho_c(\underline{x})$ is given by (see appendix F, section 4) the expectation value of the vector $\underline{x}$, i.e.

$$\underline{d} \equiv \int \underline{x}\, \rho_c(\underline{x}) d^3\underline{x} , \quad (8.45)$$

where, evidently, $\underline{d} = 0$ if $\rho_c(\underline{x})$ is spherically symmetric. Quantum mechanically, we identify the charge density with the square of the modulus of the wave-function (times the charge q_c), i.e.

$$\rho_c(\underline{x}) \equiv q_c|\psi(\underline{x})|^2 = q_c\psi^*(\underline{x})\psi(\underline{x}) \quad (8.46)$$

so that

$$\underline{d} = q_c\int \psi^*(\underline{x})\,\underline{x}\,\psi(\underline{x})\, d^3\underline{x} \quad (8.47)$$

or, equivalently

$$\underline{d} \equiv q_c\langle\psi|\underline{x}|\psi\rangle . \quad (8.48)$$

For a harmonic oscillator in the ground state, we have

$$|\psi\rangle = |0_x, 0_y, 0_z\rangle = |0_x\rangle\ |0_y\rangle\ |0_z\rangle$$

and therefore in the absence of a perturbing electric field,

$$\underline{d} \equiv q_c\ \langle 0_x, 0_y, 0_z|\underline{x}|0_x, 0_y, 0_z\rangle = 0 .$$

If we turn on the electric field, however, we find that, as a result of the perturbation, a dipole moment is induced along the z-axis, i.e.

$$\underline{d} = q_c\left[{}_o\langle 0_x|x|0_x\rangle_o\ \hat{e}_x + {}_o\langle 0_y|y|0_y\rangle_o\ \hat{e}_y + \langle 0_z|z|0_z\rangle\ \hat{e}_z\right]$$

$$= q_c\langle 0_z|z|0_z\rangle\hat{e}_z .$$

Using the eigenstates, corrected to first order (equation 8.44a), we then have

$$d_z = q_c \langle 0_z | z | 0_z \rangle$$

$$= q_c \left[{}_0\langle 0_z | + \frac{\lambda_1}{\sqrt{2}} {}_0\langle 1_z | \right] \hat{z} \left[|0_z\rangle_0 + \frac{\lambda_1}{\sqrt{2}} |1_z\rangle_0 \right] ,$$

where $\lambda_1 = q_c \mathcal{E}_z (\hbar m \omega_0^3)^{-1/2}$. Substituting for $\hat{z}$ in terms of the annihilation and creation operators, we have

$$\hat{z}\left[|0_z\rangle_0 + \frac{\lambda_1}{\sqrt{2}} |1_z\rangle_0 \right] = \left(\frac{\hbar}{2m\omega_0}\right)^{1/2} (a_z^\dagger + a_z)\left[|0_z\rangle_0 + \frac{\lambda_1}{\sqrt{2}} |1_z\rangle_0 \right]$$

$$= \left(\frac{\hbar}{2m\omega_0}\right)^{1/2} \left[|1_z\rangle_0 + \lambda_1 |2_z\rangle_0 + \frac{\lambda_1}{\sqrt{2}} |0_z\rangle_0 \right]$$

and therefore

$$d_z = q_c \left(\frac{\hbar}{2m\omega_0}\right)^{1/2} \left[{}_0\langle 0_z | + \frac{\lambda_1}{\sqrt{2}} {}_0\langle 1_z | \right] \left[|1_z\rangle_0 + \lambda_1 |2_z\rangle_0 + \frac{\lambda_1}{\sqrt{2}} |0_z\rangle_0 \right]$$

$$= q_c \left(\frac{\hbar}{2m\omega_0}\right)^{1/2} \sqrt{2}\, \lambda_1 ,$$

or, substituting for λ_1 ,

$$d_z = \frac{q_c^2}{m\omega_0^2} \mathcal{E}_z . \tag{8.49}$$

Classically, we define the *polarizability* α as the constant of proportionality between the induced dipole moment $\underline{d}$ and the applied electric field $\underline{\mathcal{E}}$, (see appendix F, section 4), i.e.

$$\underline{d} \equiv \epsilon_0 \alpha \underline{\mathcal{E}} , \tag{8.50}$$

where, in general, α is a tensor (3 x 3 matrix)#. For an *isotropically* bound charge in the ground state of a harmonic potential we therefore have that the polarizability is a *scalar* and given by

$$\alpha = \frac{q_c^2}{\epsilon_0 m \omega_0^2} . \tag{8.51}$$

Strictly speaking, we have computed the *static* polarizability. We would expect, however, this result to hold for a sinusoidally varying electric field.

$$\mathcal{E}_z(t) = \mathcal{E}_z(0) \cos \omega t ,$$

provided the frequency ω is much less than the harmonic oscillator frequency ω_0, i.e. if $\omega \ll \omega_0$.

#See Feynman, Leighton and Sands, *The Feynman Lectures on Physics*, volume II, chapter 31.

The evaluation of the sums that occur in the higher order corrections, e.g. equation 8.30, is very often a very difficult proposition. It is sometimes possible, however, to perform the calculation by the use of special techniques. By way of example, consider a sum of the type

$$\sum_{n\neq k} \frac{|G_{nk}|^2}{E_n^{(o)} - E_k^{(o)}} = \sum_{n\neq k} \frac{G_{kn}\, G_{nk}}{E_n^{(o)} - E_k^{(o)}}$$

that occurs in the computation of the second order correction to the perturbed energy of a non-degenerate state $|k\rangle$. Let us assume that it is possible to find an operator Q such that

$$G|k\rangle_o = (QH_o - H_oQ)|k\rangle_o \,. \qquad (8.52)$$

We then have

$$G_{nk} \equiv {}_o\langle n|G|k\rangle_o = {}_o\langle n|QH_o|k\rangle_o - {}_o\langle n|H_oQ|k\rangle_o$$
$$= (E_k^{(o)} - E_n^{(o)})\,{}_o\langle n|Q|k\rangle_o \,,$$

and therefore

$$-\sum_{n\neq k} \frac{G_{kn}\, G_{nk}}{E_n^{(o)} - E_k^{(o)}} = \sum_{n\neq k} {}_o\langle k|G|n\rangle_o \, {}_o\langle n|Q|k\rangle_o \,. \qquad (8.53)$$

We now note that since the states $\{|n\rangle_o\}$ form a complete set, any state $|\psi\rangle$ can be resolved in a superposition

$$|\psi\rangle = \sum_n |n\rangle_o \, {}_o\langle n|\psi\rangle \,.$$

Consequently, the state $Q|k\rangle_o$ can be written as

$$Q|k\rangle_o = \sum_n |n\rangle_o \, {}_o\langle n|Q|k\rangle_o \,, \qquad (8.54)$$

and therefore, substituting into equation 8.53, the infinite sum collapses to

$$-\sum_{n\neq k} \frac{|G_{nk}|^2}{E_n^{(o)} - E_k^{(o)}} = {}_o\langle k|GQ|k\rangle_o - {}_o\langle k|G|k\rangle_o \, {}_o\langle k|Q|k\rangle_o \,. \qquad (8.55)$$

It goes without saying that the utility of this technique rests on finding the operator Q.

Example 8.3 Polarizability of a hydrogen atom in the ground state.

The system is isotropic in this case so that the polarizability is a scalar. For an electric field along the z-axis we then have (see problem 8.7)

$$\alpha_o = \frac{2q_c^2}{\varepsilon_o} \sum_{n\neq 0} \frac{|z_{n0}|^2}{E_n^{(o)} - E_0^{(o)}} \,, \qquad (8.56)$$

where the subscript zero denotes the ground state.

It can be shown (see Appendix I), that the auxiliary operator Q for this problem (see equation 8.52 and related discussion)

$$Q = -\frac{m_r a_0}{\hbar^2}\left(a_0 + \frac{r}{2}\right) z \ , \qquad (8.57a)$$

where a_0 is the Bohr radius (equation 7.60), satisfies the equation.

$$z|0\rangle_0 = (QH_0 - H_0 Q)|0\rangle_0 \ , \qquad (8.57b)$$

where $|0\rangle_0$ is the unperturbed hydrogen atom ground state and H_0 is the unperturbed hydrogen atom Hamiltonian. Consequently, by virtue of the preceding discussion, we have from equation 8.55

$$\sum_{n\neq 0} \frac{|z_{n0}|^2}{E_n^{(0)} - E_0^{(0)}} = \frac{m_r a_0}{\hbar^2}\ {}_0\langle 0|\left(a_0 + \frac{r}{2}\right) z^2 |0\rangle_0 \qquad (8.58)$$

since ${}_0\langle 0|z|0\rangle_0 = 0$ by symmetry. Note also

$${}_0\langle 0|z^2 f(r)|0\rangle_0 = {}_0\langle 0|x^2 f(r)|0\rangle_0 = {}_0\langle 0|y^2 f(r)|0\rangle_0$$

and therefore

$${}_0\langle 0|z^2 f(r)|0\rangle_0 = \frac{1}{3}\ {}_0\langle 0|r^2 f(r)|0\rangle_0 \ .$$

Consequently, we have

$$\begin{aligned}\alpha_0 &= \frac{2q_c^2 m_r a_0}{3\varepsilon_0 \hbar^2}\ {}_0\langle 0|\left(a_0 + \frac{r}{2}\right) r^2 |0\rangle_0 \\ &= \frac{2}{3}\ \frac{q_c^2 m_r a_0{}^4}{\varepsilon_0 \hbar^2}\ {}_0\langle 0|\left[\left(\frac{r}{a_0}\right)^2 + \frac{1}{2}\left(\frac{r}{a_0}\right)^3\right]|0\rangle_0 \ ,\end{aligned}$$

or since

$$\begin{aligned}{}_0\langle 0|\left(\frac{r}{a_0}\right)^n|0\rangle_0 &= \frac{4}{a_0^3}\int_0^\infty \left(\frac{r}{a_0}\right)^n e^{-2r/a_0}\, r^2\, dr \ . \\ &= \frac{1}{2^{n+1}}\int_0^\infty \rho^{n+2}\, e^{-\rho}\, d\rho = \frac{(n+2)!}{2^{n+1}}\end{aligned} \qquad (8.59)$$

we have

$$\alpha_0 = \frac{9}{2}\ \frac{q_c^2 m_r a_0^4}{\varepsilon_0 \hbar^2}$$

and substituting for the Bohr radius ($q_c = -e$ = electron charge), we finally obtain the result,

$$\alpha_0 = \frac{9}{2}\left(4\pi a_0^3\right) \ . \qquad (8.60)$$

<u>Reference:</u> E. Merzbacher, <u>Quantum Mechanics</u>, chapter 17, section 4.

An improved perturbation expansion, that may be useful when a degeneracy is lifted to first order by the perturbation, can be obtained as follows. Let the eigenstates, for a finite value of the perturbation parameter λ, be expressed as a superposition of the unperturbed eigenstates as follows

$$|k;\lambda\rangle = [N_k(\lambda)]^{-1/2}[\,|k\rangle_0 + \lambda \sum_{n\neq k} |n\rangle_0 C_{nk}] \ , \tag{8.61a}$$

where $N_k(\lambda)$ is the normalization constant, derived from the condition that

$$\langle k;\lambda|k;\lambda\rangle = 1 = [N_k(\lambda)]^{-1}\,[1 + \lambda^2 \sum_{n} |C_{kn}|^2] ,$$

or

$$N_k(\lambda) = 1 + \lambda^2 \sum_{n\neq k} |C_{kn}|^2 . \tag{8.61b}$$

Substituting the states $|k;\lambda\rangle$, as given by equation 8.61, into the eigenvalue equation we have

$$(H_0 + \lambda\hat{V})|k;\lambda\rangle = E_k|k;\lambda\rangle$$

or

$$(H_0 + \lambda\hat{V})[\,|k\rangle_0 + \lambda\sum_{n\neq k}|n\rangle_0 \ C_{nk}] = E_k[\,|k\rangle_0 + \lambda\sum_{n\neq k}|n\rangle_0 \ C_{nk}]$$

or

$$[E_k - (E_k^{(0)} + \lambda\hat{V})]\,|k\rangle_0 + \lambda\sum_{n\neq k}[E_k - (E_n^{(0)} + \lambda\hat{V})]\,|n\rangle_0 \ C_{nk} = 0 \tag{8.62}$$

Taking the scalar product of equation 8.62 with the state ${}_0\langle k|$, we have

$$E_k - (E_k^{(0)} + \lambda V_{kk}) - \lambda^2 \sum_{n\neq k} V_{kn}C_{nk} = 0 \ ,$$

or

$$E_k = E_k^{(0)} + \lambda V_{kk} + \lambda^2 \sum_{n\neq k} V_{kn}C_{nk} \ , \tag{8.63a}$$

where

$$V_{kn} \equiv {}_0\langle k|V|n\rangle_0 \quad . \tag{8.63b}$$

We can also take the scalar product of 8.62 with a state ${}_0\langle m|$, where $m\neq k$, to obtain

$$- V_{mk} + [E_k - (E_m^{(0)} + \lambda V_{mm})]\,C_{mk} - \lambda \sum_{n\neq k} V_{mn} \ C_{nk} = 0 \ , \tag{8.64}$$

or, provided $E_k \neq E_m^{(0)} + \lambda V_{mm}$,

$$C_{mk} = - \frac{V_{mk} + \lambda \sum_{n\neq k} V_{mn}C_{nk}}{E_m^{(0)} + \lambda V_{mm} - E_k} \quad . \tag{8.65}$$

update

Note that equations 8.63 and 8.65 are exact. We can solve them approximately by an iterative procedure as follows. To lowest order, let $C_{nk} = 0$ for $n \neq k$, in equation 8.65, which yields

$$E_k \simeq E_k^{(o)} + \lambda V_{kk} = E_k^{(o)} + H'_{kk} , \tag{8.66}$$

as was obtained previously. To the next higher order, we can use 8.66 to evaluate the C_{mk}, as given by 8.65 by ignoring the higher order summation in the numerator, i.e.

$$C_{mk} \simeq \frac{-V_{mk}}{(E_m^{(o)} + H'_{mm}) - (E_k^{(o)} + H'_{kk})} \tag{8.67}$$

and therefore, from 8.63,

$$E_k \simeq E_k^{(o)} + H'_{kk} - \sum_{n\neq k} \frac{|H'_{nk}|^2}{(E_n^{(o)} + H'_{nn}) - (E_k^{(o)} + H'_{kk})} . \tag{8.68}$$

It can be seen that this expansion procedure requires that

$$|H'_{nk}| << |(E_n^{(o)} + H'_{nn}) - (E_k^{(o)} + H'_{kk})| , \tag{8.69}$$

which may be satisfied when the less stringent inequality, as given by equation 8.25, is not.

Higher order corrections may be obtained by substituting equations 8.67 and 8.68 into the right hand side of 8.65 to obtain the coefficients C_{mk}, correct to the next higher order, which can in turn be substituted into 8.63 to obtain a better estimate of the eigenvalues, and so on and so forth. It can be seen that the calculations get very complicated very quickly, if an explicit expression for E_k and C_{nk} is desired. Alternatively, a numerical scheme may be attempted. The matrix elements V_{nm} are first computed. If only relatively few elements are significant, a numerical iterative solution of equations 8.63 and 8.64 may be attempted (with an accelerated scheme like Newton's method), to obtain solutions to the problem.

A different method yet may be employed, based on a forward marching scheme starting from $\lambda = 0$ and ending at $\lambda = \lambda_1$, that is useful even when the Hamiltonian does not depend linearly on the perturbation parameter λ. See problems 8.8 and 8.9.

8.2 Perturbation of nearly degenerate states.

The preceding perturbation procedures generally fail when the perturbed state is degenerate, or has an energy that is close to some neighboring state.[#] See equations 8.25 and 8.69 and related discussion. In such a case, the contribution from the states with the same or almost the same energy is going to be large and the original idea of an asymptotic expansion in small corrections must be modified. To illustrate the method we will assume that the system of interest possesses two states, say $|1\rangle_0$ and $|2\rangle_0$, whose energies $E_1^{(o)}$ and $E_2^{(o)}$ are close to each other, but far from the remaining eigenvalues. It can be seen that the application of the preceding perturbation procedure will produce a large correction to the state $|1\rangle_0$ coming from the state $|2\rangle_0$ and vice versa. In the preceding formalism, c_{12} and c_{21} will be large. In anticipation of this result, it is then reasonable to look for solutions, at the outset, of the form

$$|\psi\rangle = a|1\rangle_0 + b|2\rangle_0 \ , \tag{8.70}$$

where

$$(H - E)|\psi\rangle = 0 \tag{8.71}$$

and

$$H = H_0 + H' \ .$$

Substituting 8.70 into 8.71, we have

$$\left.\begin{aligned} (H_{11} - E)a + H_{12}b &= 0 \\ H_{21}a + (H_{22} - E)b &= 0 \ , \end{aligned}\right\} \tag{8.72}$$

whose determinant of the coefficients must be zero. This yields two eigenvalues $E_+ \geq E_-$ given by

$$E_{\pm} = \frac{1}{2}(H_{11} + H_{22}) \pm \frac{1}{2}\left[(H_{11} - H_{22})^2 + 4|H_{12}|^2\right]^{1/2} \ , \tag{8.73}$$

where the square root is always computed as a positive quantity.

If the usual perturbation expansion were valid we would require that (see equation 8.69)

$$|H_{12}| \ll |H_{11} - H_{22}| \ , \tag{8.74}$$

which substituted into equation 8.73, yields

[#] unless, of course, it so happens that the matrix element of H' with the neighboring or degenerate state is identically zero.

$$E_{\pm} = \frac{1}{2}(H_{11} + H_{22}) \pm \frac{1}{2}\,|H_{11} - H_{22}| \left[1 + \frac{4|H_{12}|^2}{(H_{11} - H_{22})^2} \right]^{1/2}$$

$$\simeq \frac{1}{2}(H_{11} + H_{22}) \pm \frac{1}{2}\,|H_{11} - H_{22}| \pm \frac{|H_{12}|^2}{|H_{11} - H_{22}|}$$

and therefore

$$E_{+} \simeq \begin{cases} H_{11} + \dfrac{|H_{12}|^2}{|H_{11} - H_{22}|} = E_1^{(0)} + H'_{11} - \dfrac{|H_{21}|^2}{H_{22} - H_{11}}\,, & \text{if } H_{22} - H_{11} > 0 \\ \\ H_{22} + \dfrac{|H_{12}|^2}{|H_{11} - H_{22}|} = E_2^{(0)} + H'_{22} - \dfrac{|H_{12}|^2}{H_{11} - H_{22}}\,, & \text{if } H_{22} - H_{11} < 0 \end{cases} \tag{8.75a}$$

and conversely for E_-, i.e.

$$E_{-} \simeq \begin{cases} H_{22} + \dfrac{|H_{12}|^2}{|H_{11} - H_{22}|} = E_2^{(0)} + H'_{22} - \dfrac{|H_{12}|^2}{H_{11} - H_{22}}\,, & \text{if } H_{22} - H_{11} > 0 \\ \\ H_{11} + \dfrac{|H_{12}|^2}{|H_{11} - H_{22}|} = E_1^{(0)} + H'_{11} - \dfrac{|H_{21}|^2}{H_{22} - H_{11}}\,, & \text{if } H_{22} - H_{11} < 0\,, \end{cases} \tag{8.75b}$$

consistently with the results of the improved perturbation expansion. See equation 8.68 and related discussion.

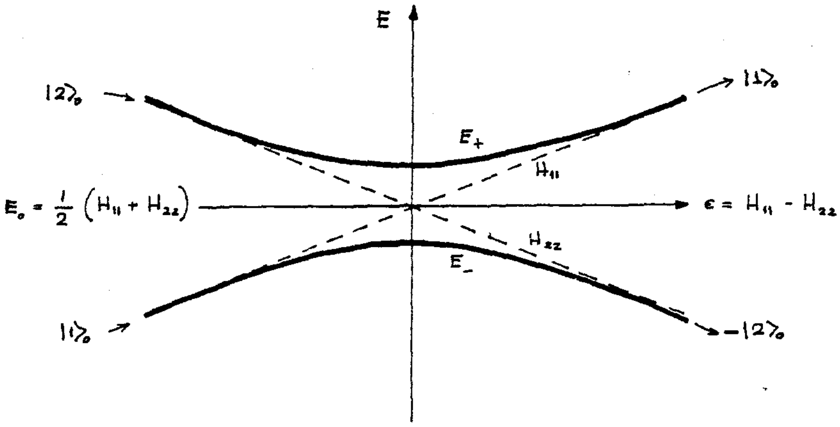
Conversely, however, if

$$|H_{11} - H_{22}| \ll |H_{12}\,, \tag{8.76}$$

then, from equation 8.73 we have,

$$E_{\pm} \simeq \frac{1}{2}(H_{11} + H_{22}) \pm |H_{12}| \left[1 + \frac{(H_{11} - H_{22})^2}{8|H_{12}|^2} \right] . \tag{8.77}$$

The eigenvalues E_+ and E_1 are sketched below, for a fixed value of $|H_{12}|$ as a function of the difference of the diagonal elements $\varepsilon = H_{11} - H_{22}$.



It can be seen that, provided $H_{12} \neq 0$, the effect of the perturbation is to lift the degeneracy by keeping the two levels apart as the energy difference, in the absence of the perturbation, approaches zero.

It is also interesting to study the resulting eigenstates as a function of $\epsilon = H_{11} - H_{22}$ and H_{12}. In particular, from equation 8.72, we have

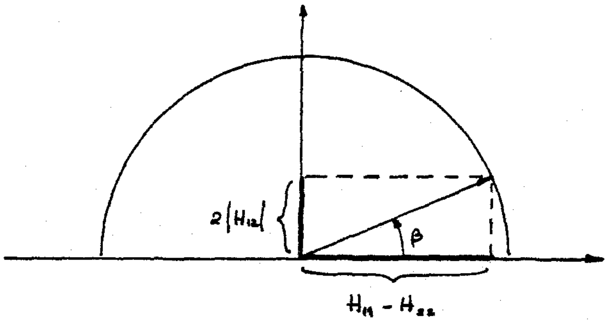
$$\left(\frac{a}{b}\right)_{\pm} = - \frac{H_{12}}{H_{11} - E_{\pm}} , \qquad (8.78)$$

where the matrix element H_{12} will in general be complex, i.e.

$$H_{12} = |H_{12}|\, e^{-i\varphi} . \qquad (8.79)$$

We now define an angle β such that

$$\tan\beta \equiv \frac{2|H_{12}|}{H_{11} - H_{22}} , \qquad (8.80)$$



where $0 < \beta < \pi$, in terms of which we have

$$\left(\frac{a}{b}\right)_{\pm} = - \frac{\sin\beta}{\cos\beta \mp 1}\, e^{-i\varphi} , \qquad (8.81)$$

and therefore

$$\left.\begin{aligned} |+\rangle &= |1\rangle_0 \cos(\beta/2) e^{-i\varphi/2} + |2\rangle_0 \sin(\beta/2) e^{i\varphi/2} \\ |-\rangle &= |1\rangle_0 \sin(\beta/2) e^{-i\varphi/2} - |2\rangle_0 \cos(\beta/2) e^{i\varphi/2} . \end{aligned}\right\} \qquad (8.82)$$

Note that we may select $\varphi = 0$ without loss of generality# and that

$$\left.\begin{aligned} |+\rangle &\to |1\rangle_0 \\ |-\rangle &\to -|2\rangle_0 \end{aligned}\right\} \quad \text{as} \quad \frac{H_{11} - H_{22}}{|H_{12}|} \to \infty \qquad (8.83a)$$

whereas

#There is an arbitrary phase difference between the unperturbed states $|1\rangle_0$ and $|2\rangle_0$ which can be chosen such that ${}_0\langle 1|H|2\rangle_0 = H_{12}$ is real.

$$\left.\begin{aligned} |+\rangle &\to |1\rangle_0 \\ |-\rangle &\to |2\rangle_0 \end{aligned}\right\} \text{ as } \frac{H_{11} - H_{22}}{|H_{12}|} \to -\infty\ , \tag{8.83b}$$

while

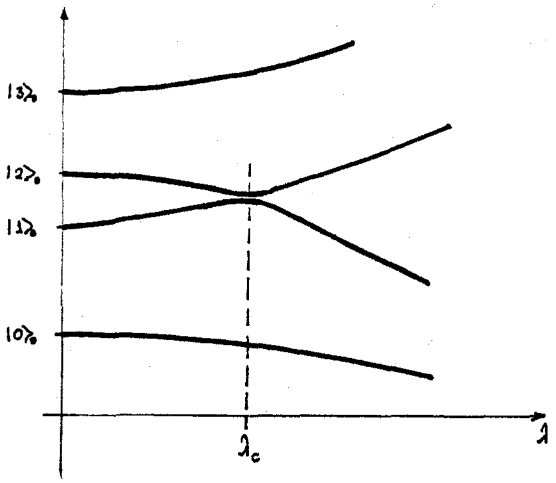
$$\left.\begin{aligned} |+\rangle &\to \frac{1}{\sqrt{2}}\,[\,|1\rangle_0 + |2\rangle_0] \\ |-\rangle &\to \frac{1}{\sqrt{2}}\,[\,|1\rangle_0 - |2\rangle_0] \end{aligned}\right\} \text{ as } \frac{H_{11} - H_{22}}{|H_{12}|} \to 0\ . \tag{8.84}$$

We can now use the states

$$|0\rangle_0,\ |-\rangle,\ |+\rangle,\ |3\rangle_0,\ |4\rangle_0,\ \ldots .$$

instead of the original set $\{|n\rangle_0\}$ as a basis in which to perform the perturbation calculation. The large corrections involving the small energy differences $(E_1 - E_2)$ in the denominator will now not occur since, for those terms, the numerator $\langle 1|H|2\rangle$ is zero by construction.

The situation described above may in fact occur as a result of a perturbation. Consider, for example, a spectrum of eigenstates that is initially discrete and non-degenerate. It is conceivable that the perturbation corrections to a pair of eigenvalues will have opposite signs so that as the value of the perturbation parameter increases, the eigenvalues approach each other or even cross over. This situation is depicted below.



The behavior of the states $|1\rangle$ and $|2\rangle$ in the vicinity of the critical value of the perturbation parameter $\lambda \simeq \lambda_c$ indicated above can only be properly described along the lines of the preceding discussion.

8.3 Perturbation of degenerate states.

We can generalize the results of the preceding discussion to the case where one or more energy levels are multiply degenerate. Again, as before, the difficulty with the ordinary perturbation theory in this case is that corrections become large and the usual perturbation procedure is invalid. Consider, for example, the effect on the degenerate states $|k,m\rangle_0$ of an (unperturbed) energy $E_k^{(o)}$, i.e.

$$H_0 |k,m\rangle_0 = E_k^{(o)} |k,m\rangle_0 \; ; \qquad m = 1,2,\ldots,N_k, \tag{8.85}$$

of a perturbation $H' = \lambda_1 V$. Following the same reasoning as in section 8.2, we seek an appropriate alternative basis for the states sharing the eigenvalue $E_k^{(o)}$, formed as a linear superposition of the states $|k,m\rangle_0$, which also diagonalizes H'. In other words, we seek the transformation

$$|k,i\rangle = \sum_m |k,m\rangle_0 \, C_{k;m,i} \tag{8.86}$$

such that

$$H|k,i\rangle = (H_0 + H')|k,i\rangle = E_{k,i}|k,i\rangle \, , \tag{8.87a}$$

where

$$E_{k,i} = \langle k,i|(H_0 + H')|k,i\rangle = E_k^{(o)} + E'_{k,i} \, , \tag{8.87b}$$

and

$$E'_{k,i} = \langle k,i|H'|k,i\rangle \, . \tag{8.87c}$$

Note that, if this transformation can be found, equation 8.87b actually gives the energies of the levels $|k,i\rangle$ correct to first order in the perturbation since they are the expectation values of the perturbed Hamiltonian computed with the unperturbed eigenstates. See equations 8.26 and 8.66 and related discussions.

Example 8.4 The linear Stark effect. Perturbation of the $n = 2$ levels of the hydrogen atom by an electric field.

There is a four-fold degeneracy of the $n = 2$ energy, corresponding to the 2s state and the three 2p states (see figure on page 7.13 and related discussion). The four $|n,\ell,m\rangle$ states are:

$$|2,0,0\rangle, \quad |2,1,1\rangle, \quad |2,1,0\rangle \quad \text{and} \quad |2,1,-1\rangle \, .$$

The perturbation H' is given by

$$H' = -q_c \mathcal{E}_z z = e\mathcal{E}_z z \, , \tag{8.88}$$

where $q_c = -e$ is the charge of the electron. Note that the perturbation does not involve the azimuthal coordinate, i.e.

$$z = a_0 (\frac{r}{a_0}) \cos\theta = (\frac{4\pi}{3})^{1/2} a_0 (\frac{r}{a_0}) Y_{1,0}(\theta,\varphi) \qquad (8.89)$$

and therefore only the states $|2,0,0\rangle$ and $|2,1,0\rangle$ are mixed by the perturbation, since

$$\langle 2,\ell,m|H'|2,1,1\rangle = 0 \ , \quad \text{for} \quad \ell=0,1 \quad \text{and} \quad m=0,\pm 1 \qquad (8.90a)$$

and

$$\langle 2,\ell,m'|H'|2,1,-1\rangle = 0 \ , \quad \text{for} \quad \ell=0,1 \quad \text{and} \quad m'=0,\pm 1 \ . \qquad (8.90b)$$

It is therefore sufficient to consider the linear combinations of $|2,0,0\rangle$ and $|2,1,0\rangle$ that diagonalize H'. This leads to the equation

$$\begin{vmatrix} \langle 2,0,0|H'|2,0,0\rangle - E' & \langle 2,0,0|H'|2,1,0\rangle \\ \langle 2,1,0|H'|2,0,0\rangle & \langle 2,1,0|H'|2,1,0\rangle - E' \end{vmatrix} = 0 \ , \qquad (8.91)$$

and since

$$\langle 2,0,0|H'|2,0,0\rangle \propto \int_0^{\pi} |Y_{00}|^2 \cos\theta \, d\cos\theta = 0$$

and

$$\langle 2,1,0|H'|2,1,0\rangle \propto \int_0^{\pi} |Y_{10}|^2 \cos\theta \, d\cos\theta = 0$$

we have that, to first order,

$$\begin{aligned} E'_{\pm} &= \pm e\mathcal{E}_z \langle 2,0,0|z|2,1,0\rangle \\ &= \pm e\mathcal{E}_z a_0 \langle 2,0,0|(\frac{r}{a_0}) \cos\theta|2,1,0\rangle \end{aligned} \qquad (8.92)$$

or, substituting (see equations 5.88 and 5.137)

$$\psi_{2,0,0} = (\frac{1}{4\pi})^{1/2} (\frac{1}{2a_0})^{3/2} (2 - \frac{r}{a}) e^{-r/2a_0}$$

$$\psi_{2,1,0} = (\frac{1}{4\pi})^{1/2} (\frac{1}{2a_0})^{3/2} (\frac{r}{a_0}) e^{-r/2a_0} \cos\theta$$

we have

$$\langle 2,0,0|(\frac{r}{a_0}) \cos\theta|2,1,0\rangle = -3 \qquad (8.93)$$

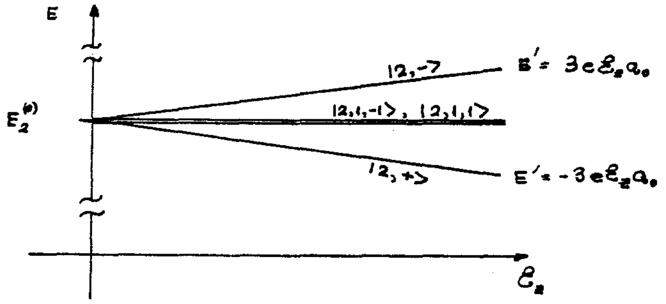
and therefore

$$E'_{\pm} = \mp 3e\mathcal{E}_z a_0 \ , \qquad (8.94)$$

corresponding to the states

$$\begin{aligned} |2,+\rangle &= \frac{1}{\sqrt{2}} [|2,0,0\rangle + |2,1,0\rangle] \\ |2,-\rangle &= \frac{1}{\sqrt{2}} [|2,0,0\rangle - |2,1,0\rangle] \ . \end{aligned} \qquad (8.95)$$

Consequently, to first order in the perturbation, the energy levels of the four $n = 2$ hydrogen atom states split into three energies (partial lifting of the degeneracy),



whose separation is linear in the magnitude of the electric field (linear Stark effect), as opposed to the energy shift of the ground state as a function of the electric field, i. e.

$$\Delta E_0 = -\frac{1}{2}\epsilon_0 \alpha_0 \mathcal{E}_z^2$$

(see example 8.3 and problem 8.4), which is quadratic in the electric field (quadratic Stark effect).

The states $|k,i\rangle$ which diagonalize the submatrix of H' corresponding to the eigenvalue $E_k^{(o)}$ (see equations 8.85, 8.86 and 8.87)[#] can now be used along with the states $|k', m_k'\rangle_0$, for $k' \neq k$, as a zeroth order basis on which the perturbation may be computed. The previously derived results now apply since the singularities have been removed. By way of example, we have for the energies $E_{k,i}$ correct to second order (see equation 8.30)

$$E_{k,i} \sim E_k^{(o)} + \langle k,i|H'|k,i\rangle - \sum_{k' \neq k} \frac{1}{E_{k'}^{(o)} - E_k^{(o)}} \sum_{m_k'} |{}_0\langle k', m_k'|H'|k,i\rangle|^2 \ . \tag{8.96}$$

[#] Note that they already diagonalize H_0 and therefore also $H = H_0 + H'$.

8.4 Time-dependent perturbation theory.

In the preceding sections we have considered the case where the perturbation V is not a function of time. If, however, $V = V(t)$, we must use an alternative scheme since the previously obtained results were derived on the basis of the time-independent Schrödinger equation.

Recall that if $H \neq H(t)$ then the general solution $|\psi, t\rangle$ can always be expressed as a linear superposition

$$|\psi, t\rangle = \sum_n |n\rangle c_n e^{-iE_n t/\hbar} \tag{8.97a}$$

where

$$H|n\rangle = E_n|n\rangle \tag{8.97b}$$

and the coefficients c_n are constants that can be evaluated from the initial conditions, i.e.

$$\sum_n |n\rangle c_n = |\psi, 0\rangle$$

and therefore, since $\langle m|n\rangle = \delta_{mn}$,

$$c_m = \langle m|\psi, 0\rangle . \tag{8.97c}$$

Consider now a Hamiltonian composed of a "large" H_0 that is time independent and a "small" V that may be a function of time, i.e.

$$H = H_0 + V(t) . \tag{8.98}$$

We can use the fact that the eigenstates of H_0 form a complete set to express the solutions of the Hamiltonian H as a linear superposition, in which the coefficients may now be functions of time, i.e.

$$|\psi, t\rangle = \sum_n |n\rangle c_n(t) e^{-iE_n t/\hbar} , \tag{8.99a}$$

where

$$H_0|n\rangle = E_n|n\rangle . \tag{8.99b}$$

To find the coefficients $c_n(t)$ we now substitute 8.99 into the time dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi, t\rangle = H|\psi, t\rangle ,$$

or

$$i\hbar \sum_n |n\rangle(\dot{c}_n - i\frac{E_n}{\hbar} c_n) e^{-iE_n t/\hbar} = [H_0 + V(t)] \sum_n |n\rangle c_n e^{-iE_n t/\hbar} ,$$

which yields

$$i\hbar \sum_n |n\rangle \dot{c}_n e^{-iE_n t/\hbar} = \sum_n V(t)|n\rangle c_n e^{-iE_n t/\hbar} .$$

Taking the scalar product with $\langle m|$, we then have

$$i\hbar\, \dot{c}_m(t) = \sum_n v_{mn}(t)\, c_n(t) \tag{8.100a}$$

where

$$v_{mn}(t) \equiv \langle m|V(t)|n\rangle\, e^{i(E_m - E_n)t/\hbar} . \tag{8.100b}$$

Equation 8.100 is, of course, exact and its solution is, the solution of the original problem, via equation 8.99, given the initial state $|\psi, 0\rangle$. It is also a set of linear homogeneous equations, i.e.

$$i\hbar \frac{d}{dt} \underline{c} - \mathcal{V}\underline{c} = 0, \tag{8.100'}$$

where $\underline{c}$ is the vector $(c_1, c_2, \ldots, c_m, \ldots)$ and $\mathcal{V}$ is the matrix (v_{mn}), and therefore possesses a solution of the form

$$c_m(t) = \sum_n \gamma_{mn}(t, t_i)\, c_n(t_i) , \tag{8.101}$$

or, equivalently

$$\underline{c}(t) = \Gamma(t, t_i)\, \underline{c}(t_i) , \tag{8.101'}$$

where t_i is some initial time. Note that

$$\mathcal{P}_{n\to m} = |\gamma_{mn}(t, t_i)|^2 \tag{8.102}$$

is the transition probability to a final state $|m\rangle$ at time t, from an initial state $|n\rangle$ at time $t = t_i$.

We can integrate equation 8.100 to obtain an integral equation for the coefficients $c_m(t)$, i.e.

$$c_m(t) = c_m(t_i) - \frac{i}{\hbar} \sum_n \int_{t_i}^{t} v_{mn}(t_1) c_n(t_1)\, dt_1 \tag{8.103}$$

which we can solve by an iterative scheme as follows. Assuming that the v_{mn} are small, we have to a zeroth approximation that

$$c_m(t) \simeq c_m(t_i) ,$$

i.e. no change. If we now substitute this in the integral equation we obtain

$$c_m(t) \simeq c_m(t_i) - \frac{i}{\hbar} \sum_n \left[\int_{t_i}^{t} v_{mn}(t_1)\, dt_1 \right] c_n(t_i) , \tag{8.104}$$

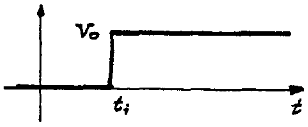
and therefore, to first order in the perturbation,

$$\boxed{\gamma_{mn}(t_f, t_i) \simeq \delta_{mn} - \frac{i}{\hbar} \int_{t_i}^{t_f} v_{mn}(t_1)\, dt_1} \tag{8.105}$$

called the <u>first Born approximation</u>.

<u>Example 8.5</u> <u>Impulsive start of a constant potential.</u>

$$V(t) = \begin{cases} V_0 & \text{for} \quad t > t_i \\ 0 & \text{for} \quad t < t_i \, . \end{cases}$$



For $t_f > t_i$ and $m \neq n$

$$\gamma_{mn}(t_f, t_i) \simeq -\frac{i}{\hbar}\int_{t_i}^{t_f} v_{mn}(t_1)\,dt_1$$

$$= -\frac{i}{\hbar}\int_{t_i}^{t_f} \langle m|V_0|n\rangle\, e^{i(E_m - E_n)t_1/\hbar}\, dt_1$$

$$= -\frac{i}{\hbar}\,\langle m|V_0|n\rangle \int_{t_i}^{t_f} e^{i(E_m - E_n)t_1/\hbar}\, dt_1$$

or

$$\gamma_{mn}(t_f, t_i) \simeq -\frac{\langle m|V_0|n\rangle}{E_m - E_n}\left[e^{i(E_m - E_n)t_f/\hbar} - e^{i(E_m - E_n)t_i/\hbar}\right] .$$

Therefore the transition probability is given by

$$\mathscr{P}_{n\to m} = |\gamma_{mn}(t_f, t_i)|^2 \simeq |\langle m|V_0|n\rangle|^2\, \frac{\sin^2[(E_m - E_n)(t_f - t_i)/2\hbar]}{[(E_m - E_n)/2]^2} . \tag{8.106}$$

Note that if $|\langle m|V_0|n\rangle| >> |E_m - E_n|$ the transition probability is small. Note also that if $E_m = E_n$

$$\mathscr{P}_{n\to m} \doteq \frac{1}{\hbar^2}|\langle m|V_0|n\rangle|^2 (t_f - t_i)^2 , \tag{8.107}$$

for small $t_f - t_i$ (where have we made this assumption?)

It is possible to arrive at a direct equation for the transition amplitude matrix $\gamma_{mn}(t_f, t_i)$ by substituting equation 6.101 into the integral equation 8.103, i.e.

$$\sum_n \gamma_{mn}(t_f, t_i)\, c_n(t_i) = c_m(t_i) - \frac{i}{\hbar}\sum_{k,n}\left[\int_{t_i}^{t_f} v_{mk}(t)\,\gamma_{kn}(t, t_i)\,dt\right] c_n(t_i)$$

or, since this equation must hold for any set of initial conditions, we must have, for all m and n,

$$\boxed{\gamma_{mn}(t_f, t_i) = \delta_{mn} - \frac{i}{\hbar} \sum_k \int_{t_i}^{t_f} v_{mk}(t)\, \gamma_{kn}(t, t_i)\, dt} \qquad (8.108)$$

We can solve equation 8.108 by an iterative scheme, as before. In particular, to zeroth order (equivalent to no perturbation), we have

$$\gamma_{mn}(t_f, t_i) \sim \delta_{mn} \, .$$

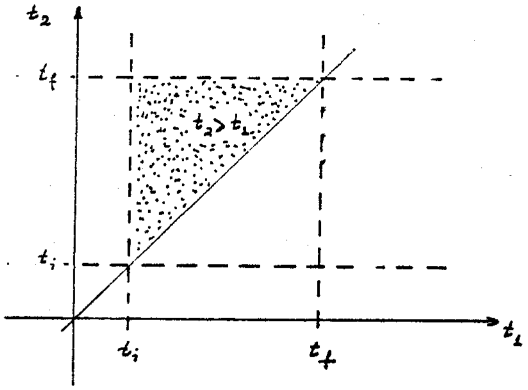
To first order, we can substitute the zeroth order solution in the left hand side to obtain equation 8.105 directly, i.e.

$$\gamma_{mn}(t_f, t_i) \sim \delta_{mn} - \frac{i}{\hbar} \int_{t_i}^{t_f} v_{mn}(t_1)\, dt_1 \, .$$

This can in turn be substituted to yield an estimate correct to second order, i.e.

$$\gamma_{mn}(t_f, t_i) \sim \delta_{mn} - \frac{i}{\hbar} \int_{t_i}^{t_f} v_{mn}(t_1)\, dt_1 + \left(\frac{-i}{\hbar}\right)^2 \sum_k \int_{t_i}^{t_f} v_{mk}(t_2) \int_{t_i}^{t_2} v_{kn}(t_1)\, dt_1\, dt_2 \, ,$$

where the integration is over a triangular region in the t_1, t_2 plane, i.e.



Second order range of integration variables.

and so forth. This yields an infinite series

$$\gamma_{mn}(t_f, t_i) \sim \gamma^{(0)}_{mn}(t_f, t_i) + \gamma^{(1)}_{mn}(t_f, t_i) + \ldots + \gamma^{(r)}_{mn}(t_f, t_i) \ldots \, , \qquad (8.109)$$

where

$$\gamma^{(0)}_{mn}(t_f, t_i) = \delta_{mn}$$

$$\gamma^{(1)}_{mn}(t_f, t_i) = -\frac{i}{\hbar}\int_{t_i}^{t_f} v_{mn}(t_1)\,dt_1$$

$$\gamma^{(2)}_{mn}(t_f, t_i) = \left(\frac{-i}{\hbar}\right)^2 \sum_{k_1}\int_{t_i}^{t_f} v_{mk_1}(t_2)\int_{t_i}^{t_2} v_{k_1 n}(t_1)\,dt_1\,dt_2 \tag{8.110}$$

$$\gamma^{(3)}_{mn}(t_f, t_i) = \left(\frac{-i}{\hbar}\right)^3 \sum_{k_1, k_2}\int_{t_i}^{t_f} v_{mk_2}(t_3)\int_{t_i}^{t_3} v_{k_2 k_1}(t_2)\int_{t_i}^{t_2} v_{k_1 n}(t_1)\,dt_1\,dt_2\,dt_3$$

$$\vdots$$

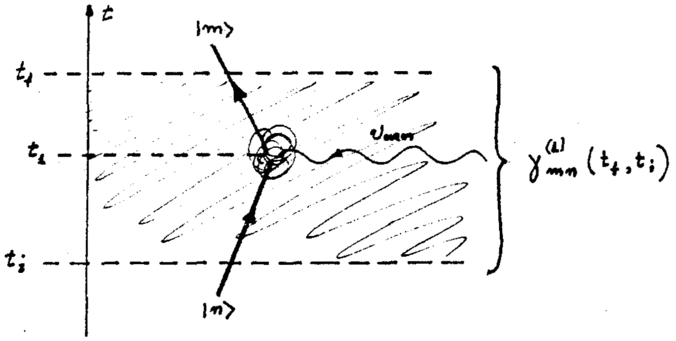
$$\gamma^{(r)}_{mn}(t_f, t_i) = \left(\frac{-i}{\hbar}\right)^r \sum_{k_1, k_2, \ldots k_{r-1}}\int_{t_i}^{t_f}\cdots\int_{t_i}^{t_f} v_{mk_{r-1}}(t_r)\ldots v_{k_2 k_1}(t_2) v_{k_1 n}(t_1)\,dt_1\,dt_2\ldots dt_r\,,$$

$$\{t_f > t_r > t_{r-1} > \ldots . > t_2 > t_1 > t_i\}$$

where the notation in the general term implies that the various factors in the integrand are to be evaluated at an array of points in time that are ordered as indicated (see for example previous two terms).

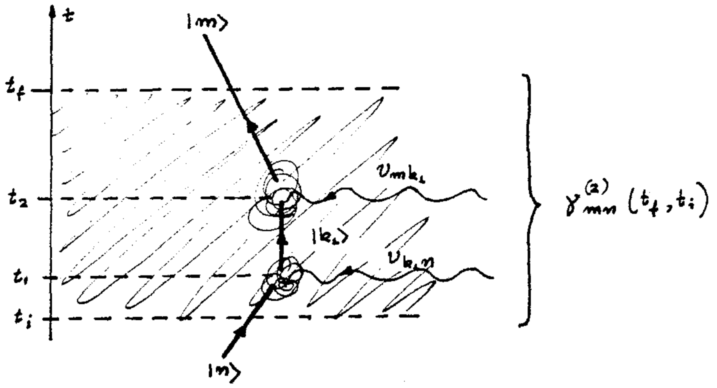
This perturbation expansion suggests a picture for the transition amplitude form a state $|n\rangle$ to a state $|m\rangle$, say for $m \neq n$, namely a superposition of the following possibilities:

(1) The system starts in the state $|n\rangle$ at time t_i and remains there until a time t_1 when the perturbation potential acts and knocks the system into the final state $|m\rangle$. This situation is depicted below:



The transition amplitude $\gamma^{(1)}_{mn}(t_f, t_i)$, corresponding to this particular sequence of events, is given by the sum (integral) over all possible intermediate times t_1, such that $t_i < t_1 < t_f$.

(2) The system starts in the state $|n\rangle$ at time t_i, remains there until a time t_1 when the perturbation acts to scatter it to a state $|k_1\rangle$, where it remains until a time t_2 when the perturbation acts again to scatter it into the final state $|m\rangle$.



The transition amplitude $\gamma^{(2)}_{mn}(t_f, t_i)$, as a result of this particular sequence of events is,

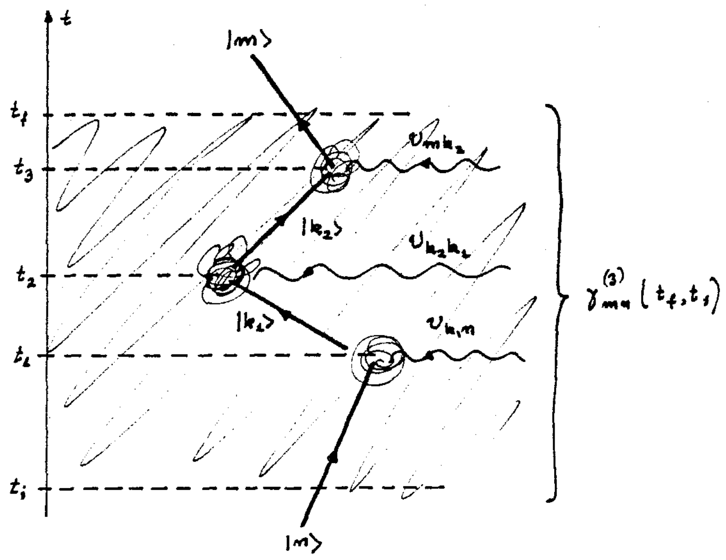
(i) the integral over all possible interaction times t_1, such that $t_i < t_1 < t_2$,

(ii) the sum over all possible intermediate states $|k_1\rangle$,

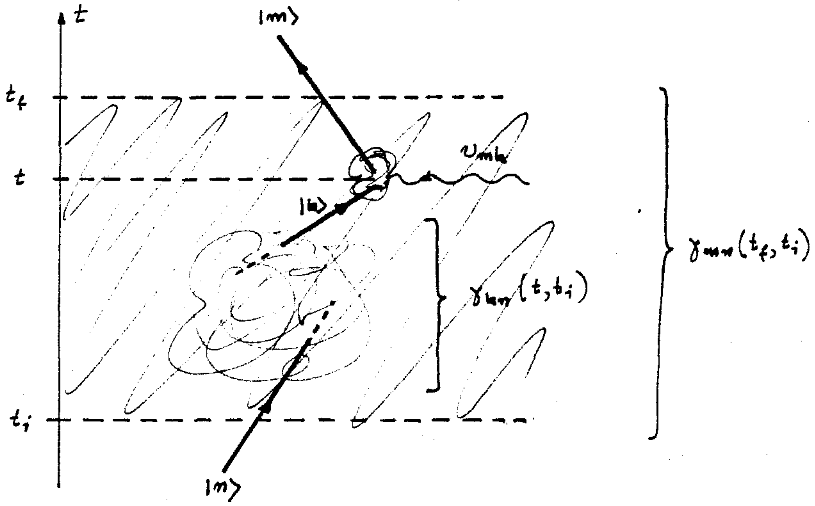
and

(iii) The integral over all possible interaction times t_2, such that $t_1 < t_2 < t_f$.

(3)



The total transition amplitude is then given by the sum in equation 6.109 over the transition amplitudes $\gamma_{mn}^{(r)}(t_f, t_i)$ representing each of the (indistinguishable) possibilities described above. The superscript r in each of the terms counts the number of times the potential has acted in the interval (t_i, t_f). Note that the exact equation 8.108 for $\gamma_{mn}(t_f, t_i)$ can also be interpreted along the same lines. Namely, by some means or other, the system has made a transition from the initial state $|n\rangle$, at a time t_i, to an intermediate state $|k\rangle$ by the time t, when the perturbation acts for the last time to knock it into the final state $|m\rangle$ where we find it at t_f.



The integral equation is then a statement of the fact that the probability amplitude to come to the state $|k\rangle$ by the time t, from the state $|n\rangle$ at time t_i, must be given by the same function appropriately evaluated.

In practical applications, one obtains a calculation formula by truncating the perturbation series of equation 8.109 after a finite number of terms.

8.4.1. Unitarity and the conservation of probability.

The fact that the modulus squared of the elements $\gamma_{mn}(t, t_i)$ represent transition probabilities, i.e. equation 8.102, imposes a very substantial restriction on the operator $\Gamma(t, t_i)$. In particular, we see that if we sum the modulus squared of $\gamma_{mn}(t, t_i)$ over all m, we have

$$\sum_m |\gamma_{mn}(t, t_i)|^2 = \sum_m \mathcal{P}_{n\to m} = 1 . \qquad (8.111)$$

In other words, the sum of the transition probabilities from the state $|n\rangle$ to all states (including $|n\rangle$) must be equal to unity, since we have covered all the possibilities.

Equation 8.111 can be shown directly from the defining equation 8.101 for the $\gamma_{mn}(t, t_i)$ as follows. We must have

$$\sum_m |c_m(t)|^2 = \sum_m c_m^*(t) c_m(t) = 1 , \qquad (8.112)$$

for all times t. Consequently, from equation 8.101, we have

$$\sum_m |c_m(t)|^2 = \sum_{m,k,n} \gamma_{mk}^*(t, t_i) \gamma_{mn}(t, t_i) c_k^*(t_i) c_n(t_i) ,$$

or

$$\sum_m |c_m(t)|^2 = \sum_{m,k,n} \gamma_{km}^\dagger(t, t_i) \gamma_{mn}(t, t_i) c_k^*(t_i) c_n(t_i) , \qquad (8.113)$$

where we have defined the adjoint matrix

$$\gamma_{km}^\dagger(t, t_i) \equiv \gamma_{mk}^*(t, t_i) \qquad (8.114)$$

as the complex conjugate, transpose of the original matrix. It is clear, however, that equation 8.113 can only be valid independently of the initial conditions if

$$\sum_m \gamma_{km}^\dagger(t, t_i) \gamma_{mn}(t, t_i) = \delta_{kn} , \qquad (8.115)$$

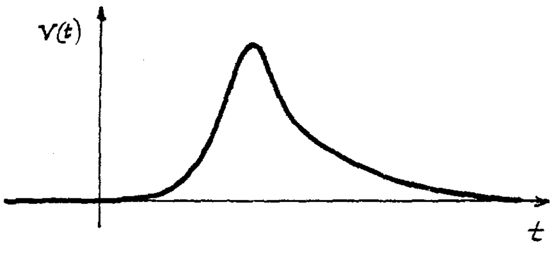
for all times t, or in matrix notation,

$$\Gamma^\dagger(t, t_i) \Gamma(t, t_i) = I , \qquad (8.115')$$

where I is the identity matrix. A matrix that satisfies equation 8.115 is called unitary.

8.4.2 Transitions as a result of perturbations localized in time. The S-matrix.

Often times the perturbation potential V(t) is localized in time, as in the sketch below.



Such a situation can occur in a (near) collision for example. In such a case one is usually interested in the response of a system, in a known state at $t = -\infty$, to the perturbation and, in particular, in the final state at $t = \infty$. From the preceding discussion, we then have

$$\mathscr{P}_{n\to m}(\infty, -\infty) = |\gamma_{mn}(\infty, -\infty)|^2 .$$

The matrix Γ, evaluated for $t_i = -\infty$ and $t_f = +\infty$ is of particular significance and is called the S-matrix

$$S \equiv \Gamma(\infty, -\infty) . \tag{8.116}$$

It is clear, from the preceding discussion, that the S-matrix is unitary.

Substituting from equation 8.109; we have to first order in the perturbation,

$$S_{mn} \sim \delta_{mn} - \frac{i}{\hbar} \int_{-\infty}^{\infty} v_{mn}(t)\, dt$$

or, substituting for $v_{mn}(t)$ from equation 8.100b,

$$S_{mn} \sim \delta_{mn} - \frac{i}{\hbar} \int_{-\infty}^{\infty} \langle m|V(t)|n\rangle\, e^{i\omega_{mn}t}\, dt , \tag{8.117a}$$

where

$$\omega_{mn} \equiv (E_m - E_n)/\hbar . \tag{8.117b}$$

In other words, the off-diagonal elements of the S-matrix are proportional to the (temporal) Fourier transform of the corresponding matrix elements of the perturbing potential, evaluated at the corresponding Bohr frequency (equation 8.117b). Note, from equation 8.116, that the transition probability to a state $|m\rangle \neq |n\rangle$ is given by

$$\mathscr{P}_{n\to m} = |S_{mn}|^2 = \frac{1}{\hbar^2} \left| \int_{-\infty}^{\infty} \langle m|V(t)|n\rangle\, e^{i\omega_{mn}t}\, dt \right|^2 . \tag{8.118}$$

Example 8.6 Response of a ground state charged particle harmonic oscillator to a Gaussian impulse electric field.

Consider a harmonic oscillator in the ground state $|0\rangle$ at $t = -\infty$ and a uniform electric field impulse

$$\mathscr{E}(t) = \frac{A}{\pi^{1/2}\tau}\, e^{-(t/\tau)^2} . \tag{8.119}$$

The perturbation potential is given by $V(t) = -q_c\mathcal{E}_x$ and the transition probabilities from the ground state are given by (equation 8.118),

$$\mathcal{P}_{0\to m} = \frac{1}{\hbar^2}\left|-\frac{q_c A}{\pi^{1/2}\tau}\langle m|x|0\rangle\int_{-\infty}^{\infty} e^{-(t/\tau)^2 + i\omega_{m0}t}\,dt\right|^2 ,$$

Now, we have

$$\langle m|x|0\rangle = \left(\frac{\hbar}{2m\omega_0}\right)^{1/2}\langle m|(a^\dagger + a)|0\rangle = \left(\frac{\hbar}{2m\omega_0}\right)^{1/2}\delta_{m1}$$

and

$$\int_{-\infty}^{\infty} e^{-\alpha t^2 + i\beta t}\,dt = \left(\frac{\pi}{\alpha}\right)^{1/2} e^{-(\beta^2/4\alpha)} = \pi^{1/2}\tau\, e^{-(\omega_0\tau/2)^2}$$

Consequently, to first order, the only non-zero transition probability is to the first excited state for which we have

$$\mathcal{P}_{0\to 1} = \frac{P^2}{2m\hbar\omega_0}\, e^{-\frac{1}{2}(\omega_0\tau)^2} , \qquad (8.120a)$$

where

$$P = q_c\int_{-\infty}^{\infty}\mathcal{E}(t)\,dt = q_c A \qquad (8.120b)$$

is the classical momentum transferred to the oscillator by the electric field.

Note that if $\tau \gg 1/\omega_0$, i.e. if the characteristic time is much longer than the classical period, the transition probability is very small. Conversely, if $\tau \ll 1/\omega_0$ the transition probability is essentially independent of time and given by the ratio of the kinetic energy imparted to the oscillator, to the energy quantum $\hbar\omega_0$, i.e.

$$\lim_{\tau\to 0}\{\mathcal{P}_{0\to 1}\} = \frac{P^2/2m}{\hbar\omega_0} .$$

Note also that for the perturbation procedure to be valid $\mathcal{P}_{0\to 1} \ll 1$ and that therefore the excitation of the harmonic oscillator is a purely quantum mechanical effect, since the classically computed energy $P^2/2m$ received by the oscillator is insufficient to allow the transition to occur. The difficulty, however, is not real because quantum mechanically $P^2/2m$ is not the energy the oscillator gains from the field.

8.4.3 Transitions as a result of harmonic perturbations.

In the present discussion we will assume that the perturbation potential $H'(t)$ is of the particular form

$$H'(t) = V^{\dagger} e^{i\omega t} + V e^{-i\omega t} , \tag{8.121}$$

where V is not a function of time and $V^{\dagger}$ is the adjoint operator to V. Note that $H'(t)$ is Hermitean by construction.

As a result of such a perturbation, the transition amplitude matrix would be given by, for $m \neq n$,

$$\gamma_{mn}(t_f, t_i) \sim -\frac{i}{\hbar} \int_{t_i}^{t_f} \langle m|H'(t)|n\rangle e^{i\omega_{mn} t} dt$$

to first order, or substituting from equation 6.121 we have,

$$\gamma_{mn}(t_f, t_i, \omega) \sim -\frac{i}{\hbar} \left[V^*_{nm} \int_{t_i}^{t_f} e^{i(\omega_{mn}+\omega)t} dt + V_{mn} \int_{t_i}^{t_f} e^{i(\omega_{mn}-\omega)t} dt \right] , \tag{8.122a}$$

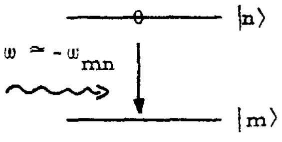
where

$$V_{mn} = \langle m|V|n\rangle \;^{\#} . \tag{8.122b}$$

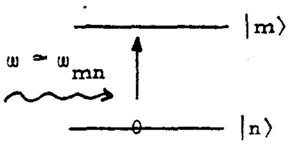
Carrying out the integration, we obtain

$$\gamma_{mn}(t_f, t_i;\omega) \sim -\frac{i}{\hbar} \left[V^*_{nm} \frac{e^{i(\omega_{mn}+\omega)t_f} - e^{i(\omega_{mn}+\omega)t_i}}{i(\omega_{mn}+\omega)} + V_{mn} \frac{e^{i(\omega_{mn}-\omega)t_f} - e^{i(\omega_{mn}-\omega)t_i}}{i(\omega_{mn}-\omega)} \right] . \tag{8.123}$$

It can be seen that the transition amplitude is composed of two parts, one of which is large when $\omega \simeq -\omega_{mn}$ while the other is large when $\omega \simeq \omega_{mn}$. The two cases correspond to the following situations (recall equation 8.117b for ω_{mn})



induced de-excitation



induced excitation .

Consider now the case $\omega \simeq \omega_{mn}$. Then, provided

$$|\omega - \omega_{mn}| << |\omega_{mn}| , \tag{8.124}$$

[#] Note that

$$\langle m|V^{\dagger}|n\rangle = \int \psi^*_m (V^{\dagger}\psi_n)dx \equiv \int (V\psi_m)^* \psi_n dx = [\int \psi^*_n (V\psi_m)dx]^* = \langle n|V|m\rangle^*$$

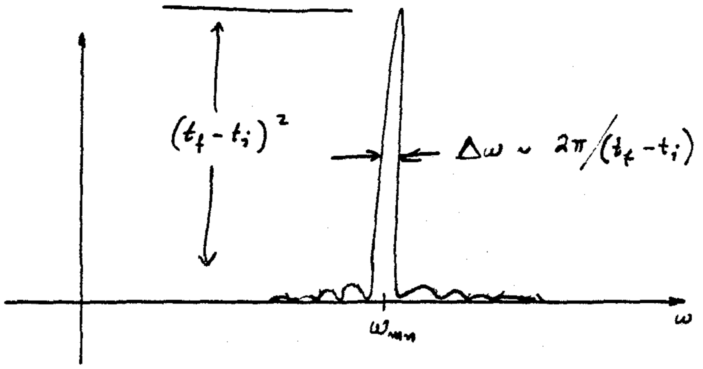
we have

$$\gamma_{mn}(t_f, t_i;\omega) \simeq -\frac{i}{\hbar} V_{mn} \frac{e^{i(\omega_{mn}-\omega)t_f} - e^{i(\omega_{mn}-\omega)t_i}}{i(\omega_{mn}-\omega)}$$

with a resulting transition probability,

$$\mathcal{P}_{n\to m}(t_f, t_i;\omega) \simeq \frac{1}{\hbar^2} |V_{mn}|^2 \frac{\sin^2[(\omega_{mn}-\omega)(t_f-t_i)/2]}{[(\omega_{mn}-\omega)/2]^2} \quad . \qquad (8.125)$$

The function multiplying the matrix element squared (over $\hbar^2$) has an appreciable magnitude only for $|\omega_{mn}-\omega| \leq \Delta\omega$,



with a width $\Delta\omega \sim 2\pi/(t_f - t_i)$ and a height given by $(t_f - t_i)^2$. Note also that

$$\int_{-\infty}^{\infty} \frac{\sin^2[(\omega_{mn}-\omega)(t_f-t_i)/2]}{[(\omega_{mn}-\omega)/2]^2} d\omega = 2(t_f-t_i) \int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} dx = 2\pi(t_f-t_i)$$

and therefore, for large $t_f - t_i$, we have

$$\frac{\sin^2[(\omega_{mn}-\omega)(t_f-t_i)/2]}{[(\omega_{mn}-\omega)/2]^2} \to 2\pi(t_f-t_i)\,\delta(\omega_{mn}-\omega) \ , \qquad (8.126)$$

where $\delta(\omega_{mn}-\omega)$ is the Dirac delta function (see Appendix D). Consequently, for large time intervals

$$\mathcal{P}_{n\to m}(t_f, t_i;\omega) \to \frac{2\pi}{\hbar^2} |V_{mn}|^2 \delta(\omega_{mn}-\omega) \cdot (t_f - t_i) \ ,$$

which allows us to define a <u>probability rate</u> (probability per unit time) for the process, given by

$$W_{n\to m} \equiv \frac{d}{dt} \mathcal{P}_{n\to m} \sim \frac{2\pi}{\hbar^2} |V_{mn}|^2 \delta(\omega_{mn}-\omega) \ ,$$

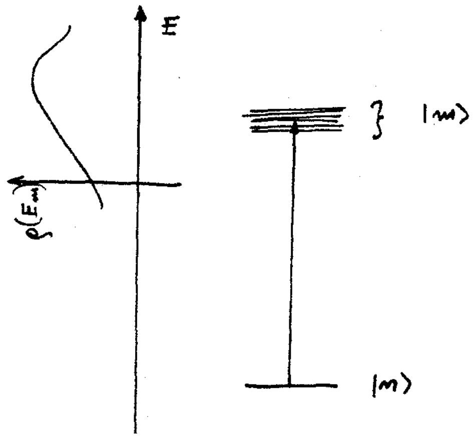
or in terms of the energies#

#Note $\delta(x/a) = |a|\delta(x)$, see Appendix D.

$$\boxed{W_{n\to m}(\omega) \sim \frac{2\pi}{\hbar} |V_{mn}|^2 \delta(E_m - E_n - \hbar\omega)} \quad . \tag{8.127}$$

This very important result is known as Fermi's golden rule. Note that a harmonic perturbation of frequency ω acts as if it possesses an energy of $\hbar\omega$.

Equation 8.127 is useful in transitions to (or from) a cluster of states of energy $E \simeq E_m$, or transitions to (or from) a continuum. In particular, consider the case where we are interested in the transition rate to a group of neighboring states $|m\rangle$.



The overall transition rate is then given by

$$W_T \sim \sum_m W_{n\to m} \sim \frac{2\pi}{\hbar} \sum_m |V_{mn}|^2 \delta(E_m - E_n - \hbar\omega) \ ,$$

or, if the final states are closely spaced, we can replace the summation over the final state index with an integral over the final density of states, i. e.

$$W_T \sim \frac{2\pi}{\hbar} \int |V_{mn}|^2 \delta(E_m - E_n - \hbar\omega)\, \rho(E_m) dE_m$$

or

$$\boxed{W_T \sim \frac{2\pi}{\hbar} |V_{mn}|^2 \rho(E_m = E_n + \hbar\omega) \ .} \tag{8.128}$$

These results can also be extended to the case where $\omega = 0$. In particular, if $H'(t) = H'$ (not a function of time), then

$$W_{n\to m} \sim \frac{2\pi}{\hbar} |H'_{mn}|^2 \delta(E_m - E_n) \quad ; \quad \omega = 0 \ . \tag{8.129}$$

See example 8.5.

8.4.4 The differential scattering cross-section in the Born Approximation.[#]

Consider a beam of particles with a well defined initial momentum $\underline{p}_0 = \hbar\underline{k}_0$, incident on a target of the origin. The particles will be deflected, as a result of the potential V(r) between the beam particles and the target particles, to emerge with a new momentum $\underline{p} = \hbar\underline{k}$. We would like to compute

the number of particles deflected into a solid angle $d\underline{\Omega}$ per second. For the purposes of calculation, consider the incident particles in an initial state of definite momentum, in a box of volume L^3, then

$$|i\rangle = \frac{1}{L^{3/2}}\, e^{i\underline{k}_0\cdot\underline{x}} \tag{8.130a}$$

whereas the emerging particles are in a final state

$$|f\rangle = \frac{1}{L^{3/2}}\, e^{i\underline{k}\cdot\underline{x}} \ . \tag{8.130b}$$

Consequently, the relevant matrix element is given by

$$\langle f|V|i\rangle = \frac{1}{L^3}\int V(r)\, e^{-i(\underline{k}-\underline{k}_0)\cdot\underline{r}}\, d^3\underline{r} \equiv \frac{1}{L^3}\,\mathcal{V}(|\underline{k}-\underline{k}_0|)\ , \tag{8.131}$$

where $\mathcal{V}(\underline{k})$ is the spatial Fourier transform of V(r). Substituting this result into equation 8.129, we have for a transition $\underline{k}_0 \rightarrow \underline{k}$

$$W_{i\rightarrow f} \sim \frac{2\pi}{\hbar L^6}\,|\mathcal{V}(|\underline{k}-\underline{k}_0|)|^2\,\delta\!\left(\frac{\hbar^2k^2}{2m} - \frac{\hbar^2k_0^2}{2m}\right) . \tag{8.132}$$

To compute now the probability per second of scattering into a particular solid angle $d\underline{\Omega}$, we integrate equation 8.132 over the density of final states $\underline{k}$ in the region included in the solid angle $d\underline{\Omega}$. The density of final states, labeled by $\underline{k}$, is given by (see section 4.3.3)

$$\rho(\underline{k})d^3\underline{k} = L^3\,\frac{d^3\underline{k}}{(2\pi)^3} = \frac{L^3}{(2\pi)^3}\,k^2\,dk\,d\underline{\Omega}\ \ , \tag{8.133}$$

[#]1. R.P. Feynman, "Advanced Quantum Mechanics" (Ph205 lecture notes, Oct. 1966).

2. See also Merzbacher, Quantum Mechanics (2nd Ed.), section 19.4.

so that

$$W_{i\to f} \sim \frac{1}{(2\pi)^2 \hbar L^3} |\mathcal{V}(\underline{k} - \underline{k}_0)|^2 \int k^2 \delta\left(\frac{\hbar^2 k^2}{2m} - \frac{\hbar^2 k_0^2}{2m}\right) dk \cdot d\underline{\Omega}$$

or#

$$W_{i\to f} \sim \frac{|\mathcal{V}(\underline{k} - \underline{k}_0)|^2}{(2\pi)^2 \hbar^3 L^3} k_0 m \, d\underline{\Omega} \quad . \tag{8.134}$$

To express this result independently of the normalization volume we use the notion of a <u>cross-section</u> denoted by dσ. This is defined as the area in the incident beam through which the flux of particles (particles per second) is equal to the number of particles per second scattered into $d\underline{\Omega}$.

If

$$v \equiv \frac{\partial E}{\partial p} = \hbar \frac{k_0}{m} \tag{8.135}$$

is the speed of the particles (note that $|\underline{k}| = |\underline{k}_0|$ as required by the delta function), we then have that in the time interval δt

$$W_{\to d\underline{\Omega}} \, \delta t = \frac{d\sigma(\underline{\Omega}) \cdot v \cdot \delta t}{L^3} \tag{8.136}$$

is the probability that a particle will be scattered into a solid angle $d\underline{\Omega}$ along the direction (θ,φ). Equation 8.136 assumes that the particles in the initial state $|\underline{k}_0\rangle$ are uniformly distributed in the volume L^3, an assumption consistent with a state of <u>definite momentum</u> (see paragraph in localization, section 1.3) and the assumed wavefunctions as given by equation 8.130.

Substituting 8.135 into 8.136 and dividing by δt we then have

$$W_{\to d\underline{\Omega}} = \frac{1}{L^3} d\sigma(\underline{\Omega}) \frac{\hbar k_0}{m} \tag{8.137}$$

which we can equate with 8.134 to obtain

$$\frac{d\sigma}{d\Omega} = \frac{m^2}{(2\pi)^2 \hbar^4} |\mathcal{V}(\underline{k} - \underline{k}_0)|^2 \quad , \tag{8.138}$$

independently of the normalization volume.

Note that the scattering described above must be elastic, i.e.

$$(\text{kinetic energy})_{\text{in}} = (\text{kinetic energy})_{\text{out}} \quad ,$$

since the scattering potential was assumed to be a fixed function of space, with no internal degrees of freedom that might have allowed inelastic scattering through an energy exchange between the system responsible for the potential and the incident particles.

$\int f(k)\, \delta[g(k)]\, dk = f(k_0)/|g'(k_0)| \quad ; \quad g(k_0) = 0 \quad .$ See Appendix D.

Example 8.7 Coulomb scattering.

The scattering potential in this case is given by

$$V(r) = \pm \frac{Z_1 Z_2 e^2}{4\pi\epsilon_0 r} \quad . \tag{8.139}$$

Strictly speaking, the function $1/r$ does not possess a Fourier transform since the integrals diverge at infinity. In a practical situation, however, the Coulomb potential of, say, a scattering nucleus is screened by some negative charge which limits the range. In any event, we will calculate the Fourier transform of

$$f(r) = \frac{e^{-\kappa r}}{r}$$

and then take the limit as $\kappa \to 0^+$. We then have

$$F(k) = \int f(r)\, e^{-i\underline{k}\cdot\underline{r}}\, d^3\underline{r}$$

$$= 2\pi \int_0^\infty r^2 f(r) \int_0^\pi e^{-ikr\cos\theta} \sin\theta \, d\theta \cdot dr \ ,$$

or

$$F(k) = \frac{4\pi}{k} \int_0^\infty r\, f(r) \sin kr \, dr \quad . \tag{8.140}$$

Substituting $f(r) = e^{-\kappa r}/r$ we then have

$$F(k) = \frac{4\pi}{k} \int_0^\infty e^{-\kappa r} \sin kr \, dr = \frac{4\pi}{k^2 + \kappa^2} \quad . \tag{8.141}$$

Taking the limit as $\kappa \to 0^+$, we then have

$$\mathcal{V}(|\underline{k} - \underline{k}_0|) = \pm \left(\frac{Z_1 Z_2 e^2}{4\pi\epsilon_0}\right) \frac{4\pi}{|\underline{k} - \underline{k}_0|^2} \tag{8.142}$$

or since $|\underline{k}| = |\underline{k}_0|$, we have

$$|\underline{k} - \underline{k}_0|^2 = k^2 + k_0^2 - 2kk_0 \cos\theta = 2k_0^2(1 - \cos\theta) = 4k^2 \sin^2\theta/2 \ ,$$

where θ is the angle subtended between $\underline{k}$ and $\underline{k}_0$ (scattering angle), and therefore

$$\mathcal{V}(k_0, \theta) = \pm \left(\frac{Z_1 Z_2 e^2}{4\pi\epsilon_0}\right) \cdot \frac{\pi}{k^2 \sin^2\frac{\theta}{2}} \quad . \tag{8.143}$$

Substituting into the differential scattering cross-section (equation 8.138), then yields

$$\frac{d\sigma}{d\Omega} = \frac{1}{4} \left[\frac{Z_1 Z_2 e^2}{(4\pi\epsilon_0) 2E_0}\right]^2 \sin^{-4}\frac{\theta}{2} \ , \tag{8.144a}$$

where

$$E_0 = \frac{1}{2m} p_0^2 = \frac{\hbar^2}{2m} k_0^2 \tag{8.144b}$$

is the kinetic energy of the scattering particle. This is a great coincidence. The first order result for the scattering cross-section for a Coulomb potential (Born approximation) not only agrees with the famous Rutheford formula, derived classically for α-particle scattering by atomic nuclei (see for example H. Goldstein, Classical Mechanics, section 3-7), but also with the exact quantum mechanical treatment (see for example Merzbacher, Quantum Mechanics, 2nd ed., section 11.8).

8.5 Transitions and the superposition of pure energy states.

We have seen that the presence of an perturbation H' has several effects on a state $|n\rangle_0$ of the original Hamiltonian. In particular,

(i) it causes a shift of the unperturbed energy $E_n^{(o)}$, by an amount

$$\Delta E_n = E_n - E_n^{(o)} \,, \tag{8.145}$$

which we can compute to second order, for example, with the aid of equation 8.30.

(ii) it results in a mixture (superposition) of the original eigenstates $\{|m\rangle_0\}$, required to represent the eigenstate corresponding to the new eigenvalue E_n, i.e.

$$|n\rangle = \sum_m |m\rangle_0 \, {}_0\langle m|n\rangle = \frac{1}{N_n^{1/2}} [|n\rangle_0 - \sum_{m\neq n} |m\rangle_0 \, c_{mn}] \,, \tag{8.146}$$

where $N_n^{1/2}$ is some appropriate normalization factor (see, for example, equation 8.61 and related discussion), and the coefficients c_{mn} are given by (to first order in the perturbation) equation 8.28.

and finally, as we have seen,

(iii) it induces transitions in and out of the state $|n\rangle$, whose rate can be determined, to lowest order in the perturbation, for a harmonic $H'(t)$, i.e.

update

$$H'(t) = V^\dagger e^{i\omega_0 t} + V e^{-i\omega_0 t} \tag{8.147a}$$

by the aid of the equation

$$W_{n\to m} \sim \frac{2\pi}{\hbar} |V_{mn}|^2 \, \delta(E_m - E_n \pm \hbar\omega_0) \,, \tag{8.147b}$$

and for a perturbation H' which is not an explicit function of time, by

$$W_{n\to m} \sim \frac{2\pi}{\hbar} |H'_{mn}|^2 \delta(E_m - E_n) \quad . \tag{8.148}$$

See equations 8.127 and 8.129 and related discussion.

Now, both the original set of unperturbed eigenstates $\{|n\rangle_0\}$, and the set of perturbed states $\{|n\rangle\}$ form a complete set, and it is therefore possible to expand any state $|\psi(t)\rangle$ as a superposition of projections along the unperturbed set $\{|n\rangle_0\}$, or the shifted set $\{|n\rangle\}$, i.e.

$$|\psi(t)\rangle = \sum_n |n\rangle_0 \, {}_0\langle n|\psi(t)\rangle = \sum_n |n\rangle \langle n|\psi(t)\rangle \quad . \tag{8.149}$$

We should realize, however, that in the presence of the perturbation, the amplitudes

$$\langle n|\psi(t)\rangle = c_n(t)\, e^{-iE_n t/\hbar} \tag{8.150}$$

do not represent projections along true eigenstates of the system. The states $\{|n\rangle\}$,[#] are no longer stationary states, since transitions in and out of these states are now possible in the presence of the perturbation. Consequently, the modulus squared of the probability amplitudes $\langle n|\psi(t)\rangle$, representing the probability of detecting the system in the state $|n\rangle$, may now be functions of time[#], i.e.

$$|\langle n|\psi(t)\rangle|^2 = |c_n(t)|^2 = \mathcal{P}_n(t) \quad .$$

This probability may be computed, at least in principle, using the transition amplitude matrix $\gamma_{nm}(t, t_i)$, since from equation 8.101 we have

$$c_n(t) = \sum_m \gamma_{nm}(t, t_i)\, c_m(t_i) \quad ,$$

and therefore

$$|c_n(t)|^2 = \sum_{m,m'} \gamma^*_{nm'}(t, t_i)\, \gamma_{nm}(t, t_i)\, c^*_{m'}(t_i)\, c_{m'}(t_i) \quad . \tag{8.151}$$

In practice, unfortunately, such sums are usually intractable and we again have to resort to methods of approximation.

We can use the transition rate result of equation 8.148 (or 8.147 for a harmonic perturbation), to compute the probability that the system will make a transition from the state $|n\rangle$ to the state $|m\rangle$ in a small time δt. In particular

---

[#] or the unperturbed states $\{|n\rangle_0\}$

[†] Compare with equation 3.83 and related discussion.

$$\delta \mathcal{P}_{n\to m} = \dot{\mathcal{P}}_{n\to m}\, \delta t = |c_n(t)|^2 \; W_{n\to m}\, \delta t \quad ,$$

which is the product of the probability that the system is found <u>in</u> $|n\rangle$ at the time t, times the probability that it undergoes the $(n\to m)$ transition in the time interval $(t, t+\delta t)$. Cancelling the δt's we then have

$$\dot{\mathcal{P}}_{n\to m} = |c_n(t)|^2 \; W_{n\to m} \quad . \tag{8.152}$$

This process results in a <u>loss</u> to $|c_n(t)|^2$, representing a transition probability <u>out</u> of the state $|n\rangle$ (in a time δt), and a <u>gain</u> to $|c_m(t)|^2$, representing the transition probability <u>to</u> the state $|m\rangle$. On balance, we then have

$$\frac{d}{dt}\, |c_n(t)|^2 = \sum_{m\neq n} (\dot{\mathcal{P}}_{m\to n} - \dot{\mathcal{P}}_{n\to m}) \quad , \tag{8.153}$$

or, substituting 8.152 for the transition probabilities,

$$\frac{d}{dt}\, |c_n(t)|^2 = -\,(\sum_{m\neq n} W_{n\to m})\, |c_n(t)|^2 + \sum_{m\neq n} W_{m\to n}\, |c_m(t)|^2 \quad ,$$

which can be written as

$$\frac{d}{dt}\, |c_n(t)|^2 + \frac{1}{\tau_n}\, |c_n(t)|^2 = \sum_{m\neq n} W_{m\to n}\, |c_m(t)|^2 \tag{8.154}$$

where the reciprocal of the characteristic time τ_n is defined by

$$\frac{1}{\tau_n} = \sum_{m\neq n} W_{n\to m} \quad . \tag{8.155}$$

Consider now a situation in which the system is in some other state for $t < 0$, and as a result of the perturbation H', or some other external influence, it is scattered into the state $|n\rangle$ at $t = 0$, i.e.

$$c_n(0^+) = 1 \quad , \qquad c_m(0^+) = 0 \qquad \text{for} \quad m \neq n \quad . \tag{8.156}$$

If we substitute these initial conditions in equation 8.154, we have for time $0 < t \leqslant \tau_n$ that

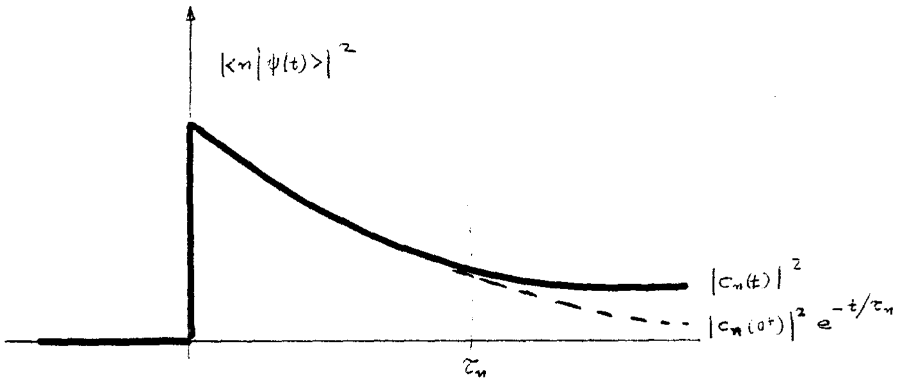
$$\frac{d}{dt}\, |c_n(t)|^2 + \frac{1}{\tau_n}\, |c_n(t)|^2 \simeq 0 \quad , \tag{8.157}$$

whose solution is a decaying time exponential, i.e.

$$|c_n(t)|^2 \simeq |c_n(0^+)|^2 \; e^{-t/\tau_n} \quad . \tag{8.158}$$

For longer times, this expression underestimates the modulus of the probability amplitude, since, as the other states begin to fill up, they will eventually start feeding back to the state $|n\rangle$, slowing down the rate of decrease of the probability $|c_n(t)|^2$, for the system to remain in the state $|n\rangle$.

update



Equation 8.158, along with the initial conditions, define the probability amplitude within a phase factor, i.e.

$$|c_n(t)| = |c_n(0^+)|\, e^{-t/2\tau_n} = e^{-t/2\tau_n} , \tag{8.159}$$

and therefore, within a constant phase factor, (see equations 8.99 and 8.149)

$$\langle n|\psi(t)\rangle \simeq \begin{cases} 0 \quad , & \text{for} \quad t < 0 \\ e^{-t/2\tau_n}\, e^{-iE_nt/\hbar} \quad , & \text{for} \quad t \geq 0 \end{cases} \tag{8.160}$$

where E_n can be computed with the aid of equation 8.30 (or 8.96 for degenerate unperturbed states), for a calculation correct to second order in the perturbation.

It can be seen that this probability amplitude does not correspond to a pure energy state, being the superposition of a continuum of energies (frequencies), i.e.

$$\langle n|\psi(t)\rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} f_n(\omega)\, e^{-i\omega t}\, d\omega \quad , \tag{8.161}$$

where the contribution $f_n(\omega)$ from the frequency ω is given by the inverse transform, i.e.

$$f_n(\omega) = \int_{-\infty}^{\infty} \langle\psi_n|\psi(t)\rangle\, e^{i\omega t}\, dt \quad ,$$

or, substituting from 8.160,

$$f_n(\omega) = \int_0^{\infty} e^{-[\frac{1}{2\tau_n} - i(\omega - \omega_n)]t}\, dt = \frac{1}{\frac{1}{2\tau_n} - i(\omega - \omega_n)} \quad . \tag{8.162}$$

If we substitute equation 8.161 in the original expansion of $|\psi(t)\rangle$ into projections along the states $|n\rangle$ (equation 8.149), we obtain

$$|\psi(t)\rangle = \sum_n |n\rangle\langle n|\psi(t)\rangle = \sum_n |n\rangle \frac{1}{2\pi} \int_{-\infty}^{\infty} f_n(\omega)\, e^{-i\omega t}\, d\omega \quad ,$$

or, interchanging the summation and integration, we have

$$|\psi(t)\rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \, \{\sum_n |n\rangle \, f_n(\omega)\} \, e^{-i\omega t} \quad . \tag{8.163}$$

This allows us to interpret $f_n(\omega)$ as the probability amplitude, of the component of $|\psi(t)\rangle$ along the state $|n\rangle$, at the frequency ω. Conversly, if $G_n(\omega - \omega_n)$ is the probability density, for the component of $|\psi(t)\rangle$ along $|n\rangle$ to have a definite frequency ω (energy $E = \hbar\omega$), we have

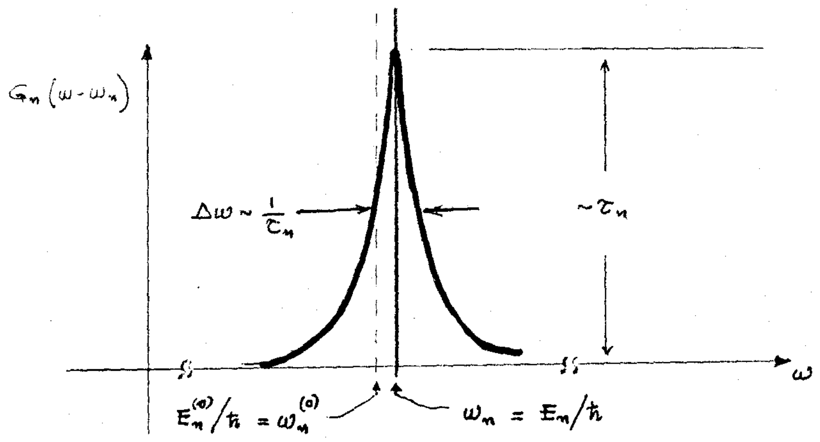
$$G_n(\omega - \omega_n) = A_n |f_n(\omega)|^2 \quad ,$$

where A_n is a normalization factor determined by the requirement that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} G_n(\omega - \omega_n) d\omega = 1 \quad ,$$

which leads to

$$\frac{1}{2\pi} G_n(\omega - \omega_n) = \frac{1}{\pi} \frac{(1/2\tau_n)}{(\omega - \omega_n)^2 + (1/2\tau_n)^2} \quad . \tag{8.164}$$



This curve is known as a Lorentzian and is seen to have a frequency width $\Delta\omega_n \sim 1/\tau_n$, corresponding to an energy width $\Delta E_n \sim \hbar/\tau_n$. This last relation is of course consistent with the energy-time uncertainty principle (see equation 2.127 and related discussion), i.e.

$$\tau_n \Delta E_n \sim \hbar \quad , \tag{8.165}$$

which tells us that the finite lifetime of the state does not allow a better definition of its energy than allowed by the uncertainty principle.

Note also that in the limit of infinite lifetimes the probability density becomes a Dirac delta function, i.e.

$$\lim_{\tau_n \to \infty} \{\frac{1}{2\pi} G_n(\omega - \omega_n)\} = \delta(\omega - \omega_n) \tag{8.166}$$

and we recover the familiar stationary state single frequency (unique energy) behavior.

Before leaving this section, we should mention that the approximation of ignoring the feedback to the state $|n\rangle$ from other states, i.e. ignoring the right hand side of equation 8.154 to obtain 8.157, is quite often a very good one. Notable examples are phenomena like radioactive decay, or spontaneous photon emission (in an infinite cavity) where because of the very large density of final states $|m\rangle$, the probability $|c_m(t)|^2$ to be in any one is very small. Alternatively, when the uranium nucleus emits an α-particle, it is not very likely ... its going to get it back, and we find that in practice no corrections are necessary to the well known exponential radioactive decay formula. Similarly, when an excited atom spontaneously emits a photon in empty space, which ... is not about to turn around to be re-absorbed, we need not worry about the reverse process for a long time and can safely use equation 8.157 instead of 8.154. On the other hand, we should also point out that it is the reverse processes, in each case, that allow the system to reach equilibrium, at which time

$$\frac{d}{dt}|c_n(t)|^2 = -\frac{1}{\tau_n}|c_n|^2 + \sum_{m\neq n} W_{m\to n}|c_m(t)|^2 = 0 \quad . \tag{8.167}$$

8.6 An improved transition rate calculation.

We are now in a position to address an unanswered question implicit in our previous derivation of the transition rate (equations 8.147, 8.148). The Dirac delta function, multiplying the transition rate was justified in anticipation of an integration over some density of states. The problem remained, however, of evaluating the rate for transitions between discrete states not describable in terms of a density of states in an obvious manner. This difficulty, we now recognize, was an artifact of having ignored the distribution around the energy "eigenvalue" (spectrum) of the discrete state, itself a consequence of the transitions. An improved calculation of the transition rate can now be done as follows.

The system is initially in the state $|n\rangle$, described by a spectrum of pure energy states given by

$$G_n(\varepsilon' - E_n)d\varepsilon' = \frac{1}{\pi}\frac{(\hbar/2\tau_n)\, d\varepsilon'}{(\varepsilon'-E_n)^2 + (\hbar/2\tau_n)^2} \quad , \tag{8.168a}$$

and makes a transition to a final state $|m\rangle$, described by a corresponding spectrum of pure energy states, given by

$$G_m(\varepsilon'' - E_m)d\varepsilon'' = \frac{1}{\pi}\frac{(\hbar/2\tau_m)\, d\varepsilon''}{(\varepsilon''-E_m)^2 + (\hbar/2\tau_m)^2} \quad . \tag{8.168b}$$

The transition rate, as was computed previously, is to be understood as a transition from ε' to ε'', **i.e.**

$$W_{\varepsilon' \to \varepsilon''} \simeq \frac{2\pi}{\hbar} |H'_{mn}|^2 \, \delta(\varepsilon' - \varepsilon'') \quad , \tag{8.169}$$

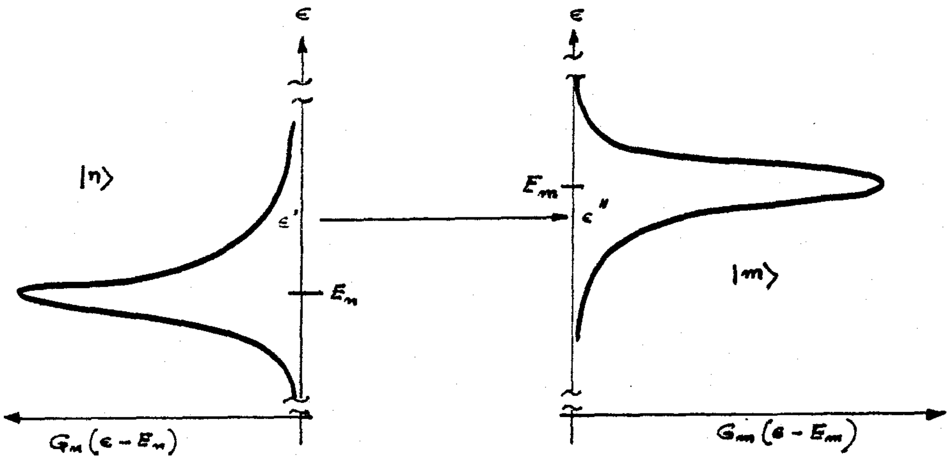
which must now be integrated over the range of ε' and ε'' in the initial and final states

$$W_{n\to m} = \iint_{-\infty}^{\infty} G_n(\varepsilon' - E_n) W_{\varepsilon' \to \varepsilon''} \, G_m(\varepsilon'' - E_m) \, d\varepsilon'' d\varepsilon'$$

or

$$W_{n\to m} = \frac{2\pi}{\hbar} |H'_{mn}|^2 \iint_{-\infty}^{\infty} G_n(\varepsilon' - E_n) \, \delta(\varepsilon' - \varepsilon'') \, G_m(\varepsilon'' - E_m) \, d\varepsilon'' d\varepsilon' \quad .$$

This situation is sketched below.



If we now perform the integration over the delta function, we obtain for the total transition rate

$$W_{n\to m} = \frac{2\pi}{\hbar} |H'_{mn}|^2 \int_{-\infty}^{\infty} G_n(\varepsilon' - E_n) \, G_m(\varepsilon' - E_m) \, d\varepsilon' \quad . \tag{8.170}$$

Substituting from equations 8.168 for the energy distributions, we obtain (see problem 8.18)

$$W_{n\to m} = \frac{2\pi}{\hbar} |H'_{mn}|^2 \frac{(\hbar/2\pi\tau_{mn})}{(E_m - E_n)^2 + (\hbar/2\tau_{mn})^2} \tag{8.171a}$$

where $1/\tau_{mn}$ is equal to the sum of the two transition rates, i.e.

$$\frac{1}{\tau_{mn}} \equiv \frac{1}{\tau_{mn}} + \frac{1}{\tau_{mn}} \quad . \tag{8.171b}$$

Note that

$$\lim_{\varepsilon\to 0} \{\frac{1}{\pi} (\frac{\varepsilon}{x^2 + \varepsilon^2})\} = \delta(x) \quad , \tag{8.172}$$

and we recover the previous result in the limit of $\tau_{mn} \to \infty$.

From equation 8.171 we see that transitions between two states of "energy" E_n and E_m are possible even though E_m is not exactly equal to E_n. No serious degradation of the transition rate would be observed for energy differences $\Delta E_{mn} = E_m - E_n$ such that

$$|\Delta E_{mn}| \leq \frac{\hbar}{2\tau_{mn}} \quad .$$

We recognize, of course, that this does not represent a violation of the conservation of energy, but rather an expression of the fact that the energy of the initial and final states is only defined within a interval

$$\Delta E_n \sim \hbar/2\tau_n \quad , \qquad \Delta E_m \sim \hbar/2\tau_m \quad ,$$

in accordance with the energy-time uncertainty relation.

Note that for a harmonic perturbation $H' = V' e^{i\omega t} + V e^{-i\omega t}$, the transition rate should be modified to

$$W_{n\to m} = \frac{2\pi}{\hbar} |H'_{mn}|^2 \frac{(\hbar/2\pi\tau_{mn})^2}{(E_m - E_n \pm \hbar\omega)^2 + (\hbar/2\tau_{mn})^2} \quad . \tag{8.173}$$

Problems

8.1 (i) Prove that the perturbed eigenstates, correct to first order in the perturbation H' (equation 8.24), form an orthonormal set, to second order in the perturbation parameter, i.e.

$$\langle k'|k\rangle = \delta_{k'k} + 0(\lambda_1{}^2)$$

(ii) Some times, it is necessary to have the eigenstates exactly normalized to all orders, i.e.

$$\langle k|k\rangle = 1 \quad \text{for all } k \ .$$

Normalize the eigenstates, as given by equation 6.24, so that this is the case.

8.2 Prove that the second order correction to the energy of the ground state is always negative.

8.3 Derive the correction to the energies of the $\ell=0$ eigenstates of a diatomic molecule, whose potential is given by

$$V(r) = V_0 + \frac{1}{2} m_r \omega_0{}^2 (r - r_0)^2 + \frac{1}{6} \gamma (r - r_0)^3 \ ,$$

using perturbation theory. Derive an appropriate expression for the perturbation parameter λ. Compare your result with the exact solution for the Morse potential fit to the interatomic potential.

8.4 Compute the energies to second order of an isotropic harmonic oscillator in a uniform electric field. Compare the ground state energy correction with the classical value.

Hint: Show that classically

$$\delta E = -\underline{\mathcal{E}} \cdot \delta \underline{d}$$

so that

$$\Delta E = -\frac{1}{2} \underline{\mathcal{E}} \cdot \underline{d} = -\frac{1}{2} \varepsilon_0 \, a \, \mathcal{E}^2 \ .$$

8.5 Compute the polarizability tensor for a charge bound anisotropically by a by a harmonic potential

$$V = \frac{1}{2} m (\omega_x^2 x^2 + \omega_y^2 y^2 + \omega_z^2 z^2) \ .$$

8.6 The matrix elements of an operator $\hat{Q}$ in the basis of the eigenstates of H_0, are given by ${}_0\langle m|\hat{Q}|n\rangle_0$. Compute the matrix elements $\langle m|\hat{Q}|n\rangle$, correct to first order in the perturbation H', if the states $|n\rangle$ are the eigenstates of $H = H_0 + H'$.

Answer:

$$\langle m|Q|n\rangle = Q_{mn} - \sum_{k'\neq m} \frac{H'_{mk'}\,Q_{k'n}}{E^{(0)}_{k'} - E^{(0)}_{m}} - \sum_{k\neq n} \frac{Q_{mk}\,H'_{kn}}{E^{(0)}_{k} - E^{(0)}_{n}}$$

where

$$Q_{ij} = {}_0\langle i|Q|j\rangle_0 \quad , \quad H'_{ij} = {}_0\langle i|H'|j\rangle_0$$

8.7 Show that the dipole moment of a bound charge q_c in a non-degenerate state $|k\rangle_0$ exposed to a uniform electric field $\underline{\mathcal{E}}$, is given by (to first order in the perturbation)

$$\underline{d}_k = \underline{d}_k^{(0)} + q_c^2 \sum_{n\neq k} \left[\frac{\underline{x}_{kn}\underline{x}_{nk} + \underline{x}_{nk}\underline{x}_{kn}}{E^{(0)}_n - E^{(0)}_k} \right] \cdot \underline{\mathcal{E}} \quad ,$$

where $\underline{d}_k^{(0)}$, the permanent dipole moment of the unperturbed state $|k\rangle_0$, is given by

$$\underline{d}_k^{(0)} \equiv q_c\; {}_0\langle k|\underline{x}|k\rangle_0 \quad ,$$

and $\underline{x}_{nm}$ is the matrix element

$${}_0\langle n|\underline{x}|m\rangle_0 \quad .$$

Note that the polarizability of the system found in the state k is a tensor, i.e.

$$\alpha_k \equiv \frac{q_c^2}{\epsilon_0} \sum_{n\neq k} \left[\frac{\underline{x}_{kn}\underline{x}_{nk} + \underline{x}_{nk}\underline{x}_{kn}}{E^{(0)}_n - E^{(0)}_k} \right] .$$

8.8 Hellmann-Feynman theorem. Show that if the Hamiltonian depends on a parameter λ, i.e.

$$H = H(\lambda)$$

and $E_k(\lambda_1)$, $|k;\lambda_1\rangle$ are the eigenvalues and eigenstates corresponding to a particular value of $\lambda = \lambda_1$, then

$$\frac{dE_k(\lambda_1)}{d\lambda} = \langle k;\lambda_1|\frac{\partial H(\lambda_1)}{\partial\lambda}|k;\lambda_1\rangle \quad .$$

8.9 Show that if $H = H(\lambda)$, then

$$\frac{d}{d\lambda}|k;\lambda_1\rangle = -\sum_{n\neq k} |n;\lambda_1\rangle \frac{\langle n;\lambda_1|\frac{\partial H(\lambda_1)}{\partial\lambda}|k;\lambda_1\rangle}{\langle n;\lambda_1|H(\lambda_1)|n;\lambda_1\rangle - \langle k;\lambda_1|H(\lambda_1)|k;\lambda_1\rangle} \quad .$$

update

The results of problems 8.8 and 8.9 may be used to provide a forward marching scheme to compute the energies and eigenstates for large values of the perturbation parameter, provided the denominators in the expression above do not become zero along the way.

Problems (continued)

8.10 Consider a particle of mass m and charge q_c in a one-dimensional potential (see section 4.2.5)

$$U(z) = \begin{cases} \infty & \text{for} \quad |z| > L/2 \\ 0 & \text{for} \quad |z| < L/2 \end{cases} .$$

(i) Compute the dipole moment of the eigenstates of the system.

(ii) Compute the induced dipole moment of the eigenstates in response to a weak uniform electric field $\underline{\mathscr{E}} = \mathscr{E}_z \hat{e}_z$.

(iii) Compute the energy corrections as a result of the perturbation.

8.11 Show that equation 8.81 is valid for both positive and negative $H_{11} - H_{22}$.

8.12 Compute the effect of a perturbing uniform electric field on the $n = 3$ states of a hydrogen atom.

Variational methods. It is possible to use the fact that the ground state energy is the lowest eigenvalue of the Hamiltonian to obtain an approximate value for it even if the wavefunctions cannot be solved for exactly, i.e.

$$E_{\text{ground}} = \min_{\text{all } \psi} \{\langle\psi|H|\psi\rangle\} \quad , \quad \text{where} \quad \langle\psi|\psi\rangle = 1 \quad .$$

8.13 Estimate the binding energy of the electron in a hydrogen atom using the trial function

$$R(r) = \begin{cases} A(1 - r/a) \; ; & r < a \\ 0 \; ; & r > a \end{cases} ,$$

where a is an adjustable parameter.

8.14 Prove the statement of equation 8.102.

8.15 Show that the operator $\Gamma(t, t_i)$ satisfies the differential equation

$$i\hbar \frac{\partial}{\partial t} \Gamma(t, t_i) = e^{iH_0 t} \, V(t) \, e^{-iH_0 t} \, \Gamma(t, t_i) \quad .$$

8.16 For the problem described in example 8.6, compute the transition probability $\mathscr{P}_{0\to 2}$ to lowest order.

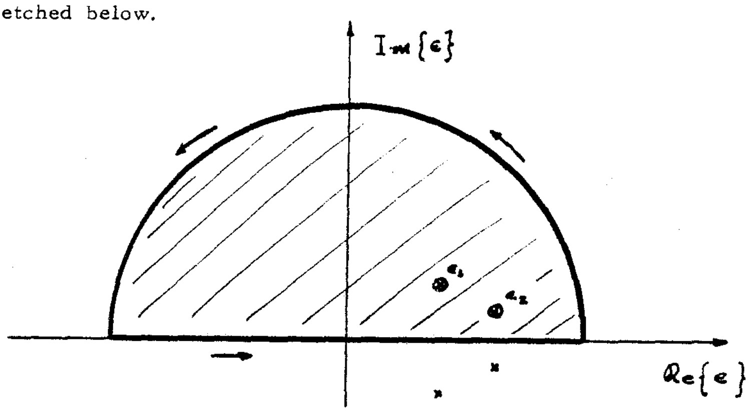
8.17 Derive the corresponding expression for Fermi's Golden Rule (equation 8.127), for the case $\omega \simeq -\omega_{mn}$ (induced de-excitation).

Problems (continued)

8.18 Compute the integral in equation 8.170 using complex integration. Close the contour on the upper half plane, noting that the integrand has two simple poles inside the contour at ε_1 and ε_2, where

$$\varepsilon_1 = E_n + \frac{i\hbar}{2\tau_n} \quad , \quad \varepsilon_2 = E_m + \frac{i\hbar}{2\tau_m} \quad ,$$

as sketched below.

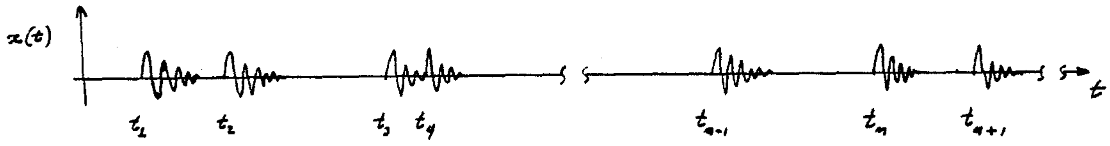


8.19 Derive equation 8.173.

8.20 Find the spectrum[#] of a (classical) process $x(t)$ described by

$$x(t) = a \sum_{n=1}^{\infty} U(t-t_n)\, e^{-(t-t_n)/2\tau_0} \sin\omega_0 (t-t_n)$$

where $U(t)$ is the unit step function and the times $\{t_n\}$, $n=1, \infty$ are random.



[#] The spectrum is defined as the Fourier transform of the <u>autocorrelation function</u>, which is defined in turn, as a function of the time delay τ, as the average value of the product $x(t)\, x(t+\tau)$, i.e.

$$S(\omega) \equiv \int_{-\infty}^{\infty} R(\tau)\, e^{i\omega\tau}\, d\tau \quad , \quad R(\tau) \equiv \langle x(t)\, x(t+\tau) \rangle \quad .$$

Note that $R(\tau) = R(-\tau)$.

update

8.21 Find the spectrum of a (classical) periodic process

$$x(t) = a \sin(\omega_0 t + \delta) \quad ,$$

whose phase δ *is interrupted and changes at random times* $\{t_n\}$, $n=1,\infty$, by a random (large) amount, i.e.

$$x(0 < t \leq t_1) = a \sin(\omega_0 t + \delta_1)$$

$$x(t_1 < t \leq t_2) = a \sin(\omega_0 t + \delta_2)$$

$$\vdots$$

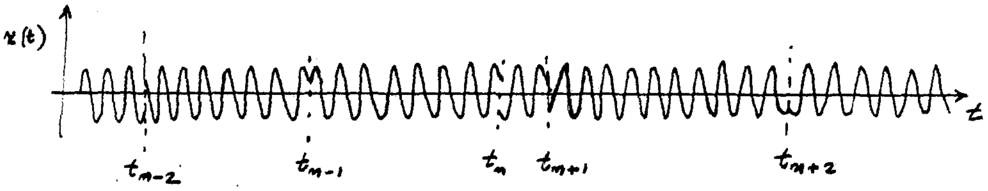
$$x(t_{n-1} < t < t_n) = a \sin(\omega_0 t + \delta_n)$$

$$\vdots$$

etc.

Assume that the time intervals $(t_n - t_{n-1})$ are Poisson distributed with a mean of τ, i.e.

$$\mathcal{P}\{(t_n - t_{n-1}) > t\} = e^{-t/\tau} \quad .$$



9. THE ELECTROMAGNETIC FIELD

9.1 The Hamiltonian

From classical electromagnetism, we have that the energy of an electromagnetic field filling a volume V, containing no charges and currents, is given by (see for example R. P. Feynman, R. B. Leighton and M. Sands, The Feynman Lectures on Physics, v. II, chapter 27).

$$E = \frac{\varepsilon_0}{2} \int [|\underline{\mathcal{E}}(\underline{x}, t)|^2 + c^2 |\underline{\mathcal{B}}(\underline{x}, t)|^2] d^3\underline{x} \tag{9.1}$$

where $\underline{\mathcal{E}}(\underline{x}, t)$ and $\underline{\mathcal{B}}(\underline{x}, t)$ are the electric and magnetic fields respectively and c is the speed of light. We can simplify this expression by the introduction of the scalar and vector potentials, $\varphi(\underline{x}, t)$ and $\underline{A}(\underline{x}, t)$ respectively, in terms of which we have

$$\underline{\mathcal{E}}(\underline{x}, t) = - \frac{\partial}{\partial \underline{x}} \varphi(\underline{x}, t) - \frac{\partial}{\partial t} \underline{A}(\underline{x}, t)$$

and

$$\underline{\mathcal{B}}(\underline{x}, t) = \frac{\partial}{\partial \underline{x}} \times \underline{A}(\underline{x}, t) .$$

(See appendix F). You will note that the fields φ and $\underline{A}$ are not uniquely determined. In fact, the (gauge) transformation

$$\underline{\tilde{A}}(\underline{x}, t) = \underline{A}(\underline{x}, t) + \frac{\partial}{\partial \underline{x}} \chi(\underline{x}, t) \tag{9.2a}$$

$$\tilde{\varphi}(\underline{x}, t) = \varphi(\underline{x}, t) - \frac{\partial}{\partial t} \chi(\underline{x}, t) \tag{9.2b}$$

results in a new scalar and vector field $\tilde{\varphi}$ and $\underline{\tilde{A}}$ which yield the same $\underline{\mathcal{E}}$ and $\underline{\mathcal{B}}$ fields, as can be verified by direct substitution. It will prove convenient in the subsequent discussion, to transform the potentials φ and $\underline{A}$ in such a way as to have a divergence-free vector potential, i. e.,

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{A}(\underline{x}, t) = 0 , \tag{9.3a}$$

as well as zero scalar potential, i. e.,

$$\varphi(\underline{x}, t) = 0 . \tag{9.3b}$$

This can always be done (see Appendix J) and results in a choice of gauge for the electromagnetic potentials that is called the radiation gauge.

In this gauge, the energy of the system can then be expressed in terms of the vector potential alone, i. e.,

$$E = \frac{\varepsilon_0}{2} \int \left\{ |\underline{\dot{A}}(\underline{x}, t)|^2 + c^2 \left|\frac{\partial}{\partial \underline{x}} \times \underline{A}(\underline{x}, t)\right|^2 \right\} d^3\underline{x} . \tag{9.4}$$

We now assume that the vector potential satisfies periodic boundary conditions at the edges of the volume V, i.e.,

$$\underline{A}(-L_x/2, y, z, t) = \underline{A}(L_x/2, y, z, t) \tag{9.5a}$$

$$\underline{A}(x, -L_y/2, z, t) = \underline{A}(x, L_y/2, z, t) \tag{9.5b}$$

$$\underline{A}(x, y, -L_z/2, t) = \underline{A}(x, y, L_z/2, t) \tag{9.5c}$$

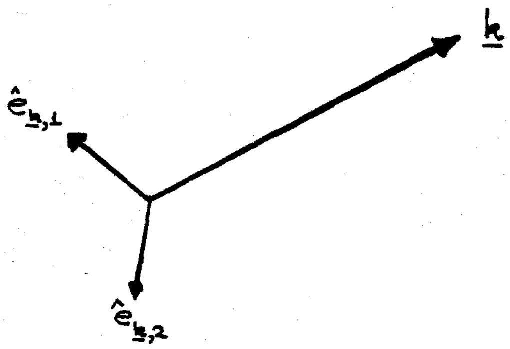
and expand into the Fourier components for each of the admissible wavevectors (see problem F.1), i.e.

$$\underline{A}(\underline{x}, t) = \sum_{\underline{k}, \ell=1,2} A_{\underline{k},\ell}\, e^{i\underline{k}\cdot\underline{x}}\, \hat{e}_{\underline{k},\ell} \tag{9.6}$$

where

$$A_{\underline{k},\ell}(t) = \frac{1}{V}\int [\hat{e}_{\underline{k},\ell}\cdot\underline{A}(\underline{x}, t)]\, e^{-i\underline{k}\cdot\underline{x}}\, d^3\underline{x} \tag{9.7}$$

and where $\hat{e}_{\underline{k},1}$ and $\hat{e}_{\underline{k},2}$ are the two orthogonal unit vectors to the wavevector $\underline{k}$. See figure below.



Note that the expansion of $\underline{A}(\underline{x}, t)$, as given by equation 9.6, is divergence free.

We now compute

$$\dot{\underline{A}}(\underline{x}, t) = \sum_{\underline{k},\ell} \dot{A}_{\underline{k},\ell}\, e^{i\underline{k}\cdot\underline{x}}\, \hat{e}_{\underline{k},\ell} \tag{9.8}$$

and

update

$$\frac{\partial}{\partial\underline{x}} \times \underline{A}(\underline{x}, t) = i\sum_{\underline{k},\ell} A_{\underline{k},\ell}\, e^{i\underline{k}\cdot\underline{x}}\, (\underline{k} \times \hat{e}_{\underline{k},\ell}), \tag{9.9}$$

where since

$$\hat{e}_{\underline{k},1} \times \hat{e}_{\underline{k},2} = \hat{e}_{\underline{k}} \tag{9.10a}$$

we have

$$\underline{k} \times \hat{e}_{\underline{k},1} = k\,\hat{e}_{\underline{k},2} \tag{9.10b}$$

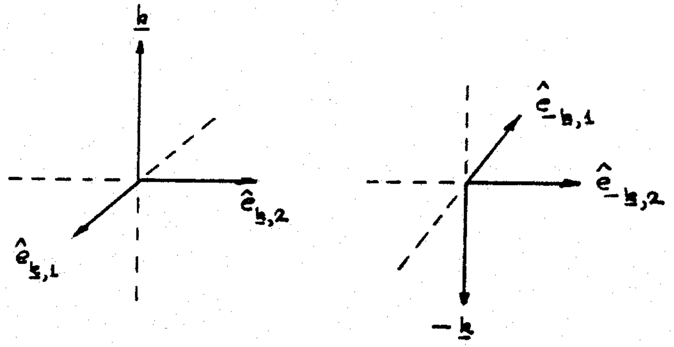
and

$$\underline{k} \times \hat{e}_{\underline{k},2} = -k\,\hat{e}_{\underline{k},1} \tag{9.10c}$$

where $k = |\underline{k}|$ and $\hat{e}_{\underline{k}} = \underline{k}/k$. We also adopt the convention

$$\hat{e}_{\underline{k},\ell} = (-1)^{\ell}\,\hat{e}_{-\underline{k},\ell}\,, \tag{9.11}$$

so as to keep both $(\hat{e}_{\underline{k}}, \hat{e}_{\underline{k},1}, \hat{e}_{\underline{k},2})$ and $(\hat{e}_{-\underline{k}}, \hat{e}_{-\underline{k},1}, \hat{e}_{-\underline{k},2})$ right handed. See figure below.



Substituting 9.8 and 9.9 into 9.4 we then have

$$E = \frac{1}{2}(\epsilon_0 V) \sum_{\underline{k},\ell} \left(|\dot{A}_{\underline{k},\ell}|^2 + c^2 k^2 |A_{\underline{k},\ell}|^2 \right). \tag{9.12}$$

We now decompose the vector potential Fourier components ,

$$A_{\underline{k},\ell} = \frac{1}{\sqrt{2}} \left(Q_{\underline{k},2\ell-1} - i\,Q_{\underline{k},2\ell} \right), \tag{9.13a}$$

where

$$Q_{\underline{k},2\ell-1} = \frac{2^{1/2}}{V} \int [\hat{e}_{\underline{k},\ell} \cdot \underline{A}(\underline{x},t)] \cos(\underline{k}\cdot\underline{x})\, d^3\underline{x} \tag{9.13b}$$

and

$$Q_{\underline{k},2\ell} = \frac{2^{1/2}}{V} \int [\hat{e}_{\underline{k},\ell} \cdot \underline{A}(\underline{x},t)] \sin(\underline{k}\cdot\underline{x})\, d^3\underline{x}\ . \tag{9.13c}$$

Note that, as a consequence of the right handed coordinate convention

$$A_{\underline{k},\ell} = (-1)^{\ell} A^{*}_{-\underline{k},\ell}\,, \tag{9.14}$$

and also

$$Q_{\underline{k},2\ell-1} = (-1)^{\ell} Q_{-\underline{k},2\ell-1}, \qquad Q_{\underline{k},2\ell} = (-1)^{\ell-1} Q_{-\underline{k},2\ell}. \tag{9.15}$$

Note also that, as in the case of the liquid (see section 5.6), not all $Q_{\underline{k},j}$ are independent and we must restrict the summation over some half-space of $\underline{k}$, i.e.

$$E = \frac{1}{2}(\varepsilon_0 V) \sum_{+\underline{k},\, j=1,4} \left(\dot{Q}^2_{\underline{k},j} + \omega^2_{\underline{k}} Q^2_{\underline{k},j} \right) \tag{9.16}$$

where

$$\omega^2_{\underline{k}} = c^2 k^2 \tag{9.17}$$

and where $+\underline{k}$ denotes that the summation runs over some selected half space of $\underline{k}$.

In an analogous manner, we can now define conjugate momenta

$$\Pi_{\underline{k},j} = (\varepsilon_0 V) \dot{Q}_{\underline{k},j}, \tag{9.18}$$

to each coordinate $Q_{\underline{k},j}$,

which we can use to write the energy as the classical Hamiltonian of the system in terms of its proper coordinates, i.e.

$$H = \sum_{+\underline{k},\, j} \left[\frac{1}{2M} \Pi^2_{\underline{k},j} + \frac{1}{2} M\omega^2_{\underline{k}} Q^2_{\underline{k},j}\right], \tag{9.19}$$

where

$$M \equiv \varepsilon_0 V . \tag{9.20}$$

We see that the system possesses a Hamiltonian corresponding to a system of uncoupled harmonic oscillators of "mass" $M = \varepsilon_0 V$ and frequency $\omega_{\underline{k}} = c|\underline{k}|$.

We can see that the canonical equations derived from this Hamiltonian yield the classical equations of motion, which in this case reduce to the (source-free) wave equation for the vector potential[#], i.e.

$$-\left(\frac{\partial}{\partial \underline{x}} \cdot \frac{\partial}{\partial \underline{x}}\right) \underline{A} + \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \underline{A} = 0 . \tag{9.21}$$

See appendix F, equations F.15 and F.16 and related discussion. If we express the vector potential in terms of its Fourier expansion (equation 9.6), we have

update

[#] no charges and currents, gauge $\varphi = 0$, $\frac{\partial}{\partial \underline{x}} \cdot \underline{A} = 0$.

$$\sum_{\underline{k},\ell} \left(+\, k^2 A_{\underline{k},\ell} + \frac{1}{c^2} \ddot{A}_{\underline{k},\ell}\right) e^{i\underline{k}\cdot\underline{x}}\, \hat{e}_{\underline{k},\ell} = 0 \quad , \tag{9.22}$$

or since this must hold for every $\underline{x}$, we have

$$k^2 A_{\underline{k},\ell} + \frac{1}{c^2} \ddot{A}_{\underline{k},\ell} = 0 \quad , \tag{9.23}$$

for every $\underline{k}$ and $\ell = 1,2$.

Since the coefficients $A_{\underline{k},\ell}$ are complex (see equations 9.13), equation 9.23 must hold for the real and imaginary parts separately, in other words

$$k^2 Q_{\underline{k},j} + \frac{1}{c^2} \ddot{Q}_{\underline{k},j} = 0 \quad , \tag{9.24}$$

for all $\underline{k}$ and $j = 1,2,3,4$. These equations are then Maxwell's equations, in Fourier space, for the vector potential.

Consider now the canonical equations from the Hamiltonian of equation 9.19, or

$$\dot{Q}_{\underline{k},j} = \frac{\partial H}{\partial \Pi_{\underline{k},j}} = \frac{1}{M} \Pi_{\underline{k},j} \tag{9.25a}$$

(consistently with our definition of $P_{\underline{k},j}$), and

$$\dot{\Pi}_{\underline{k},j} = -\frac{\partial H}{\partial Q_{\underline{k},j}} = -M \omega_{\underline{k}}^2 Q_{\underline{k},j} \quad . \tag{9.25b}$$

Differentiating 9.25a with respect to time, we have

$$\dot{\Pi}_{\underline{k},j} = M \ddot{Q}_{\underline{k},j} \quad , \tag{9.26}$$

which substituted in 9.25b yields the required relation (equation 9.24; recall $\omega_{\underline{k}}^2 = c^2 k^2$).

9.2 Transition to Quantum Mechanics.

Armed with the proper Hamiltonian, we now make the transition to Quantum Mechanics in our usual way, namely by selecting a pair of variables (operators) for the conjugate coordinates of the system $(Q_{\underline{k},j}, \Pi_{\underline{k},j})$, such that the standard commutation relations are obeyed. In particular, we must have

$$[Q_{\underline{k}',j'}, Q_{\underline{k},j}] = 0 \tag{9.27a}$$

$$[Q_{\underline{k}',j'}, \Pi_{\underline{k},j}] = i\hbar\, \delta_{\underline{k}'\underline{k}}\, \delta_{j'j} \tag{9.27b}$$

update #2

and

$$[\Pi_{\underline{k}',j'}, \Pi_{\underline{k},j}] = 0 \quad , \tag{9.27c}$$

for $\underline{k}$ in the selected half-space of the admissible wavenumbers and $j = 1,2,3,4$.

The most convenient way of achieving this for a system of harmonic oscillators is, as we have seen, to define ladder operators for each independent oscillator, in particular,

$$b_{\underline{k},j} \equiv \left(\frac{1}{2\hbar}\right)^{1/2} \left[(\epsilon_0 V\omega_{\underline{k}})^{1/2} Q_{\underline{k},j} + i(\epsilon_0 V\omega_{\underline{k}})^{-1/2} \Pi_{\underline{k},j}\right] \tag{9.28a}$$

and

$$b^{\dagger}_{\underline{k},j} \equiv \left(\frac{1}{2\hbar}\right)^{1/2} \left[(\epsilon_0 V\omega_{\underline{k}})^{1/2} Q_{\underline{k},j} - i(\epsilon_0 V\omega_{\underline{k}})^{-1/2} \Pi_{\underline{k},j}\right] \ . \tag{9.28b}$$

See equations 5.46 and related discussion. These operators obey the commutations relations

$$\left.\begin{aligned} &[b_{\underline{k}',j'},\ b_{\underline{k},j}] = 0 \\ &[b_{\underline{k}',j'},\ b^{\dagger}_{\underline{k},j}] = \delta_{\underline{k}',\underline{k}}\ \delta_{j',j} \\ &[b^{\dagger}_{\underline{k}',j'},\ b^{\dagger}_{\underline{k},j}] = 0 \quad , \end{aligned}\right\} \tag{9.29}$$

and can be used to express the Hamiltonian for the system, i.e.

$$H = \sum_{+\underline{k},j} \hbar\omega_{\underline{k}} \left(b^{\dagger}_{\underline{k},j}\, b_{\underline{k},j} + \frac{1}{2}\right) \ , \tag{9.30}$$

which we recognize as the Hamiltonian for a system of uncoupled oscillators expressed in terms of the corresponding ladder operators. See equation 5.48 and related discussion.

The eigenstates of the system can be specified in the occupation (quantum) number representation of the harmonic oscillators for each mode (+ $\underline{k}$, j), i.e.

$$|\underline{m}\rangle = |m_{\underline{k}_1,1},\ \ldots\ldots\ ,\ m_{\underline{k},j},\ \ldots\ldots\ \rangle \ , \tag{9.31}$$

analogously to equation 5.45. The result of the operation on the state $|\underline{m}\rangle$ with the annihilation or creation operation operators $b_{\underline{k},j}$, $b^{\dagger}_{\underline{k},j}$ of one mode (+ $\underline{k}$, j) is then given by as before (equations 5.47),

$$b_{\underline{k},j}|\underline{m}\rangle = m^{1/2}_{\underline{k},j}\ |\ldots\ldots\ ,\ m_{\underline{k},j} - 1\ ,\ \ldots\ldots\ \rangle \tag{9.32a}$$

and

$$b^{\dagger}_{\underline{k},j}|\underline{m}\rangle = (m_{\underline{k},j} + 1)^{1/2}|\ \ldots\ldots\ ,\ m_{\underline{k},j} + 1\ ,\ \ldots\ldots\ \rangle \ . \tag{9.32b}$$

The energy of the eigenstates $|\underline{m}\rangle$ can be computed in the usual manner, namely as the expectation value of the Hamiltonian, or

$$E_{\underline{m}} = \langle\underline{m}|H|\underline{m}\rangle = \sum_{+\underline{k},j} E_{\underline{k},j} \tag{9.33a}$$

where

$$E_{\underline{k},j} = \hbar\omega_{\underline{k}}\left(m_{\underline{k},j} + \frac{1}{2}\right) \ , \tag{9.33b}$$

update

or

$$E_{\underline{m}} = \sum_{+\underline{k},j} \hbar\omega_{\underline{k}}\left(m_{\underline{k},j} + \frac{1}{2}\right) \ . \tag{9.34}$$

9.3 The Quantum Mechanical vector potential operator.

We can invert the defining equations 9.28 to express $Q_{\underline{k},j}$ and $\Pi_{\underline{k},j}$ in terms of the annihilation and creation operators for the corresponding mode. This yields (see also equations 5.27),

$$Q_{\underline{k},j} = \left(\frac{\hbar}{2\varepsilon_0 V\omega_{\underline{k}}}\right)^{1/2} (b^\dagger_{\underline{k},j} + b_{\underline{k},j}) \tag{9.35a}$$

and

$$\Pi_{\underline{k},j} = (\varepsilon_0 V)\,\dot{Q}_{\underline{k},j} = i\left(\frac{\varepsilon_0 V\hbar\omega_{\underline{k}}}{2}\right)^{1/2} (b^\dagger_{\underline{k},j} - b_{\underline{k},j}) \;, \tag{9.35b}$$

where, it should be noted, that the $(b_{\underline{k},j}, b^\dagger_{\underline{k},j})$ are only defined for $\underline{k} > 0$ (in the selected positive half-space). We can use these operator equations to express the vector potential Fourier coordinates $A_{\underline{k},\ell}$ and their time derivatives $\dot{A}_{\underline{k},\ell}$. In particular, from the defining equations 9.12 we have

$$A_{\underline{k},\ell} = \frac{1}{\sqrt{2}} (Q_{\underline{k},2\ell-1} - i\,Q_{\underline{k},2\ell})$$

in which we can substitute 9.35a to obtain for $\underline{k} > 0$

$$A_{\underline{k},\ell} = \left(\frac{\hbar}{2\varepsilon_0 V\omega_{\underline{k}}}\right)^{1/2} \left[\frac{1}{\sqrt{2}} (b^\dagger_{\underline{k},2\ell-1} - i\,b^\dagger_{\underline{k},2\ell}) + \frac{1}{\sqrt{2}} (b_{\underline{k},2\ell-1} - i\,b_{\underline{k},2\ell})\right] \;. \tag{9.36a}$$

We can now use equations 9.15 to extend $A_{\underline{k},\ell}$ to the negative half space of $\underline{k}$. In particular, we have

$$\begin{aligned} A_{-\underline{k},\ell} &= \frac{1}{\sqrt{2}} (Q_{-\underline{k},2\ell-1} - i\,Q_{-\underline{k},2\ell}) \\ &= \frac{(-1)^\ell}{\sqrt{2}} (Q_{\underline{k},2\ell-1} + i\,Q_{\underline{k},2\ell}) \;. \end{aligned}$$

You will note that the right hand side is expressed in terms of the field sine and cosine Fourier components for the positive half space of $\underline{k}$. This allows us to use the ladder operators $b_{\underline{k}j}$, $b^\dagger_{\underline{k},j}$ again, which yield for $\underline{k} > 0$

$$A_{-\underline{k},\ell} = (-1)^\ell \left(\frac{\hbar}{2\varepsilon_0 V\omega_{\underline{k}}}\right)^{1/2} \left[\frac{1}{\sqrt{2}} (b^\dagger_{\underline{k},2\ell-1} + i\,b^\dagger_{\underline{k},2\ell}) + \frac{1}{\sqrt{2}} (b_{\underline{k},2\ell-1} + b_{\underline{k},2\ell})\right] \;. \tag{9.36b}$$

Note that operator equations 9.36a and 9.36b for the Fourier components of the vector potential, satisfy equation 9.14.

update

Equations 9.36 suggest a new set of ladder operators. In particular, for $\underline{k} > 0$,

$$a_{\underline{k},\ell} = \frac{1}{\sqrt{2}} (b_{\underline{k},2\ell-1} - i\,b_{\underline{k},2\ell}) \tag{9.37a}$$

$$a^\dagger_{\underline{k},\ell} = \frac{1}{\sqrt{2}} (b^\dagger_{\underline{k},2\ell-1} + i\,b^\dagger_{\underline{k},2\ell}) \;, \tag{9.38a}$$

(which can be seen to be Hermitean adjoints of each other) and also

$$a_{-\underline{k},\ell} = \frac{(-1)^{\ell}}{\sqrt{2}} \left(b_{\underline{k},2\ell-1} + i\, b_{\underline{k},2\ell}\right) \tag{9.37b}$$

$$a^{\dagger}_{-\underline{k},\ell} = \frac{(-1)^{\ell}}{\sqrt{2}} \left(b^{\dagger}_{\underline{k},2\ell-1} - i\, b^{\dagger}_{\underline{k},2\ell}\right) \ . \tag{9.38b}$$

These expressions allow for the definition of the new operators $a_{\underline{k},\ell}$ and $a^{\dagger}_{\underline{k},\ell}$ <u>throughout $\underline{k}$-space.</u> It can be verified directly, that the vector potential Fourier component along $\underline{k},\ell$ is correctly given by, <u>for all $\underline{k}$,</u>

$$\hat{A}_{\underline{k},\ell} = \left(\frac{\hbar}{2\varepsilon_0 V \omega_{\underline{k}}}\right)^{1/2} \left[a_{\underline{k},\ell} + (-1)^{\ell} a^{\dagger}_{-\underline{k},\ell}\right] \ , \tag{9.39}$$

which substituted in the Fourier expansion for $\underline{A}(\underline{x},t)$, yields the Quantum Mechanical operator for the vector potential, i.e.

$$\hat{\underline{A}}(\underline{x},t) = \left(\frac{\hbar}{\varepsilon_0 V}\right)^{1/2} \sum_{\underline{k}} \frac{e^{i\underline{k}\cdot\underline{x}}}{\omega_{\underline{k}}^{1/2}} \sum_{\ell=1,2} \frac{1}{\sqrt{2}} \left[a_{\underline{k},\ell} + (-1)^{\ell} a^{\dagger}_{-\underline{k},\ell}\right] \hat{e}_{\underline{k},\ell} \ . \tag{9.40}$$

Using a similar procedure we also have, for $\underline{k} > 0$,

$$(\varepsilon_0 V)\, \dot{A}_{\underline{k},\ell} = \frac{1}{\sqrt{2}} \left(\Pi_{\underline{k},2\ell-1} - i\, \Pi_{\underline{k},2\ell}\right) \tag{9.41a}$$

and

$$(\varepsilon_0 V)\, \dot{A}_{-\underline{k},\ell} = \frac{(-1)^{\ell}}{\sqrt{2}} \left(\Pi_{\underline{k},2\ell-1} + i\, \Pi_{\underline{k},2\ell}\right) \ . \tag{9.41b}$$

Substituting 9.35b and the defining equations for $a_{\underline{k},\ell}$ and $a^{\dagger}_{\underline{k},\ell}$ (equations 9.37 and 9.38), then yields the operator for $\dot{A}_{\underline{k},\ell}$, i.e.

$$\hat{\dot{A}}_{\underline{k},\ell} = -i \left(\frac{\hbar\omega_{\underline{k}}}{2\varepsilon_0 V}\right)^{1/2} \left[a_{\underline{k},\ell} - (-1)^{\ell} a^{\dagger}_{-\underline{k},\ell}\right] \ , \tag{9.42}$$

and therefore also the operator for $\dot{\underline{A}}(\underline{x},t)$, i.e.

$$\hat{\dot{\underline{A}}}(\underline{x},t) = -i\left(\frac{\hbar}{\varepsilon_0 V}\right)^{1/2} \sum_{\underline{k}} \omega_{\underline{k}}^{1/2}\, e^{i\underline{k}\cdot\underline{x}} \sum_{\ell=1,2} \frac{1}{\sqrt{2}} \left[a_{\underline{k},\ell} - (-1)^{\ell} a^{\dagger}_{-\underline{k},\ell}\right] \hat{e}_{\underline{k},\ell} \ . \tag{9.43}$$

Note that this is the operator for (minus) the electric field, since $\underline{\mathcal{E}} = -\dot{\underline{A}}$ in this gauge.

What are the new operators $a_{\underline{k},\ell}$ and $a^{\dagger}_{\underline{k},\ell}$? Well, we can show by direct substitution in the commutation relations for $b_{\underline{k},j}$ and $b^{\dagger}_{\underline{k},j}$, as given by equation 9.29, that they are also suitable ladder operators, valid for all of $\underline{k}$-space, i.e.

update

$$[a_{\underline{k}',\ell'},\ a_{\underline{k},\ell}] = 0 \tag{9.44a}$$

$$[a_{\underline{k}',\ell'},\ a^{\dagger}_{\underline{k},\ell}] = \delta_{\underline{k}',\underline{k}}\ \delta_{\ell',\ell} \tag{9.44b}$$

$$[a^{\dagger}_{\underline{k},\ell'}, a^{\dagger}_{\underline{k},\ell}] = 0 \quad . \tag{9.44c}$$

We can invert equations 9.37 to obtain

$$b_{\underline{k},2\ell-1} = \frac{1}{\sqrt{2}}[a_{\underline{k},\ell} + (-1)^{\ell} a_{-\underline{k},\ell}] \tag{9.45a}$$

$$b_{\underline{k},2\ell} = \frac{i}{\sqrt{2}}[a_{\underline{k},\ell} - (-1)^{\ell} a_{-\underline{k},\ell}] \quad , \tag{9.45b}$$

and also equations 9.38 to obtain

$$b^{\dagger}_{\underline{k},2\ell-1} = \frac{1}{\sqrt{2}}[a^{\dagger}_{\underline{k},\ell} + (-1)^{\ell} a^{\dagger}_{-\underline{k},\ell}] \tag{9.46a}$$

$$b^{\dagger}_{\underline{k},2\ell} = \frac{-i}{\sqrt{2}}[a^{\dagger}_{\underline{k},\ell} - (-1)^{\ell} a^{\dagger}_{-\underline{k},\ell}] \quad , \tag{9.56b}$$

and use these expressions in the Hamiltonian (equation 9.30),

$$H = \sum_{+\underline{k},j} \hbar\omega_{\underline{k}} (b^{\dagger}_{\underline{k},j} b_{\underline{k},j} + \frac{1}{2})$$

$$= \sum_{+\underline{k},\ell} \hbar\omega_{\underline{k}} (b^{\dagger}_{\underline{k},2\ell-1} b_{\underline{k},2\ell-1} + b^{\dagger}_{\underline{k},2\ell-1} b_{\underline{k},2\ell} + 1)$$

to obtain

$$H = \sum_{+\underline{k},\ell} \hbar\omega_{\underline{k}} (a^{\dagger}_{\underline{k},\ell} a_{\underline{k},\ell} + a^{\dagger}_{-\underline{k},\ell} a_{-\underline{k},\ell} + 1) \quad .$$

If we now allow the summation to run over both half spaces of $\underline{k}$, we finally have

$$\boxed{H = \sum_{\underline{k},\ell} \hbar\omega_{\underline{k}} (a^{\dagger}_{\underline{k},\ell} a_{\underline{k},\ell} + \frac{1}{2})} \quad . \tag{9.47}$$

It can be seen that this Hamiltonian is of the same type as the one obtained previously (equation 9.30) but in this representation involves a summation of uncoupled harmonic oscillators corresponding to all wavenumbers $\underline{k}$ (both half-spaces) and to two polarizations per wavenumber.# This is a much more satisfactory state of affairs, which oversteps the artificial partition of wavenumber space into two halves.

#As opposed to half the wavenumbers, but four modes per wavenumber [(two phases per polarization) × (two polarizations)]. The total number of modes is of course the same.

The eigenstates corresponding to the new Hamiltonian are now expressible in terms of an alternate set of orthonormal basis vectors

$$|\underline{n}\rangle \equiv |n_{\underline{k}_1,1}, n_{\underline{k}_1,2}, n_{\underline{k}_2,2}, \ldots\ldots, n_{\underline{k},\ell} \ldots\ldots \rangle \ , \tag{9.48}$$

such that $\langle \underline{n}'|\underline{n}\rangle = \delta_{\underline{n}',\underline{n}}$, and

$$a_{\underline{k},\ell}|\underline{n}\rangle = n_{\underline{k},\ell}^{1/2}\,|\ \ldots\ldots, n_{\underline{k},\ell} - 1, \ \ldots\ldots\ \rangle \tag{9.49a}$$

$$a^{\dagger}_{\underline{k},\ell}|\underline{n}\rangle = (n_{\underline{k},\ell} + 1)^{1/2}|\ \ldots\ldots, n_{\underline{k},\ell} + 1, \ \ldots\ldots\ \rangle \ , \tag{9.49b}$$

with corresponding eigenvalues,

$$E_{\underline{n}} = \sum_{\underline{k},\ell} \hbar\omega_{\underline{k}}\,(n_{\underline{k},\ell} + \tfrac{1}{2}) \ . \tag{9.50}$$

At this point we will open a parenthesis by noting that the energy of the ground state, i.e.

$$E_0 = \frac{1}{2}\sum_{\underline{k},\,\ell=1,2} \hbar\omega_{\underline{k}} = \sum_{\underline{k}} \hbar\omega_{\underline{k}} \ , \tag{9.51}$$

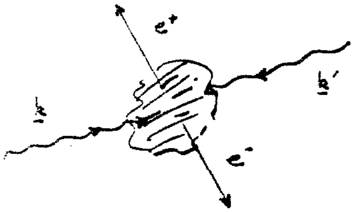
is infinite. This is a consequence of the fact (among other things) that we are assuming that this formalism is valid in the limit of infinite wavevectors. We have seen this situation before, when analyzing the quantum behavior of a fluid at low temperatures (section 5.6). There we also had a Hamiltonian composed of uncoupled harmonic oscillators extending to wavenumbers of infinite magnitude. In computing the expected value of the energy at thermal equilibrium, we subtracted the (infinite) ground state energy (equation 5.95), which is a constant anyway and does not affect the heat capacity calculation. In any event, it was possible in that case to see first of all why the Hamiltonian representing uncoupled harmonic oscillators, as well as the extension to infinite wavevectors was an approximation. On the one hand, we kept only terms of second order when expanding both the kinetic energy of the fluid (equations 5.69 and 5.70) and the potential energy (equations 5.71 and 5.72), and on the other, we assumed that our continuum approximation is valid to infinite wavevectors, in obvious violation of our original requirement of large length scales (equation 5.61). See also discussion at the bottom of page 5.19 and 5.20 as well as problem 5.18. In the case of the electromagnetic field, however, it is not quite so obvious what is the nature of the difficulty that results in this infinity. The Hamiltonian was derived from the energy in the field, as given by equation 9.1, this time with no approximations.

Even if we assume, for the moment, that the framework of our formalism# is not where

---

#i.e. replacing the conjugate (q, p) pairs in the classical Hamiltonian by the appropriate quantum operators.

the problem lies, it is not necessary to wander too far from where we are at this point, to stray into unchartered territory. It is clear, for example, that viewing the behavior of the electromagnetic field on a stage where charges and currents plan no role, removes from the Hamiltonian, as derived from our field energy equation 9.1, an entire set of degrees of freedom (coordinates), dissallowing charges and currents to appear, if initially absent. Yet we know from experiment that radiation of sufficiently short wavelength (γ-rays) can create an electron-position pair, as indicated below.



This process not only allows charges and currents to appear, even if absent at some point in time, but also provides a mechanism for coupling electromagnetic modes of different wavevectors $\underline{k}$. Both of these observations render the exactness of our Hamiltonian suspect and underscore the need for a description of the electromagnetic field in a manner that includes charges, currents, and creation and annihilation of elementary particles from the start. This, however, is beyond the scope of these lectures if not beyond the honest reach of present day science.#

Nevertheless, if we restrict ourselves to small wavenumbers, i.e.

$$\hbar c\underline{k} << m_e c^2 \simeq 0.5 \text{ MeV} \quad , \tag{9.52}$$

a limitation quite acceptable to the discussion of atomic phenomena (energies of a few eV) and the accompanying interactions with the electromagnetic field, we would expect an adequate description of processes that are consistent with these limitations. The analogy with the quantum description of a fluid at low temperatures is useful. The infinity of the energy of the ground state, even though disquieting in itself, actually poses no problems in a theory valid in the non-relativistic limit in which energy can only be defined within an arbitrary additive constant or, equivalently, be measured from an arbitrary origin. In our case, we will take the non-relativistic limit to be expressed by inequality 9.52, and we will close the parenthesis by measuring field energies using the energy of the ground state as the origin. This is equivalent to using a Hamiltonian with the ground state energy subtracted from it, or, in operator form,

---

#For a discussion along the lines of the evolution of these ideas see the preface to the book by J. Schwinger, on Selected Papers on Quantum Electrodynamics, (Dover 1958).

$$H \to (H - E_0) = \sum_{\underline{k}, \ell} \hbar\omega_{\underline{k}}\, a^\dagger_{\underline{k}, \ell}\, a_{\underline{k}, \ell} \qquad (9.53)$$

9.4 The Electromagnetic field in thermal equilibrium. Black body radiation.

Let us now assume that the electromagnetic field in the volume V is in thermal equilibrium with the surroundings which are at a temperature T. We can compute the expected energy of the whole system using the results of section 5.5 of a system of uncoupled harmonic oscillators in thermal equilibrium. In particular, from equation 5.56,

$$\langle H - E_0 \rangle = \int_0^\infty \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1}\, \rho(\omega)\, d\omega, \qquad (9.54)$$

where, in this case, $\rho(\omega)$ is the number of modes $(\underline{k}, \ell)$ of the electromagnetic field with frequencies between ω and $\omega + d\omega$. In particular, since the total number of modes whose frequency ω' is less than ω is given by

$$\begin{aligned} N(\omega' < \omega) &= 2\, \frac{\frac{4}{3}\pi k^3(\omega)}{(2\pi)^3/V} \\ &= 2V\, \frac{\frac{4}{3}\pi\omega^3}{(2\pi c)^3} , \end{aligned}$$

where the factor of 2 in front arises from the fact that there are two modes $(\ell = 1, 2)$ per $\underline{k}$ wavevector. Therefore, since

$$\rho(\omega) = \frac{d}{d\omega} N(\omega' < \omega) ,$$

we have

$$\rho(\omega) = \frac{V\omega^3}{\pi^2 c^3}. \qquad (9.55)$$

Substituting into equation 9.54, we have

$$\langle H - E_0 \rangle = \frac{V\hbar}{\pi^2 c^3} \int_0^\infty \frac{\omega^3}{e^{\hbar\omega/k_B T} - 1}\, d\omega , \qquad (9.56)$$

and therefore we have a <u>spectral energy density</u> [energy in $(\omega, \omega + d\omega)$ per unit volume] of

$$S(\omega)\, d\omega = \left(\frac{\hbar}{\pi^2 c^3}\right)\left(\frac{\omega^3}{e^{\hbar\omega/k_B T} - 1}\right) d\omega , \qquad (9.57)$$

It can be readily verified, that we have derived the Black Body spectrum of Planck (equation 2.130). The total energy density can then be computed from 9.56, i.e.,

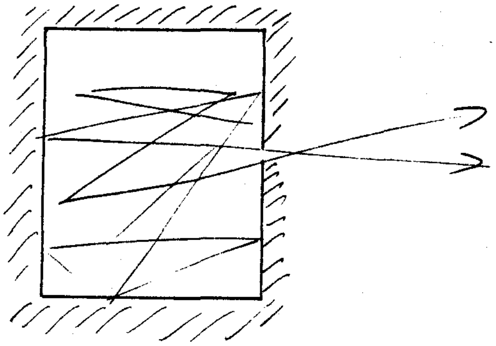
$$\epsilon \equiv \frac{1}{V}\langle H - E_0\rangle = \frac{\hbar}{\pi^2 c^3}\int_0^\infty \frac{\omega^3\, d\omega}{e^{\hbar\omega/k_B T} - 1}$$

$$= \frac{(k_B T)^4}{\pi^2(\hbar c)^3}\int_0^\infty \frac{x^3\, dx}{e^x - 1} \, ,$$

or, substituting for the integral from equation 5.98, we have

$$\epsilon = \frac{\pi^2 (k_B T)^4}{15(\hbar c)^3} \, . \tag{9.58}$$

We can also compute the energy flux (Watts per unit area) radiated out of a small opening in a cavity



which, for an isotropic gas of photons all travelling with speed c, is given by

$$\mathcal{J} = \frac{1}{4} c\epsilon \tag{9.59}$$

or

$$\mathcal{J} = \sigma T^4 \quad , \tag{9.60}$$

where

$$\sigma \equiv \frac{\pi^2 k_B{}^4}{60\, \hbar^3 c^2} \tag{9.61}$$

is the Stefan-Boltzmann constant that had been determined empirically to be equal to

$$\sigma = 5.67 \times 10^{-8}(\text{Watts/m}^2)/{}^\circ\text{K}^4 \, ,$$

computed here from first principles.

9.5 The eigenstates of the electromagnetic field.

If $|0\rangle$ is the ground state of the whole system, what kind of state is represented by

$$|0, \ldots\ldots, 0, 1_{\underline{k},\ell}, 0, \ldots\ldots\rangle = a^{\dagger}_{\underline{k},\ell}|0\rangle \quad , \tag{9.62}$$

the first excited state of the mode $(\underline{k},\ell)$? Alternatively, what is the result of the action of the creation operator $a^{\dagger}_{\underline{k},\ell}$ on the ground state?

We know that this state is an eigenstate of the Hamiltonian and therefore has an energy $\epsilon_{1_{\underline{k},\ell}}$ given by

$$(H - E_0)(a^{\dagger}_{\underline{k},\ell}|0\rangle = \epsilon_{1_{\underline{k},\ell}}(a^{\dagger}_{\underline{k},\ell}|0\rangle \quad ,$$

or substituting for $H - E_0$ in terms of the annihilation and creation operators (equation 9.53), we have

$$(H - E_0)(a^{\dagger}_{\underline{k},\ell}|0\rangle = \hbar\omega_{\underline{k}}(a^{\dagger}_{\underline{k},\ell}|0\rangle) \tag{9.63}$$

or

$$\epsilon_{1_{\underline{k},\ell}} = \hbar\omega_{\underline{k}} \quad . \tag{9.64}$$

We see that <u>the result of the operation of</u> $a^{\dagger}_{\underline{k},\ell}$ <u>on the ground state is to create an excitation of energy</u> $\hbar\omega_{\underline{k}}$, where $\omega_{\underline{k}} = c|\underline{k}|$.

How can we relate such a state with a time varying electric and magnetic field? We will illustrate this through an example.

<u>Example 9.1</u> <u>Compute the expectation values of the electric and magnetic fields corresponding to the quantum state of the field which at $t = 0$ is given by</u>

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}[|0\rangle + |0, \ldots\ldots, 0, 1_{\underline{k},\ell}, 0, \ldots\rangle] \quad .$$

Note that this state can be written as

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}(1 + a^{\dagger}_{\underline{k},\ell})|0\rangle$$

and has a time dependence given by

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}(1 + e^{-i\omega_{\underline{k},\ell}t}\, a^{\dagger}_{\underline{k},\ell})|0\rangle\, e^{-iE_0 t/\hbar} \quad ,$$

where we have factored out the phase factor corresponding to the energy of the ground state.

We first compute the (time-dependent) Fourier components $A_{\underline{k}',\ell'}(t)$ of the vector potential. These are given in terms of the expectation value of the corresponding operator (equation 9.39). In particular, we have

$$\langle A_{\underline{k}',\ell'}(t)\rangle = \left(\frac{\hbar}{2\epsilon_0 V\omega_{\underline{k}'}}\right)^{1/2} \langle\psi(t)|[a_{\underline{k}',\ell'} + (-1)^{\ell'} a^{\dagger}_{-\underline{k}',\ell'}]|\psi(t)\rangle$$

$$= \left(\frac{\hbar}{2\epsilon_0 V\omega_{\underline{k}'}}\right)^{1/2} [\langle\psi(t)|a_{\underline{k}',\ell'}|\psi(t)\rangle + (-1)^{\ell'} \langle\psi(t)|a^{\dagger}_{-\underline{k}',\ell'}|\psi(t)\rangle]$$

Now

$$a_{\underline{k}',\ell'}|\psi\rangle = \frac{1}{\sqrt{2}} (a_{\underline{k}',\ell'} + e^{-i\omega_{\underline{k},\ell}t} a_{\underline{k}',\ell'} a^{\dagger}_{\underline{k},\ell})|0\rangle e^{-iE_0 t/\hbar}$$

$$= \frac{1}{\sqrt{2}} e^{-i\omega_{\underline{k},\ell}t}|0\rangle \delta_{\underline{k}',\underline{k}} \delta_{\ell',\ell} e^{-iE_0 t/\hbar}$$

and therefore we have,

$$\langle\psi|a_{\underline{k}',\ell'}|\psi\rangle = \frac{1}{\sqrt{2}} e^{iE_0 t/\hbar} \langle 0|(1 + a_{\underline{k},\ell})\cdot\frac{1}{\sqrt{2}} e^{-i\omega_{\underline{k},\ell}t}|0\rangle \delta_{\underline{k}',\underline{k}} \delta_{\ell',\ell} e^{-iE_0 t/\hbar}$$

$$= \frac{1}{2} e^{-i\omega_{\underline{k},\ell}t} \delta_{\underline{k}',\underline{k}} \delta_{\ell',\ell} \quad .$$

We can compute the second matrix element by noting that $a^{\dagger}_{-\underline{k}',\ell'}$ is the hermitean adjoint of $a_{-\underline{k}',\ell'}$, i.e.

$$\langle\psi|, a^{\dagger}_{-\underline{k}',\ell'}|\psi\rangle = \langle\psi|a_{-\underline{k}',\ell'}, |\psi\rangle$$

$$= \langle\psi|, a_{-\underline{k}',\ell'}|\psi\rangle^{*}$$

and therefore, from the previous calculation,

$$\langle\psi|a^{\dagger}_{-\underline{k}',\ell'}|\psi\rangle = \frac{1}{2} e^{i\omega_{\underline{k},\ell}t} \delta_{-\underline{k}',\underline{k}} \delta_{\ell',\ell} \quad .$$

Consequently, we have for the expectation value of $A_{\underline{k}',\ell'}$

$$\langle A_{\underline{k}',\ell'}\rangle = \frac{1}{2} \left(\frac{\hbar}{2\epsilon_0 V\omega_{\underline{k}'}}\right)^{1/2} [e^{-i\omega_{\underline{k},\ell}t} \delta_{\underline{k}',\underline{k}} + (-1)^{\ell'} e^{i\omega_{\underline{k},\ell}t} \delta_{-\underline{k}',\underline{k}}] \delta_{\ell',\ell}$$

and therefore, substituting for the vector potential $\underline{A}(\underline{x},t)$, we have

$$\langle\underline{A}(\underline{x},t)\rangle = \sum_{\underline{k}',\ell'} e^{i\underline{k}\cdot\underline{x}} \langle A_{\underline{k}',\ell'}\rangle \hat{e}_{\underline{k}',\ell'}$$

$$= \frac{1}{2} \left(\frac{\hbar}{2\epsilon_0 V\omega_{\underline{k}}}\right)^{1/2} [e^{i(\underline{k}\cdot\underline{x}-\omega_{\underline{k}}t)} + e^{-i(\underline{k}\cdot\underline{x}-\omega_{\underline{k}}t)}] \hat{e}_{\underline{k},\ell}$$

[note that $\hat{e}_{-\underline{k},\ell} = (-1)^{\ell} \hat{e}_{\underline{k},\ell}$], or

$$\langle \underline{A}(\underline{x}, t)\rangle = \left(\frac{\hbar}{2\varepsilon_0 V\omega_{\underline{k}}}\right)^{1/2} \cos(\underline{k}\cdot\underline{x} - \omega_{\underline{k}} t)\ \hat{e}_{\underline{k},\ell} \quad .$$

Using the vector potential, we can now compute the expectation value of the electric field, i. e.

$$\langle \underline{\mathcal{E}}(\underline{x}, t)\rangle = -\frac{\partial}{\partial t}\langle \underline{A}(\underline{x}, t)\rangle = -\left(\frac{\hbar\omega_{\underline{k}}}{2\varepsilon_0 V}\right)^{1/2} \sin(\underline{k}\cdot\underline{x} - \omega_{\underline{k}} t)\ \hat{e}_{\underline{k},\ell}$$

and the magnetic field

$$\langle \underline{\mathcal{B}}(\underline{x}, t)\rangle = \frac{\partial}{\partial \underline{x}} \times \langle \underline{A}(\underline{x}, t)\rangle = -\frac{1}{c}\left(\frac{\hbar\omega_{\underline{k}}}{2\varepsilon_0 V}\right)^{1/2} \sin(\underline{k}\cdot\underline{x} - \omega_{\underline{k}} t)\ \hat{e}_{\underline{k}} \times \hat{e}_{\underline{k},\ell} \quad .$$

We can see that this is a propagating wave, with a speed c $(= \omega_{\underline{k}}/|\underline{k}|)$, in the direction of the wavevector $\underline{k}$, of polarization $-\hat{e}_{\underline{k},\ell}$, and of uniform amplitude throughout the volume V.

Without repeating the algebra, can you see what the resulting $\langle\underline{\mathcal{E}}\rangle$ and $\langle\underline{\mathcal{B}}\rangle$ would have been if the quantum state was given by

$$|\psi_1(0)\rangle = \frac{1}{\sqrt{2}}\left[|0\rangle + |0,\ldots\ldots,0,2_{\underline{k},\ell},0,\ldots\ldots\rangle\right] \quad ,$$

as compared to

$$|\psi_2(0)\rangle = \frac{1}{\sqrt{2}}\left[|0\rangle + |0,\ldots\ldots,0,1_{2\underline{k},\ell},0,\ldots\ldots\rangle\right] \quad ?$$

In other words, what is the difference between a quantum state involving the second excited level of the mode $(\underline{k}, \ell)$, as compared to the first excited state of the mode $(2\underline{k}, \ell)$? Both have the same energy, and as we shall see, the same momentum.

9.6 The electromagnetic field momentum.

From Maxwell's equations, we have that the total momentum of the electromagnetic field is given by[#]

$$\underline{P} = \varepsilon_0 \int_V (\underline{\mathcal{E}} \times \underline{\mathcal{B}})\, d^3\underline{x} \quad . \tag{9.65}$$

It should be emphasized that the integral of equation 9.65 represents the (linear) momentum of the electromagnetic field and should not be confused with the (generalized) momenta $\Pi_{\underline{k},j}$, conjugate to the vector potential Fourier components $Q_{\underline{k},j}$.

Expressing $\underline{\mathcal{E}}$ and $\underline{\mathcal{B}}$ in terms of the vector potential, we have

$$\underline{P} = -\varepsilon_0 \int_V \dot{\underline{A}} \times \left(\frac{\partial}{\partial \underline{x}} \times \underline{A}\right) d^3\underline{x} \quad , \tag{9.66}$$

and substituting the Fourier expansion of equation 9.6, we have

---

[#] See, for example, Feynman, Leighton and Sands, v. II, ch. 27.

$$\underline{P} = -\, i(\varepsilon_0 V) \sum_{\underline{k},\, \ell} (-1)^{\ell}\, \underline{k}\, \dot{A}_{-\underline{k},\, \ell}\, A_{\underline{k},\, \ell} \quad . \tag{9.67}$$

where we have used that for the admissible wavenumbers (periodic boundary conditions at the edges of the volume V), we have

$$\int_V e^{i(\underline{k}'+\underline{k})\cdot \underline{x}}\, d^3\underline{x} = V\, \delta_{-\underline{k}',\, \underline{k}} \quad ,$$

and also that

$$\hat{e}_{\underline{k},\, \ell'} \times (\underline{k} \times \hat{e}_{\underline{k},\, \ell}) = \underline{k}\, \delta_{\ell',\, \ell} \quad .$$

If we now substitute the operator expressions for $A_{\underline{k},\, \ell}$ and $\dot{A}_{-\underline{k},\, \ell}$ from equations 9.39 and 9.42, we have

$$\hat{\underline{P}} = -\frac{\hbar}{2} \sum_{\underline{k},\, \ell} (-1)^{\ell}\, \underline{k}\, [a_{-\underline{k},\, \ell} - (-1)^{\ell} a^{\dagger}_{\underline{k},\, \ell}]\, [a_{\underline{k},\, \ell} + (-1)^{\ell} a^{\dagger}_{-\underline{k},\, \ell}]$$

$$= -\frac{\hbar}{2} \sum_{\underline{k},\, \ell} (-1)^{\ell}\, \underline{k}\, [a_{-\underline{k},\, \ell}\, a_{\underline{k},\, \ell} - a^{\dagger}_{\underline{k},\, \ell}\, a^{\dagger}_{-\underline{k},\, \ell} + (-1)^{\ell} a_{-\underline{k},\, \ell}\, a^{\dagger}_{-\underline{k},\, \ell} - (-1)^{\ell} a^{\dagger}_{\underline{k},\, \ell}\, a_{\underline{k},\, \ell}]$$

or

$$\hat{\underline{P}} = \frac{1}{2} \sum_{\underline{k},\, \ell} \hbar\, \underline{k} (a^{\dagger}_{\underline{k},\, \ell}\, a_{\underline{k},\, \ell} - a^{\dagger}_{-\underline{k},\, \ell}\, a_{-\underline{k},\, \ell} - 1) + \frac{1}{2} \sum_{\underline{k},\, \ell} (-1)^{\ell}\, \hbar\, \underline{k} (a^{\dagger}_{\underline{k},\, \ell}\, a^{\dagger}_{-\underline{k},\, \ell} - a_{\underline{k},\, \ell}\, a_{-\underline{k},\, \ell}) \quad .$$

It can be seen that the second sum is identically zero, since the factor in the parentheses does not change sign as $\underline{k} \to -\underline{k}$, whereas $\hbar\, \underline{k}$ does. By the same token the -1 in the parentheses of the first series does not contribute to the sum either and we have

$$\hat{\underline{P}} = \frac{1}{2} \sum_{\underline{k},\, \ell} \hbar\, \underline{k}\, (a^{\dagger}_{\underline{k},\, \ell}\, a_{\underline{k},\, \ell} - a^{\dagger}_{-\underline{k},\, \ell}\, a_{-\underline{k},\, \ell}) \quad ,$$

or, equivalently

$$\boxed{\hat{\underline{P}} = \sum_{\underline{k},\, \ell} \hbar\, \underline{k}\, a^{\dagger}_{\underline{k},\, \ell}\, a_{\underline{k},\, \ell}} \quad .$$

This is a very appealling result. We find that the field momentum associated with the mode $(\underline{k}, \ell)$ is equal to $\hbar\, \underline{k}$ times the occupation number operator for that mode, i.e.

$$\hat{n}_{\underline{k},\, \ell} = a^{\dagger}_{\underline{k},\, \ell}\, a_{\underline{k},\, \ell} \quad . \tag{9.69}$$

We can use the number operator to express both the field Hamiltonian and the field momentum. In particular,

$$\hat{H} - E_0 = \sum_{\underline{k},\ell} \hbar\omega_{\underline{k}}\, \hat{n}_{\underline{k},\ell} \qquad (9.70)$$

and

$$\hat{\underline{P}} = \sum_{\underline{k},\ell} \hbar\, \underline{k}\, \hat{n}_{\underline{k},\ell} \quad . \qquad (9.71)$$

9.7 The photon.

We have already seen that the operator $a^{\dagger}_{\underline{k},\ell}$ creates an excitation of energy $\hbar\omega_{\underline{k}}$. We can now use the field momentum operator of equation 9.68 to find its momentum. In particular, since $a^{\dagger}_{\underline{k},\ell}|0\rangle$ is an eigenstate of the number operator $\hat{n}_{\underline{k},\ell}$, with an eigenvalue of unity, we have that the momentum of the field in the first excited state of the mode $(\underline{k},\ell)$ is given by

$$\langle\hat{\underline{P}}\rangle_{1_{\underline{k},\ell}} = \langle 0|a_{\underline{k},\ell}\, \hat{\underline{P}}\, a^{\dagger}_{\underline{k},\ell}|0\rangle = \hbar\, \underline{k} \qquad (9.72)$$

(note the use of the annihilation operator - Hermitean adjoint - for the bra vector).

It is now clear that the operator $a^{\dagger}_{\underline{k},\ell}$ creates an excitation of energy $\hbar\omega_{\underline{k}}$, momentum $\hbar\,\underline{k}$, and (plane) polarization along $\hat{e}_{\underline{k},\ell}$. This excitation we may now identify with the corresponding photon, which served as the conceptial basis for Planck's black body spectrum theory, Einstein's photoelectric theory and Compton's scattering theory. We can also show that this excitation has zero rest mass as follows. We can define a four-vector momentum

$$p_{\mu} = (\frac{1}{c}\, E, \underline{p}) \quad , \qquad (9.73)$$

where E is the energy and $\underline{p}$ is the 3-vector linear momentum, and a quantum operator corresponding to the relativistic 4-momentum, which for the excitations of the field would then be given by

$$\hat{p}_{\mu} = [\frac{1}{c}(\hat{H} - E_0), \hat{\underline{P}}] \qquad (9.78)$$

(note that since we are interested in the excitation energy, we have removed the ground state energy from the Hamiltonian). We can now compute the expectation value of this operator in the state $a^{\dagger}_{\underline{k},\ell}|0\rangle$, i.e.

$$\langle\hat{p}_{\mu}\rangle_{1_{\underline{k},\ell}} = [\frac{1}{c}\,\langle\hat{H} - E_0\rangle_{1_{\underline{k},\ell}}, \langle\hat{\underline{P}}\rangle_{1_{\underline{k},\ell}}]$$

or, combining the results of the previous calculations for a field in the state $a^{\dagger}_{\underline{k},\ell}|0\rangle$, we have

$$\langle\hat{p}_{\mu}\rangle_{1_{\underline{k},\ell}} = (\frac{1}{c}\,\hbar\omega_{\underline{k}}, \hbar\,\underline{k}) = \frac{\hbar}{c}\,(\omega_{\underline{k}}, c\,\underline{k}) \quad . \qquad (9.79)$$

It can be seen that this corresponds to zero rest mass, since

$$m_0^2 = \frac{1}{c^2} \| \langle \hat{p}_\mu \rangle_{1_{\underline{k},\ell}} \|^2 = \frac{\hbar^2}{c^4} (\omega_{\underline{k}}^2 - c^2 k^2) = 0 \quad . \qquad (9.80)$$

It is perhaps interesting that the ground state of the field looks like an object of infinite rest mass (and no momentum).

Note that a field eigenstate with $n_{\underline{k},\ell}$ photons in the mode $(\underline{k},\ell)$ is represented by setting the corresponding harmonic oscillator to the $n_{\underline{k},\ell}^{th}$ excitation level, i.e.

$$\{n\ (\underline{k},\ell)\text{-photons}\} \Rightarrow |\ldots\ldots, n_{\underline{k},\ell}, \ldots\ldots\rangle \quad .$$

If no other photons are present, this state can be constructed from the ground state by repeated application of the $a_{\underline{k},\ell}^\dagger$ operator,

$$|0,\ldots\ldots,0,n_{\underline{k},\ell},0,\ldots\ldots\rangle = (n_{\underline{k},\ell}!)^{-1/2}\, (a_{\underline{k},\ell}^\dagger)^{n_{\underline{k},\ell}} |0\rangle \quad . \qquad (9.81)$$

In a similar fashion, we can represent any eigenstate of the electromagnetic field by repeated application of the corresponding creation operators on the field ground state, i.e.

$$|\underline{n}\rangle = |n_{\underline{k}_1,1}, n_{\underline{k}_1,2}, \ldots\ldots, n_{\underline{k},\ell}, \ldots\ldots\rangle = \prod_{\underline{k},\ell} (n_{\underline{k},\ell}!)^{-1/2} (a_{\underline{k},\ell}^\dagger)^{n_{\underline{k},\ell}} |0\rangle \quad . \qquad (9.82)$$

Note also that the vector potential

$$\underline{A}(\underline{x},t) = \sum_{\underline{k},\ell} e^{i\underline{k}\cdot\underline{x}} A_{\underline{k},\ell}\, \hat{e}_{\underline{k},\ell} \quad ,$$

expressed as a quantum mechanical operator (equation 9.39),

$$\hat{\underline{A}}(\underline{x},t) = \left(\frac{\hbar}{2\varepsilon_0 V}\right)^{1/2} \sum_{\underline{k}} \omega_{\underline{k}}^{-1/2}\, e^{i\underline{k}\cdot\underline{x}} \cdot \sum_{\ell=1,2} [a_{\underline{k},\ell} + (-1)^\ell a_{-\underline{k},\ell}^\dagger]\, \hat{e}_{\underline{k},\ell} \quad ,$$

has an interesting operator interpretation. The function $e^{i\underline{k}\cdot\underline{x}}$, which we recognize as a space wavefunction of momentum $\hbar\underline{k}$, multiplies the sum $[a_{\underline{k},\ell} + (-1)^\ell a_{-\underline{k},\ell}^\dagger]$ of operators. The first one <u>annihilates</u> a photon of momentum $\hbar\underline{k}$, while the second one <u>creates</u> a photon of momentum $-\hbar\underline{k}$. In both cases, the momentum of the system is decreased by $\delta\underline{p} = \hbar\underline{k}$, which is offset by the momentum $\hbar\underline{k}$ of $e^{i\underline{k}\cdot\underline{x}}$. The net effect is no change in momentum and we can read equation 9.39 as a superposition of all the possible ways of altering the field leaving the momentum unchanged.

Note also that the first set of ladder operators $b_{\underline{k},j}^\dagger$, which also generate excitations of energy $\hbar\omega_{\underline{k}}$, create states which we recognize to be standing wave states of zero momentum. See problem 9.5. The operators $b_{\underline{k},j}^\dagger$ also generate a complete set of orthogonal eigenstates, and it is of course possible, even though not very practical, to express everything in terms of those states. See problem 9.6.

9.8 Photons and wave-particle duality.

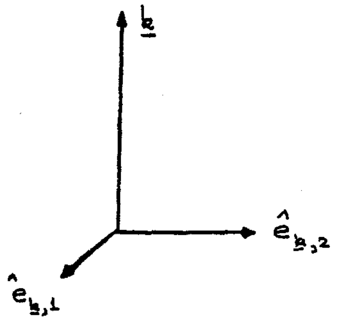
We have come a long way and in the process we have resolved, I believe, at least in one case, the riddle of wave-particle duality. The quantization of the electromagnetic field, the most manifestly wave-like phenomenon in nature, affords us an understanding of how it is possible for a wave-like entity to behave as an object, or a superposition of objects, of discrete energy and momentum. As we will see later on, the electromagnetic field when interacting with matter (charged particles) can only make transitions up or down the ladder of harmonic oscillator levels corresponding to each normal mode. This always results in a loss or a gain by the field of discrete energy and momentum in units of $\hbar\omega_{\underline{k}}$ and $\hbar\underline{k}$ respectively. There is no contradiction in accepting this "dual" picture, which the formalism of the quantization of the field describes in an explicit manner. It is perhaps interesting that our first success at explaining this riddle came from a theory that was relativistically correct, since the field energy as given by equation 9.1, was derived from Maxwell's equations which are relativistically correct, remaining invariant under a Lorentz transformation[#]. Our resulting quantum theory, having been derived from the relativistically correct Hamiltonian, is of course also relativistically correct.

9.9 The angular momentum of the photon.

If we explicitly write out the terms in the Hamiltonian corresponding to the two polarizations $\ell=1,2$ for each $\underline{k}$, we have from equation 9.47

$$H = \sum_{\underline{k}} \hbar\omega_{\underline{k}} (a^{\dagger}_{\underline{k},1}\, a_{\underline{k},1} + a^{\dagger}_{\underline{k},2}\, a_{\underline{k},2} + 1) \quad , \tag{9.83}$$

where, you recall, for each wavevector $\underline{k}$, the ladder operators $a_{\underline{k},\ell}$, $a^{\dagger}_{\underline{k},\ell}$ create excitations polarized along $\hat{e}_{\underline{k},1}$ and $\hat{e}_{\underline{k},2}$, for $\ell=1,2$. The unit vector geometry is sketched below



[#] See for example Feynman, Leighton and Sands, v. II, ch. 26.

It can be seen that the Hamiltonian as given by equation 9.83 corresponds to a system of uncoupled, two-dimensional, harmonic oscillators, one for each wavevector $\underline{k}$. See the Hamiltonian of equation 6.9, in section 6.1 on the two-dimensional harmonic oscillator. Note that, for each wavector $\underline{k}$, we have that the direction $\hat{e}_{\underline{k}}$ acts as the z-axis whereas the polarization direction $\hat{e}_{\underline{k},1}$ and $\hat{e}_{\underline{k},2}$ act like the x and y axes respectively, i.e.

$$\hat{e}_x \rightarrow \hat{e}_{\underline{k},1}, \quad \hat{e}_y \rightarrow \hat{e}_{\underline{k},2}, \quad \hat{e}_z \rightarrow \hat{e}_{\underline{k}} \quad . \tag{9.84}$$

What about angular momentum? The classical angular momentum of the field is given by

$$\underline{J} = \varepsilon_0 \int_V \underline{x} \times (\underline{\mathcal{E}} \times \underline{\mathcal{B}})\, d^3\underline{x}^{\#} \, .$$

If we substitute for $\underline{\mathcal{E}}$ and $\underline{\mathcal{B}}$ in terms of the vector potential $\underline{A}$, we find that the total angular momentum can be split into two parts, i.e.[†]

$$\underline{J} = \underline{L} + \underline{S} \quad , \tag{9.86a}$$

where the i^{th} component of $\underline{L}$ is given by

$$L_i = -\varepsilon_0 \sum_{j=1}^{3} \int_V \dot{A}_j \left(\underline{x} \times \frac{\partial}{\partial \underline{x}}\right)_i A_j \, d^3\underline{x} \quad , \qquad i = 1,2,3 \tag{9.86b}$$

and

$$\underline{S} = \varepsilon_0 \int_V (\underline{A} \times \dot{\underline{A}})\, d^3\underline{x} \quad . \tag{9.86c}$$

We can now show that $\underline{L}$ is transverse (zero Fourier components along $\underline{k}$) while $\underline{S}$ is longitudinal. In particular,

$$\underline{S} = \sum_{\underline{k}} S_{\underline{k}}\, \hat{e}_{\underline{k}} \quad , \tag{9.87}$$

where, substituting the field operators for $\underline{A}$ and $\dot{\underline{A}}$ (equations 9.40 and 9.43), we have

$$S_{\underline{k}} = i\hbar\,(a_{\underline{k},1}\, a^{\dagger}_{\underline{k},2} - a^{\dagger}_{\underline{k},1}\, a_{\underline{k},2}) \quad . \tag{9.88}$$

Comparing $S_{\underline{k}}$ with the z-component of the two-dimensional Harmonic oscillator angular momentum (equation 6.13), we see that the analogy is complete. We also see, as in the case of the two-dimensional harmonic oscillator, that states of definite excitation level along $\hat{e}_{\underline{k},1}$ and $\hat{e}_{\underline{k},2}$ are not states of definite angular momentum along $\hat{e}_{\underline{k}}$ (not eigenstates of $\hat{S}_{\underline{k}}$). Since, however, H and $\hat{\underline{S}}$ commute, i.e.

[#] Comparing with equation 9.65 for the total angular momentum $\underline{P}$ in the field, we see that this is the integral of $\underline{x} \times \underline{p}$ per unit volume.

[†] This is a non-trivial calculation. For an outline of the derivation see A. Messia, Quantum Mechanics, chapter 21, § 23 and § 28, note that $\hbar = 1$ and cgs units are used.

$$[H, \hat{\underline{S}}] = 0 \quad , \tag{9.89}$$

we are motivated to seek simultaneous eigenstates of both H and $\hat{\underline{S}}$.

We have already solved this problem in the case of the two-dimensional harmonic oscillator. In an analogous fashion, we define new ladder operators for each wavevector mode $\underline{k}$,

$$a_{\underline{k},R} = \frac{1}{\sqrt{2}} (a_{\underline{k},1} - i a_{\underline{k},2}) \tag{9.90a}$$

$$a^\dagger_{\underline{k},R} = \frac{1}{\sqrt{2}} (a^\dagger_{\underline{k},1} + i a^\dagger_{\underline{k},2}) \tag{9.90b}$$

and

$$a_{\underline{k},L} = \frac{1}{\sqrt{2}} (a_{\underline{k},1} + i a_{\underline{k},2}) \tag{9.90c}$$

$$a^\dagger_{\underline{k},L} = \frac{1}{\sqrt{2}} (a^\dagger_{\underline{k},1} - i a^\dagger_{\underline{k},2}) \quad . \tag{9.90d}$$

These are proper ladder operators, which obey the commutation relations (σ = R, L)

$$[a_{\underline{k}',\sigma'}, a_{\underline{k},\sigma}] = 0 \tag{9.91a}$$

$$[a_{\underline{k}',\sigma'}, a^\dagger_{\underline{k},\sigma}] = \delta_{\underline{k}',\underline{k}} \, \delta_{\sigma',\sigma} \tag{9.91b}$$

$$[a^\dagger_{\underline{k}',\sigma'}, a^\dagger_{\underline{k},\sigma}] = 0 \quad , \tag{9.91c}$$

(See equations 6.15. The algebra is identical) and can be used to express the Hamiltonian, linear momentum and $\underline{S}$-angular momentum operators, i.e.

$$H = \sum_{\underline{k},\sigma} \hbar\omega_{\underline{k}} (a^\dagger_{\underline{k},\sigma} a_{\underline{k},\sigma} + \tfrac{1}{2}) \tag{9.92}$$

$$\underline{P} = \sum_{\underline{k},\sigma} \hbar \underline{k} \, a^\dagger_{\underline{k},\sigma} a_{\underline{k},\sigma} \tag{9.93}$$

and

$$\underline{S} = \sum_{\underline{k}} \hbar \, (a^\dagger_{\underline{k},R} a_{\underline{k},R} - a^\dagger_{\underline{k},L} a_{\underline{k},L}) \quad . \tag{9.94}$$

This Hamiltonian has a new set of eigenvectors

$$|\underline{n}\rangle = |n_{\underline{k}_1,R}, n_{\underline{k}_1,L}, \ldots , n_{\underline{k},\sigma}, \ldots \rangle \quad , \tag{9.95}$$

which are also eigenvectors of $\underline{P}$ and $\underline{S}$. Note that

$$a^\dagger_{\underline{k},\sigma} a_{\underline{k},\sigma} = \hat{n}_{\underline{k},\sigma} \quad ; \quad \sigma = R, L \tag{9.96}$$

is the corresponding number operator for the excitation number $n_{\underline{k},\sigma}$ in the (circular) polarization $(\underline{k}, \sigma)$.

A field eigenstate with $n_{\underline{k},R}$, $n_{\underline{k},L}$ excitations, i.e.

$$|\psi_{\underline{k}}\rangle = |0,\ldots,0,n_{\underline{k},R},n_{\underline{k},L},0,\ldots\rangle \qquad (9.97a)$$

has an excitation energy,

$$\langle H-E_0\rangle_{\underline{k}} = \langle\psi_{\underline{k}}|(H-E_0)|\psi_{\underline{k}}\rangle = \hbar\omega_{\underline{k}}(n_{\underline{k},R}+n_{\underline{k},L}) \quad , \qquad (9.97b)$$

a linear momentum

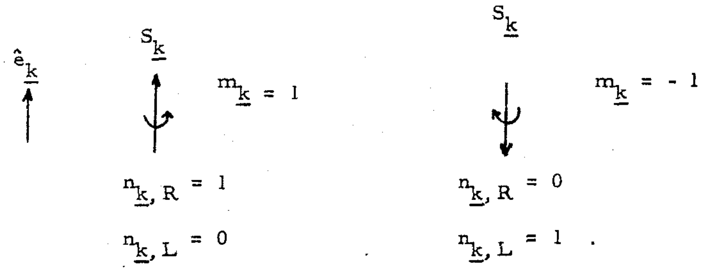
$$\langle\underline{P}\rangle_{\underline{k}} = \langle\psi_{\underline{k}}|\underline{P}|\psi_{\underline{k}}\rangle = \hbar\,\underline{k}(n_{\underline{k},R}+n_{\underline{k},L}) \quad , \qquad (9.97c)$$

and a component of angular momentum along $\underline{k}$ given by

$$\langle\underline{S}\rangle_{\underline{k}} = \hbar(n_{\underline{k},R}-n_{\underline{k},L}) \quad . \qquad (9.97d)$$

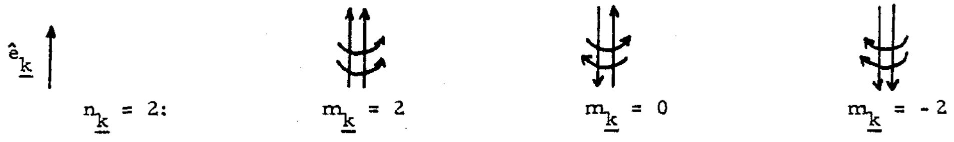
We see that the component of the angular momentum along $\underline{k}$, the direction of propagation, is quantized in units of $\hbar$. The energy state $n_{\underline{k}} = n_{\underline{k},R} + n_{\underline{k},L} = 1$ (one photon of wavevector $\underline{k}$) can have values of angular momentum along $\underline{k}$ of $+\hbar$ or $-\hbar$ but not zero, corresponding to,

$$m_{\underline{k}} = n_{\underline{k},R} - n_{\underline{k},L} = \pm 1 \quad .$$



The energy state $n_{\underline{k}} = n_{\underline{k},R} + n_{\underline{k},L} = 2$ (two photons of wavenumber $\underline{k}$) can have $\underline{k}$-angular momentum components of $+2\hbar$, 0, $-2\hbar$ corresponding to

$$m_{\underline{k}} = n_{\underline{k},R} - n_{\underline{k},L} = +2,\ 0,\ -2 \ .$$



It is clear from these pictures that this behavior can be explained completely by assuming that a photon has an intrinsic angular momentum of $\hbar$ which is always either parallel or anti-parallel to its wavevector $\underline{k}$. Note also that the quantum mechanical description of a two-dimensional harmonic oscillator explains completely why a single photon ($n_{\underline{k}} = 1$) cannot have a projection of its angular momentum, along its wavevector $\underline{k}$, which is zero ($m_{\underline{K}} = 0$).

update

It should be mentioned that a single photon can have a component of its total angular momentum along a particular axis, which is any multiple of $\hbar$ (including zero). The preceeding discussion refers to the intrinsic angular momentum, which we will subsequently identify with the spin of the photon. The total angular momentum, however, as given by equations 9.86, is composed of two parts, the orbital angular momentum $\underline{L}$ and the spin part $\underline{S}$. As we will see in the discussion of the interaction of radiation with matter, electric dipole transitions change charged particle angular momentum by one unit of $\hbar$, which is carried away by the photon. Electric quadrupole transitions, however, changes the charged particle angular momentum by two units of $\hbar$, or leaves it unchanged, emitting or absorbing a photon of one unit of orbital angular momentum, resulting in a total angular momentum of zero or two units of $\hbar$.[#] Transitions corresponding to higher moments can result in (single) photons of yet larger total angular momentum. These higher angular momentum single photon states can be obtained by considering the $e^{i\underline{k}\cdot\underline{x}}$ plane waves in our original expansion of the vector potential (equation 9.6), as a superposition of definite angular momentum wavefunctions. This results in a different labeling of the normal modes, from $(k_x, k_y, k_z;\sigma)$, corresponding to the $e^{i\underline{k}\cdot\underline{x}}$ spatial functions, to $(k,\lambda,\mu;\sigma)$, corresponding to the spatial functions of definite angular momentum

$$j_\lambda(kr)\, Y^*_{\lambda,\mu}(\theta_{\underline{k}},\varphi_{\underline{k}})\, Y_{\lambda,\mu}(\theta,\varphi) \quad ; \qquad \hat{e}_{\underline{k}} = (\theta_{\underline{k}},\varphi_{\underline{k}}) \quad , \quad \hat{e}_{\underline{x}} = (\theta,\varphi)$$

where j_λ is the λ^{th} spherical Bessel function and $Y_{\lambda,\mu}$ are the (λ,μ) spherical harmonics. See problem 7.18.

9.10 The vector potential in terms of the circular polarization operators.

We can invert equations 9.90 to express the linear polarization operators in terms of the circular polarization operators. This yields,

$$a_{\underline{k},1} = \frac{1}{\sqrt{2}}\left(a_{\underline{k},R} + a_{\underline{k},L}\right) \tag{9.98a}$$

$$a^\dagger_{\underline{k},1} = \frac{1}{\sqrt{2}}\left(a^\dagger_{\underline{k},R} + a^\dagger_{\underline{k},L}\right) \tag{9.98b}$$

$$a_{\underline{k},2} = \frac{i}{\sqrt{2}}\left(a_{\underline{k},R} - a_{\underline{k},L}\right) \tag{9.98c}$$

$$a^\dagger_{\underline{k},2} = \frac{-i}{\sqrt{2}}\left(a^\dagger_{\underline{k},R} - a^\dagger_{\underline{k},R}\right) \tag{9.98d}$$

update

[#] $j = \ell - s = 0$, or $j = \ell + s = 2$. See discussion on addition of angular momentum, section 10.2

if we substitute these into the operator expression for $\underline{A}(\underline{x}, t)$, as given by equation 9.40, we have,

$$\underline{A}(\underline{x}, t) = \left(\frac{\hbar}{2\varepsilon_0 V}\right)^{1/2} \sum_{k} \omega_{\underline{k}}^{-1/2} e^{i\underline{k}\cdot\underline{x}} \sum_{\sigma=R,L} (a_{\underline{k},\sigma} - a^{\dagger}_{\underline{k},\sigma})\, \hat{e}_{\underline{k},\sigma} \quad , \tag{9.99}$$

where the unit vectors $\hat{e}_{\underline{k},\sigma}$ $(\sigma = R, L)$ are defined by

$$\hat{e}_{\underline{k},R} \equiv \frac{1}{\sqrt{2}} (\hat{e}_{\underline{k},1} + i\,\hat{e}_{\underline{k},2}) \tag{9.100a}$$

and

$$\hat{e}_{\underline{k},L} \equiv \frac{1}{\sqrt{2}} (\hat{e}_{\underline{k},1} - i\,\hat{e}_{\underline{k},2}) \quad , \tag{9.100b}$$

as required by the algebra.

The new complex unit vectors, corresponding to a wavevector $\underline{k}$, can be seen to be orthonormal, i.e.

$$\hat{e}_{\underline{k},\sigma'} \cdot \hat{e}_{\underline{k},\sigma} = \delta_{\sigma',\sigma} \tag{9.101}$$

and have the properties

$$\hat{e}^{*}_{\underline{k},R} = \hat{e}_{\underline{k},L} \quad , \qquad \hat{e}^{*}_{\underline{k},L} = \hat{e}_{\underline{k},R} \tag{9.102}$$

and also

$$\hat{e}_{-\underline{k},R} = -\,\hat{e}_{\underline{k},L} \quad , \qquad \hat{e}_{-\underline{k},L} = -\,\hat{e}_{\underline{k},R} \quad , \tag{9.103}$$

or, combining 9.102 and 9.103, we have

$$\hat{e}_{-\underline{k},\sigma} = -\,\hat{e}^{*}_{\underline{k},\sigma} \quad . \tag{9.104}$$

We recognize then to be the (classical) complex unit vectors for right and left circular polarization, and could have been used for the Fourier decomposition of the vector potential at the outset, i.e.

$$\underline{A}(\underline{x}, t) = \sum_{\underline{k},\sigma} e^{i\underline{k}\cdot\underline{x}} A_{\underline{k},\sigma}(t)\, \hat{e}_{\underline{k},\sigma} \quad , \tag{9.105}$$

where, using the orthonormality of the $\hat{e}_{\underline{k},\sigma}$, we have the inversion formula, i.e.

$$A_{\underline{k},\sigma} = \frac{1}{V} \int_{V} [\hat{e}^{*}_{\underline{k},\sigma} \cdot \underline{A}(\underline{x}, t)]\; e^{-i\underline{k}\cdot\underline{x}}\, d^3\underline{x} \quad . \tag{9.106}$$

Note that, as opposed to the plane polarization unit vectors, both the circular polarization unit vectors behave the same way as $\underline{k} \to -\underline{k}$. This property is also shared by the two circular polarization $\underline{k}$-components of $\underline{A}(\underline{x}, t)$, i.e.

update

$$A_{-\underline{k},\sigma} = -\,A^{*}_{\underline{k},\sigma} \quad , \qquad \sigma = R, L \quad . \tag{9.107}$$

Note also that, as a result, the annoying factor of $(-1)^{\ell}$ in the operator expression for $\underline{A}(\underline{x}, t)$ has disappeared[#]. The necessity for it was dictated by the requirement to keep the coordinate system right handed as $\underline{k} \to -\underline{k}$ (see equation 9.11 and related discussion. Nevertheless, since the choice of which direction to call $\hat{e}_{\underline{k},1}$ (which then gave $\hat{e}_{\underline{k},2}$) was arbitrary, it seemed unsymmetrical to make $\hat{e}_{\underline{k},1}$ antisymmetric under an inversion of $\underline{k}$ $(\underline{k} \to -\underline{k})$, leaving $\hat{e}_{\underline{k},2}$ symmetric. This asymmetry, of course, does not plague the circular polarization unit vectors and all the resulting formulae. Things always look better in their natural coordinates!

9.11 Photon modes.

It should be clear by now that there is no such thing as *the* photon. We have been able to decompose the vector potential into a variety of normal modes:

(i) plane polarization, sine and cosine standing plane wave spatial modes

$$[\sin(\underline{k}\cdot\underline{x})\ ,\ \cos(\underline{k}\cdot\underline{x})]\ \hat{e}_{\underline{k},\ell}\ ;\ (\underline{k} > 0,\ \ell=1,2)\ , \tag{9.108a}$$

(ii) plane polarization, travelling plane wave modes

$$e^{i\underline{k}\cdot\underline{x}}\ \hat{e}_{\underline{k},\ell}\ ;\quad (\text{all }\ \underline{k},\ \ell=1,2)\ , \tag{9.108b}$$

(iii) circular polarization, travelling plane wave modes

$$e^{i\underline{k}\cdot\underline{x}}\ \hat{e}_{\underline{k},\sigma}\ ;\quad (\text{all }\ \underline{k},\ \sigma=R,L)\ , \tag{9.108c}$$

and finally,

(iv) circular polarization, radial wave modes

$$j_{\lambda}(kr)Y^{*}_{\lambda,\mu}(\theta_{\underline{k}},\varphi_{\underline{k}})Y_{\lambda,\mu}(\theta,\varphi)\ \hat{e}_{\underline{k},\sigma}\ .\quad (\text{all }\ \underline{k},\ \sigma=R,L) \tag{9.108d}$$

This is of course by no means an exhaustive list, since the energy (Hamiltonian) is always separable into any set of appropriate normal modes, in terms of which we should always recover our familiar uncoupled harmonic oscillators. In any such set, "the photon" of a particular mode is to be understood as representing one unit of excitation of the corresponding harmonic oscillator. The choice of the set of modes (basis states) is of course arbitrary, and one which might as well be made to suit the special properties of the particular problem

update

[#] Compare equations 9.40 and 9.99.

of interest. Plane wave modes are best suited in describing phenomena with well-defined (eigenstates of) photon linear momentum. Circular polarization modes are the natural choice in describing phenomena involving well-defined photon spin (component of photon angular momentum along linear momentum). Radial wave modes correspond to photon states of well-defined angular momentum, etc. Finally, these decompositions should be regarded as choices of convenience, since they all correspond to complete sets of spatial and polarization functions, excitations (photons) in any one set being expressible as linear combinations of excitations (photons) in any other.

Problems

9.1 Derive equation 9.42 for the time rate of change of the vector potential Fourier component along the mode $(\underline{k}, \ell)$.

9.2 Compute the expectation values of

$$\mathcal{E}^2(\underline{x}, t) = \underline{\mathcal{E}}(\underline{x}, t) \cdot \underline{\mathcal{E}}(\underline{x}, t) \ , \qquad \mathcal{B}^2(\underline{x}, t) = \underline{\mathcal{B}}(\underline{x}, t) \cdot \underline{\mathcal{B}}(\underline{x}, t)$$

for an electromagnetic field in the state $|\psi\rangle$ of example 9.1.

9.3 Using the results of the preceding problem compute the classical energy in the field using equation 9.1. Compare with the quantum mechanical result.

9.4 Compute $\langle\underline{\mathcal{E}}(\underline{x}, t)\rangle$, $\langle\underline{\mathcal{B}}(\underline{x}, t)\rangle$, $\langle\mathcal{E}^2(\underline{x}, t)\rangle$ and $\langle\mathcal{B}^2(\underline{x}, t)\rangle$ for an electromagnetic field in the state

$$|0, \ldots\ldots, 0, 1_{\underline{k}, \ell}, 0, \ldots\ldots\rangle = a^{\dagger}_{\underline{k}, \ell}|0\rangle \ .$$

Discuss the results.

9.5 Find the energy and momentum of the state of the electromagnetic field given by

$$b^{\dagger}_{\underline{k}, j}\,|0\rangle \ .$$

9.6 Using the expansions of the running wave mode ladder operators $a_{\underline{k}, \ell}$, $a^{\dagger}_{\underline{k}, \ell}$ in terms of the standing wave mode ladder operators $b_{\underline{k}, j}$, $b^{\dagger}_{\underline{k}, j}$, express the states

$$a_{\underline{k}, 1}\,|0\rangle \ , \quad \frac{1}{\sqrt{2}}\,(a^{\dagger}_{\underline{k}, 1})^2\,|0\rangle$$

as a superposition of the standing wave mode eigenstates of the Hamiltonian of equation 9.30.

9.7 Prove equation 9.88 by substituting equations 9.40 and 9.43 in equation 9.86c.

9.8 Prove equation 9.89.

10. SPIN

In our discussion of Quantum Mechanics thus far, we have described particles in terms of their position coordinates

$$\underline{x} = (x, y, z) \qquad (10.1a)$$

and associated momenta

$$\underline{p} = (p_x, p_y, p_z) \ . \qquad (10.1b)$$

The wavefunctions could then be expressed in terms of $\underline{x}$, i.e.

$$\psi = \psi(x, y, z;t) \ , \qquad (10.2a)$$

or $\underline{p}$, i.e.

$$\Psi = \Psi(p_x, p_y, p_z;t) \ , \qquad (10.2b)$$

the two being equivalent through the Fourier transform (see chapter 2). It should be realized that equation 10.2 tacitly assumes that the particles can be completely described in terms of the three coordinate variables and, in particular, possess no additional degrees of freedom. This is almost always an invalid assumption and it is remarkable, in a sense, that so many correct conclusions can be derived without consideration of the additional degrees of freedom that have been ignored. The reason for our success thus far (apart from the careful selection of topics!) is that oftentimes the energies available in the interactions are small enough, or the types of interactions are such so that the internal degrees of freedom do not participate in the phenomena of interest. A notable example along these lines is our discussion of diatomic molecules in which the two atoms were treated as point particles. It is clear that this is an approximation, which however is justified if the energies of interest are not large enough to alter the electronic configuration of the participant atoms. How an increase in the energy would engage such degrees of freedom is illustrated by the behavior of a diatomic molecule as a function of temperature (section 7.4.1) where,

(i) if the temperature is below the characteristic rotational temperature θ_r, the whole molecule acts like a point particle with eigenstates expressed in terms of the wavevectors (momenta) of the center of mass, i.e.

$$|\psi\rangle = |K_x, K_y, K_z\rangle \quad \text{for} \quad T \ll \theta_r$$

and a resulting heat capacity given by $(3/2)k_B T$,

(ii) if the temperature is above the rotational temperature θ_r but below the vibrational temperature θ_v, the molecule acts like a rigid rotator and the eigenfunctions now require two more quantum numbers, namely the eigenvalues of L^2 and L_z, i.e.

$$|\psi\rangle = |K_x, K_y, K_z, \ell, m\rangle \qquad \text{for} \quad \theta_r < T \ll \theta_v$$

and a resulting heat capacity given by $(5/2)k_B T$,

(iii) if the temperature is above the vibrational temperature but below a characteristic electronic temperature θ_e[#], the molecule will now also engage the vibrational degree of freedom and the eigenstates now need <u>six</u> quantum numbers to be specified, i.e.

$$|\psi\rangle = |K_x, K_y, K_z, v, \ell, m\rangle \qquad \text{for} \quad \theta_v < T \ll \theta_e$$

and a resulting heat capacity of $(7/2)k_B T$.

This process, of course, does not end here. At higher energies yet we would have to consider the <u>internal</u> degrees of freedom of each of the atoms (electronic excitation) with all the complications.

In a similar vein, we have ignored any internal degrees of freedom in the α-particles, which are found to obey the Coulomb scattering law (Rutheford formula) of example 8.7, when we treated them as point particles describable by the plane wave states of equations 8.130. A notable exception to these simplifications was our discussion of the electromagnetic field where we found that each mode is characterized, in addition to the three wavevector coordinates, by an additional degree of freedom namely the two possible states of circular polarization (parallel or anti-parallel to the respective wavevector). In that case, a particular mode requires <u>four</u> variables to be specified, i.e.

$$(k_x, k_y, k_z, \sigma)$$

where $\sigma = \pm 1$ for right and left circularly polarized light respectively. This description is exact[†] and no additional degrees of freedom are found necessary to render the description of photons complete. Interestingly enough, most particles that are considered elementary also possess a degree of freedom in addition to the three position coordinates (or conjugate momenta). In much the same way as the photon , they are found to possess <u>intrinsic angular momentum of</u>

[#] Say, $\theta_e \simeq -E_R/k_B$ for a hydrogen molecule.

[†] ... provided, of course, that the energies are not so high as to allow processes such as pair production.

of fixed magnitude, called spin. The additional degree of freedom is then associated with the projection of the spin angular momentum along a particular axis of quantization, say, the z-axis. The *spin* angular momentum $\underline{S}$ must then be considered along with the *orbital* angular momentum $\underline{L}$ in describing the *total* angular momentum $\underline{J}$, given by the vector sum

$$\underline{J} = \underline{L} + \underline{S} \ . \tag{10.3}$$

10.1 The total angular momentum operator.

The total angular momentum operator is also a vector and as such can be resolved along three orthogonal directions, i.e.

$$\underline{J} = J_x \hat{e}_x + J_y \hat{e}_y + J_z \hat{e}_z \ . \tag{10.4}$$

The three components J_x, J_y, J_z also obey the angular momentum commutation relations

$$[J_x, J_y] = i\hbar J_z \tag{10.5a}$$

$$[J_y, J_z] = i\hbar J_x \tag{10.5b}$$

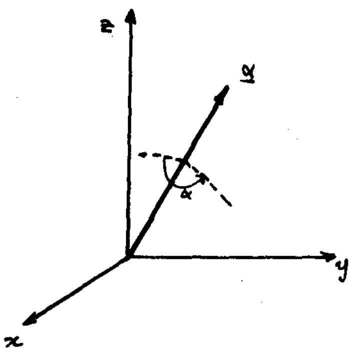
and

$$[J_z, J_x] = i\hbar J_y \tag{10.5c}$$

which were derived for orbital angular momentum, i.e. equation 6.48, from the defining equation 6.46 for $\underline{L}$, on the basis of the commutation relations of $\underline{x}$ and $\underline{p}$. They are, however, more fundamental and survive generalizations of angular momentum to include spin. They can be shown to be a direct consequence of the association of the angular momentum operator $\underline{J}/\hbar$ as the generator of rotations. In particular, the operator

$$\mathcal{R}(\underline{\alpha}) = e^{-i\underline{\alpha}\cdot\underline{J}/\hbar} \tag{10.6}$$

performs a rotation of angle $\alpha = |\underline{\alpha}|$ about an axis along $\underline{\alpha}$.



If $\mathcal{R}(\underline{a})$ represents such a rotation, the operator $\underline{J}$ must obey the commutation relations of equation 10.5 which can be abbreviated by the "vector" equation[#].

$$\underline{J} \times \underline{J} = i\hbar \underline{J} . \qquad (10.7)$$

Most of the results derived for the orbital angular momentum $\underline{L}$ were derived from the commutation relations and so also apply to the total angular momentum $\underline{J}$. In particular, the magnitude squared of $\underline{J}$, i.e.

$$J^2 = J_x^2 + J_y^2 + J_z^2 \qquad (10.8)$$

commutes with each of the three components of $\underline{J}$, or

$$[J^2, \underline{J}] = 0 \qquad (10.9)$$

(see equation 6.51 and related discussion).

We can also define raising and lowering operators, in complete analogy to the orbital angular momentum case, i.e.

$$J_{\pm} = J_x \pm i J_y , \qquad (10.10)$$

which satisfy the commutation relations

$$[J_z, J_{\pm}] = \pm \hbar J_{\pm} \qquad (10.11a)$$

and

$$[J_+, J_-] = 2\hbar J_z , \qquad (10.11b)$$

in terms of which we can also express J^2, i.e.

$$J^2 = J_+J_- + J_z(J_z - \hbar) = J_-J_+ + J_z(J_z + \hbar) . \qquad (10.12)$$

We now select the operators J^2 and J_z, as in the case of orbital angular momentum, and seek eigenstates and eigenvalues such that

$$J^2 |j, m_j\rangle = \hbar^2 j(j+1) |j, m_j\rangle \qquad (10.13a)$$

and

$$J_z |j, m_j\rangle = \hbar m_j |j, m_j\rangle . \qquad (10.13b)$$

Again, as in the case of orbital angular momentum we find that we must have

$$-j \le m_j \le j \qquad (10.14)$$

(see equation 6.69 and related discussion), and that

$$J_{\pm} |j, m_j\rangle = \hbar \sqrt{j(j+1) - m_j(m_j \pm 1)} \; |j, m_j \pm 1\rangle , \qquad (10.15)$$

i.e. operation with J_+ on a state $|j, m_j\rangle$, generates a state proportional to $|j, m_j + 1\rangle$, while operation with J_- on a state $|j, m_j\rangle$, generates a state

---

[#] Note that no ordinary vector obeys this equation.

proportional to $|j, m_j - 1\rangle$. Therefore by repeated application we have

$$(J_+)^{m'} |j, m_j\rangle \propto |j, m_j + m'\rangle \qquad (10.16a)$$

while

$$(J_-)^{m''} |j, m_j\rangle \propto |j, m_j - m''\rangle \ , \qquad (10.16b)$$

where m' and m'' are integers. In order to satisfy equation 10.14, however, the two sequences must terminate and that will only happen (see equation 10.15) if, for some m' and m'',

$$m_j + m' = j \qquad (10.17a)$$

and

$$m_j - m'' = -j \ . \qquad (10.17b)$$

Eliminating m_j from these equations, we then have

$$m' + m'' = 2j \ , \qquad (10.18)$$

which admits both integral and half-integral solutions for j. Consequently, j can take on the values

$$j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \ \ldots . \qquad (10.19)$$

with correspondingly integer or half-integer values for m_j. The total angular momentum eigenstates $|j, m_j\rangle$ are then,

$$\begin{array}{c|ccccccccc}
 & m_j \longrightarrow & & & & & & & & \\
 & & & & & |0,0\rangle & & & & \\
 & & & & |1/2,-1/2\rangle & & |1/2,1/2\rangle & & & \\
j & & & |1,-1\rangle & & |1,0\rangle & & |1,1\rangle & & \\
\downarrow & & |3/2,-3/2\rangle & & |3/2,-1/2\rangle & & |3/2,1/2\rangle & & |3/2,3/2\rangle & \\
 & |2,-2\rangle & & |2,-1\rangle & & |2,0\rangle & & |2,1\rangle & & |2,2\rangle \\
 & & & & & \vdots & & & & \\
 & & & & & \text{etc.} & & & &
\end{array}$$

Note that this derivation for the possible values of the angular momentum quantum number j is identical up to equation 10.17a with the corresponding analysis for the orbital angular momentum quantum number ℓ. The difference there, however, was that since $L_z = -i\hbar \frac{\partial}{\partial\varphi}$ we knew, ahead of time that the eigenvalues of L_z had to be integer multuples of $\hbar$.

10.2 Addition of angular momenta. Clebsch-Gordan coefficients.

Consider now a system composed of two parts represented by independent coordinates and possessing an angular momentum $\underline{J_1}$, $\underline{J_2}$ respectively. Note that the independence of the coordinates of systems {1} and {2} requires that

$$[J_{1k}, J_{2\ell}] = 0 \quad \text{for} \quad k, \ell = 1, 2, 3 \ . \qquad (10.20)$$

The total angular momentum of the system is then given by the vector sum

$$\underline{J} = \underline{J_1} + \underline{J_2} \ , \tag{10.21}$$

with corresponding eigenstates $|j, m\rangle$ such that

$$J^2 |j, m\rangle = \hbar^2 j(j+1)|j, m\rangle \tag{10.22a}$$

and

$$J_z|j, m\rangle = \hbar m|j, m\rangle \ . \tag{10.22b}$$

We also have, however, that since the system is composed of two independent systems, with corresponding angular momenta, the eigenstates

$$|j_1, j_2; m_1, m_2\rangle \equiv |j_1, m_1\rangle|j_2, m\rangle \ , \tag{10.23}$$

which form a complete set for the combination of the two subsystems, can serve as a basis for J^2 and J_z. We now note that, since

$$\begin{aligned} J_z|j_1, j_2; m_1, m_2\rangle &= (J_{1_z} + J_{2_z})|j_1, m_1\rangle|j_2, m_2\rangle \\ &= \hbar(m_1 + m_2)|j_1, m_1\rangle|j_2, m_2\rangle \ , \end{aligned}$$

the states $|j_1, j_2; m_1, m_2\rangle$ must be eigenstates of J_z corresponding to the eigenvalue

$$m = m_1 + m_2 \ . \tag{10.24}$$

Note also that the operator

$$J^2 = J_1^2 + J_2^2 + 2\underline{J_1} \cdot \underline{J_2} = J_1^2 + J_2^2 + 2(J_{1_x}J_{2_x} + J_{1_y}J_{2_y} + J_{1_z}J_{2_z})$$

commutes with the operators J_1^2 and J_2^2 and can therefore be specified at the same time as the square of the angular momentum of each subsystem. The states $|j_1, j_2; m_1, m_2\rangle$, however, are not eigenstates of J^2, since the scalar product $\underline{J_1} \cdot \underline{J_2}$ involves the x and y components of $\underline{J_1}$ and $\underline{J_2}$ which are not diagonal in this basis. Using the fact, however, that j, j_1, j_2 may be simultaneously specified, we seek linear combinations of the states $|j_1, j_2; m_1, m_2\rangle$, which are eigenstates of J^2 (and therefore also of J_z). In particular we want to find the coefficients $\langle j_1, j_2; m_1, m_2|j, m\rangle$ of the eigenstates

$$|j_1, j_2, j, m\rangle = \sum_{m_1, m_2} |j_1, j_2; m_1 m_2\rangle\langle j_1, j_2; m_1, m_2|j, m\rangle^{\#} \ . \tag{10.25}$$

[#] The notation is unfortunately confusing. The convention adopted here is that a state vector <u>with</u> a semi-colon separating the first pair from the second pair of indeces refers to the (tensor) product states of equation 10.23, whereas a state vector <u>without</u> a semi-colon refers to the eigenstates of J^2 and J_z resulting from the linear superposition of a given j_1, j_2 pair.

These coefficients are called vector addition coefficients or Clebsch-Gordan coeficients.

These coefficients must be zero unless $m = m_1 + m_2$ (equation 10.24) and therefore the summation is actually over a single index, i. e.

$$|j_1, j_2, j, m\rangle = \sum_{m_2} |j_1, j_2; m - m_2, m_2\rangle \langle j_1, j_2; m - m_2, m_2 | j, m\rangle \quad . \qquad (10.26)$$

It also follows that the maximum value of m is given by

$$\max\{m\} = \max\{m_1 + m_2\} = j_1 + j_2$$

and therefore also

$$\max\{j\} = j_1 + j_2 \quad . \qquad (10.27)$$

The next lowest value of m is therefore $j_1 + j_2 - 1$, which is expressible as the linear combination of the states

$$|j_1, j_2; j_1 - 1, j_2\rangle \quad \text{and} \quad |j_1, j_2; j_1, j_2 - 1\rangle \ .$$

It can be verified that the two orthonormal combinations of these two states correspond to the two values for j,

$$j = j_1 + j_2 \quad \text{and} \quad j = j_1 + j_2 - 1 \ .$$

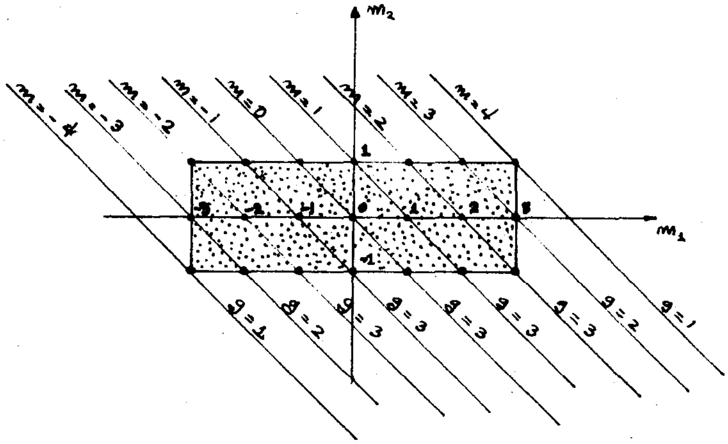
The next lower value for m is given by $j_1 + j_2 - 2$, which must be a linear combination of all the states which satisfy $m = m_1 + m_2$ or

$$|j_1, j_2; j_1 - 2, j_2\rangle \ , \quad |j_1, j_2; j_1 - 1, j_2 - 1\rangle \quad \text{and} \quad |j_1, j_2; j_1, j_2 - 2\rangle \ .$$

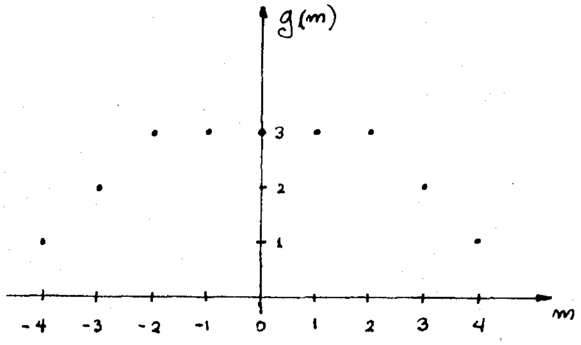
It can again be verified that the three orthogonal linear combinations of these states correspond to three values for j, i. e.

$$j_1 + j_2 , \quad j_1 + j_2 - 1 \quad \text{and} \quad j_1 + j_2 - 2 \ .$$

This process can be represented pictorially as follows. Consider the plane m_1, m_2. Each point on this plane corresponding to integer values of the coordinates represents a state of a given m. An example for which $j_1 = 3$ and $j_1 = 1$ is sketched below.



A line of constant m is evidently a line at - 45° given by $m_2 = m - m_1$. The largest (and smallest) m has a degeneracy $g_{j_1,j_2}(m)$ of unity (occurs only once) and corresponds to the state with an angular momentum $j = j_1 + j_2$. The next lowest m $(j_1 + j_2 - 1)$ has a degeneracy of two and must therefore involve an additional (orthogonal) state corresponding to $j = j_1 + j_2 - 1$ which possesses such a z-projection. The next lowest m has a value $j_1 + j_2 - 2$ and (in the example sketched above) a degeneracy of 3, involving thereby a third value for j. This process ends when $m_1 = -j_1$ or $m_2 = -j_2$ at which time no additional increase in degeneracy is observed,



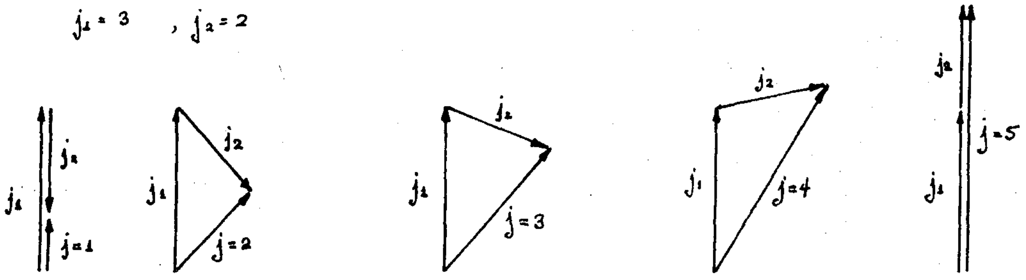
Consequently, the lowest value of j required by this sequence is

$$\min\{j\} = |j_1 - j_2| \ . \tag{10.28}$$

Combining equations 7.27 and 7.28 we arrive at the <u>triangular inequality</u>

$$|j_1 - j_2| \le j \le j_1 + j_2 \ , \tag{10.29}$$

since the possible values of j correspond to all possible integer solutions for the length of the third sides of triangles whose other two sides are given by the integers j_1 and j_2.



Note that the degeneracy of an eigenvalue j is given by (2j+1), corresponding to the (2j+1) possible values of m, and that

$$\sum_{j=|j_1-j_2|}^{j_1+j_2} (2j+1) = (2j_1+1)(2j_2+1) \ , \tag{10.30}$$

consistently with the number of independent states given by equation 10.23.

From the preceding discussion, we see that

$$\langle j_1, j_2; m_1, m_2 | j, m \rangle = 0 \quad \text{if} \quad m \neq m_1 + m_2 \ . \tag{10.31}$$

Applying J_+ and J_- to equation 10.25 we also find the recursion relations

$$\begin{aligned} &[(j\pm m')(j\pm m'+1)]^{1/2} \langle j_1, j_2; m_1, m_2 | j, m'\mp 1\rangle = \\ &= [(j_1\mp m_1)(j_1\pm m_1+1)]^{1/2} \langle j_1, j_2; m_1\pm 1, m_2 | j, m'\rangle \\ &\quad + [(j_2\mp m_2)(j_2\pm m_2+1)]^{1/2} \langle j_1, j_2; m_1, m_2\pm 1 | j, m'\rangle \ . \end{aligned} \tag{10.32}$$

From these equations, we can generate the required coefficients. By way of example, if $m_1 = j_1$ and $m' = j$ in the left hand side of 10.32, then (only the top sign survives) and we have

$$\begin{aligned} &(2j)^{1/2} \langle j_1, j_2; j_1, j-j_1-1 | j, j-1\rangle = \\ &= [(j_2-j+j_1+1)(j_2+j-j_1)]^{1/2} \langle j_1, j_2; j_1, j-j_1 | j, j\rangle \ , \end{aligned} \tag{10.33}$$

where we have used equation 10.31, i.e.

$$m = m'-1 = m_1+m_2 \Rightarrow m_2 = m'-m_1-1 = j-j_1-1 \ .$$

Consequently, $\langle j_1, j_2; j_1, j-j_1-1 | j, j-1\rangle$ can be computed if $\langle j_1, j_2; j_1, j-j_1 | j, j\rangle$ is known. Setting $m_1 = j_1$ and $m' = j-1$, the lower sign of 10.32 yields a relationship between

$$\langle j_1, j_2; j_1, j-j_1 | j, j\rangle \text{ and } \langle j_1, j_2; j_1-1, j-j_1 | j, j-1\rangle, \ \langle j_1, j_2; j_1, j-j_1-1 | j, j-1\rangle \ ,$$

which, combined with 10.33, yields one more coefficient in terms of $\langle j_1, j_2; j_1, j-j_1 | j, j\rangle$, and so on and so forth. Finally, the process is closed by the requirement that the states are normalized and that the Clebsch-Gordan coefficients are real numbers[#]. The orthogonality and normalization conditions require that

$$\sum_{m_1, m_2} \langle j_1, j_2; m_1, m_2 | j_1, j_2, j, m\rangle \langle j_1, j_2; m_1, m_2 | j_1, j_2, j', m'\rangle = \delta_{mm'}\delta_{jj'}$$

and

$$\sum_{j,m} \langle j_1, j_2; m_1, m_2 | j_1, j_2, j, m\rangle \langle j_1, j_2; m_1', m_2' | j_1, j_2, j, m\rangle = \delta_{m_1m_1'}\delta_{m_2m_2'} \ .$$

[#] Evidently, if one of them is real, the rest of them are also real by virtue of equation 10.32.

These relationships express the fact that the matrix of the coefficients, form a unitary matrix, which (since it is real) also performs the inverse transformation, i.e.

$$|j_1, j_2; m_1, m_2\rangle = \sum_{j,m} |j_1, j_2, j, m\rangle\langle j_1, j_2; m_1, m_2 | j_1, j_2, j, m\rangle \qquad (10.35)$$

The Clebsch-Gordan coefficients for the common cases of $j_2 = 1/2$ and $j_2 = 1$ are listed below.

Table 10.1a Clebsch-Gordan coefficients for $j_2 = 1/2$, $j_1 \geq j_2$.

	$m_2 = -1/2$	$m_2 = 1/2$
$j = j_1 - 1/2$	$\left(\frac{j_1+m+1/2}{2j_1+1}\right)^{1/2}$	$-\left(\frac{j_1-m+1/2}{2j_1+1}\right)^{1/2}$
$j = j_1 + 1/2$	$\left(\frac{j_1-m+1/2}{2j_1+1}\right)^{1/2}$	$\left(\frac{j_1+m+1/2}{2j_1+1}\right)^{1/2}$

Table 10.1b Clebsch-Gordan coefficients for $j_2 = 1$, $j_1 \geq j_2$.

	$m_2 = -1$	$m_2 = 0$	$m_2 = 1$
$j = j_1 - 1$	$\sqrt{\frac{(j_1+m+1)(j_1+m)}{2j_1(2j_1+1)}}$	$-\sqrt{\frac{(j_1-m)(j_1+m)}{j_1(2j_1+1)}}$	$\sqrt{\frac{(j_1-m)(j_1-m+1)}{2j_1(2j_1+1)}}$
$j = j_1$	$\sqrt{\frac{(j_1-m)(j_1+m+1)}{2j_1(j_1+1)}}$	$\sqrt{\frac{m^2}{j_1(j_1+1)}}$	$-\sqrt{\frac{(j_1+m)(j_1-m+1)}{2j_1(j_1+1)}}$
$j = j_1 + 1$	$\sqrt{\frac{(j_1-m)(j_1-m+1)}{(2j_1+1)(2j_1+2)}}$	$\sqrt{\frac{(j_1-m+1)(j_1+m+1)}{(2j_1+1)(j_1+1)}}$	$\sqrt{\frac{(j_1+m)(j_1+m+1)}{(2j_1+1)(2j_1+2)}}$

update

The Clebsch-Gordan coefficients are also useful in expressing products of the spherical harmonics. In particular, we have

$$Y_{\ell_2,m_2}(\underline{\Omega})\, Y_{\ell_1,m_1}(\underline{\Omega}) = \sum_{\ell=|\ell_2-\ell_1|}^{\ell_2+\ell_1} \left[\frac{(2\ell_2+1)(2\ell_1+1)}{4\pi(2\ell+1)}\right]^{1/2} \langle \ell_2, \ell_1; 0, 0 | \ell, 0\rangle\langle \ell_2, \ell_1; m_2, m_1 | \ell, m\rangle Y_{\ell,m}(\underline{\Omega}) \qquad (10.36a)$$

where $\underline{\Omega} = (\theta, \varphi)$ and $m = m_1 + m_2$, which can be used in turn to evaluate integrals of products of three spherical harmonics, i.e.

$$I(\ell_3, m_3 \,|\, \ell_2, m_2; \ell_1, m_1) \equiv \int_{\underline{\Omega}} Y^*_{\ell_3, m_3}(\underline{\Omega})\, Y_{\ell_2, m_2}(\underline{\Omega})\, Y_{\ell_1, m_1}(\underline{\Omega})\, d\underline{\Omega} \tag{10.36b}$$

where $d\underline{\Omega} = \sin\theta\, d\theta d\varphi$, which we can compute from equation 10.36a and are given by[#]

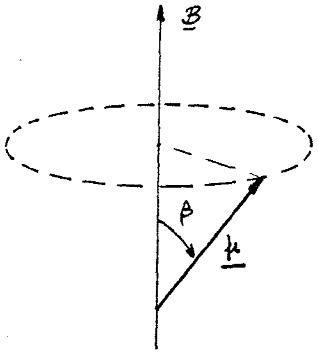
$$I(\ell_3, m_3 \,|\, \ell_2, m_2; \ell_1, m_1) = \left[\frac{(2\ell_2+1)(2\ell_1+1)}{4\pi(2\ell_3+1)}\right]^{1/2} \langle \ell_2, \ell_1; 0, 0 \,|\, \ell_3, 0\rangle \langle \ell_2, \ell_1; m_2, m_1 \,|\, \ell_3, m_3\rangle \quad . \tag{10.36c}$$

These integrals occur quite frequently in perturbation calculations. See, for example, equations 8.89 and 8.92 of example 8.4.

It should also be noted, before leaving this section, that since *orbital* angular momentum contributes integer quantum numbers, of necessity (see equations 10.24 and 10.29), *the half-integer values must be contributed by the spin angular momentum*.

10.3 The spin of the electron.

In 1922 Stern and Gerlach[†] in a set of experiments with striking results measured the possible values of the magnetic moment of silver atoms by sending a beam through a magnetic field with a strong gradient. If the atoms possess a magnetic moment $\underline{\mu}$, there is an energy associated with their orientation in a magnetic field given by (see Appendix F)



$$U = -\underline{\mu}\cdot\underline{\mathcal{B}} = -\mu\mathcal{B}\cos\beta \tag{10.37}$$

(energy is lowest when $\underline{\mu}$ and $\underline{\mathcal{B}}$ are parallel). If the magnetic field is non-

---

[#] See for example C. Claude-Tannoudji, B. Diu and F. Laloë, *Quantum Mechanics* (1977), volume II, complement C_X.

[†] see R. Eisberg and R. Resnick, *Quantum Physics*, section 8-3 for a historical account.

update #2

uniform there is a net force acting on such a dipole given by

$$\underline{F} = -\nabla U = \mu \cos\beta \, \nabla\mathcal{B} \tag{10.38}$$

(for a magnetic field along a fixed particular direction). If it is so arranged that the gradient of $\mathcal{B}$ is perpendicular to the beam, the vertical deflection measures the magnitude of the force and, by extension, the projection of the magnetic moment on the direction of the magnetic field, or $\mu\cos\beta$. This analysis would predict a range of displacements corresponding to all possible relative orientations, subtended angles β, i.e.

$$-\mu \leq \mu\cos\beta \leq \mu$$

in marked contradiction to the experimental results, which yielded two well-defined, symmetric (about zero deflection) displacements, corresponding to $\pm\mu$, with no intermediate values. These results may be taken as in support of the quantization of the z-component of angular momentum, since the orbital motion of a charge q_c in a magnetic field $\underline{\mathcal{B}}$ gives rise to an interaction

$$H' = -\frac{q_c}{2m} \underline{\mathcal{B}} \cdot \underline{L} \tag{10.39}$$

(see example 6.1 and problems 7.5, 7.7) or if $\underline{\mathcal{B}} = \mathcal{B}\hat{e}_z$ we have for an electron ($q_c = -e$)

$$H' = \frac{e}{2m} \mathcal{B} L_z = \mu_B \mathcal{B} m_\ell \,, \tag{10.40a}$$

where

$$\mu_B \equiv \frac{e\hbar}{2m} = 9.27 \times 10^{-24} \text{ amp} \cdot \text{m}^2 \tag{10.40b}$$

is the Bohr magneton (9.27×10^{-21} erg/gauss in c.g.s) and m_ℓ is the z-component orbital angular momentum quantum number. Even though, however, this is in qualitative agreement, the quantitative description of the behavior, as given by equation 10.40, does not agree with the experimental results. The situation, in the case of the silver atoms is actually quite complicated because there are several electrons involved. It becomes clearer if hydrogen atoms in their ground state are used in the Stern-Gerlach apparatus (Phipps and Taylor, 1927)[#], for which equation 10.40 predicts no effect since, for the ground state, $\ell = m_\ell = 0$. The ground state hydrogen atoms, however, also yielded two distinct spots leaving no doubt that a magnetic moment can arise for reasons other than can be accounted for by orbital angular momentum.

These difficulties (and numerous others) were removed by the proposal put

---

[#] Notably, after the proposal by Uhlenbeck and Goudsmit that the electron possesses spin.

forth by Uhlenbeck and Goudsmit in 1925 (interestingly enough, on the basis of spectroscopic data, quite independently of the Stern-Gerlach experiments which were not immediately appreciated) that the electron possesses intrinsic angular momentum (spin) whose magnitude squared is given by

$$\langle S^2 \rangle = \frac{3}{4} \hbar^2 \tag{10.41a}$$

and whose z-component is given by

$$\langle S_z \rangle = \pm \frac{1}{2} \hbar \ , \tag{10.41b}$$

corresponding, of course, to an $s = 1/2$ and an $m_s = \pm 1/2$.

Associated with this spin, an electron also possesses a magnetic moment proportional to $\underline{S}$, i.e.

$$\mu_e^{(s)} = - g_e \left(\frac{\mu_B}{\hbar}\right) \underline{S} \ , \tag{10.42}$$

where μ_B is the Bohr magneton. The constant of proportionality g_e is called the spin g-factor and found to be equal to#

$$g_e \simeq 2.00 \quad . \tag{10.43}$$

Note, incidentally that any projection of $\mu_e^{(s)}$ will have the values

$$\mu_{e_z}^{(s)} \simeq \pm \mu_B \tag{10.44}$$

since $$S_z = \pm \hbar/2 \ .$$

Other elementary particles also possess spin and magnetic moments. In particular, the proton and the neutron are also spin 1/2 particles with corresponding magnetic moments

$$\mu_p = + g_p \left(\frac{\mu_N}{\hbar}\right) \underline{S} = + 5.59 \left(\frac{\mu_N}{\hbar}\right) \underline{S} \tag{10.44a}$$

and

$$\mu_n = + g_n \left(\frac{\mu_N}{\hbar}\right) \underline{S} = + 3.83 \left(\frac{\mu_N}{\hbar}\right) \underline{S} \tag{10.44b}$$

where μ_N is the nuclear magneton,

$$\mu_N = \frac{e\hbar}{2M_p} = 0.505 \times 10^{-26} \ \mathrm{A \cdot m^2} \ . \tag{10.45}$$

The fact that the constant of proportionality g (apart from the dimensional factors) differs from unity should not be surprising. Even classically, the spin

---

#This result can be derived from the relativistic Dirac theory (e.g. Merzbacher, Quantum Mechanics, 2nd ed., chapter 24.2). Precise spectroscopic data (Lamb) yield a value $g_e = 2.00232$ in complete agreement with the theoretical result $g_e = 2(1 + \alpha/2\pi)$, when radiative corrections are included ($\alpha \simeq 1/137$ = fine structure constant).

angular momentum is related to the distribution of mass of the spinning body, whereas the magnetic moment of a spinning charge depends on the charge distribution within the body. It is interesting that classically a uniformly dense sphere, with all the charge residing on the surface, would also have a g of 5/3. See appendix F, problem F. 15.

It should be noted that the theoretical justification for the magnetic moments of the proton and especially the neutron are in considerably worse shape.

10.4 Spin 1/2.

10.4.1 Spinors and the Pauli spin matrices

If the eigenstates of S^2 and S_z are given by $|s, m_s\rangle$, i.e.

$$S^2 |s, m_s\rangle = \hbar^2 s(s+1) |s, m_s\rangle \qquad (10.46a)$$

$$S_z |s, m_s\rangle = \hbar m_s |s, m_s\rangle \quad , \qquad (10.46b)$$

then in the subspace $s = 1/2$ there are only two states

$$|1/2,\ 1/2\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$|1/2, -1/2\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

which are often denoted $|+\rangle$ and $|-\rangle$ for short, corresponding to "spin up" or "spin down". A wavefunction for spin 1/2 particle, is then a two-dimensional vector, i.e.

$$|\Psi\rangle = |1/2, 1/2\rangle\ \psi_+(x, y, z) + |1/2, -1/2\rangle\ \psi_-(x, y, z) \qquad (10.47a)$$

or, equivalently

$$|\Psi\rangle = \begin{pmatrix} \psi_+(x, y, z) \\ \psi_-(x, y, z) \end{pmatrix} . \qquad (10.47b)$$

The states represented by equations 10.47 are called spinors. They are the natural wavefunctions in the relativistic Dirac equation, but must be introduced in an ad hoc way in the Schrödinger equation, which was derived on the basis of a scalar Hamiltonian.

The three components of $\underline{S}$

$$\underline{S} = S_x \hat{e}_x + S_y \hat{e}_y + S_z \hat{e}_z , \qquad (10.48)$$

being components of an angular momentum operator, also satisfy the commutation relations of equation 10.5. In addition, however, for the special case of

$s = 1/2$ we also have

$$S_x^2 = S_y^2 = S_z^2 = \frac{1}{4}\hbar^2 \tag{10.49}$$

and also

$$S_+^2 = S_-^2 = 0 \tag{10.50}$$

from which we deduce, since

$$S_+^2 = (S_x + iS_y)^2 = \cancelto{0}{(S_x^2 - S_y^2)} + i(S_xS_y + S_yS_x) = 0 \ ,$$

that

$$S_xS_y + S_yS_x = 0 \tag{10.51}$$

or that the operators S_x, S_y, S_z anticommute.

The operators S_x, S_y and S_z are conveniently expressed in terms of the Pauli matrices, defined by (for spin 1/2)

$$\underline{S} = \frac{\hbar}{2}\underline{\sigma} \tag{10.52}$$

and given explicitly by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -0 \end{pmatrix} \ . \tag{10.53}$$

Their properties are summarized below:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1 \tag{10.54a}$$

$$\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z \qquad \text{et cycl.} \tag{10.54b}$$

$$\sigma_x\sigma_y\sigma_z = i \tag{10.54c}$$

$$\mathrm{Tr}(\sigma_x) = \mathrm{Tr}(\sigma_y) = \mathrm{Tr}(\sigma_z) = 0^{\#} \tag{10.54d}$$

and

$$\det(\sigma_x) = \det(\sigma_y) = \det(\sigma_z) = -1 \ . \tag{10.54e}$$

Note also that the three Pauli matrices, augmented by the identity matrix (2 x 2)

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \sigma_0 \tag{10.55}$$

form a complete basis in which any 2 x 2 matrix can be expressed. In other words, if C is any 2 x 2 matrix

$$C = c_0\sigma_0 + c_x\sigma_x + c_y\sigma_y + c_z\sigma_z = \sum_{i=0}^{4} c_i\sigma_i \tag{10.56}$$

[#] Tr = Trace = sum of diagonal elements.

where

$$C = \begin{pmatrix} c_0 + c_z & c_x - ic_y \\ c_x + ic_y & c_0 - c_z \end{pmatrix} \qquad (10.57)$$

10.4.2 Rotation of spin 1/2 states.

Note that, using equation 10.6 for the rotation operator, we have for a system in an $\ell = 0$ state

$$\mathcal{R}(\underline{\alpha}) = e^{-i\underline{\alpha} \cdot \underline{S}/\hbar} \quad . \qquad (10.58)$$

or, using the Pauli matrices for a spin 1/2 system,

$$\mathcal{R}_{1/2}(\underline{\alpha}) = e^{-i\underline{\alpha} \cdot \underline{\sigma}/2} \quad . \qquad (10.59)$$

Since any component of σ has the property

$$\sigma_i^{2n} = 1 \Rightarrow \sigma_i^{2n+1} = \sigma_i \, , \qquad i = 1, 2, 3 \qquad (10.60)$$

we can compute the rotation operator of equation 10.59 in closed form. This yields:

$$\mathcal{R}_{1/2}(\underline{\alpha}) = \cos(\tfrac{\alpha}{2})\, I - i \sin(\tfrac{\alpha}{2})\, \sigma_\alpha \, , \qquad (10.61a)$$

where I is the identity matrix and σ_α is the projection of $\underline{\sigma}$ along the direction of the rotation vector $\underline{\alpha}$, i.e.

$$\sigma_\alpha = \frac{1}{\alpha}(\sigma_x \alpha_x + \sigma_y \alpha_y + \sigma_z \alpha_z) \; . \qquad (10.61b)$$

Note that, therefore, for spin 1/2 particles[#] a rotation of 2π about any axis changes the sign of the wavefunction, i.e.

$$\mathcal{R}_{1/2}(2\pi) = -I \qquad (10.62)$$

or

$$\mathcal{R}_{1/2}(2\pi)\, |\Psi\rangle = \mathcal{R}_{1/2}(2\pi) \begin{pmatrix} \psi_+(\underline{x}) \\ \psi_-(\underline{x}) \end{pmatrix} = - \begin{pmatrix} \psi_+(\underline{x}) \\ \psi_-(\underline{x}) \end{pmatrix} \quad . \qquad (10.63)$$

This may appear strange since a rotation of 360° should return the system in its original state. The minus sign, however, is only a phase factor which will not enter in computing expectation values of operators corresponding to physical observables.

[#] or half-integer in general, in fact.

10.4.3 Spin 1/2 wavefunctions

From the preceding discussion we saw that the wavefunctions of spin 1/2 particles are not scalars, but can be represented as two-dimensional vectors (spinors). We can express these eigenfunctions in the basis of the simultaneous eigenstates of L^2 and L_z, and S^2 and S_z. In particular, using the result of the addition of angular momentum we have, for $\underline{J_1} = \underline{L}$ and $\underline{J_2} = S$, for a particle of spin 1/2 in a total angular momentum state (j, m), we have

$$\begin{aligned}|\ell, 1/2, j, m\rangle = {} & |\ell, 1/2; m-1/2, 1/2\rangle\langle\ell, 1/2; m-1/2, 1/2|j, m\rangle \\ & + |\ell, 1/2; m+1/2, -1/2\rangle\langle\ell, 1/2; m+1/2, -1/2|j, m\rangle\end{aligned} \tag{10.64}$$

(see equation 10.26). The matrix elements multiplying the two ket vectors, on the right hand side of the equation, are the corresponding Clebsch-Gordan coefficients tabulated in Table 10.1a for the case $j_2 = 1/2$. From the restriction on the resulting angular momentum sum, we see that there are only two possible values for j (see equation 10.29), namely

$$j = \ell \pm 1/2\ , \tag{10.65}$$

Therefore, we have for the total angular momentum wavefunctions, in the coordinate representation,

$$\begin{aligned}\Psi(\theta,\varphi) &= \langle\theta,\varphi|\ell, 1/2, \ell\pm 1/2, m\rangle \\ &= \langle\theta,\varphi|\ell, 1/2; m-1/2, 1/2\rangle\langle\ell, 1/2; m-1/2, 1/2|\ell\pm 1/2, m\rangle \\ &+ \langle\theta,\varphi|\ell, 1/2; m+1/2, -1/2\rangle\langle\ell, 1/2; m+1/2, 1/2|\ell\pm 1/2, m\rangle\end{aligned} \tag{10.66}$$

or, substituting the spherical harmonics and the corresponding Clebsch-Gordan coefficients,

$$\langle\theta,\varphi|\ell, \underset{\substack{\uparrow \\ s}}{1/2}, \underset{\substack{\uparrow \\ j}}{\ell\pm 1/2}, m\rangle = \sqrt{\frac{1}{2\ell+1}}\begin{pmatrix}\pm\,(\ell\pm m\pm 1/2)^{1/2}\,Y_{\ell,m-1/2}(\theta,\varphi) \\ \\ (\ell\mp m+1/2)^{1/2}\,Y_{\ell,m+1/2}(\theta,\varphi)\end{pmatrix} . \tag{10.67}$$

These states are designated in spectroscopic notation by the convention

$${}^{2s+1}(\ell)_j\ ,$$

where

ℓ	letter[#]	
0	S	
1	P	
2	D	
3	F	
4	G	alphabetical .
.	.	
.	.	
.	.	

By way of example, a $^2P_{3/2}$ state represents $s = 1/2$, $\ell = 1$ and $j = \ell + 1/2 = 3/2$, a $^3D_1$ state represents an $s = 1$ (possible in a two electron configuration), $\ell = 2$ and $j = 1$.

10.5 The spinning electron in a Coulomb field. Spin-Orbit interactions.

In our discussion of the hydrogenic wavefunctions (section 7.3), we solved the Schrödinger equation with a Hamiltonian consisting of the kinetic energy (in the center of mass coordinate system) and the electrostatic Coulomb attraction between the nucleus of charge Ze and the electron of charge -e, i.e.

$$H_0 = \frac{1}{2m_r} p^2 - \frac{Ze^2}{4\pi\varepsilon_0 r} \quad . \tag{10.68}$$

While this is appropriate for a description of a spinless electron, additional terms must be taken into account if the spin (and resulting magnetic moment) are included.

Strictly speaking, equation 10.68 is not correct relativistically because, in writing the potential energy in terms of the (static) Coulomb potential, *we have ignored the fact that the electron is moving*. In a frame moving with the electron, a *magnetic* field will appear, resulting from the relativistic transformation of the *electric* field generated by the (~ stationary) nucleus. From the transformation equations of the E&M field[†], we have (for the component of the magnetic field perpendicular to the velocity between the two frames) for $(v/c)^2 << 1$ classically

$$\{\underline{\mathcal{B}}\}_{el} = -\frac{1}{c^2}\, \underline{v} \times \{\underline{\mathcal{E}}\}_{rest} \, , \tag{10.69}$$

[#] The $\ell = 0$ states were always Sharp (non-degenerate), the $\ell = 1$ states were found to contribute the Principal (strongest) lines in the spectrum ($\Delta\ell = 1$ for strong transitions to the $\ell = 0$ ground state), the $\ell = 2$ states gave Diffuse transitions (high degeneracy, low instrument resolution in early days). "E" was omitted (reserved for energy) and the rest are alphabetical.

[†] see for example, *The Feynman Lectures on Physics*, II, ch. 26.

where $\underline{\mathcal{E}}$ is the electric field in the stationary (nucleus) frame, $\underline{v}$ is the velocity between the two frames (electron-proton relative velocity) and $\underline{\mathcal{B}}'$ is the magnetic field in the frame of the electron. Making the substitutions

$$\underline{v} = \frac{1}{m}\,\underline{p}'$$

and

$$\underline{\mathcal{E}} = -\frac{1}{e}\,\underline{F} = \frac{1}{e}\,\frac{\partial V(r)}{\partial r}\,\frac{\underline{r}}{r}$$

we have

$$\{\underline{\mathcal{B}}\}_{el} = \frac{1}{emc^2}\,\frac{1}{r}\,\frac{\partial V(r)}{\partial r}\,\underline{L} \tag{10.70}$$

where V(r) is the Coulomb potential and $\underline{L} = \underline{r} \times \underline{p}$ is the orbital angular momentun. Consequently there is an interaction between the magnetic moment and this magnetic field which, is the electron rest frame, is given by

$$\{E_{LS}\}_{el} = -\underline{\mu}^{(s)} \cdot \{\underline{\mathcal{B}}\}_{el} ,$$

or, substituting for the magnetic moment (equation 10.42),

$$\{E_{LS}\}_{el} = \frac{g_e}{2m^2c^2}\left(\frac{1}{r}\,\frac{\partial V}{\partial r}\right)\underline{L}\cdot\underline{S} \tag{10.71}$$

The transformation of the ($\underline{L}$, $\underline{S}$) coupling energy back to the nuclear rest frame requires some care because the two frames are not merely moving with respect to each other but also in relative acceleration, by virtue of the circular motion. This contributes a kinematic effect, as a result of the way velocities transform relativistically, called the Thomas precession. If two frames are in a state of relative translation with a velocity $\underline{v}$ and a relative acceleration of magnitude $\underline{a}$, then the two coordinate systems appear to be precessing relative to each other at a rate given by the vector angular velocity vector[#]

$$\underline{\omega}_T = -\frac{1}{2c^2}\,\underline{v} \times \underline{a} \ , \tag{10.22}$$

which can be evaluated in the nuclear rest frame to give

$$\underline{\omega}_T = -\frac{1}{2c^2}\,\underline{v} \times \underline{a} = -\frac{1}{2mc^2}\,\underline{v} \times \underline{F} = -\frac{e}{2mc^2}\,\underline{v} \times \underline{\mathcal{E}} \ . \tag{10.73}$$

Now a classical magnetic moment $\underline{\mu}$ precesses about a constant magnetic field $\underline{\mathcal{B}}$ in such a way as to maintain a constant subtended angle β (see sketch on page 10.11) so that, in a non-dissipative environment $U = -\underline{\mu}\cdot\underline{\mathcal{B}}$ is constant. This precession frequency is called the Larmor precession and given by setting the torque equal to the time rate of change of the angular momentum, i.e.

[#] See, for example, R. Eisberg and R. Resnick, Quantum Physics, appendix J.

$$\underline{\mu} \times \underline{B} = \frac{d\underline{S}}{dt} = \underline{\omega}_L \times \underline{S}$$

so that

$$\underline{\omega}_L = \frac{1}{m^2 c^2} \left(\frac{1}{r}\frac{\partial V}{\partial r}\right) \underline{L} \quad . \qquad (10.74)$$

Comparison of this frequency with the Thomas precession frequency shows that it is exactly twice as large, so that we have to reduce the energy $\{E_{LS}\}_{el}$ by a factor of two to bring it to the nuclear rest frame. Consequently, the appropriate interaction Hamiltonian is given by

$$H'_{LS} = \frac{g_e Z(-E_R)a_0}{2m^2 c^2 r^3} \, \underline{L}\cdot\underline{S} \quad , \qquad (10.75)$$

where $-E_R$ is the hydrogen atom ground state binding energy (equation 7.64).

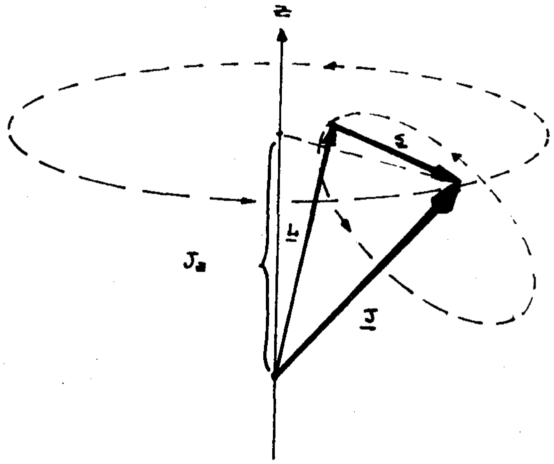
The most important effect of this term, is to destroy the conservation of orbital angular momentum. In particular,

$$\frac{d}{dt} L_z = \frac{i}{\hbar}[H_0 + H_{LS}, L_z] = \frac{i}{\hbar}[H_{LS}, L_z] \neq 0 \ ,$$

whereas the total angular momentum is conserved, i.e.

$$\begin{aligned}\frac{d}{dt} J_z \propto [\underline{L}\cdot\underline{S}, J_z] &= [\underline{L}, J_z]\cdot\underline{S} + \underline{L}\cdot[\underline{S}, J_z] \\ &= [\underline{L}, L_z]\cdot\underline{S} + \underline{L}\cdot[\underline{S}, S_z] \\ &= [L_x, L_z]S_x + [L_y, L_z]S_y + L_x[S_x, S_z] + L_y[S_y, S_z] \\ &= 0 \ .\end{aligned}$$

Note, that J^2, L^2 and S^2 are all conserved, i.e. the magnitudes of the vectors $|\underline{J}|$, $|\underline{L}|$ and $|\underline{S}|$ are constant. The projection of $\underline{L}$ and $\underline{S}$ on a particular axis (say, the z-axis), however, are not constant, whereas the projection of $\underline{J}$ on the z-axis is constant. It is easy to see that $\underline{L}\cdot\underline{J} = L^2 + \underline{L}\cdot\underline{S}$ and $\underline{S}\cdot\underline{J} = S^2 + \underline{S}\cdot\underline{L}$ are also constant. This suggests the following picture that is consistent with the above. The total angular momentum $\underline{J}$ precesses about the z-axis at a fixed inclination. The orbital angular momentum $\underline{L}$ and the spin angular momentum $\underline{S}$ precess about the total angular momentum $\underline{J}$, both at fixed angles. This situation is sketched below.



We can compute the effects of H'_{LS} using perturbation theory, since the magnitude of this interaction scales with the ratio E_R/mc^2 and is therefore small. In particular, the first order correction, is given by

$$\Delta E_{LS} = \langle n;\ell, s, j, m|H'|n;\ell, s, j, m\rangle$$

$$= \frac{g_e Z(-E_R)a_0}{2m^2c^2} \langle n, \ell|\frac{1}{r^3}|n, \ell\rangle\langle \ell, s, j, m|\underline{L}\cdot\underline{S}|\ell, s, j, m\rangle \ . \tag{10.76}$$

The radial integral is tabulated in Pauling and Wilson, Introduction to Quantum Mechanics[#] for various powers of r (see table 7.2 below). To evaluate the $\langle\underline{L}\cdot\underline{S}\rangle$ matrix element, we note that

$$\underline{L}\cdot\underline{S} = \frac{1}{2}(J^2 - L^2 - S^2)$$

$$= \frac{\hbar^2}{2}[j(j+1) - \ell(\ell+1) - 3/4]$$

or

$$\langle\underline{L}\cdot\underline{S}\rangle \begin{cases} = \dfrac{\hbar^2}{2}\ell\ , & \text{for } j = \ell + 1/2 \\ = -\dfrac{\hbar^2}{2}(\ell+1)\ , & \text{for } j = \ell - 1/2,\ \ell \neq 0 \end{cases} \tag{10.77}$$

[#] pages 144-145.

TABLE 10.2 Hydrogenic radial integrals.

$$\langle r^2 \rangle = \frac{a_0^2 n^4}{Z^2} \left\{ 1 + \frac{3}{2} \left[1 - \frac{\ell(\ell+1) - 1/3}{n^2} \right] \right\}$$

$$\langle r \rangle = \frac{a_0 n^2}{Z} \left\{ 1 + \frac{1}{2} \left[1 - \frac{\ell(\ell+1)}{n^2} \right] \right\}$$

$$\langle \frac{1}{r} \rangle = \frac{Z}{a_0 n^2}$$

$$\langle \frac{1}{r^2} \rangle = \frac{Z^2}{a_0^2 n^3 (\ell + 1/2)}$$

$$\langle \frac{1}{r^3} \rangle = \frac{Z^3}{a_0^3 n^3 \ell(\ell + 1/2)(\ell + 1)}$$

$$\langle \frac{1}{r^4} \rangle = \frac{Z^4 [3 - \ell(\ell+1)/n^2]}{2 a_0^4 n^3 \ell(\ell + 3/2)(\ell+1)(\ell^2 - 1/4)}$$

Combining the two expectation values, we then have from equation 10.76

$$\Delta E_{LS} \begin{cases} = \dfrac{Z^2 |E_n| \alpha^2}{n} \cdot \dfrac{1}{(2\ell+1)(\ell+1)} & ; \quad j = \ell + 1/2 \\[2ex] = - \dfrac{Z^2 |E_n| \alpha^2}{n} \cdot \dfrac{1}{\ell(2\ell+1)} & ; \quad j = \ell - 1/2, \ \ell \neq 0 \ . \end{cases} \tag{10.78}$$

The constant α is a dimensionless constant called the fine structure constant and given by (see also equation 7.63 and related discussion),

$$\alpha = \frac{e^2}{4\pi\varepsilon_0 \hbar c} = \frac{1}{137.0377} \simeq \frac{1}{137} \ . \tag{10.79}$$

10.6 Relativistic corrections to the kinetic energy.

Before leaving this area, it should be pointed out that it would be inconsistent to include the spin-orbit coupling as a perturbation term and omit the relativistic correction to the kinetic energy, to the same order. In particular, since

$$\begin{aligned} K &= (p^2 c^2 + m^2 c^4)^{1/2} - mc^2 \\ &\simeq \frac{p^2}{2m} - \frac{1}{2mc^2} \left(\frac{p^2}{2m}\right)^2 + \ldots \ , \end{aligned} \tag{10.80}$$

we should include a second correction term

$$H'_{\mathrm{rel}} = -\frac{1}{2mc^2}\left(\frac{p^2}{2m}\right)^2 \qquad (10.81)$$

so that

$$H = H_0 + H'_{LS} + H'_{\mathrm{rel}} \; . \qquad (10.82)$$

The energy levels then, corrected to first order are given by

$$\Delta E = \langle H'_{LS}\rangle_0 + \langle H'_{\mathrm{rel}}\rangle_0 \; .$$

Note that

$$H'_{\mathrm{rel}} = -\frac{1}{2mc^2} K^2 = -\frac{1}{2mc^2}(H_0 - V)^2$$

$$= -\frac{1}{2mc^2}(H_0^2 - 2H_0V + V^2)$$

so that

$$\langle H'_{\mathrm{rel}}\rangle_0 = -\frac{1}{2mc^2}[E_n^{(o)} - 2E_n^{(o)}\langle V\rangle_0 + \langle V^2\rangle_0] \; .$$

The matrix elements $\langle V\rangle_0$ and $\langle V^2\rangle_0$ can be computed by direct integration, and using the results of Table 7.2, yield

$$\Delta E_{\mathrm{rel}} = \langle H'_{\mathrm{rel}}\rangle_0 = \frac{Z^2|E_n|\alpha^2}{4n^2}\left(3 - \frac{4n}{\ell + 1/2}\right) \; . \qquad (10.83)$$

Combining this result with the spin-orbit calculations to the same order we have

$$\Delta E = \Delta E_{LS} + \Delta E_{\mathrm{rel}} = \frac{Z^2|E_n|\alpha^2}{4n^2}\left(3 - \frac{4n}{j + \frac{1}{2}}\right) \; . \qquad (10.84)$$

for both $j = \ell + 1/2$ and $j = \ell - 1/2$. By way of comparison, the energy levels, as given by the relativistic Dirac equation, are[#]

$$\frac{E}{mc^2} = \left\{ 1 + \frac{(Z\alpha)^2}{[\sqrt{(j+1/2)^2 - (Z\alpha)^2} + (n-j-1/2)]^2} \right\}^{-1/2} \; . \qquad (10.85)$$

[#] Merzbacher, Quantum Mechanics, (2nd Ed.) ch. 24.4.

11. INTERACTION OF RADIATION WITH MATTER. EMISSION AND ABSORPTION.

11. 1 The Hamiltonian.

We have seen that the Hamiltonian of a single charged particle moving in the presence of a potential $U(\underline{x})$ is given by

$$H_{\text{particle}} = \frac{1}{2m}\, \underline{p}\cdot\underline{p} + U(\underline{x}) \quad . \tag{11.1}$$

We have also seen that in the presence of an externally imposed electromagnetic field, this Hamiltonian must be altered through the change in the momentum, i. e.[#]

$$\underline{p} \rightarrow \underline{p} - q_c \underline{A} \quad , \tag{11.2}$$

and the addition of the interaction with the external electrostatic field, i. e.

$$U(\underline{x}) \rightarrow U(\underline{x}) + q_c \varphi(\underline{x}, t) \quad , \tag{11.3}$$

where $\underline{A}(\underline{x}, t)$ and $\varphi(\underline{x}, t)$ are the external vector and scalar potentials respectively. The resulting Hamiltonian can then be written, as we have seen,

$$H = \frac{1}{2m} (\underline{p} - q_c\underline{A})\cdot(\underline{p} - q_c\underline{A}) + U(\underline{x}) + q_c\varphi \quad . \tag{11.4}$$

This is in fact the Hamiltonian we used to solve a variety of problems[†] in which we ignored the effect <u>of</u> the charge <u>on</u> the electromagnetic field.

On the other hand, <u>in the absence of charges (and currents)</u>, we have seen that, starting from the classical energy of the electromagnetic field (see equation 9. 4)

$$E = \frac{1}{2}\,\epsilon_0 \int_V [\,|\dot{\underline{A}}|^2 + c^2 \,|\frac{\partial}{\partial \underline{x}} \times \underline{A}|^2]\, d^3\underline{x} \quad , \tag{11.5}$$

we were able to define a Hamiltonian corresponding to a system of uncoupled harmonic oscillators in each of the wavenumber and polarization modes, which is given, in terms of the corresponding ladder operators, by[‡]

$$H_{\text{field}} = \sum_{\underline{k},\,\ell} \hbar\omega_{\underline{k}}\, (a^{\dagger}_{\underline{k},\,\ell}\; a_{\underline{k},\,\ell} + \frac{1}{2}) \quad . \tag{11.6}$$

[#] See equation 1. 43 and related discussion.

[†] See example 6. 1 on the motion of a charged particle in a uniform magnetic field $[\underline{A} = -\frac{1}{2}(\underline{x} \times \underline{\mathcal{B}}),\ \varphi = 0]$, and the discussion of the hydrogenic wavefunctions $[\underline{A} = 0,\ \varphi = \frac{e}{4\pi\epsilon_0 r}]$.

[‡] or, in terms of the circular polarizations

$$H_{\text{field}} = \sum_{\underline{k},\,\sigma} \hbar\omega_{\underline{k}}\, (a^{\dagger}_{\underline{k},\,\sigma}\; a_{\underline{k},\,\sigma} + \frac{1}{2}) \quad . \tag{11.6'}$$

If we now combine the Hamiltonian of equation 11.4 for a charged particle in the presence of an electromagnetic field, with the Hamiltonian of equation 11.6, we have

$$H = \frac{1}{2m} (\underline{p} - q_c\underline{A}) \cdot (\underline{p} - q_c\underline{A}) + U(\underline{x}) + \sum_{\underline{k}, \ell} \hbar\omega_{\underline{k}} (a^\dagger_{\underline{k}, \ell}\, a_{\underline{k}, \ell} + \tfrac{1}{2}) \tag{11.7}$$

In writing equation 11.7, we have included the electrostatic field energy of $q_c\varphi(\underline{x}, t)$ into $U(\underline{x})$. You will recall that the Hamiltonian of equation 11.6 was derived from the energy of the field, as given by equation 11.5, assuming that $\varphi = 0$ and $\frac{\partial}{\partial \underline{x}} \cdot \underline{A} = 0$ (radiation gauge), as is always possible through a gauge transformation in the absence of charges and currents (appendix J). This, of course, does not disallow an electrostatic field, which can always be superimposed on the electromagnetic field described by the radiation gauge. It is only that that part of the field is not described by the Hamiltonian of equation 11.6. It should also be noted, that the vector potential $\underline{A}(\underline{x}, t)$, as it appears in equation 11.7, is to be understood as an operator in the sense of equation 9.40 (or equation 9.98), and was not written in that form in the interest of brevity.

If we now expand the first term of the Hamiltonian of equation 11.7, we have

$$\frac{1}{2m} (\underline{p} - q_c\underline{A}) \cdot (\underline{p} - q_c\underline{A}) \equiv \frac{1}{2m} p^2 - \frac{q_c}{2m} (\underline{p} \cdot \underline{A} + \underline{A} \cdot \underline{p}) + \frac{q_c^2}{2m} A^2 \quad . \tag{11.8}$$

Note that

$$\underline{A} \cdot \underline{p} \equiv \underline{p} \cdot \underline{A} + [\underline{A} \cdot \underline{p} - \underline{p} \cdot \underline{A}] \quad ,$$

or

$$\underline{A} \cdot \underline{p} = \underline{p} \cdot \underline{A} + \sum_i \, [A_i, p_i]$$

and substituting for $\underline{p} = - i\hbar (\frac{\partial}{\partial \underline{x}} \cdot \underline{A})$, we have

$$\underline{A} \cdot \underline{p} = \underline{p} \cdot A + i\hbar \, (\frac{\partial}{\partial \underline{x}} \cdot \underline{A}) \quad . \tag{11.9}$$

But in the radiation gauge, the vector potential is divergence free and therefore

$$\underline{A} \cdot \underline{p} = \underline{p} \cdot \underline{A} \quad , \tag{11.10}$$

or $\underline{A}$ and $\underline{p}$ commute and therefore,

$$\frac{1}{2m} (\underline{p} - q_c\underline{A}) \cdot (\underline{p} - q_c\underline{A}) = \frac{1}{2m} p^2 - \frac{q_c}{m} \underline{p} \cdot \underline{A} + \frac{q_c^2}{2m} A^2 \quad . \tag{11.11}$$

Substituting this expression in our Hamiltonian, we see that it breaks into three parts,

$$H = H_{\text{particle}} + H_{\text{field}} + H_{\text{interaction}} \quad , \tag{11.12}$$

where

$$H_{\text{particle}} = \frac{1}{2m} p^2 + U(\underline{x}) \quad , \tag{11.13a}$$

is the Hamiltonian of the particle alone,

$$H_{\text{field}} = \sum_{\underline{k}, \ell} \hbar\omega_{\underline{k}} (a^\dagger_{\underline{k}, \ell}\, a_{\underline{k}, \ell} + \tfrac{1}{2}) , \tag{11.13b}$$

is the Hamiltonian of the electromagnetic field, in the absence of charges and currents, and

$$H_{\text{interaction}} = -\frac{q_c}{m}\,\underline{p}\cdot\underline{A} + \frac{q_c^2}{2m}\,\underline{A}\cdot\underline{A} \qquad (11.13c)$$

is the interaction Hamiltonian.

The problem posed by this total Hamiltonian is quite intractable in the general case, and one which, as a consequence, we will tackle through the machinery of perturbation theory.

11.2 The unperturbed Hamiltonian and eigenstates.

Of the total Hamiltonian of equation 11.12 we will consider as our unperturbed Hamiltonian H_0, in the sense of the discussion of equation 8.1, as given by

$$H_0 = H_{\text{particle}} + H_{\text{field}} \quad . \qquad (11.14)$$

This Hamiltonian includes the coordinates and momenta (operators) for both the particle motion and the dynamics of the (charge-free) electromagnetic field, i.e.

$$H_0 = H_0\{\underline{x}, \underline{p};\ [a_{\underline{k},\ell}, a^{\dagger}_{\underline{k},\ell}]_{\text{all }(\underline{k},\ell)}\} \quad , \qquad (11.15)$$

possessing all the degrees of freedom for the description of the combined system. The unperturbed eigenstates of this Hamiltonian are then (tensor) products of the unperturbed eigenstates of each subsystem (particle, field) separately, since H_0 is the sum of the two partial Hamiltonians (no cross terms), which describes uncoupled subsystems (see equation 4.6 and related discussion). We therefore have

$$|\psi\rangle_0 = |\psi(\underline{x})\rangle_0\,|\underline{n}\rangle_0 = |\psi(\underline{x});\,\underline{n}\rangle_0 \quad , \qquad (11.16)$$

where $\psi(\underline{x})$ is the wavefunction corresponding to the (unperturbed) particle motion and $|\underline{n}\rangle$ is the field eigenstate (i.e. equation 9.48).

In the absence of a coupling between the particle (atomic) Hamiltonian and the field Hamiltonian, these states would be stationary states of the combined system. In particular, the atomic subsystem, if initially in an excited state would always remain there, that being an eigenstate of the system. It is therefore, only through the unavoidable action of the interaction Hamiltonian that an atom in an excited state can transition to the ground state emitting a photon or, conversly, be raised to an excited state through the absorption of a photon. Alternatively, the states $|\psi\rangle_0$ of equation 11.16 may be eigenstates of H_0, but not of the total Hamiltonian

$$H = H_0 + H_{\text{interaction}} \qquad (11.17)$$

and are not therefore (exact) eigenstates of the combined particle-field system. We will find that the presence of the electromagnetic field induces transitions between the atomic (quasi-) eigenstates. In fact, it is not even necessary for the electromagnetic field to have any excitations (even though ... they help) for such transitions to occur. Even in the "absence" of an external field (read: electromagnetic field in the ground state), the atomic system in an excited state can exchange energy with the field, raising one (or more of <u>its</u> modes to an excited state, as the charged particle transitions down to a lower atomic state. This is the phenomenon of <u>spontaneous emission</u>, recognized by Einstein as necessary for the treatment of the problem of thermal equilibrium between electromagnetic radiation and the walls of a surrounding cavity,[#] which cannot be understood unless both the atom <u>and</u> the radiation field are treated as interacting quantum systems.

11.3 <u>Radiation transitions in the weak field approximation.</u>

The eigenstates of the unperturbed Hamiltonian form a complete set of basis states of the combined system, and it is of course possible to express any state of the total Hamiltonian as a linear combination of those states.

In what follows, we will treat the case of transitions of the total system (particle + field), from some initial total state $|\psi_i\rangle$ to a final state $|\psi_f\rangle$[†]. Such transitions include emission of radiation, where

$$\left|\begin{matrix}\text{initial system}\\ \text{state}\end{matrix}\right\rangle = \left|\begin{matrix}\text{excited atomic}\\ \text{system state}\end{matrix}\right\rangle \left|\underline{n}\ \text{field state}\right\rangle$$

and

$$\left|\begin{matrix}\text{final system}\\ \text{state}\end{matrix}\right\rangle = \left|\begin{matrix}\text{lower atomic}\\ \text{system state}\end{matrix}\right\rangle \left|\begin{matrix}\underline{n} + \underline{m}\ \text{photon}\\ \text{field state}\end{matrix}\right\rangle$$

and <u>absorption</u>,

$$\left|\begin{matrix}\text{initial system}\\ \text{state}\end{matrix}\right\rangle = \left|\begin{matrix}\text{lower atomic}\\ \text{system state}\end{matrix}\right\rangle \left|\underline{n}\ \text{field state}\right\rangle$$

and

$$\left|\begin{matrix}\text{final system}\\ \text{state}\end{matrix}\right\rangle = \left|\begin{matrix}\text{excited atomic}\\ \text{system state}\end{matrix}\right\rangle \left|\begin{matrix}\underline{n} - \underline{m}\ \text{photon}\\ \text{field state}\end{matrix}\right\rangle .$$

We can compute the transition rate of such events, using the results of time dependent perturbation theory, in particular, if we can ignore the finite energy spread in the initial and final states[‡], we have to lowest order,

$$W_{i\to f} = \frac{2\pi}{\hbar} \left|\langle\psi_f|H'|\psi_i\rangle\right|^2 \delta(E_f - E_i) \quad , \qquad (11.17)$$

where the delta function is to be understood in the sense of a subsequent integration over some density of states.

update #2

To stay consistent with the order of the perturbation calculation, we should exclude the

---

[#] See, for example Feynman, Leighton, Sands, <u>The Feynman Lectures on Physics</u>, v. I, section 42-5.

[†] The subscripts $_0$ denoting eigenstates of the unperturbed Hamiltonian are implied.

[‡] See sections 8.5 and 8.6.

second term in the interaction Hamiltonian of equation 11.13c since it is higher order in the vector potential.[#] We therefore have,

$$H' \simeq -\frac{q_c}{m}\,\hat{\underline{p}}\cdot\hat{\underline{A}} \tag{11.18}$$

which substituted in equation 11.17 yields

$$W_{i\to f} = \frac{2\pi}{\hbar}\left|-\frac{q_c}{m}\langle\psi_f|\underline{p}\cdot\underline{A}|\psi_i\rangle\right|^2\delta(E_f - E_i) \tag{11.19}$$

We see that the transition matrix is proportional to the matrix element of the component of the particle momentum in the direction of polarization of the local value of the vector potential.

Using the commutation relations of the coordinate $\underline{x}$ of the atomic particle with the particle Hamiltonian, i.e.

$$[\underline{x}, H_{part}] = \underline{x}\,H_{part} - H_{part}\,\underline{x} = \frac{i\hbar}{m}\,\underline{p} \quad , \tag{11.20}$$

we can re-write the matrix element in a way that is more useful when doing computations when using atomic wavefunctions in the coordinate representation. In particular

$$-\frac{q_c}{m}\langle\psi_f|\underline{p}\cdot\underline{A}|\psi_i\rangle = -\frac{iq_c}{\hbar}\langle\psi_f|(\underline{x}\,H_p - H_p\,\underline{x})\cdot\underline{A}|\psi_i\rangle$$

or, since H_p operates on the (unperturbed) atomic system eigenstate only,

$$-\frac{q_c}{m}\langle\psi_f|\underline{p}\cdot\underline{A}|\psi_i\rangle = \frac{i}{\hbar}\,q_c(\epsilon_f - \epsilon_i)\,\langle\psi_f|\underline{x}\cdot\underline{A}|\psi_i\rangle \quad ,$$

where ϵ_i and ϵ_f are the initial and final state energies of the atomic subsystem. If we now define the Bohr frequency of the transition, i.e.

$$\omega_{fi} \equiv \frac{1}{\hbar}(\epsilon_f - \epsilon_i) \quad ,$$

we have

$$-\frac{q_c}{m}\langle\psi_f|\underline{p}\cdot\underline{A}|\psi_i\rangle = i\omega_{fi}\,q_c\,\langle\psi_f|\underline{x}\cdot\underline{A}|\psi_i\rangle \tag{11.21}$$

or, using the electric dipole moment operator

$$\hat{\underline{d}} \equiv q_c\hat{\underline{x}} \quad , \tag{11.22}$$

we have

$$-\frac{q_c}{m}\langle\psi_f|\underline{p}\cdot\underline{A}|\psi_i\rangle = i\omega_{fi}\,\langle\psi_f|\underline{d}\cdot\underline{A}|\psi_i\rangle \quad . \tag{11.23}$$

update

Substituting into our original expression we then have for the transition rate

$$\boxed{W_{i\to f} = \frac{2\pi\omega_{fi}^2}{\hbar}\,|\langle\psi_f|\underline{d}\cdot\underline{A}|\psi_i\rangle|^2\,\delta(E_f - E_i)} \quad , \tag{11.24}$$

[#] This (weak field) approximation is valid as long as $q_c|\underline{A}| << |\underline{p}|$.

To compute the transition matrix element, we now need to substitute the quantum operator expression for the vector potential[#]. Using plane polarization modes, this yields

$$\langle \psi_f | \underline{d} \cdot \underline{A} | \psi_i \rangle = \left(\frac{\hbar}{2\varepsilon_0 V}\right)^{1/2} \sum_{\underline{k}', \ell'} \omega_{k'}^{-1/2} \langle \psi_f | \underline{d}\, e^{i\underline{k}' \cdot \underline{x}} \left[a_{\underline{k}', \ell'} + (-1)^{\ell'} a^{\dagger}_{-\underline{k}', \ell'}\right] | \psi_i \rangle \cdot \hat{e}_{\underline{k}', \ell'} \; . \tag{11.25}$$

We can now separate the matrix element into a product of a matrix element (integral) of operators over the particle coordinates and a matrix element of operators over the field coordinates, i.e.

$$\langle \psi_f | \underline{d}\, e^{i\underline{k}' \cdot \underline{x}} \left[a_{\underline{k}', \ell'} + (-1)^{\ell'} a^{\dagger}_{-\underline{k}', \ell'}\right] | \psi_i \rangle = \underbrace{\langle \psi_f | \underline{d}\, e^{i\underline{k}' \cdot \underline{x}} | \psi_i \rangle}_{\text{particle}} \langle \underline{n}_f | \underbrace{\left[a_{\underline{k}', \ell'} + (-1)^{\ell} a^{\dagger}_{-\underline{k}', \ell}\right]}_{\text{field}} | \underline{n}_i \rangle \tag{11.26}$$

The first matrix element,

$$\underline{M}_{fi}(-\underline{k}') \equiv q_c \langle \psi_f | \underline{x}\, e^{i\underline{k}' \cdot \underline{x}} | \psi_i \rangle = q_c \int_V \left[\psi_f^*(\underline{x})\; \underline{x}\; \psi_i(\underline{x})\right] e^{i\underline{k}' \cdot \underline{x}}\, d^3\underline{x} \tag{11.27}$$

we recognize as the (spatial) Fourier transform of the quantity $(\underline{x}\, \psi_f^* \psi_i)$, evaluated at $-\underline{k}'$ (see appendix B, equation B.2).

The second matrix element, can be seen to connect an initial field state $|\underline{n}_i\rangle$ to a final field state $|\underline{n}_f\rangle$ through the annihilation or creation of a single excitation. Consequently, the matrix element is zero (and therefore the transition rate also) unless the final field state has one more or one less photon, in some mode $(\underline{k}, \ell)$, than the initial field state, with the number of photons in all other modes the same, i.e. we must have:

$$\boxed{\begin{array}{ll} \text{case a:} & n^{(f)}_{\underline{k}, \ell} = n^{(i)}_{\underline{k}, \ell} + 1 \quad \& \quad n^{(f)}_{\underline{k}'', \ell''} = n^{(i)}_{\underline{k}'', \ell''} \quad \text{for} \quad (\underline{k}'', \ell'') \neq (\underline{k}, \ell) \\ \text{case b:} & n^{(f)}_{\underline{k}, \ell} = n^{(i)}_{\underline{k}, \ell} - 1 \quad \& \quad n^{(f)}_{\underline{k}'', \ell''} = n^{(i)}_{\underline{k}'', \ell''} \quad \text{for} \quad (\underline{k}'', \ell'') \neq (\underline{k}, \ell) \end{array}} \tag{11.28}$$

[Unfortunately, it is difficult to arrest the proliferation of mode indeces in this calculation; they really are three:

(i) The mode $(\underline{k}, \ell)$ denotes the one mode whose occupation index (number of photons) changes by one, up or down, between the initial and final field states.

(ii) The mode index $(\underline{k}', \ell')$ denotes a (dummy) summation index over the Fourier components of the vector potential operator.

update #2

[#] i.e. equation 9.40 for plane polarization modes, or equation 9.99 for circular polarization modes.

(iii) The mode index ($\underline{k}''$, ℓ'') denotes a running index over the modes of the initial and final field states.]

A restrictive condition of the type expressed by equation 11.28 is known as a selection rule. This particular one tells us that transitions described by equation 11.24 can only involve the emission (case a) or absorption (case b) of a single photon.[#] Or, in other words, not only must the final field state, in a particular mode ($\underline{k}$, ℓ), have exactly one more or one less photon, but that that can occur for exactly one mode.

The summation of equation 11.25, over the vector potential Fourier components along the field modes ($\underline{k}'$, ℓ') now mercifully collapses to a single term, for initial and final field states obeying the selection rule of equation 11.28, or is identically equal to zero, i. e.

$$\langle \psi_f | \underline{d} \cdot \underline{A} | \psi_i \rangle = \left(\frac{\hbar}{2 \varepsilon_0 V \omega_k} \right)^{1/2} \begin{cases} (n^{(i)}_{\underline{k},\ell} + 1)^{1/2} \; \hat{e}_{\underline{k},\ell} \cdot \underline{M}_{fi}(\underline{k})^{\ddagger} \, , & \text{case a (emission)} \\ (n^{(i)}_{\underline{k},\ell})^{1/2} \; \hat{e}_{\underline{k},\ell} \cdot \underline{M}_{fi}(-\underline{k}) \, , & \text{case b (absorption)} \\ 0 \, , & \text{otherwise} \, . \end{cases} \tag{11.29}$$

Note also, that since the final state has one more or one less photon in the mode ($\underline{k}$, ℓ), the energy difference between the system (particle + field) initial and final states is given by,

$$E_f - E_i = (\varepsilon_f - \varepsilon_i) \pm \hbar\omega_k = \begin{cases} (\varepsilon_f + \hbar\omega_k) - \varepsilon_i \, , & \text{emission} \\ \varepsilon_f - (\varepsilon_i + \hbar\omega_k) \, , & \text{absorption} \, . \end{cases}$$

[#] This is not to say that multiphoton transitions are dissallowed. They do occur but as a consequence of higher order terms of the perturbation $-\frac{q_c}{m} \underline{p} \cdot \underline{A}$, or as a consequence of the second order term (in the field) $\frac{q_c^2}{2m} \underline{A} \cdot \underline{A}$.

[‡] note that, for plane polarization, $\hat{e}_{\underline{k},\ell} = (-1)^{\ell} \, \hat{e}_{-\underline{k},\ell}$.

11.4 The long wavelength approximation.

When computing transitions between atomic system states, the order of magnitude of various quantities involved, allow another approximation to be made which greatly simplifies the calculations. In particular, in computing the matrix element

$$\underline{M}_{fi}(\underline{k}) = q_c \int (\psi_f^* \, \underline{x} \, \psi_i) e^{-i\underline{k}\cdot\underline{x}} \, d^3\underline{x} \quad ,$$

we see that the exponential is multiplied by the square of the atomic wavefunctions which diminish very rapidly outside the region of the extent of the atom, of the order of the Bohr radius (equation 7.60), i.e.

$$|\underline{x}| \leqslant a_B \simeq \left(\frac{4\pi\epsilon_0}{e^2}\right) \frac{\hbar^2}{m_e} = 0.529 \times 10^{-8} \text{ cm} \quad .$$

On the other hand, the photon wavevectors involved are of the order of

$$k = \frac{\omega_{fi}}{c} \leqslant \frac{|E_R|}{\hbar c} \quad ,$$

where E_R is the hydrogen atom binding energy (Rydberg; see equation 7.64). Therefore

$$k \leqslant \frac{|E_R|}{\hbar c} = \frac{1}{2} \left(\frac{e^2}{4\pi\epsilon_0 \hbar c}\right) \frac{1}{a_B} = \frac{\alpha}{2a_B} \quad ,$$

where α is the fine structure constant ($\alpha \simeq 1/137$). Therefore the phase factor in the exponential is less than or of the order of

$$|\underline{k}\cdot\underline{x}| \leqslant ka_B \lesssim \frac{\alpha}{2} \sim 4 \times 10^{-3} \quad , \tag{11.31}$$

and actually more like $10^{-4} - 10^{-3}$ for transitions in the infrared, visible and near ultra-violet. We are therefore justified in expanding the exponential in a Taylor series about $\underline{x} = 0$ and keeping the leading terms, i.e.

$$e^{-i\underline{k}\cdot\underline{x}} = 1 - i\underline{k}\cdot\underline{x} + \ldots \quad ,$$

which substituted in the $\underline{M}_{fi}$ matrix element yields

$$\underline{M}_{fi}(\underline{k}) \simeq q_c [\langle\psi_f|\underline{x}|\psi_i\rangle - i\underline{k}\cdot\langle\psi_f|\underline{x}\,\underline{x}|\psi_i\rangle + \ldots] \quad . \tag{11.32}$$

It can be seen that the first matrix element is the (vector) <u>electric dipole matrix element</u>

$$\underline{d}_{fi} = q_c \langle\psi_f|\underline{x}|\psi_i\rangle \quad , \tag{11.33a}$$

which contributes to the mode $(\underline{k}, \ell)$ to the extent of its projection on the polarization unit vector $\hat{e}_{\underline{k},\ell}$ of the mode $(\underline{k}, \ell)$, i.e.

$$\begin{aligned} \underline{d}_{fi}\cdot\hat{e}_{\underline{k},\ell} = q_c [&\langle\psi_f|x|\psi_i\rangle (\hat{e}_x\cdot\hat{e}_{\underline{k},\ell}) + \langle\psi_f|y|\psi_i\rangle (\hat{e}_y\cdot\hat{e}_{\underline{k},\ell}) \\ &+ \langle\psi_f|z|\psi_i\rangle (\hat{e}_z\cdot\hat{e}_{\underline{k},\ell})] \quad . \end{aligned} \tag{11.33b}$$

update

The second matrix element is the (rank 2 tensor) <u>electric quadrupole matrix element,</u>

$$\underline{\underline{Q}}_{fi} = \langle \psi_f | \underline{x}\,\underline{x} | \psi_i \rangle \quad , \tag{11.34a}$$

which contributes to the mode $(\underline{k}, \ell)$ to the extent of its projection on the wavevector $\underline{k}$ and the mode polarization unit vector $\hat{e}_{\underline{k},\ell}$, i.e.

$$\underline{k} \cdot \underline{\underline{Q}}_{fi} \cdot \hat{e}_{\underline{k},\ell} = q_c \sum_{m,n=1}^{3} k_m \langle \psi_f | x_m \, x_n | \psi_i \rangle \, (\hat{e}_n \cdot \hat{e}_{\underline{k},\ell}) \quad . \tag{11.34b}$$

Higher omitted terms in the expansion of $e^{-i\underline{k}\cdot\underline{x}}$ contribute projections of higher moments of the spatial function $\psi_f^*(\underline{x})\,\psi_i(\underline{x})$. Since each successive term is dimensionally of the order of $ka_B \sim \alpha/2$, this expansion can be seen to be a <u>multipole expansion</u>[#], <u>in powers of the fine structure constant</u>.

The leading term in this expansion is the electric dipole term, which yields for the matrix element,

$$\underline{M}_{fi}\,(\underline{k}) \sim q_c \langle \psi_f | \underline{x} | \psi_i \rangle \quad . \tag{11.35}$$

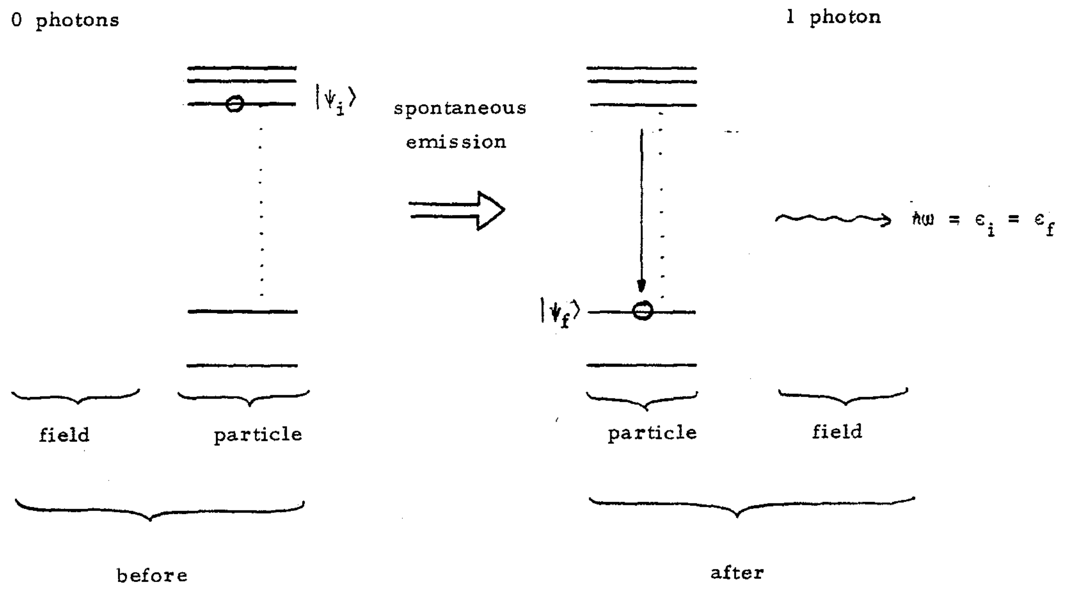
Transitions which result from the contribution of this matrix element are called <u>electric dipole transitions</u>. Transitions due to the next term would be <u>electric quadrupole transitions</u>, etc. Unless the electric dipole contributions vanish identically, they generally dominate all other radiative modes by a factor of the order $(ka_B)^{-1} \gtrsim 2/\alpha \sim 250$. This sequence is denoted by $\mathcal{E}1$, $\mathcal{E}2$, ... etc. in spectroscopic notation, where the $(\mathcal{E}n)^{\text{th}}$ transition rate is proportional to the square of the n^{th} moment of the particle wavefunction product $(\psi_f^*\,\psi_i)$, i.e.

$$(W_{i\to f})_{\mathcal{E}n} \propto |\langle \psi_f | \underbrace{\underline{x}\;\underline{x}\;\ldots\;\underline{x}}_{\text{n-factors}} | \psi_i \rangle|^2 \quad .$$

11.5 <u>Spontaneous lifetime of a charged particle in an excited state.</u>

Consider now a charged particle in an excited state in the absence of any external illumination. As hinted previously, however, even in the absence of any external fields, the excited particle may still interact with the ground state of the field to emit radiation. In the context of our previous discussion, we would like to compute the transition probability from an excited atomic state and an electromagnetic field ground state, to a lower atomic state and an electromagnetic field excited state.

[#] See appendix D, equation D.25 and related discussion.



Using the results of the preceding discussion, we may write the transition matrix element (equation 11.29; emission),

$$\langle\Psi_f|\underline{d}\cdot\underline{A}|\Psi_i\rangle = \left(\frac{\hbar}{2\varepsilon_0 V\omega_{\underline{k}}}\right)^{1/2} \hat{e}_{\underline{k},\ell}\cdot\underline{M}_{fi}(\underline{k}) \tag{11.36}$$

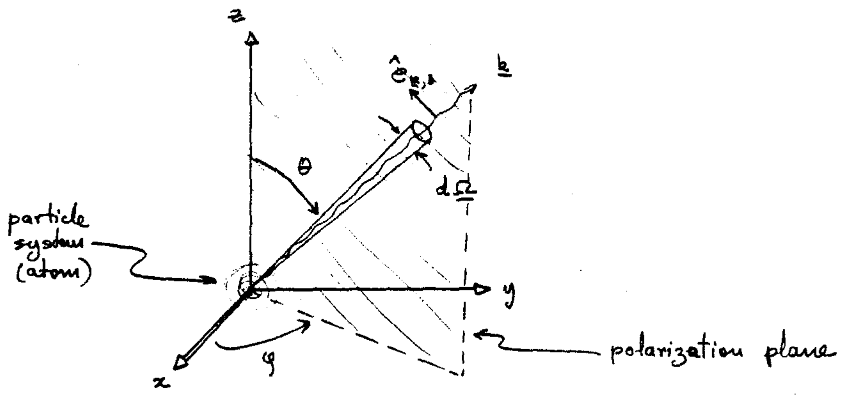
(note that $n^{(i)}_{\underline{k},\ell} = 0$). Substituting in equation 11.24 for the transition probability, then yields

$$W_{i\to f} = \frac{\pi\omega^2_{fi}}{\varepsilon_0 V\omega_{\underline{k}}}\ |\hat{e}_{\underline{k},\ell}\cdot\underline{M}_{fi}(\underline{k})|^2\ \delta(E_f - E_i) \quad . \tag{11.37}$$

To obtain a numerical value, we now need to integrate over the density of final states of the system. Since the particle is going to a single lower state, the system density of final states is dominated by the density of final states of the emitted photon, which is given by

$$\rho(E_f)\ dE_f = \rho(k)\ dk = V\ \frac{k^2 dk}{(2\pi)^3}\ \underline{d\Omega} \quad , \tag{11.38}$$

where $\underline{d\Omega}$ is the solid angle differential element around the wavevector $\underline{k}$ of the emitted photon.



Now since $k = \omega_k/c$, we have

$$\rho(E_f)\, dE_f = V \frac{\omega_k^2\, d\omega_k}{(2\pi c)^3}\, d\underline{\Omega} \qquad (11.38')$$

(where $d\underline{\Omega} = \sin\theta d\theta d\emptyset$), which yields

$$W_{i\to f} = \frac{\omega_{fi}^2}{(4\pi\varepsilon_0)2\pi c^3} \int\limits_{\text{final states}} \omega_k |\hat{e}_{\underline{k},\ell} \cdot \underline{M}_{fi}(\underline{k})|^2\, \delta[\hbar(\omega_{fi} + \omega_k)]\, d\omega_k\, d\underline{\Omega} \quad ,$$

or [recall $\delta(\alpha x) = \delta(x)/|\alpha|$] integrating over the final photon frequencies, we have ($\omega_k = |\omega_{fi}| = \omega$)

$$W_{i\to f} = \frac{\omega^3}{(4\pi\varepsilon_0)2\pi\hbar c^3} \int\limits_{\underline{\Omega}} |\hat{e}_{\underline{k},\ell} \cdot \underline{M}_{fi}(\underline{k})|^2\, d\underline{\Omega} \quad . \qquad (11.39)$$

Now, if θ is the angle between the wavevector $\underline{k}$ and the z-axis, which we may select to be in the plane formed by $\underline{k}$ and $\hat{e}_{\underline{k},\ell}$, we have[#]

$$\hat{e}_{\underline{k},\ell} \cdot \underline{M}_{fi}(\underline{k}) = |\underline{M}_{fi}(\underline{k})| \sin\theta \quad ,$$

and therefore

$$W_{i\to f} = \frac{1}{4\pi\varepsilon_0} \left(\frac{k^3}{h}\right) \int\limits_{\underline{\Omega}} |\underline{M}_{fi}(\underline{k})|^2 \sin^2\theta\, d\underline{\Omega} \quad , \qquad (11.40)$$

where, $k = \omega/c$ and $h = 2\pi\hbar$. Note that the particle coordinate system z-axis was chosen *to lie in the photon polarization plane* $(\underline{k}, \hat{e}_{\underline{k},\ell})$.

If the particle system is isotropic[†], we find that $|\underline{M}_{fi}(\underline{k})|^2$ will not be a function of orientation, and therefore we have

$$\int\limits_{\underline{\Omega}} |\underline{M}_{fi}(\underline{k})|^2 \sin^2\theta\, d\underline{\Omega} = |\underline{M}_{fi}(\underline{k})|^2 \int\limits_{\underline{\Omega}} \sin^2\theta\, d\underline{\Omega} = \frac{8\pi}{3} |\underline{M}_{fi}(\underline{k})|^2, \qquad (11.41)$$

which yields for the spontaneous transition rate,

$$\boxed{W_{i\to f} = \frac{1}{4\pi\varepsilon_0} \left(\frac{4k^3}{3\hbar}\right) |\underline{M}_{fi}(\underline{k})|^2} \quad , \qquad (11.42)$$

for an isotropic system.

---

update

[#] note that $\underline{k}$ and $\hat{e}_{\underline{k},\ell}$ are orthogonal.

[†] and only if. This may not be true in the presence of a magnetic field, or in an anisotropic crystal, for example.

These results can be simplified in the case of electric dipole radiation (long wavelength approximation), where

$$\underline{M}_{fi}(\underline{k}) \simeq \underline{M}_{fi}(0) = q_c \langle \psi_f | \underline{x} | \psi_i \rangle \equiv q_c \, \underline{x}_{fi} \equiv \underline{d}_{fi} \quad ,$$

so that we may write the transition rate (for an electron, in particular; $q_c = -e$) as,

$$\boxed{(W_{i \to f})_{\mathcal{E}1} = \alpha \left(\frac{\omega^3}{2\pi c^2}\right) \int_{\underline{\Omega}} |\underline{x}_{fi}|^2 \sin^2\theta \, d\underline{\Omega}} \quad , \qquad (11.43)$$

which, for an isotropic system, reduces to,

$$\boxed{(W_{i \to f})_{\mathcal{E}1} = \alpha \left(\frac{4\omega^3}{3c^2}\right) |\underline{x}_{fi}|^2} \quad , \qquad (11.44)$$

where α is the fine structure constant,

$$\alpha = \frac{1}{4\pi\varepsilon_0} \left(\frac{e^2}{\hbar c}\right) \simeq \frac{1}{137} \quad .$$

We should point out that this transition rate is in fact the A_{if} coefficient of spontaneous emission in the radiation theory of Einstein (see first footnote at the bottom of page 11.4 and related discussion), explicitly computed here from first principles.

It is convenient to express the final result as a characteristic reciprocal time for electric dipole radiation, times remaining dimensionless quantities. In particular, for electric dipole radiation of an isotropic system (equation 11.44), we have

$$\begin{aligned}(W_{i \to f})_{\mathcal{E}1} &= \alpha \left(\frac{4\omega^3}{3c^2}\right) |\underline{x}_{if}|^2 \\ &= \frac{4\alpha c}{3} \left(\frac{E_R}{\hbar c}\right) \left|\frac{E_R}{\hbar c} \underline{x}_{fi}\right|^2 \left(\frac{\hbar\omega}{E_R}\right)^3 \\ &= \frac{\alpha^4 c}{6 a_B} |\underline{x}_{fi}/a_B|^2 \left(\frac{\hbar\omega}{E_R}\right)^3 \quad ,\end{aligned}$$

where E_R is the Rydberg (hydrogen aton binding energy; $E_R/\hbar c = \alpha/2a_B$), α is the fine structure constant, and a_B is the Bohr radius.

The characteristic time that emerges for electric dipole radiation is then equal to

$$\tau_{\mathcal{E}1} \equiv 6\left(\frac{a_B}{c}\right) \alpha^{-4} = 0.3735 \times 10^{-9} \text{ sec} \quad , \qquad (11.45)$$

in terms of which, the electric dipole spontaneous emission rate from an excited isotropic system becomes

$$\boxed{(W_{i\to f})_{\mathcal{E}1} = |\underline{x}_{fi}/a_B|^2 \; (\hbar\omega/E_R)^3 \; \tau_{E1}^{-1}} \quad , \tag{11.46a}$$

where

$$\tau_{\mathcal{E}1}^{-1} = 2.677 \times 10^9 \text{ sec}^{-1} \quad . \tag{11.46b}$$

<u>Example 11. 1</u> <u>Rate of spontaneous emission from an excited hydrogen atom in the 2p state.</u>

We need to compute the matrix elements for $m' = 0, \pm 1$, i.e.

$$\begin{aligned} \underline{x}_{fi}/a_B &= \langle n'', \ell'', m'' | \underline{x}/a_B | n', \ell', m' \rangle \\ &= \langle 1, 0, 0 \, | \underline{x}/a_B | 2, 1, m' \rangle \\ &= \frac{1}{a_B} \int_0^\infty R^*_{1,0} \, (r R_{2,1}) \, r^2 dr \cdot \int_{\underline{\Omega}} Y^*_{0,0} \, (\underline{x}/r) \, Y_{1,m'} d\underline{\Omega} \quad , \end{aligned}$$

where (see section 7. 3),

$$|n, \ell, m\rangle = R_{n,\ell}(r) \; Y_{\ell,m}(\theta, \varphi)$$

are the hydrogen atom eigenstates.

The radial integral is given by

$$\frac{1}{a_B} \int_0^\infty R^*_{1,0} \, (r R_{2,1}) \, r^2 dr = \int_0^\infty R^*_{1,0}(\rho) \; R_{2,1}(\rho) \; \rho^3 d\rho$$

where (see equations 7. 74),

$$\rho = r/a_B \; , \qquad R_{2,1}(\rho) = \frac{1}{2\sqrt{6}} \, \rho \, e^{-\rho/2} \quad , \qquad R_{1,0}(\rho) = 2e^{-\rho}$$

and therefore

$$\frac{1}{a_B} \int_0^\infty R^*_{1,0} \, (r R_{2,1}) \, r^2 dr = \frac{1}{\sqrt{6}} \int_0^\infty \rho^4 \; e^{-3\rho/2} \; d\rho^{\#} = \frac{1}{\sqrt{6}} \left(\frac{2}{3}\right)^5 4! = 1.290 \quad . \tag{11.47}$$

The integral over the spherical harmonics can be evaluated by noting that $Y_{0,0}(\theta, \varphi) = (4\pi)^{-1/2}$ is a constant and that $\underline{x}/r$ can be expressed in terms of the $\ell = 1$ spherical harmonics, i.e.

[#] $$\int_0^\infty x^n \, e^{-x} \, dx = n!$$

$$\underline{x}/r = \sin\theta(\cos\varphi\ \hat{e}_x + \sin\varphi\ \hat{e}_y) + \cos\theta\ \hat{e}_z$$

or

$$\underline{x}/r = (\tfrac{4\pi}{3})^{1/2}\ [\frac{1}{\sqrt{2}}\ Y_{1,-1} - Y_{1,1})\ \hat{e}_x + \frac{i}{\sqrt{2}}(Y_{1,-1} + Y_{1,1})\ \hat{e}_y + Y_{1,0}\ \hat{e}_z] \quad , \tag{11.48}$$

so that

$$\int_{\underline{\Omega}} Y^*_{0,0}\ (\underline{x}/r) Y_{1,m'} d\underline{\Omega} = \frac{1}{\sqrt{3}}\ [\frac{1}{\sqrt{2}}\int (Y_{1,-1} - Y_{1,1})^*\ Y_{1,m'}\ d\underline{\Omega}\ \hat{e}_x$$

$$- \frac{i}{\sqrt{2}}\int (Y_{1,-1} + \underline{Y_{1,1}})^*\ Y_{1,m'} d\underline{\Omega}\ \hat{e}_y$$

$$+ \int\ Y^*_{1,0}\ Y_{1,\underline{m}'}\ d\underline{\Omega}\ \hat{e}_z] \quad ,$$

[note that $(\underline{x}/r) = (\underline{x}/r)^*$] and therefore, using the orthonormality of the spherical harmonics,

$$\int_{\underline{\Omega}} Y^*_{0,0}(\underline{x}/r)\ Y_{1,m'} d\underline{\Omega} = \frac{1}{\sqrt{3}}\ [\frac{1}{\sqrt{2}}(\delta_{m',1} - \delta_{m',1})\hat{e}_x - \frac{i}{\sqrt{2}}(\delta_{m',-1} + \delta_{m',1})\hat{e}_y + \delta_{m',0}\ \hat{e}_z] \quad . \tag{11.49}$$

Consequently, we have, <u>independently of m'</u> (why?).

$$|\int_{\underline{\Omega}} Y^*_{0,0}(\underline{x}/r)\,Y_{1,m'}\,d\underline{\Omega}\,|^2 = \frac{1}{3}\ (\delta_{m',-1} + \delta_{m',0} + \delta_{m',1}) = \frac{1}{3} \quad . \tag{11.50}$$

Combining equations 11.47 and 11.50, we then compute the (dimensionless) matrix element of $\underline{x}$,

$$|\langle 1s|\underline{x}/a_B|2p\rangle|^2 = 1.290^2/3 = 0.5549 \quad . \tag{11.51}$$

We also need to compute the photon energy (in units of the Rydberg), i.e.

$$\frac{\hbar\omega}{E_R} = \frac{\varepsilon_{2p} - \varepsilon_{1s}}{E_R} = -\frac{1}{4} + 1 = \frac{3}{4} \quad .$$

Substituting these two results into the electric dipole spontaneous emission equation (11.47), we finally have

$$(W_{2p\to 1s})_{\mathcal{E}1} = 0.5549\ (\tfrac{3}{4})^3\ \tau^{-1}_{\mathcal{E}1} = 0.2341\ \tau^{-1}_{\mathcal{E}1} \tag{11.52}$$

and, using the value of the characteristic time for electric dipole emission, we obtain

$$(W_{2p\to 1s})_{\mathcal{E}1} = 0.6267 \times 10^9\ \text{sec}^{-1} \quad ,$$

corresponding to a lifetime with respect to spontaneous emission of

$$(\tau_s)_{th} \equiv W^{-1}_{i\to f} = 1.596\ \text{nsec}\ ; \quad (2p\to 1s) \tag{11.53}$$

update

The experimentally[#] determined value for this quantity is equal to

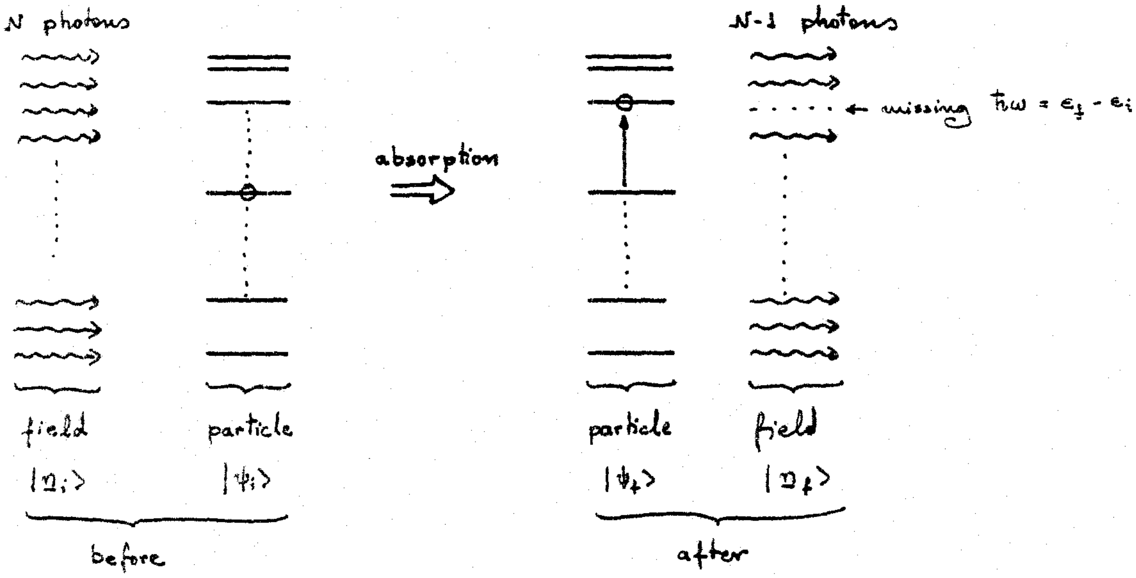
---

[#] W.S. Bickel and A.S. Goodman (1966), <u>Phys. Rev. 148</u>, page 1.

$$(\tau_s)_{\exp} = (1.600 \pm 0.004)\ \text{nsec}\ ;\quad (2p\to 1s)\ . \tag{11.54}$$

11.6 Induced absorption and emission.

Let us consider a charged particle in a cavity filled with electromagnetic radiation. The charged particle can undergo transitions both up and down, absorbing or emitting photons respectively. Consider first the case of absorption by the charged particle subsystem.



The initial and final system states are then given by:

$$|\Psi_i\rangle = |\psi_i\rangle|n^{(i)}_{\underline{k}_1,1},\ \ldots\ldots,\ n^{(i)}_{\underline{k},\ell},\ \ldots\ldots\ \rangle$$
$$|\Psi_f\rangle = |\psi_f\rangle|n^{(i)}_{\underline{k}_1,1},\ \ldots\ldots,\ (n^{(i)}_{\underline{k},\ell} - 1),\ \ldots\ldots\ \rangle \tag{11.55}$$

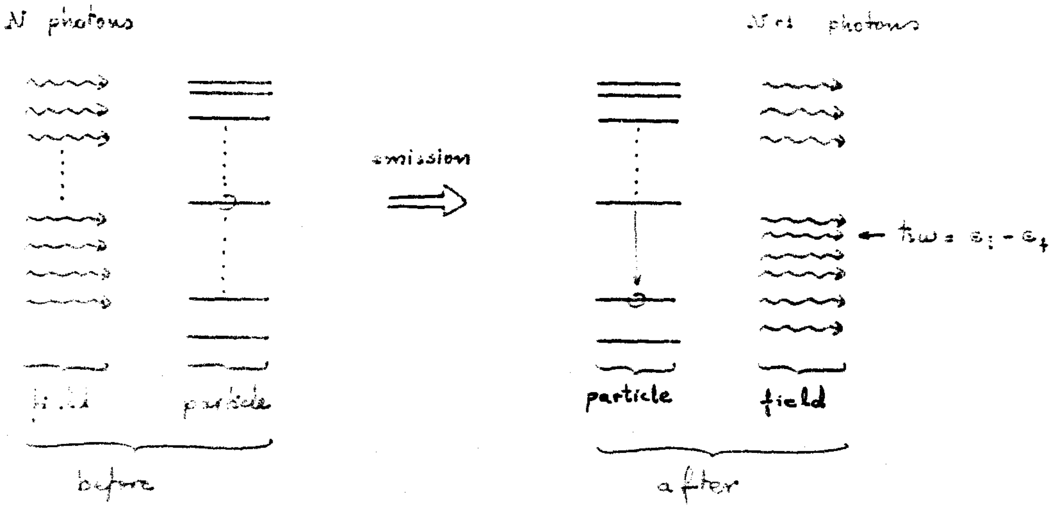
taking the particle to a higher energy state $|\psi_f\rangle$ from $|\psi_i\rangle$, at the loss of one photon in the mode $(\underline{k},\ell)$ by the field. The transition matrix element is now given by equation 11.29 (case b), or

$$\langle\Psi_f|\underline{d}\cdot\underline{A}|\Psi_i\rangle = \left(\frac{\hbar}{2\epsilon_0 V\omega_k}\right)^{1/2} [\hat{e}_{\underline{k},\ell}\cdot\underline{M}_{fi}(-\underline{k})]\ \sqrt{n^{(i)}_{\underline{k},\ell}}\ ,$$

which, substituted in equation 11.24 for the transition rate, yields

$$\boxed{W_{i\to f} = \frac{\pi\omega^2_{fi}}{\epsilon_0 V\omega_k}\ |\hat{e}_{\underline{k},\ell}\cdot\underline{M}_{fi}(-\underline{k})|^2\ n^{(i)}_{\underline{k},\ell}\ \delta(\epsilon_f - \epsilon_i - \hbar\omega_k)}\ . \tag{11.56}$$

In the case of emission,



the initial and final states are given by

$$\begin{aligned} |\Psi_i\rangle &= |\psi_i\rangle |n^{(i)}_{\underline{k}_1,1}, \ldots\ldots, n^{(i)}_{\underline{k},\ell}, \ldots\ldots\rangle \\ |\Psi_f\rangle &= |\psi_f\rangle |n^{(i)}_{\underline{k}_1,1}, \ldots\ldots, (n^{(i)}_{\underline{k},\ell}+1), \ldots\ldots\rangle , \end{aligned} \qquad (11.57)$$

where $|\psi_f\rangle$ is now a <u>lower</u> energy particle state, with the field <u>gaining</u> a photon in the mode $(\underline{k}, \ell)$. The transition matrix element, in this case is given by (equation 11.29, case b),

$$\langle\Psi_f|\underline{d}\cdot\underline{A}|\Psi_i\rangle = \left(\frac{\hbar}{2\varepsilon_0 V\omega_k}\right)^{1/2} [\hat{e}_{\underline{k},\ell}\cdot\underline{M}_{fi}(\underline{k})] \sqrt{n^{(i)}_{\underline{k},\ell}+1} \quad ,$$

resulting in a transition rate given by

$$\boxed{W_{i\to f} = \frac{\pi\omega^2_{fi}}{\varepsilon_0 V\omega_k} |\hat{e}_{\underline{k},\ell}\cdot\underline{M}_{fi}(\underline{k})|^2 (n^{(i)}_{\underline{k},\ell}+1)\, \delta(\varepsilon_f - \varepsilon_i + \hbar\omega_k)} \quad . \qquad (11.59)$$

For both absorption and emission, it is difficult to derive general formulae, since the final result depends in an explicit way on the details of each particular problem. It is preferable to compute the transition rate in each case, starting from equations 11.57 and 11.59, which hold in general,[#] for absorption and emission respectively. Their use will be illustrated through examples in the discussion below.

Example 11.2 Particle transitions as a result of diffuse cavity radiation.

update

Consider <u>absorption</u> first. This represents a transition <u>from</u> a cluster of states for the <u>absorbed</u> photon. The total transition rate is then given by equation 11.56, integrated over

---

[#] provided the spread in ε_f and ε_i can be ignored. See sections 8.5 and 8.6.

the absorbed photon density of states, which for an electromagnetic field in a cavity of volume V, is given by,

$$\rho(\underline{k})\, d\underline{k} = \frac{V}{(2\pi)^3}\, d^3\underline{k} = \frac{V}{(2\pi)^3}\, k^2 dk\, d\underline{\Omega} = \frac{V}{(2\pi c)^3}\, \omega_k^2 d\omega_k\, d\underline{\Omega} \quad , \tag{11.60}$$

and therefore

$$[W_{i\to f}]_{abs} = \frac{\pi\omega_{fi}^2}{\epsilon_0 V} \sum_{\ell} \int_{\underline{k}} \frac{1}{\omega_k}\, n^{(i)}_{\underline{k},\ell} |\hat{e}_{\underline{k},\ell} \cdot \underline{M}_{fi}(-\underline{k})|^2\ \delta(\epsilon_f - \epsilon_i - \hbar\omega_{\underline{k}})\ \rho(\underline{k})\, d\underline{k}$$

$$= \frac{1}{4\pi\epsilon_0}\left(\frac{\omega_{fi}^2}{2\pi c^3}\right) \sum_{\ell} \int_{\underline{k}} \omega_k\, n^{(i)}_{\underline{k},\ell} |\hat{e}_{\underline{k},\ell} \cdot \underline{M}_{fi}(-\underline{k})|^2\ \delta[\hbar(\omega_{fi} - \omega_k)]\, d\omega_k\, d\underline{\Omega}$$

or

$$[W_{i\to f}]_{abs} = \frac{1}{4\pi\epsilon_0}\left(\frac{k^3}{2\pi\hbar}\right) \sum_{\ell} \int_{\underline{\Omega}} n^{(i)}_{\underline{k},\ell} |\hat{e}_{\underline{k},\ell} \cdot \underline{M}_{fi}(-\underline{k})|^2\, d\underline{\Omega} \quad , \tag{11.61}$$

where $\omega_k = \omega_{fi} = \omega$ and $k = \omega/c$.

Now <u>if the electromagnetic field is isotropic</u>, i.e.

$$n^{(i)}_{\underline{k},\ell} = n^{(i)}_{\ell}(\omega) \tag{11.62}$$

or, the modes for a given frequency ω and polarization ℓ are uniformly occupied with respect to the direction of the corresponding wavevectors $\underline{k} = \frac{\omega}{c}\hat{e}_{\underline{k}}$, we have

$$[W_{i\to f}]_{abs} = \frac{1}{4\pi\epsilon_0}\left(\frac{k^3}{2\pi\hbar}\right) \sum_{\ell} n^{(i)}_{\ell}(\omega) \int_{\underline{\Omega}} \underline{M}_{fi}(-\underline{k})|^2\ \sin^2\theta_{\ell}\, d\underline{\Omega} \quad , \tag{11.63}$$

which, in the case of an <u>isotropic particle system</u>, can be written as (see equation 11.41 and related discussion)

$$[W_{i\to f}]_{abs} = \frac{1}{4\pi\epsilon_0}\left(\frac{4k^3}{3\hbar}\right) |\underline{M}_{fi}(\underline{k})|^2\ n^{(i)}(\omega) \quad , \tag{11.64a}$$

where

$$n^{(i)}(\omega) = n_1^{(i)}(\omega) + n_2^{(i)}(\omega) \tag{11.64b}$$

is the total number of photons (both polarizations) of frequency ω. In the case of electric dipole transitions, this simplifies further to

$$[W_{i\to f}]_{\mathcal{E}1\ abs} = \alpha\left(\frac{4\omega^3}{3c^2}\right) |\underline{x}_{fi}|^2\ n^{(i)}(\omega) \quad , \tag{11.65}$$

where α is the fine structure constant.

To compute the <u>emission</u> rate, note that this represents a transition <u>to</u> a cluster of states available for the emitted photon, which in the case of a diffuse[#] excitation field is given by equation 11.60. Integrating the transition rate to a single state (equation 11.59) over the density of final states, we then have (see equation 11.61 and related discussion),

$$[W_{i\to f}]_{em} = \frac{1}{4\pi\epsilon_0} \left(\frac{k^3}{2\pi\hbar}\right) \sum_{\ell} \int_{\underline{\Omega}} (n^{(i)}_{\underline{k},\ell} + 1) |\hat{e}_{\underline{k},\ell} \cdot \underline{M}_{fi}(\underline{k})|^2 \, d\underline{\Omega} \quad , \tag{11.66}$$

which under the assumptions of isotropy of both the field and the particle system, reduces, in a manner similar to the case of absorption (see equation 11.64 and related discussion), to

$$[W_{i\to f}]_{em} = \frac{1}{4\pi\epsilon_0} \left(\frac{4k^3}{3\hbar}\right) |\underline{M}_{fi}(\underline{k})|^2 \, [n^{(i)}(\omega) + 1] \quad , \tag{11.67}$$

and in the case of electric dipole radiation,

$$[W_{i\to f}]_{\mathcal{E}1\,em} = \alpha \left(\frac{4\omega^3}{3c^2}\right) |\underline{x}_{fi}|^2 \, [n^{(i)}(\omega) + 1] \quad . \tag{11.68}$$

There are several noteworthy features of those solutions:

1. The probability per unit time for the particle to absorb a photon is proportional to the available number of photons to be absorbed.

2. The probability per unit time for the particle to emit a photon is proportional to the number of photons, of the right frequency, inducing it to make the transition, plus one. If there are no photons initially in the right modes $[n^{(i)}(\omega) = 0]$, we recover the spontaneous emission formula.

3. The difference of one photon between the two rates is essential for detailed balance of the rates of the two processes (emission and absorption), as required for the combined [particle + field] system to be able to attain equilibrium. Consider an initial field state with n photons of the proper frequency and the particle in some lower state. One photon is now absorbed, raising the particle to a higher state, and leaving the field with (n - 1) photons. The rate for this process is n/τ_s, where τ_s is the spontaneous lifetime corresponding to the reverse transition. The particle now is in the higher state and may come back to the lower state emitting a photon. The rate for this process is now the number of photons in the field (n - 1), plus one, divided by the spontaneous lifetime, i.e. $[(n - 1) + 1]/\tau_s = n/\tau_s$, equal, in other words, to the reverse process, as required by the equilibrium condition. It was this reasoning that

---

[#] as opposed to a laser, for example which may have coherent spatial and frequency modes, to be discussed later.

led Einstein to conclude that the mechanism of spontaneous emission was required for equilibrium to be possible.[#]

11.7 Electric multipole radiative transition selection rules.

We have already seen that, in the weak field approximation (first order in the vector potential), the properties of the matrix element between the initial and final field states, restrict the transitions to the emission and absorption of a single photon[†]. See equations 11.28 and related discussion. It is possible to deduce general conclusions about the particle states that can be connected radiative transitions (within the weak field approximation), from the matrix element $\underline{M}_{fi}(\underline{k})$ over the corresponding particle states. See equation 11.26, section 11.4 and related discussions. In particular, for electric dipole transitions, i.e.

$$[\underline{M}_{fi}(\underline{k})]_{\mathcal{E}1} \simeq q_c \langle \psi_f | \underline{x} | \psi_i \rangle \quad ,$$

we have to compute the matrix element of

$$\underline{x} = r(\frac{4\pi}{3})^{1/3} [\frac{1}{\sqrt{2}} (Y_{1,-1} - Y_{1,1}) \hat{e}_x + \frac{i}{\sqrt{2}} (Y_{1,-1} + Y_{1,1}) \hat{e}_y + Y_{1,0} \hat{e}_z] \quad ,$$

see equation 11.48), or using the right and left circular unit vectors

$$\hat{e}_R = \frac{1}{\sqrt{2}} (\hat{e}_x + i \hat{e}_y) \quad , \qquad \hat{e}_L = \frac{1}{\sqrt{2}} (\hat{e}_x - i \hat{e}_y) \quad , \tag{11.69}$$

we also have[‡]

$$\underline{x} = r(\frac{4\pi}{3})^{1/3} (Y_{1,-1} \hat{e}_R - Y_{1,1} \hat{e}_L + Y_{1,0} \hat{e}_z) \quad , \tag{11.70}$$

between the initial and final particle states, which can be expressed as products of radial and angular parts[*], i.e. $[\underline{\Omega} = (\theta, \varphi)]$

$$|\psi_i\rangle = R_{n'\ell'}(r) Y_{\ell', m'}(\underline{\Omega}) \quad , \qquad |\psi_f\rangle = R_{n''\ell''}(r) Y_{\ell'', m''}(\underline{\Omega}) \quad . \tag{11.71}$$

This results in a matrix element which is a product of radial and angular parts[*], i.e.

[#] See Eisberg and Resnick, Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles (John Wiley, 1974), pp. 426-429.

[†] Second order in the vector potential involves two photons, third order three, etc. See sketches on pages 8.30 and 8.31.

[‡] note how the circular unit vectors are selected in a natural way, even though we started with cartesian unit vectors. See also section 9.10.

[*] or linear superpositions of products of radial and angular parts.

update

$$\underline{M} = \int_0^\infty R^*_{n''\ell''}(r)\; R_{n'\ell'}(r)\; r^3\, dr \cdot \int_{\underline{\Omega}} Y^*_{\ell'',m''}(\underline{\Omega})\; Y_{1,m}(\underline{\Omega})\; Y_{\ell' m'}(\underline{\Omega})\; d\underline{\Omega} \quad . \tag{11.72}$$

It can be seen that the angular integral restricts the possible values of (ℓ'', m'') of the final state. In particular, we must have[#]

$$\ell'' = \ell' \pm 1 \quad , \tag{11.73a}$$

and

$$m'' = m' + m \quad ,$$

or since $m = 0, \pm 1$,

$$m'' = m' \qquad \text{or} \qquad m'' = m' \pm 1 \quad . \tag{11.73b}$$

The change from ℓ' to $\ell' \pm 1$, corresponding to a change of one unit of $\hbar$ in the angular momentum of the charged particle, is carried away by the photon, which has a spin angular momentum of $\pm\hbar$ along its direction of propagation (wavevector $\underline{k}$), and is therefore emitted in a state of zero <u>orbital momentum</u>. See discussion on photon angular momentum.

The next higher order transitions[†] are <u>electric quadrupole</u> ($\mathcal{E}2$) transitions, involving the quadrupole (tensor) moment (equation 11.34a)

$$\underline{\underline{Q}}_{fi} \equiv \langle \psi_f | \underline{x}\; \underline{x} | \psi_i \rangle \quad ,$$

involving elements like

$$(\underline{\underline{Q}}_{fi})_{xx} = \langle \psi_f | x^2 | \psi_i \rangle \quad , \quad (\underline{\underline{Q}}_{fi})_{xy} = \langle \psi_f | xy | \psi_i \rangle \quad , \quad (\underline{\underline{Q}}_{fi})_{yz} = \langle \psi_f | yz | \psi_i \rangle \quad \text{etc.}$$

From the expression of $\underline{x}$ in terms of the spherical harmonics, we see that the quadrupole matrix involves matrix elements of products of pairs of the $\ell = 1$ spherical harmonics, i.e.

$$(\underline{\underline{Q}}_{fi}) \propto \sum_{m_1, m_2} c_{m_1, m_2} \langle \psi_f | r\; Y_{1,m_1}\; Y_{1,m_2} | \psi_i \rangle \quad . \tag{11.74}$$

These products can in turn be expressed as linear combinations of single spherical harmonics (see equation 10.36a), and we deduce the corresponding selection rules for electric quadrupole radiation, i.e.

$$\ell'' = \begin{cases} \ell' \quad \text{or} \quad |\ell' \pm 2| \quad , & \text{if} \quad \ell' \neq 0 \\ 2 \quad , \quad \text{if} \quad \ell' = 0 \end{cases} \tag{11.75a}$$

and

$$m'' = m' \; , \quad m' \pm 1 \; , \quad m' \pm 2 \; , \tag{11.75b}$$

[#] see equation 10.36 and related discussion on addition of angular momenta, section 10.2.

[†] which become important if electric dipole transitions are forbidden by the previously established selection rules,

where $|\psi_i\rangle$ and $|\psi_f\rangle$ are given by equations 11.71.

Consequently, electric quadrupole photons can change the angular momentum of the charged particle by two units of $\hbar$, or leave it unchanged [but $(\ell'=0) \not\to (\ell''=0)$]. They consequently correspond to excitations of photon modes of higher angular momentum. See discussion of photon angular momentum.

11.8 Magnetic radiative transitions.

In discussing radiative transitions, we have so far ignored the possible spin of the charged particle, in the presence of which the interaction Hamiltonian must be augmented to include the associated spin magnetic moment coupling to the magnetic field, i.e.[#]

$$H_{int} = -\frac{q_c}{m}\,\underline{p}\cdot\underline{A} + \frac{q_c^2}{2m}\,\underline{A}\cdot\underline{A} - \underline{\mu}^{(s)}\cdot\underline{\mathcal{B}} \;, \tag{11.76}$$

where $\underline{\mu}^{(s)}$ is the spin magnetic moment of the particle, or, in terms of the vector potential

$$H_{int} = -\frac{q_c}{m}\,\underline{p}\cdot\underline{A} + \frac{q_c^2}{2m}\,\underline{A}\cdot\underline{A} - \underline{\mu}^{(s)}\cdot\left(\frac{\partial}{\partial\underline{x}} \times \underline{A}\right) \;. \tag{11.77}$$

It can be seen that this term also contributes to first order in the vector potential[†], which must therefore also be considered when computing first order (single photon) transitions in the weak field approximation. In particular,

$$H_{int} = H' + H'' + 0(A^2) \tag{11.78}$$

where H' is proportional to the charge q_c (equation 11.18) and leads to the electric dipole, quadrupole, etc., transitions we have already discussed, and

$$H'' = -\,\underline{\mu}^{(s)}\cdot\left(\frac{\partial}{\partial\underline{x}} \times \underline{A}\right) \;, \tag{11.79}$$

is proportional to the spin magnetic moment $\underline{\mu}^{(s)}$ and is responsible for magnetic radiative transitions.[‡] In particular, substituting for the spin magnetic moment,

$$\underline{\mu}^{(s)} = g\left(\frac{q_c}{2m}\right)\underline{S} \tag{11.80}$$

where g is the spin g-factor, m is the mass of the particle and $\underline{S}$ is the particle spin angular momentum, we have

---

[#] see equations 10.42, 10.44 and related discussion, and appendix F, equation F.46 and related discussion.

[†] even though as we will see leads to considerably weaker transitions,

[‡] note that a neutron which is neutral but possesses a spin magnetic moment (equation 10.44b) may participate in such transitions.

$$H'' = -\, g\left(\frac{q_c}{2m}\right) \underline{S} \cdot \left(\frac{\partial}{\partial \underline{x}} \times \underline{A}\right) \tag{11.81}$$

(note $q_c = -e$ for an electron).

Transitions caused by this interaction are then characterized by a rate

$$(W_{i\to f})_m = \frac{2\pi}{\hbar} \left|-\, g\left(\frac{q_c}{2m}\right) \langle \psi_f | \underline{S} \cdot \left(\frac{\partial}{\partial \underline{x}} \times \underline{A}\right) | \psi_i \rangle\right|^2 \delta(E_f - E_i)$$

or

$$(W_{i\to f})_m = \frac{2\pi}{\hbar} \left(\frac{g\, q_c}{2m}\right)^2 \left|\langle \psi_f | \underline{S} \cdot \left(\frac{\partial}{\partial \underline{x}} \times \underline{A}\right) | \psi_i \rangle\right|^2 \delta(E_f - E_i) \quad . \tag{11.82}$$

The transition matrix element can now be expressed using the vector potential quantum operator (equation 9.99), i. e.[#]

$$\langle \psi_f | \underline{S} \cdot \left(\frac{\partial}{\partial \underline{x}} \times \underline{A}\right) | \psi_i \rangle = \left(\frac{\hbar}{2\varepsilon_0 V}\right)^{1/2} \sum_{\underline{k},\sigma} \mathrm{sign}(\sigma) \frac{k}{\omega_k^{1/2}} \langle \psi_f | \underline{S}\, e^{i\underline{k}\cdot\underline{x}} (a_{\underline{k},\sigma} - a^{\dagger}_{-\underline{k},\sigma}) | \psi_i \rangle\, \hat{e}_{\underline{k},\sigma} \tag{11.83}$$

which can be separated, as before, into a product of a matrix element over the particle coordinates and a matrix element over the field coordinates, i. e.

$$\langle \psi_f | \underline{S}\, e^{i\underline{k}\cdot\underline{x}} (a_{\underline{k},\sigma} - a^{\dagger}_{-\underline{k},\sigma}) | \psi_i \rangle = \underbrace{\langle \psi_f | \underline{S}\, e^{i\underline{k}\cdot\underline{x}} | \psi_i \rangle}_{\text{particle}} \; \underbrace{\langle \underline{n}_f | (a_{\underline{k},\sigma} - a^{\dagger}_{-\underline{k},\sigma}) | \underline{n}_i \rangle}_{\text{field}} \quad . \tag{11.84}$$

The matrix element over the field coordinates is similar to the corresponding matrix element for the $(\underline{p}\cdot\underline{A})$ interactions (see equations 11.26 and 11.28, and related discussion). The matrix element over the particle coordinates can be factored further into a product of a matrix element over the particle position coordinates and a matrix element involving the particle spin states, i. e. if

$$|\psi_i\rangle = |n', \ell', m'_\ell; m'_s\rangle$$
$$|\psi_f\rangle = |n'', \ell'', m''_\ell; m''_s\rangle^{\dagger} \quad , \tag{11.85}$$

we have,

$$\langle \psi_f | \underline{S}\, e^{i\underline{k}\cdot\underline{x}} | \psi_i \rangle = \langle n'', \ell'', m''_\ell | e^{i\underline{k}\cdot\underline{x}} | n', \ell', m'_\ell \rangle \; \langle m''_s | \underline{S} | m'_s \rangle \quad . \tag{11.85}$$

[#] note that $(\sigma = R, L)$

$$i\underline{k} \times \hat{e}_{\underline{k},\sigma} = k\ \mathrm{sign}(\sigma)\ \hat{e}_{\underline{k},\sigma} \quad ,$$

where $\mathrm{sign}(R) \equiv 1$ and $\mathrm{sign}(L) \equiv -1$.

[†] we had not included the particle spin states in the previous calculations, since they are unaffected by the $(\underline{p}\cdot\underline{A})$ interactions.

It can be seen that in the long wavelength approximation[#], corresponding to magnetic dipole ($\mathcal{M}1$) transitions, the position state of the particle is unaffected, i.e. we must have, for a non-zero matrix element,

$$n'' = n' \ , \quad \ell'' = \ell' \ , \quad m''_\ell = m'_\ell \ , \tag{11.87}$$

while from the matrix element over the spin states we have, from the matrix elements of S_z,

$$m''_s = m'_s \ , \tag{11.88a}$$

or, from the matrix elements of S_x and S_y,

$$m''_s = m'_s \pm 1 \ . \tag{11.88b}$$

It can be seen, however, that these matrix elements can only result in transitions in the presence of perturbations that can lift the energy degeneracy between different spin states sharing the same spatial wavefunctions (no energy left for the photon!). This may be the consequence of coupling to additional degrees of freedom, e.g. nuclear spin [nuclear magnetic resonance (NMR) spectroscopy], or any external influences that lift the spin degeneracy, e.g. Zeeman splitting of different spin states of a charged particle in a magnetic field, or internal perturbations, e.g. spin-orbit coupling, relativistic corrections, etc.

Magnetic dipole transitions are generally even weaker than electric quadrupole transitions and in practice need only be considered when both electric dipole and electric quadrupole matrix elements vanish identically. From a simple scaling analysis, the ratio of magnetic dipole to electric dipole transition rate is of the order of

$$\frac{(W_{i\to f})_{\mathcal{M}1}}{(W_{i\to f})_{\mathcal{E}1}} \propto \frac{|\langle\psi_f|H''|\psi_i\rangle|^2}{|\langle\psi_f|H'|\psi_i\rangle|^2} \sim \left(\frac{\hbar k}{mc}\right)^2 = \left(\frac{\hbar\omega}{mc^2}\right)^2 \tag{11.89}$$

which, for electron transitions ($mc^2 \simeq 0.5$ MeV) in the visible ($\hbar\omega \sim 2$eV) is of the order of 10^{-11}.

Before leaving this section, it should be noted that all these results and selection rules we have discussed so far should not be considered as absolute rigid rules. Particle motion perturbations, such as spin-orbit coupling (section 10.5), other relativistic effects (e.g. section 10.6), coupling between the electron magnetic moment with the nuclear magnetic moment, as well as electron-electron interactions in multi-electron systems, introduce sufficient complexity into the analysis of real life radiative transition problems to suggest a

---

[#] i.e. $e^{i\underline{k}\cdot\underline{x}} \simeq 1$

certain degree of caution in using these results. They nevertheless remain a very useful starting point for further refinements and can always be relied on to provide a very powerful set of guidelines. For a more detailed treatment, the interested reader should consult specialized texts in spectroscopy.#

#Alan Corney, Atomic and Laser Spectroscopy (Oxford University Press, 1977), for example, as well as the texts by Gerhard Herzberg on molecular spectroscopy, published by Van Nostrand, to name a few.

APPENDIX A: Calculus of Variations.

Given a scalar function

$$L = L[\underline{q}(t), \dot{\underline{q}}(t), t] \ , \qquad (A.1)$$

the problem of determining the (vector) path $\underline{q}(t)$ on which the integral

$$I\{\underline{q}\} = \int_{t_1}^{t_2} L(\underline{q}, \dot{\underline{q}}, t)\, dt \qquad (A.2)$$

has an <u>extremum</u> is a problem in the Calculus of Variations.

To simplify the thinking, consider first a function of x and $y(x)$ of the form

$$F = F[y(x), y'(x), x] \ , \qquad (A.3)$$

where

$$y'(x) = \frac{d}{dx} y(x) \ .$$

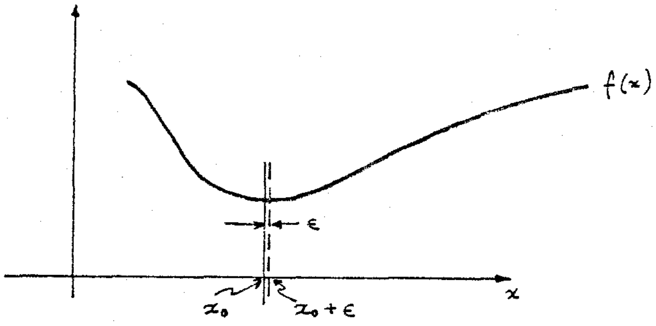
We wish to find the curve $y(x)$ that extremizes (minimizes or maximizes) the definite integral,

$$J\{y\} = \int_{x_1}^{x_2} F[y(x), y'(x), x]\, dx \ . \qquad (A.4)$$

What does it mean to extremize the integral J with respect to the curve $y(x)$? Well, let us recall what it means to minimize (or maximize) any function $f(x)$ with respect to a variable, x in this case. If we have

$$\min\{f(x)\} = f(x_0) \ ,$$

i.e., if the minimum is attained for $x = x_0$, then a small variation $x \rightarrow x_0 + \epsilon$ leaves the function $f(x)$ unchanged to first order in ϵ.



In other words, if $\min\{f(x)\} = f(x_0)$, then

$$\delta f \Big|_{x_0} = f(x_0 + \epsilon) - f(x_0) = 0(\epsilon^2) \ . \qquad (A.5)$$

We can generalize this idea to find the extremum of $J\{y\}$. Assume that the extremum is attained on a particular path $y_0(x)$. In that case, if we consider a neighboring path

$$y(x) = y_0(x) + \varepsilon\,\eta(x)\ , \qquad (A.6)$$

such that the function $\varepsilon\eta(x)$ is small (i.e., $\varepsilon\,\eta(x)/y(x) << 1$ everywhere in $x_1 < x < x_2$) and vanishes at the endpoints x_1, x_2, i.e.

$$\eta(x_1) = \eta(x_2) = 0\ , \qquad (A.7)$$

we require that

$$\delta J = J\{y_0 + \varepsilon\eta\} - J\{y_0\} = 0(\varepsilon^2) \qquad (A.8)$$

or, since

$$J\{y_0 + \varepsilon\eta\} = \int_{x_1}^{x_2} F[y_0(x) + \varepsilon\eta(x),\ y_0'(x) + \varepsilon\eta'(x), x]\,dx \qquad (A.9)$$

we can perform a Taylor expansion of $F(y_0 + \varepsilon\eta, y_0' + \varepsilon\eta', x)$ about y_0 to obtain

$$F(y_0 + \varepsilon\eta, y_0' + \varepsilon\eta', x) = F(y_0, y_0', x) + $$
$$+ \varepsilon[\eta \frac{\partial}{\partial y} F(y_0, y_0', x) + \eta' \frac{\partial}{\partial y'} F(y_0, y_0', x)]$$
$$+ 0(\varepsilon^2)\ .$$

Substituting in equation A.9, we obtain

$$J\{y_0 + \varepsilon\eta\} = J\{y_0\} + \varepsilon \int_{x_1}^{x_2} [\eta \frac{\partial}{\partial y} F(y_0, y_0', x) + \eta' \frac{\partial}{\partial y'} F(y_0, y_0', x)]\,dx$$
$$+ 0(\varepsilon^2)\ .$$

Integration of the second term in the integrand by parts yields

$$\int_{x_1}^{x_2} \eta' \frac{\partial}{\partial y'} F(y_0, y_0', x)\,dx = \eta(x) \frac{\partial}{\partial y'} F(y_0, y_0', x) \Bigg|_{x_1}^{x_2}$$
$$- \int_{x_1}^{x_2} \eta(x) \frac{d}{dx} \left(\frac{\partial F}{\partial y'}\right)_{y_0} dx\ .$$

The first term vanishes by virtue of the imposed condition (equation A.7) on the possible choices of $\eta(x)$, and therefore we have

$$\delta J\{y_0\} = J\{y_0 + \varepsilon\eta\} - J\{y_0\}$$
$$= \varepsilon \int_{x_1}^{x_2} \eta(x) \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]_{y=y_0} dx + 0(\varepsilon^2)\ .$$

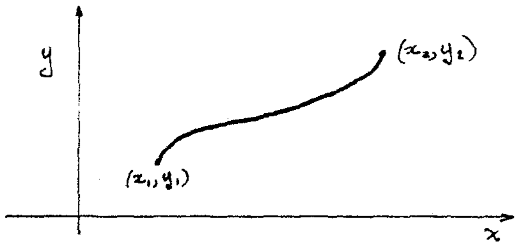
Consequently, if $J\{y\}$ possesses an extremum at $y(x) = y_0(x)$, then equation A.8 must be satisfied for any $\eta(x)$ that vanishes at the endpoints and therefore $y_0(x)$ must be the solution to the differential equation, known as Euler's equation,

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0\ . \qquad (A.10)$$

The difference between the two types of derivatives that appear in equation A.10

should be emphasized. The *partial* derivatives $\partial F/\partial y$ and $\partial F/\partial y'$ arose in the Taylor expansion of F considered as a function $F(y, y', x)$ of three variables. The *total* derivative $d(\partial F/\partial y')/dx$ arises from the integration by parts *with respect to* x and is therefore a derivative with respect to x of the function $F[y(x), y'(x), x]$ considered as a function of a single variable x.

Example A.1 Shortest distance between two points



The distance is given by the integral of the arc length s, i.e.

$$s = \int_{x_1}^{x_2} \sqrt{1 + y'^2(x)}\, dx \quad . \tag{A.11}$$

The function F in this case is given by

$$F(y, y', x) = F(y') = \sqrt{1 + y'^2} \tag{A.12}$$

Therefore, Euler's equation becomes

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = -\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0$$

or

$$\frac{\partial F}{\partial y'} = \text{const.}$$

Differentiating A.12 with respect to y', we have

$$\frac{\partial}{\partial y'}(1 + y'^2)^{1/2} = y'(1 + y'^2)^{-1/2} = c$$

or

$$y' = \left(\frac{c^2}{1 - c^2}\right)^{1/2} = \text{const} \ ,$$

i.e., a straight line.

The extension to paths of higher dimensions is straight-forward. If, for example, the integrand is a function of a *vector* function

$$\underline{y}(x) = [y_1(x), y_2(x), \ldots\ldots, y_N(x)] \ ,$$

i.e.

$$F = F[\underline{y}(x), \underline{y}'(x), x]$$

we then consider small variations about the extremal $\underline{y}_0(x)$, of the form

$$\underline{y}(x) = \underline{y}_0(x) + \epsilon\underline{\eta}(x) \ ,$$

where $\underline{\eta}(x)$ is a vector function

$$\underline{\eta}(x) = [\eta_1(x), \eta_2(x), \ldots\ldots, \eta_N(x)]$$

that vanishes at the end-points, i. e.

$$\underline{\eta}(x_1) = \underline{\eta}(x_2) = 0 \ .$$

Following the procedure in the one dimensional case, we Taylor expand about $\underline{y}_0(x)$ to obtain

$$\begin{aligned} F(\underline{y}_0 + \epsilon\underline{\eta}, y_0' + \epsilon\underline{\eta}', x) &= F(\underline{y}_0, y_0', x) + \\ &+ \epsilon \sum_i [\eta_i \frac{\partial}{\partial y_i} F(\underline{y}_0, y_0', x) + \eta_i' \frac{\partial}{\partial y_i'} F(\underline{y}_0, y_0', x)] \\ &+ 0(\epsilon^2) \ , \end{aligned}$$

which we can substitute and integrate by parts, as before, to obtain

$$\delta J\{y\} = \epsilon \int_{x_1}^{x_2} \sum_i \eta_i(x) \left[\frac{\partial F}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial F}{\partial y_i'}\right)\right]_{\underline{y}_0} dx + 0(\epsilon^2)$$

and therefore for an extremum the vector path must satisfy the N differential equations,

$$\frac{\partial F}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial F}{\partial y_i'}\right) = 0 \quad \text{for} \quad i=1,2,\ldots,N \ . \tag{A. 13}$$

The problem of selecting the vector path of the configuration vector $\underline{q}(t)$ in the action integral of equation A. 2 is a problem of this type and therefore the Euler equations become

$$\boxed{\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0 \ , \qquad i=1,2,\ldots,N} \tag{A. 14}$$

<u>Example A. 2</u> <u>The ballistic trajectory</u> (see example 1. 1).

The Lagrangian is given by

$$L(\underline{q}, \underline{\dot{q}}, t) = m[\tfrac{1}{2}(\dot{x}^2 + \dot{y}^2) - gy] \quad .$$

We therefore have <u>two</u> Euler equations, since

$$\underline{q}(t) = [x(t), y(t)] \ ,$$

namely

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = -\frac{d}{dt}\dot{x} = 0 \tag{A. 15a}$$

$$\frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) = -g - \frac{d}{dt}\dot{y} = 0 \ . \tag{A. 15b}$$

update

Therefore

$$\dot{x} = \text{const} \quad , \quad \ddot{y} = -g \ , \tag{A. 16}$$

which together with the boundary conditions yield the parabolic trajectory of equation 1. 14.

References

1. I. M. Gelfand and S. V. Fomin, Calculus of Variations (Prentice Hall, 1963).

2. P. M. Morse and H. Feshbach, Methods of Theoretical Physics (McGraw-Hill, 1953), part I, chapter 3.

3. J. Mathews and R. L. Walker, Mathematical Methods of Physics (Benjamin, 1965), chapter 12.

Problems

A. 1 Find the equation of the curve y(x) for which the functional

$$J\{y\} = \int_a^b F(y', y, x)\, dx$$

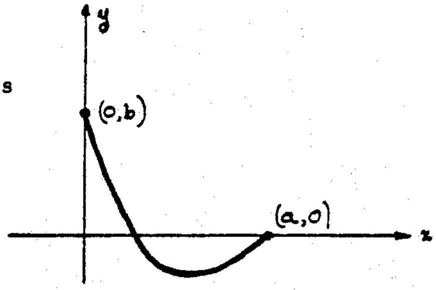
has an extremum subject to the constraints

$$y(a) = \alpha, \quad y(b) = \beta, \quad K\{y\} = \int_a^b G(y', y, x)\, dx = \gamma \quad .$$

Hint: Consider the functional

$$L\{y\} = J\{y\} + \lambda[\gamma - K\{y\}] \; .$$

A. 2 Find the equation of a long heavy chain of length ℓ strung between the two points (0, b) and (a, 0).



APPENDIX B: The Fourier Transform.

If a function f(x) can be expressed as

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k)\, e^{ikx}\, dk \quad , \tag{B.1}$$

then the function F(k) is given by

$$F(k) = \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx \quad , \tag{B.2}$$

and called the Fourier transform of f(x). Both f(x) and F(k) can be complex. In particular if

$$f(x) = f_R(x) + i\, f_I(x) \tag{B.3}$$

and

$$F(k) = F_R(k) + i\, F_I(k) \tag{B.4}$$

we have ($e^{ikx} \equiv \cos kx + i \sin kx$),

$$F_R(k) = \int_{-\infty}^{\infty} [f_R(x) \cos kx + f_I(x) \sin kx]\, dx \tag{B.5a}$$

$$F_I(k) = - \int_{-\infty}^{\infty} [f_R(x) \sin kx + f_I(x) \cos kx]\, dx \tag{B.5b}$$

and also

$$f_R(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [F_R(k) \cos kx - F_I(k) \sin kx]\, dk \tag{B.6a}$$

$$f_I(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [F_R(k) \sin kx + F_I(k) \cos kx]\, dk \ . \tag{B.6b}$$

The pair of functions f(x) and F(k) are called a Fourier transform pair, and have the following properties :

§1. <u>If f(x) is real, then $F^*(k) = F(-k)$.</u>

Proof: If $f_I(x) = 0$ then

$$F_R(k) = \int_{-\infty}^{\infty} f(x) \cos kx\, dx \tag{B.7a}$$

and

$$F_I(k) = - \int_{-\infty}^{\infty} f(x) \sin kx\, dx \tag{B.7b}$$

and therefore $F_R(k) = F_R(-k)$ and $F_I(k) = - F_I(-k)$, which proves the theorem.

§2. <u>If f(x) is purely imaginary, $F^*(k) = - F(-k)$.</u>

§3. If f(x) is even then F(k) is also even.

Proof: If $f(x) = f(-x)$, then

$$F(k) = \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx = \int_{-\infty}^{\infty} f(x) \cos kx\, dx$$

and, in fact

$$F(k) = 2 \int_{0}^{\infty} f(x) \cos kx\, dx ,$$

and therefore $F(k) = F(-k)$.

§4. If f(x) is real and even, F(k) is real and even.

§5. Even-odd decomposition. A function f(x) can always be decomposed into an even function plus an odd function. In particular if

$$f_e(x) = \frac{1}{2} [f(x) + f(-x)] \tag{B.8a}$$

and

$$f_o(x) = \frac{1}{2} [f(x) - f(-x)] , \tag{B.8b}$$

then

$$f(x) = f_e(x) + f_o(x) , \tag{B.8c}$$

where $f_e(x) = f_e(-x)$ and $f_o(x) = - f_o(x)$. Then if we decompose F(k), the Fourier transform of f(x), in a similar fashion into an even function $F_e(k)$ and an odd function $F_o(k)$, we have

$$f_e(x) = \frac{1}{\pi} \int_{0}^{\infty} F_e(k) \cos kx\, dk \tag{B.9a}$$

$$f_o(x) = \frac{i}{\pi} \int_{0}^{\infty} F_o(k) \sin kx\, dk \tag{B.9b}$$

and correspondingly

$$F_e(k) = 2 \int_{0}^{\infty} f_e(x) \cos kx\, dx \tag{B.10a}$$

$$i\, F_o(k) = 2 \int_{0}^{\infty} f_o(x) \sin kx\, dx . \tag{B.10b}$$

Note that if f(x) is real then $F_e(k) = F_R(k)$ and $F_o(k) = i\, F_I(k)$, where $F(k) = F_R(k) + i\, F_I(k)$, the real and imaginary parts of F(k).

§6. Linearity: If $F_1(k)$ and $F_2(k)$ are the Fourier transforms of $f_1(x)$ and $f_2(x)$ then $F(k) = a_1 F_1(k) + a_2 F_2(k)$ is the Fourier transform of $f(x) = a_1 f_1(x) + a_2 f_2(x)$.

§7. Symmetry: If F(k) is the Fourier transform of f(x), then $2\pi f(-k)$ is the Fourier transform of F(x).

§8. Scaling: If F(k) is the Fourier transform of f(x) then the Fourier transform of f(ax) is given by $|a|^{-1} F(k/a)$, where a is a real constant.

Proof: If $a > 0$,

$$G(k) = \int_{-\infty}^{\infty} f(ax)e^{-ikx}\,dx = \frac{1}{a}\int_{-\infty}^{\infty} f(x')e^{-i(k/a)x'}\,dx' = \frac{1}{a}F(k/a) .$$

If $a < 0$ the reversal in the order of integration changes the sign and the theorem is proven.

§9. Complex conjugate: If F(k) is the Fourier transform of f(x), then $F^*(-k)$ is the Fourier transform of $f^*(x)$.

§10. Cordinate shift: If F(k) is the Fourier transform of f(x), then $F(k)e^{-ikx_0}$ is the Fourier transform of $f(x - x_0)$

Proof:

$$f(x - x_0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(k)e^{ik(x - x_0)}\,dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} [F(k)e^{-ikx_0}]e^{ikx}\,dk .$$

Similarly $F(k - k_0)$ is the Fourier transform of $e^{ik_0x} f(x)$.

§11. Differentiation: If F(k) is the Fourier transform of f(x), then $(ik)^n F(k)$ is the Fourier transform of $d^n f(x)/dx^n$. Similarly, $d^n F(k)/dk^n$ is the Fourier transform of $(-ix)^n f(x)$.

§12. Moments: If F(k) is the Fourier transform of f(x), then $d^n F(o)/dx^n$ is equal to the Fourier transform of

$$(-i)^n \int_{-\infty}^{\infty} x^n f(x) dx ,$$

where the integral is the n^{th} moment of $f(x)$. Also, using a Taylor expansion about $k = 0$

$$F(k) = \sum_{m=0}^{m} \frac{(-ik)^n}{n!} \int_{\infty}^{\infty} x^n f(x)\, dx .$$

§13. Convolution: If $F(k)$ and $G(k)$ are the Fourier transforms of $f(x)$ and $g(x)$ respectively, then $F(k)\cdot G(k)$ is the Fourier transform of $u(x)$, where

$$u(x) = \int_{-\infty}^{\infty} f(x')\, g(x - x')\, dx'$$

is the convolution of $f(x)$ and $g(x)$. Conversly, the Fourier transform of $f(x)\cdot g(x)$ is given by $V(k)$, where

$$V(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k')\, G(k - k')\, dk' .$$

§14. Parseval's theorems: If $F(k)$ and $G(k)$ are the Fourier transforms of $f(x)$ and $g(x)$ respectively, then

$$\int_{-\infty}^{\infty} f(x)\cdot g(x)\, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k)\cdot G(-k)\, dk . \qquad \text{(B.11)}$$

as a special case

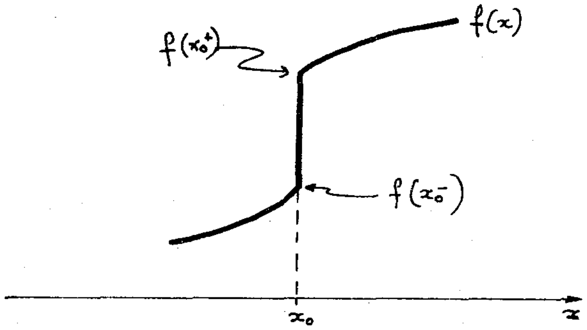
$$\int_{-\infty}^{\infty} |f(x)|^2\, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(k)|^2\, dk . \qquad \text{(B.12)}$$

Proof: From the result of §13 we have

$$\int_{-\infty}^{\infty} f(x)\cdot g(x) e^{-ikx}\, dx = \frac{1}{2\pi} \int_{\infty}^{\infty} F(k')\cdot G(k - k')\, dk' .$$

Setting $k = 0$ we obtain equation B.11. Equation B.12 follows if we choose $g(x) = f^*(x)$ and also use the results of §9.

The function $f(x)$ may be discontinuous,



i.e., $f(x_0^-) \neq f(x_0^+)$. In that case, at x_0 the Fourier integral of equation B.1 assumes the mean value of $f(x_0^-)$ and $f(x_0^+)$, i.e., if

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k)\, e^{ikx}\, dk$$

then

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} F(k)\, e^{ikx_0}\, dk = \frac{1}{2} \left[f(x_0^-) + f(x_0^+) \right] .$$

Similarly for discontinuous $F(k)$.

These results can be generalized to more dimensions in a straightforward manner. In particular, for two dimensions we have that if

$$f(x,y) = \left(\frac{1}{2\pi}\right)^2 \iint_{-\infty}^{\infty} F(k_x, k_y)\, e^{i(k_x x + k_y y)}\, dk_x\, dk_y \qquad (B.13)$$

then

$$F(k_x, k_y) = \iint_{-\infty}^{\infty} f(x,y)\, e^{-i(k_x x + k_y y)}\, dx\, dy \qquad (B.14)$$

and if

$$f(\underline{x}) = \left(\frac{1}{2\pi}\right)^3 \iiint_{-\infty}^{\infty} F(\underline{k})\, e^{i\underline{k}\cdot\underline{x}}\, d^3\underline{k} \qquad (B.15)$$

then

$$F(\underline{k}) = \iiint_{-\infty}^{\infty} f(\underline{x})\, e^{-i\underline{k}\cdot\underline{x}}\, d^3\underline{x} , \qquad (B.16)$$

where $\underline{x} = (x, y, z)$ and $\underline{k} = (k_x, k_y, k_z)$.

References

1. A. Papoulis, The Fourier Integral and its Applications (McGraw-Hill, 1962).

Problems

B.1 Prove that if f(x) is imaginary and

$$F(k) = \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx$$

then $$F^{*}(k) = -F(-k) \; .$$

B.2 Prove the complex conjugate property (§9).

B.3 Prove the differentiation properties (§11).

Hint: assuming that $f^{(n)}(\pm\infty) = F^{(n)}(\pm\infty) = 0$, integrate by parts repeatedly.

Notation:

$$f^{(n)}(x) \equiv \frac{d^n}{dx_n} f(x) \; .$$

B.4 Extend the differentiation properly (§11) to three dimensions by proving that if $F(\underline{k})$ is the Fourier transform of $f(\underline{x})$, then $i\underline{k}\, F(\underline{k})$ is the Fourier transform of the gradient of $f(\underline{x})$, i.e.,

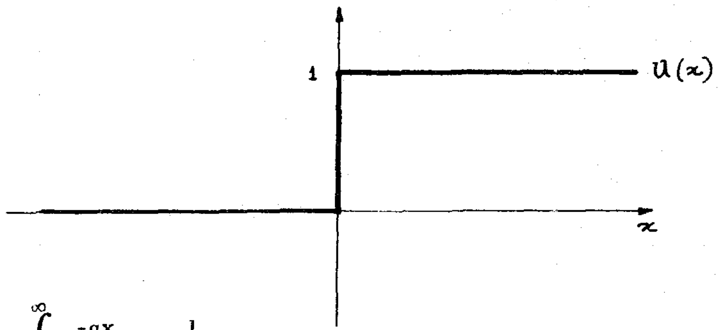
$$i\underline{k}\, F(\underline{k}) = \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial \underline{x}} f(\underline{x}) \right] e^{-i\underline{k}\cdot\underline{x}}\, d^3\underline{x}$$

Notation:

$$\frac{\partial}{\partial \underline{x}} \equiv \hat{e}_x \frac{\partial}{\partial x} + \hat{e}_y \frac{\partial}{\partial y} + \hat{e}_z \frac{\partial}{\partial z} \; .$$

B.5 Find the Fourier transform of U(x), the unit step function, defined by

$$U(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \; . \end{cases}$$



Hint: $$\int_{-\infty}^{\infty} e^{-\alpha x}\, dx = \frac{1}{\alpha}$$

APPENDIX C: The minimum joint spread wave packet. Hermite polynomials.

We have to solve the equation ($\beta > 0$)

$$v''(\xi) + (\beta - \xi^2)\, v(\xi) = 0 \qquad \text{(C. 1)}$$

subject to the boundary conditions

$$v(\pm\infty) = 0 \qquad \text{and} \qquad \int_{-\infty}^{\infty} v^2(\xi)\, d\xi = 1$$

From the differential equation for $v(\xi)$, we note that

$$v''(\xi) \to \xi^2 v(\xi) \qquad \text{as} \qquad |\xi| \to \infty\,. \qquad \text{(C. 2)}$$

We therefore need a function whose asymptotic behavior for large $|\xi|$ satisfies the condition C. 2. Noting that

$$\frac{d^2}{d\xi^2}\left(\xi^{\mu} e^{\pm\xi^2/2}\right) = \left[\xi^2 \pm (2\mu + 1) + \frac{\mu(- 1)}{\xi^2}\right]\left(\xi^{\mu} e^{\pm\xi^2/2}\right)$$

has the right behavior for any finite μ, we try a solution of the form

$$v(\xi) = A\; h(\xi)\, e^{-\xi^2/2}\,, \qquad \text{(C. 3)}$$

where the negative exponent was selected to satisfy the boundary condition at $\pm\infty$, and the constant A may be required for normalization. Substituting in the original equation, we obtain an equation for $h(\xi)$, which is given by

$$h''(\xi) - 2\xi h'(\xi) + (\beta - 1)\, h(\xi) = 0\,. \qquad \text{(C. 4)}$$

We now try a power series expansion for $h(\xi)$, i.e.,

$$h(\xi) = \sum_{m=0}^{\infty} a_m\, \xi^{m+s}\,, \qquad \text{(C. 5)}$$

where $a_0 \neq 0$. We then have

$$h'(\xi) = \sum_{m=0}^{\infty} (m + s)\, a_m\, \xi^{m+s-1} \qquad \text{(C. 6)}$$

and therefore

$$2\xi h'(\xi) = \sum_{m=0}^{\infty} 2(m + s)\, a_m\, \xi^{m+s} \qquad \text{(C. 7)}$$

and also

$$\begin{aligned} h''(\xi) &= \sum_{m=0}^{\infty} (m + s)(m + s - 1)\, a_m\, \xi^{m+s-2} \\ &= s(s - 1)\, a_0\, \xi^{s-2} + (1 + s)\, s\, a_1\, \xi^{s-1} \\ &\quad + \sum_{m=0}^{\infty} (m + s + 2)(m + s + 1)\, a_{m+2}\, \xi^{m+s}. \end{aligned} \qquad \text{(C. 8)}$$

Substituting equations C.5, C.7 and C.8 into C.4 we obtain

$$s(s-1)a_0\,\xi^{s-2} + (1+s)s\,a_1\,\xi^{s-1} + \\ + \sum_{m=0}^{\infty}\left[(m+s+2)(m+s+1)a_{m+2} - (2m+2s+1-\beta)a_m\right]\xi^{m+s} = 0 \qquad (C.9)$$

Equation C.9 can only be satisfied if

$$s(s-1)a_0 = 0 \qquad (C.10a)$$

$$(1+s)s\,a_1 = 0 \qquad (C.10b)$$

and

$$(m+s+2)(m+s+1)\,a_{m+2} = (2m+2s+1-\beta)\,a_m \;;\; m = 0,1,2,\ldots \qquad (C.10c)$$

From equation C.10a we see that $s=0$ or $s=1$ are acceptable values for s. From equation C.10b we see that either $s = 0$ or $a_1 = 0$ (or both). Then, from equation C.10c we get a_2 from a_0, a_4 from a_2 etc., a_3 from a_1, a_5 from a_3, etc. We now note that

$$\frac{a_{m+2}}{a_m} \to \frac{2}{m} \quad \text{as} \quad m \to \infty,$$

therefore, for large ξ, the series tends to $\xi^{\nu}e^{\xi^2}$ where ν is a finite number and consequently $v(\xi)$, as given by equation C.3, cannot satisfy the boundary conditions at $\pm\infty$. The only possibility is if the series terminates to yield a polynomial for $h(\xi)$. That will happen if the coefficient of a_m in equation (C.10c) is zero for some m, which will occur if the constant β is equal to

$$\beta = 2(m+s) + 1 \qquad (C.11)$$

where $m = 0,1,2,\ldots$ and $s = 0$ or 1. Since this must happen for even m (the recursion relation C.10c relates a_{m+2} to a_m and $a_0 \neq 0$) we must choose $a_1 = 0$. Consequently we have an even or odd power polynomial depending on whether $s = 0$ or 1 respectively. β must then be of the form

$$\beta = 2n + 1 \;; \qquad n = 0,1,2,\ldots \qquad (C.12)$$

to yield in an even or odd polynomial $h_n(\xi)$, depending on whether n is even or odd. The resulting polynomials are called Hermite polynomials. The first few of these are listed below:

$$\begin{aligned} h_0(\xi) &= 1 \\ h_1(\xi) &= 2\xi \\ h_2(\xi) &= 4\xi^2 - 2 \\ h_3(\xi) &= 8\xi^3 - 12\xi \\ h_4(\xi) &= 16\xi^4 - 48\xi^2 + 12 \end{aligned} \qquad (C.13)$$

$$h_5(\xi) = 120\xi - 160\xi^3 + 32\xi^5$$

$$h_6(\xi) = 120 - 720\xi^2 + 480\xi^4 - 64\xi^6 \ .$$

and solve the equation

$$h_n''(\xi) - 2\xi h_n'(\xi) + 2n\ h_n(\xi) = 0 \ , \qquad (C.\ 14)$$

where we have substituted C. 12 into C. 4.

The study of the properties of the Hermite polynomials is facilitated by the <u>generating function</u>, defined by

$$\sum_{n=0}^{\infty} \frac{s^n}{n!}\ h_n(\xi) \equiv \phi(s, \xi) \qquad (C.\ 15)$$

and equal to

$$\phi(s, \xi) = e^{-s^2 + 2s\xi} \ . \qquad (C.\ 16)$$

We now differentiate $\phi(s, \xi)$ with respect to s to get

$$\frac{\partial}{\partial s}\phi(s, \xi) = (-2s + 2\xi)\,\phi(s, \xi)$$

$$= \sum_{n=0}^{\infty} \frac{(-2s + 2\xi)s^n}{n!}\ h_n(\xi) = \sum_{n=0}^{\infty} \frac{s^{n-1}}{(n-1)!}\ h_n(\xi) \ ,$$

where the last series was obtained by differentiating the defining series directly. Equating the coefficients of equal powers of s we obtain

$$2\xi h_n(\xi) = h_{n+1}(\xi) + 2n\ h_{n-1}(\xi) \ . \qquad (C.\ 17)$$

Similarly, differentiating with respect to ξ we have

$$h_n'(\xi) = 2n\ h_{n-1}(\xi) \ . \qquad (C.\ 18)$$

It is easy to show that the sequence of functions which are related by equations C. 17 and C. 18 are solutions to equation C. 14, proving thereby that the generating function given by equation C. 16 is indeed the correct one.

Using the fact that for any function f(s), we have

$$f(s) = \sum_{n=0}^{\infty} \frac{s^n}{n!}\ f^{(n)}(0) \ ,$$

where

$$f^{(n)}(a) \equiv \frac{d^n}{ds^n}\ f(s)\Big|_{s=a} \ ,$$

and consequently

$$h_n(\xi) = \frac{\partial^n}{\partial s^n}\ \phi(s, \xi) \ .$$

Noting that

$$g(s,\xi) = e^{-s^2+2s\xi} = e^{\xi^2-(s-\xi)^2}$$

we have

$$\frac{\partial^n}{\partial s^n} g(s,\xi) = e^{\xi^2}\frac{\partial^n}{\partial s^n} e^{-(s-\xi)^2} = (-1)^n e^{\xi^2}\frac{\partial^n}{\partial \xi^n} e^{-(s-\xi)^2}$$

and therefore

$$\boxed{h_n(\xi) = (-1)^n e^{\xi^2}\frac{d^n}{d\xi^n} e^{-\xi^2}}\ , \tag{C. 19}$$

a much simpler way of generating the Hermite polynomials.

Integrals of Hermite polynomials. The generating function also helps in evaluating the normalization integral corresponding to the wave functions

$$v_n(\xi) = A_n\, h_n(\xi)\, e^{-\xi^2/2} \tag{C. 20}$$

(see equation C. 3). In particular, consider the integral

$$\begin{aligned} I_{nm} &= \int_{-\infty}^{\infty} v_n(\xi)\, v_m(\xi)\, d\xi \\ &= A_n\, A_m \int_{-\infty}^{\infty} h_n(\xi)\, h_m(\xi)\, e^{-\xi^2}\, d\xi\ . \end{aligned}$$

To generate the integrals of the cross products we integrate the products of two generating functions, i. e.,

$$\begin{aligned} J(s_1, s_2) \equiv \int_{-\infty}^{\infty} g(s_1,\xi)\, g(s_2,\xi)\, e^{-\xi^2}\, d\xi &= \int_{-\infty}^{\infty} e^{-s_1^2+2s_1\xi}\, e^{-s_2^2+2s_2\xi}\, e^{-\xi^2}\, d\xi \\ &= \pi^{1/2}\, e^{2s_1 s_2} \\ &= \pi^{1/2} \sum_n \frac{(2s_1 s_2)^n}{n!}\ , \end{aligned}$$

but, integrating the product of the two series corresponding to $g(s_1,\xi)$ and $g(s_2,\xi)$ we also have

$$J(s_1, s_2) = \sum_{n,m} \frac{s_1^n s_2^m}{n!\, m!} \int_{-\infty}^{\infty} h_n(\xi)\, h_m(\xi)\, e^{-\xi^2}\, d\xi\ .$$

Equating terms, we then see that

$$\boxed{\int_{-\infty}^{\infty} h_n(\xi)\, h_m(\xi)\, e^{-\xi^2}\, d\xi = \pi^{1/2}\, 2^n\, n!\, \delta_{nm}} \tag{C. 21}$$

where δ_{nm} is the Kronecker delta (see Appendix D) and given by

$$\delta_{nm} = \begin{cases} 1 & , \text{ if } n=m \\ 0 & , \text{ if } n\neq m \end{cases} .$$

Using equation C.21 it is easy to see that the proper normalization constant A_n for the $v_n(\xi)$ in equation C.20 is given by

$$A_n = (\pi^{1/2}\, 2^n\, n!)^{-1/2} . \qquad (C.22)$$

Consequently, the sequence of functions

$$v_n(\xi) = (\pi^{1/2}\, 2^n\, n!)^{-1/2}\, h_n(\xi)\, e^{-\xi^2/2} \qquad (C.23)$$

forms an <u>orthonormal</u> set, i.e.,

$$\int_{-\infty}^{\infty} v_n(\xi)\, v_m(\xi)\, d\xi = \delta_{nm} \quad , . \qquad (C.24)$$

since the v_n's have been normalized and are orthogonal to each other (solutions to a Sturm-Liouville problem).

Problems

C.1 Generate the first four Hermite polynomials using the recursion relation C.10c.

C.2 Prove equations C.17 and C.18 and show that the sequence of functions $h_n(\xi)$ that satisfies them is a solution to equation C.14.

C.3 Generate the first four Hermite polynomials using equation C.19.

C.4 Using equations C.17, C.23 and C.24, compute the matrix element

$$\langle m|\xi|n\rangle \equiv \int_{-\infty}^{\infty} v_m(\xi)\ \xi\ v_n(\xi)\, d\xi\ .$$

C.5 If κ is the conjugate coordinate to ξ and $\kappa = -i\partial/\partial\xi$ is the corresponding operator in ξ-space, compute the matrix element

$$\langle m|\kappa|n\rangle \doteq \int_{-\infty}^{\infty} v_m(\xi)\ \kappa\ v_n(\xi)\, d\xi$$

using equation C.18.

APPENDIX D: The Kronecker and Dirac delta functions.

Consider the problem of the scalar product of two vectors

$$\underline{u} = \sum_i \hat{e}_i u_i \qquad \text{(D. 1)}$$

and

$$\underline{v} = \sum_j \hat{e}_j v_j \qquad \text{(D. 2)}$$

where the $\hat{e}_i$ are the basis vectors. We then have

$$\underline{u} \cdot \underline{v} = \left(\sum_i \hat{e}_i u_i \right) \cdot \left(\sum_j \hat{e}_j u_j \right) = \sum_{i,\,j,} (\hat{e}_i \cdot \hat{e}_j)\, u_i u_j \; . \qquad \text{(D. 3)}$$

In the important special case where the $\hat{e}_i$ are <u>orthonormal</u>, i.e., orthogonal unit vectors, we have

$$\hat{e}_i \cdot \hat{e}_j = \begin{cases} 1 & \text{if} \quad i=j \\ & \qquad . \\ 0 & \text{if} \quad i \neq j \end{cases}$$

It is convenient to define a function δ_{ij} that has this property, i.e.,

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i=j \\ \\ 0 & \text{if} \quad i \neq j \end{cases} \qquad \text{(D. 4)}$$

known as the <u>Kronecker delta</u>. Evidently if a set of $\hat{e}_i$ form an orthonormal basis then

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij} \qquad \text{(D. 5)}$$

Substituting equation D. 5 in equation D. 3 we have (for orthonormal basis vectors)

$$\underline{u} \cdot \underline{v} = \sum_{ij} u_i v_j \delta_{ij} = \sum_i u_i \left(\sum_j v_j \delta_{ij} \right) = \sum_i u_i v_i \qquad \text{(D. 6)}$$

It can be seen that the role of the function δ_{ij} is to <u>select</u> one of the terms of a summation, i.e.,

$$\sum_j v_j \delta_{ij} = v_i \qquad \text{(D. 7)}$$

From the defining equation D. 4, we can see that δ_{ij} is really a matrix (tensor of rank two) and is in fact the identity matrix

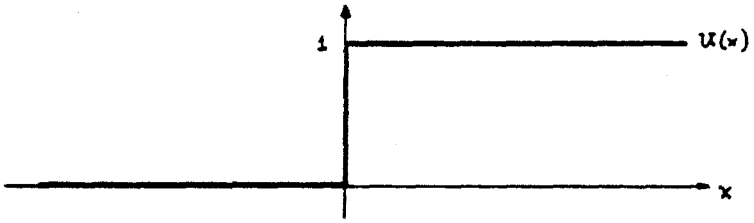
$$I = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}$$

or

$$I_{ij} = \delta_{ij}$$

To introduce the Dirac delta function, it is convenient to define the unit step function

$$U(x) = \begin{cases} 0 \text{ for } x < 0 \\ 1 \text{ for } x > 0 \end{cases} .$$



With it we can express any integral over finite limits as an integral over the whole x-axis, i.e.,

$$\int_a^b f(x)\,dx = \int_{-\infty}^{\infty} \left[U(x-a) - U(x-b)\right] f(x)\,dx .$$

Consider now the integral

$$\int_{-h/2}^{h/2} f(x)\,dx = \int_{-\infty}^{\infty} \left[U(x+\frac{h}{2}) - U(x-\frac{h}{2})\right] f(x)\,dx$$

where h is a small length. This we can write in terms of the mean value of the function f(x)

$$\int_{-\infty}^{\infty} \left[U(x+\frac{h}{2}) - U(x-\frac{h}{2})\right] f(x)\,dx = h\,f(\xi) ,$$

where

$$-\frac{h}{2} < \xi < \frac{h}{2}$$

or

$$\int_{-\infty}^{\infty} \frac{1}{h}\left[U(x+\frac{h}{2}) - U(x-\frac{h}{2})\right] f(x)\,dx = f(\xi) .$$

Taking the limit as h → 0

$$\lim_{h\to 0}\left\{\int_{-\infty}^{\infty} \frac{1}{h}\left[U(x+\frac{h}{2}) - U(x-\frac{h}{2})\right] f(x)\,dx\right\} = \lim_{h\to 0}\{f(\xi)\}$$

or

$$\int_{-\infty}^{\infty} \lim_{h\to 0}\left\{\frac{1}{h}\left[U(x+\frac{h}{2}) - U(x-\frac{h}{2})\right]\right\} f(x)\,dx = f(0) ,$$

if f(x) is continuous at x = 0.

We can now define

$$\boxed{\delta(x) = \lim_{h \to 0} \left\{ \frac{1}{h} \left[U(x + \frac{h}{2}) - U(x - \frac{h}{2}) \right] \right\}} \tag{D.8}$$

which is the Dirac delta function.

Using this definition, we can show the following properties of $\delta(x)$.

(i)
$$\int_{-\infty}^{x} \delta(x)\,dx = U(x) \tag{D.9}$$

(ii)
$$\delta(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ \infty & \text{for } x = 0 \end{cases} \tag{D.10}$$

(iii)
$$\int_{-\infty}^{\infty} \delta(x)\,dx = 1 \tag{D.11}$$

(iv)
$$\delta(x) = \delta(-x) \tag{D.12}$$

(v)
$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx}\,dk \tag{D.13}$$

Proof:

We essentially wish to show that the Fourier transform of $\delta(x)$ is unity. From the definition,

$$\mathcal{F}\{\delta(x)\} = \lim_{h \to 0} \int_{-\infty}^{\infty} \frac{1}{h} \left[U(x+\frac{h}{2}) - U(x - \frac{h}{2}) \right] e^{-ikx}\,dx$$

or

$$\mathcal{F}\{\delta(x)\} = \lim_{h \to 0} \left\{ \frac{1}{h} \int_{-h/2}^{h/2} e^{-ikx}\,dx \right\} = \lim_{h \to 0} \left\{ \frac{2}{kh} \sin \frac{kh}{2} \right\} = 1$$

We now take the inverse transform to get

$$\delta(x) = \lim_{h \to 0} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{2}{kh} \sin \frac{kh}{2}\, e^{ikx}\,dk \right\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{h \to 0} \left\{ \frac{2}{kh} \sin \frac{kh}{2} \right\} e^{ikx}\,dk$$

from which the desired result follows by taking the limit.

(vi)
$$\int_{-\infty}^{\infty} f(x)\,\delta\left[g(x)\right] dx = \sum_i \frac{f(x_i)}{|g'(x_i)|} \tag{D.14}$$

where $g(x_i) = 0$.

It is this property that is the claim to fame of $\delta(x)$. Some important special cases are

$$\int_{-\infty}^{\infty} f(x)\,\delta(x)\,dx = f(0) \tag{D.15}$$

$$\int_{-\infty}^{\infty} f(x)\,\delta(x-a)\,dx = f(a) \tag{D.16}$$

$$\int_{-\infty}^{\infty} f(x)\,\delta(\alpha x)\,dx = \frac{f(0)}{|\alpha|} \;. \tag{D.17}$$

The similarity of $\delta(x-a)$ and δ_{ij} is apparent from the above relations. We see that it is a property of $\delta(x-a)$ to select the value of $f(x)$ at $x = a$ by integrating over x. i.e.,

$$f(a) = \int_{-\infty}^{\infty} f(x)\,\delta(x-a)\,dx \;.$$

In turn, δ_{ij} selects the value of f_j at $i = j$ by summing over j. i.e.,

$$f_i = \sum_i f_j \delta_{ij} \;.$$

The two are entirely equivalent. Indeed it is possible to define $\delta(x-a)$ in terms of δ_{ij} by taking the limit of infinitesimally spaced indeces in the summation.

We can now make the extension to three dimensions as follows. We would like a property of the sort

$$\int_V f(\underline{r})\,\delta(\underline{r})\,d^3\underline{r} = f(0) \tag{D.18}$$

or, equivalently

$$\int_V f(\underline{r})\,\delta(\underline{r}-\underline{a})\,d^3\underline{r} = f(\underline{a})$$

We can see now how to write $\delta(\underline{r}-\underline{a})$ if we express the integral in Cartesian coordinates.

i.e., since

$$\int_V f(x,y,z)\,\delta(\underline{r}-\underline{a})\,dxdydz = f(a_x, a_y, a_z) \tag{D.19}$$

it is clear that the function that does the job is

$$\text{(vii)} \qquad \delta(\underline{r} - \underline{a}) = \delta(x - a_x)\,\delta(y - a_y)\,\delta(z - a_z) \;. \qquad \text{(D.20)}$$

We can define the derivative of the delta function as follows. Since, at least formally,

$$\int_{-\infty}^{\infty} f(x)\,\delta'(x)\,dx = \cancelto{0}{f(x)\,\delta(x)\Big|_{-\infty}^{\infty}} - \int_{-\infty}^{\infty} f'(x)\,\delta(x)\,dx$$

we have

$$\int_{-\infty}^{\infty} f(x)\,\delta'(x)\,dx = -f'(0)\,,$$

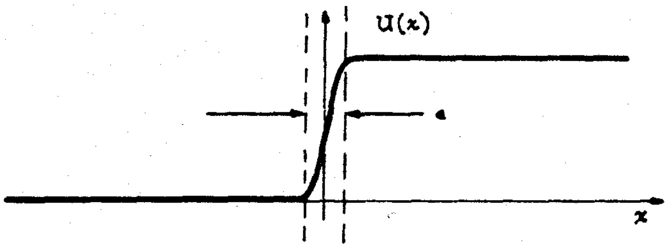
or, by repeated application of integration by parts

$$\text{(viii)} \quad \int_{-\infty}^{\infty} f(x)\,\delta^{(n)}(x)\,dx = (-1)^n f^{(n)}(0)\;. \qquad \text{(D.21)}$$

We can get an intuitive feeling for what $\delta'(x)$ and $\delta''(x)$ look like by trying to draw them. From the definition and the definition of the derivative of a function we can see that

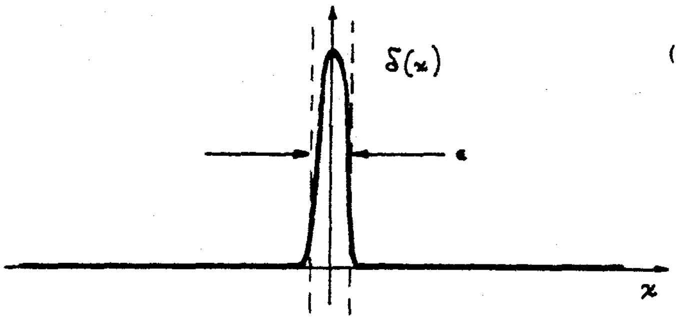
$$\delta(x) = \frac{d}{dx}\,U(x) \qquad \text{(D.22)}$$

Let us now imagine that $U(x)$ had slightly rounded corners over an interval ϵ,



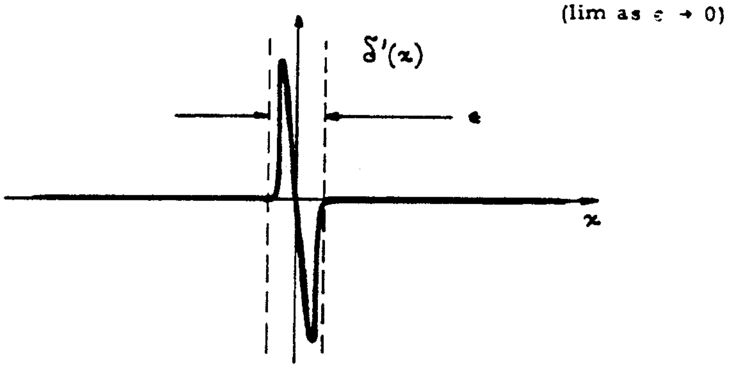
(lim as $\epsilon \to 0$)

since the delta function is the derivative of the above, we can get it by plotting the slope of the rounded $U(x)$.

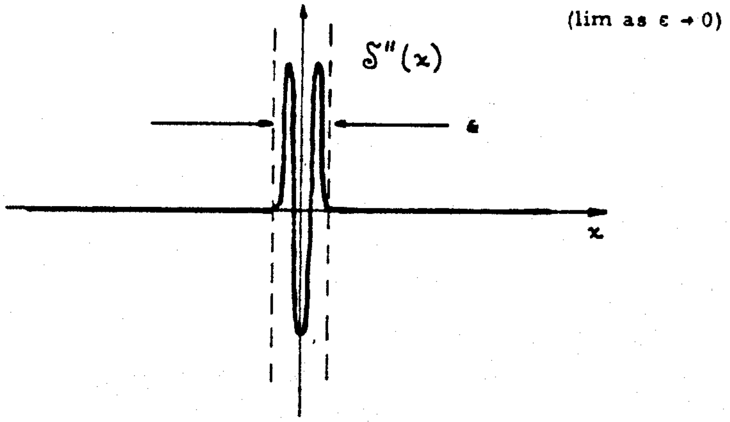


(lim as $\epsilon \to 0$)

We differentiate that in a similar manner to get



and



etc.

These we can use to express any function that satisfies relatively weak conditions that need not concern us here, as a series of derivatives of the delta function, i.e.,

$$f(x) = f_0\,\delta(x) - f_1\,\delta'(x) + \tfrac{1}{2} f_2\,\delta''(x) - \;.\;.\;.$$

or

$$f(x) = \sum_{n=0}^{\infty} (-1)^n \frac{f_n}{n!}\,\delta^{(n)}(x) \tag{D.23}$$

where

$$f_0 = \int_{-\infty}^{\infty} f(x)\,dx$$

$$f_1 = \int_{-\infty}^{\infty} x f(x)\,dx$$

$$f_2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$

$$f_n = \int_{-\infty}^{\infty} x^n f(x)\,dx \tag{D.24}$$

are the various moments of f(x) about x = 0. We can make a plausibility argument at this point as to the validity of this expansion as a representation of f(x). If we multiply both sides of the equation by x^m and integrate over the x - axis we get

$$\int_{-\infty}^{\infty} x^m f(x)\,dx = \int_{-\infty}^{\infty} x^m \left\{ \sum_{n=0}^{\infty} (-1)^n \frac{fn}{n!}\, \delta^{(n)}(x) \right\} dx$$

or

$$\int_{-\infty}^{\infty} x^m f(x)\,dx = \sum_{n=0}^{\infty} (-1)^n \frac{fn}{n!} \int_{-\infty}^{\infty} x^m \delta^{(n)}(x)\,dx$$

but

$$\int_{-\infty}^{\infty} x^m \delta^{(n)}(x) dx = (-1)^n \frac{d^n}{dx^n}\, x^n \delta_{mn} = (-1)^n \frac{\delta_{mn}}{n!}$$

where δ_{mn} is the Kronecker delta. Therefore we must have

$$\int_{-\infty}^{\infty} x^m\, f(x)\,dx = f_m$$

which is how we defined f_m. We have therefore reduced the problem to showing that two functions (subject to certain restrictions) are equal if all their moments are equal, which is a plausible statement. As an additional "proof", compare the Fourier transform of both sides of equation D.23 and §12 of Appendix A. That, in fact, is the easiest way to see whether it is possible to represent a function f(x) in terms of an expansion of the type given by equation D.23. In particular, if the Fourier transform of f(x) can be represented as a series of integral powers of k.

To those familiar in multipole expansions, this is exactly what we have done. A glance at the picture of $\delta'(x)$ shows that it is dipole, whereas f_1 is the dipole moment of f(x) about the origin, etc.

The extension to higher dimensions is straight forward. In particular a function $f(\underline{x})$, localized about $\underline{x} = 0$ can be expanded in a series

$$f(\underline{x}) = \sum_{k,\ell,m} (-1)^{k+\ell+m} \frac{f_{k\ell m}}{k!\,\ell!\,m!} \frac{\partial^k}{\partial x^k} \frac{\partial^\ell}{\partial y^\ell} \frac{\partial^m}{\partial z^m}\, \delta(\underline{x}) \,, \tag{D.25}$$

where

$$f_{k\ell m} = \int_{\underline{x}} x^k y^\ell z^m f(\underline{x}) d^3\underline{x} \qquad (D.26)$$

are the various moments of $f(\underline{x})$ about the origin along each direction.

To demonstrate how this machinery can be very useful, let us consider a specific problem. Let us consider an f(x) which is localized about the origin and bounded by a constant,

i.e.,

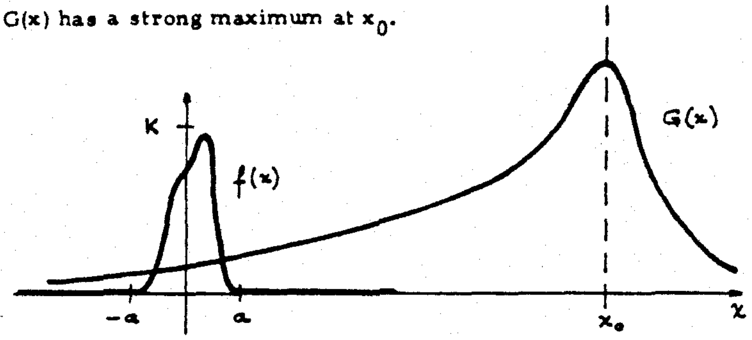
$$f(x) = 0 \quad \text{for } |x| > a$$

$$f(x) < K$$

Let us suppose we wish to calculate the integral

$$I = \int_{-\infty}^{\infty} f(x)G(x)dx$$

where G(x) has a strong maximum at x_0.



It is very easy to see that such a thing cannot be calculated in the general case. We can approximate the answer, however, in the following way. Substituting our delta function expansion for f(x) we get

$$I = \int_{-\infty}^{\infty} \left\{ \sum_{n=0}^{\infty} (-1)^n \frac{f_n}{n!} \delta^{(n)}(x) \right\} G(x)\, dx$$

or

$$I = \sum_{n=0}^{\infty} (-1)^n \frac{f_n}{n!} \int_{-\infty}^{\infty} \delta^{(n)}(x) G(x)\, dx$$

therefore

$$I = \sum_{n=0}^{\infty} \left\{ \frac{f_n}{n!} G^{(n)}(0) \right\}$$

where

$$G^{(n)}(0) = \frac{d^n}{dx^n} G(x) \bigg|_{x=0}$$

is the n^{th} derivative of G(x) evaluated at the origin.

Let us now make the assumption that G(x) goes to zero for large distances from its maximum like

$$G(x) \sim \frac{G_0}{|x-x_0|^r} \text{ for large } |x-x_0|;\ r \geq 1$$

then

$$G^{(n)}(x) = O\left(\frac{n!}{|x-x_0|^{n+r}}\right) \quad \text{as } |x-x_0| \to \infty$$

and

$$G^{(n)}(0) = O\left(\frac{n!}{x_0^{\,n+r}}\right)$$

We can also make an upper bound estimate for f_n as follows.

$$f_n = \int_{-\infty}^{\infty} x^n f(x)\,dx$$

$$= \int_{-a}^{a} x^n f(x)\,dx$$

$$= a^{n+1} \int_{-1}^{1} \xi^n f(a\xi)\,d\xi \qquad ; \qquad \xi = \frac{x}{a}$$

$$\leq \frac{2Ka^{n+1}}{n+1}$$

$$= O\left(\frac{a^{n+1}}{n+1}\right)$$

That means that the terms in the series go like

$$(\text{const.}) \left(\frac{1}{n+1}\right)\left(\frac{a}{x_0}\right)^{n+1}$$

and the series converges since it is bounded term by term by

$$-(\text{const}) \left(\frac{a}{x_0}\right)^{r-1} \ln\left(1 - \frac{a}{x_0}\right)$$

provided

$$\left(\frac{a}{x_0}\right) < 1$$

which we are assuming for the purposes of this discussion.

Taking the first few terms. we get (let r = 1; the worst case)

$$I = f_0 G(0) + f_1 G'(0) + \tfrac{1}{2} f_2 G''(0) + O\left(\left|\frac{a}{x_0}\right|^4\right)$$

If x_0 is a few times a we can see that we have a very good estimate for the integral. (If x_0 = 10a, the error after three terms is of the order of 0.01%). In addition we see that the exact details of $f(x)$ do not affect the solution and are not needed. Indeed we can often estimate or measure the moments of a function without knowing its exact form. In particular, if $f(x)$ is an even function of x, i. e., if

$$f(x) = f(-x)$$

all the odd moments vanish identically and we have

$$I = f_0 G(0) + \tfrac{1}{2} f_2 G''(0) + 0\left[\left(\frac{a}{x_0}\right)^4\right]$$

(~ 0.001% if $x_0 = 10a$, not bad for just two terms!)

Problems

D. 1 Prove equation D. 14. Hint: prove D. 17 first.

D. 2 Prove the Fourier transform formula, i. e., if

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k) e^{ikx} \, dk$$

then

$$F(k) = \int_{-\infty}^{\infty} f(x) e^{-ikx} dx$$

by direct substitution. Hint: Careful with integration variables!

D. 3 Prove Parseval's theorems (equations A. 11 and A. 12) by direct substitution.

D. 4 From Maxwell's equations we have that, for a stationary charge density $\rho(\underline{r})$

$$\phi(\underline{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\underline{r}')}{|\underline{r}' - \underline{r}|} d^3 \underline{r}'$$

where $\phi(\underline{r})$ is the electrostatic potential. If the charge density is confined to a region $|\underline{r}| < a$ and in addition has no net charge, i. e.,

$$O = \int \rho(\underline{r}) d^3 \underline{r} = 0$$

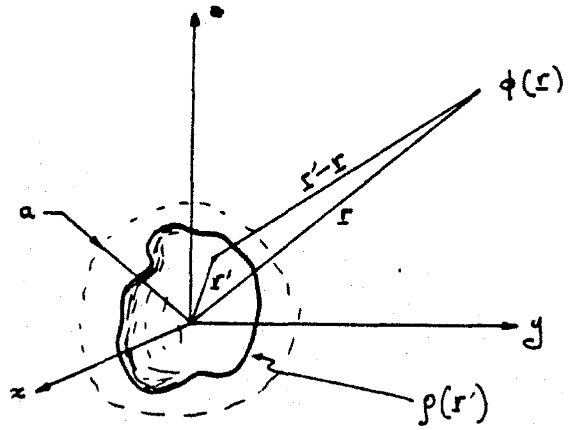
but has a dipole moment along the z - axis given by

$$d_z = \int z \rho(\underline{r}) d^3 \underline{r} \ ,$$

with all other moments equal to zero, i.e.

$$\rho_{k\ell m} = \int x^k y^\ell z^m \rho(\underline{r}) d^3 \underline{r} = 0$$

for $(k, \ell, m) \neq (0, 0, 1)$, find the potential $\phi(\underline{r})$ for $|\underline{r}| > a$.



D.5 Find the behavior of $\phi(\underline{r})$ in the previous problem if $|\underline{r}/a| << 1$, and

$$Q = \int \rho(\underline{r}) d^3 \underline{r} = 0$$

but

$$\rho_{k\ell m} = \int x^k y^\ell z^m \rho(\underline{r}) d^3 \underline{r} \neq 0$$

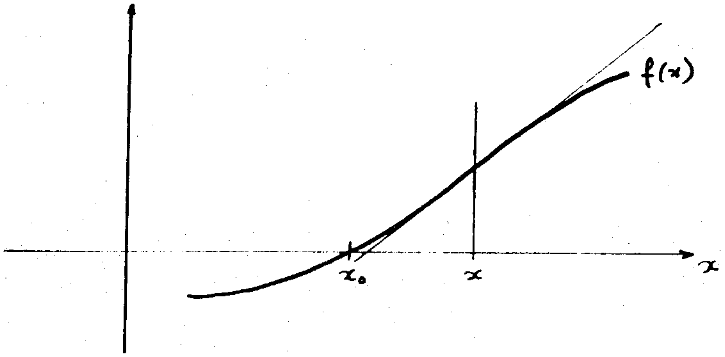
for $k + \ell + m \geq 1$.

APPENDIX E: Newton's method.

Consider a root $x = x_0$ of the equation

$$f(x) = 0 \quad , \tag{E. 1}$$

where $f(x)$ is well-behaved in the vicinity of $x = x_0$.



If we expand $f(x)$ in the vicinity of the root, we have

$$f(x_0) = f(x) + (x_0 - x)\, f'(x) + 0[(x_0 - x)^2\, f''(x)]$$

or, since $f(x_0) = 0$, we have

$$x_0 - x = - \frac{f(x)}{f'(x)} + 0\left[(x_0 - x)^2\, \frac{f''(x)}{f'(x)}\right] \quad ,$$

provided $f'(x) \neq 0$. Consequently, in the vicinity of $x = x_0$ we have,

$$x_0 \simeq x - \frac{f(x)}{f'(x)} \quad . \tag{E. 2}$$

Equation E. 2 can serve as the basis of a very efficient iterative scheme, i. e.

$$x^{(n+1)} = x^{(n)} - \frac{f^{(n)}}{f'^{(n)}} \quad , \tag{E. 3}$$

where $f^{(n)} = f[x^{(n)}]$ and $f'^{(n)} = f'[x^{(n)}]$, known as Newton's method. Using equation E. 3 and an initial guess $x^{(0)}$, we obtain $x^{(1)}$. From $x^{(1)}$ we obtain $x^{(2)}$, etc. This sequence usually converges extremely rapidly to the root x_0.

Example E. 1. Square root algorithm.

We have

$$f(x) = x^2 - a = 0$$

$$f'(x) = 2x \quad .$$

Therefore

$$x^{(n+1)} = x^{(n)} - \frac{x^{(n)^2} - a}{2x^{(n)}}$$

$$= x^{(n)} - \frac{x^{(n)}}{2} + \frac{a}{2x^{(n)}}$$

or

$$x^{(n+1)} = \frac{x^{(n)} + a/x^{(n)}}{2} \quad .$$

For $a = 2$, and an initial guess of $x^{(0)} = 1$ we have

$$x^{(0)} = 1.0$$

$$x^{(1)} = 1.50$$

$$x^{(2)} = 1.4167$$

$$x^{(3)} = 1.414216$$

$$x^{(4)} = 1.41421356 \quad ,$$

which is equal to $\sqrt{2}$, to as many places.

Quite frequently, it may be advantageous to avoid having to evaluate both $f(x)$ and $f'(x)$ at every iteration of equation E. 3, by using two successive estimates of $f(x)$ to obtain an estimate for $f'(x)$, i. e.

$$f'^{(n)} \simeq \frac{f^{(n)} - f^{(n-1)}}{x^{(n)} - x^{(n-1)}} \tag{E. 4}$$

Substituting E. 4 in E. 3, we then have,

$$x^{(n+1)} = x^{(n)} - \left[\frac{x^{(n)} - x^{(n-1)}}{f^{(n)} - f^{(n-1)}}\right] f^{(n)} \quad . \tag{E. 5}$$

The iterative algorithm based on E. 5 is usually referred to as the chord method. Note, that it requires two guesses, $x^{(0)}$ and $x^{(1)}$ to start.

Example E. 2. Energy levels of a rectangular potential well.

We must solve equations 4. 59

$$\cot \xi = \frac{\xi}{\sqrt{\beta^2 - \xi^2}} \quad ; \quad \text{even parity}$$

$$\tan \xi = \frac{-\xi}{\sqrt{\beta^2 - \xi^2}} \quad ; \quad \text{odd parity} \quad .$$

Useful to scale $\xi = \beta x$. This yields

$$f_e(x) = (1 - x^2)^{1/2} \cos \beta x - x \sin \beta x = 0 \tag{E. 6a}$$

$$f_0(x) = (1 - x^2)^{1/2} \sin \beta x + x \cos \beta x = 0 \quad , \qquad \text{(E. 6b)}$$

where we have multiplied by the denominators to avoid infinities.

We first solve for the lowest eigenvalue $\xi_0 = \beta x_0$, for $\beta = \sqrt{30} = 5.4772$.

$$\xi_0 = \frac{\pi}{2} \Rightarrow x_0^{(o)} = \frac{\pi}{2\beta} = 0.2868$$

and as a second guess, using 4.62

$$\xi_0 = \frac{\pi\beta}{2(\beta + 1)} \Rightarrow x_0^{(1)} = \frac{\pi}{2(\beta + 1)} = 0.2425 \quad .$$

We then iterate, for the lowest eigenvalue

n	$x_0^{(n)}$	$f_e^{(n)}$
0	0.2868	-0.2868
1	0.2425	-0.0024
2	0.2421	0.0000 .

For the next even eigenvalue, we have

$$\xi_2 = \frac{3\pi}{2} \Rightarrow x_2^{(o)} = \frac{3\pi}{2\beta} = 0.8604 \quad ,$$

and as a second guess,

$$\xi_2 = \frac{3\pi\beta}{2(\beta + 1)} \Rightarrow x_2^{(1)} = \frac{3\pi}{2(\beta + 1)} = 0.7275 \quad .$$

We then iterate to obtain

n	$x_2^{(n)}$	$f_e^{(n)}$
0	0.8604	0.8604
1	0.7275	0.0871
2	0.7125	-0.0165
3	0.7149	0.0000 .

For the lowest odd eigenvalue, we have

$$\xi_1 \simeq \pi \Rightarrow x_1^{(o)} = \frac{\pi}{\beta} = 0.5736 \quad .$$

For the next estimate, we have

$$\xi_1 = \frac{\pi\beta}{\beta + 1} \Rightarrow x_1^{(1)} \frac{\pi}{\beta + 1} = 0.4850 \quad .$$

We now iterate for x_1,

n	$x_1^{(n)}$	$f_0^{(n)}$
0	0.5736	-0.5736
1	0.4850	-0.0212
2	0.4816	0.0013
3	0.4818	0.0000 .

For the next higher odd eigenvalue, we have

$$\xi_3 \simeq 2\pi \Rightarrow x_3^{(o)} \simeq \frac{2\pi}{\beta} = 1.1471 \quad .$$

This value, however, is greater than unity and the radical in equation E. 6b will be negative. So, we will use $x_3^{(o)} = 0.99$ and $x_3^{(1)} = 0.95$ just to start the iteration. We then have,

n	$x_3^{(n)}$	$f_0^{(n)}$
0	0.9900	0.5384
1	0.9500	0.1725
2	0.9311	0.0144
3	0.9294	0.0003
4	0.9294	0.0000 .

This work is greatly facilitated if a programable calculator (or a computer) is available.

Appendix F: Electromagnetic fields.

F. 1. The electric and magnetic fields.

The electromagnetic field is described by six scalar fields, the three components of the electric field

$$\underline{\mathcal{E}} = (\mathcal{E}_x, \mathcal{E}_y, \mathcal{E}_z) \qquad \text{(F. 1a)}$$

and the three components of the magnetic field

$$\underline{\mathcal{B}} = (\mathcal{B}_x, \mathcal{B}_y, \mathcal{B}_z) \quad . \qquad \text{(F. 1b)}$$

These fields exert a force on a moving charge q_c, given by

$$\underline{F} = q_c(\underline{\mathcal{E}} + \dot{\underline{x}} \times \underline{\mathcal{B}}) \quad , \qquad \text{(F. 2)}$$

(Coulomb force) where $\dot{\underline{x}}$ is the charge velocity.

In the MKSA system of units, the force is measured in Newtons,

$$\begin{aligned}[F] &= \text{Newton} \equiv 1 \text{ kg} \cdot \text{meter/sec}^2 \\ &= 10^5 \text{ dynes} \\ &(= \text{weight of } 0.102 \text{ by}) \ ,\end{aligned} \qquad \text{(F. 3a)}$$

the electric field in Volts per meter,

$$[\underline{\mathcal{E}}] = \text{Volts/meter} \ , \qquad \text{(F. 3b)}$$

the magnetic field in Webers per square meter,

$$\begin{aligned}[\mathcal{B}] &= \text{Weber/meter}^2 \equiv 1 \text{ Newton/ampere} \cdot \text{meter} \\ &= 10^4 \text{ gauss} \ ,\end{aligned} \qquad \text{(F. 3c)}$$

(note that a Weber is a unit of magnetic flux), the charge in Coulombs

$$[q_c] = \text{Coulomb} \equiv \text{Newton} \cdot \text{meter/Volt} \ , \qquad \text{(F. 3d)}$$

and the current in amperes

$$[I_c] = \text{ampere} \equiv \text{Coulomb/sec} \ . \qquad \text{(F. 3e)}$$

Charge is conserved, and as a consequence, charge density

$$\rho_c = \rho_c(\underline{x}, t) \quad \text{Coulombs/m}^3 \ , \qquad \text{(F. 4a)}$$

and current density

$$\underline{j}_c = \underline{j}_c(\underline{x}, t) \quad \text{amperes/meter}^2 \ , \qquad \text{(F. 4b)}$$

obey a local conservation equation,

$$\frac{\partial}{\partial t}\rho_c(\underline{x}, t) + \frac{\partial}{\partial \underline{x}} \cdot \underline{j}_c(\underline{x}, t) = 0 \ . \qquad \text{(F. 5)}$$

F. 2. Maxwell's equations.

The electric and magnetic fields are related to the charge and current density by means of Maxwell's equations,

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{\mathcal{E}} = \frac{1}{\varepsilon_0} \rho_c \tag{F. 6a}$$

$$\frac{\partial}{\partial \underline{x}} \times \underline{\mathcal{E}} = - \frac{\partial}{\partial t} \underline{\mathcal{B}} \tag{F. 6b}$$

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{\mathcal{B}} = 0 \tag{F. 6c}$$

$$\frac{\partial}{\partial \underline{x}} \times \underline{\mathcal{B}} = \frac{1}{c^2} \left(\frac{1}{\varepsilon_0} \underline{j}_c + \frac{\partial}{\partial t} \underline{\mathcal{E}}\right) , \tag{F. 6d}$$

where c^2 is the square of the speed of light,

$$c = 2.998 \times 10^8 \text{ m/sec} \tag{F. 7}$$

and ε_0 is the permittivity of vacuum and equal to

$$\varepsilon_0 = 8.854 \times 10^{-12} \text{ Farad/meter}^{\#} , \tag{F. 8}$$

such that,

$$4\pi\varepsilon_0 c^2 \equiv 10^7 \text{ Farad} \cdot \text{meters/sec}^2 . \tag{F. 9}$$

It may seem strange that equation F. 9 is written as a definition. There exists an arbitrariness, however, in the scale of the units in which we measure electromagnetic qualities which is removed by equation F. 9. This constant of proportionality relates electromagnetic quantities and mechanical forces. For example, the force between two point charges q_1, q_2 spaced by a distance r, is given by

$$F_e = \left(\frac{1}{4\pi\varepsilon_0}\right) \frac{q_1 q_2}{r^2} , \tag{F. 10}$$

in much the same way as we have another arbitrary constant G[†] to relate our units of mass to our units of force, i. e. for two point masses m_1, m_2

$$F_g = - G \frac{m_1 m_2}{r^2} . \tag{F. 11}$$

If the units of force and charge are chosen independently, then we have no choice about $1/4\pi\varepsilon_0$, just as if the units of force and mass are chosen independently, we have no choice about the gravitational constant. To claim that we have no right to assign arbitrary units to charge, i. e. Coulombs, as is done in the c. g. s. (Gaussian) electromagnetic units since they are related by the Coulomb force, is equivalent to claiming that we have no business in assigning arbitrary units to mass since it is related to force and distance through the gravitational force.

[#] A Farad is the unit of capacitance given by

$$1 \text{ Farad} = 1 \text{ Coulomb/Volt}$$

[†] $G = 6.673 \times 10^{-11}$ Nt · meters^2/kg^2

F.3. The electromagnetic potentials.

The electric and magnetic fields can be solved in terms of the charge and current density through the use of the electromagnetic potentials. More specifically, we define a scalar potential

$$\varphi = \varphi(\underline{x}, t) \quad \text{Volts} \quad ,$$

and a vector potential

$$\underline{A} = \underline{A}(\underline{x}, t) \quad \text{Webers/meter} \quad ,$$

in terms of which,

$$\underline{\mathcal{E}} = -\frac{\partial}{\partial \underline{x}} \varphi - \frac{\partial}{\partial t} \underline{A} \tag{F. 12a}$$

and

$$\underline{\mathcal{B}} = \frac{\partial}{\partial \underline{x}} \times \underline{A} \quad . \tag{F. 12b}$$

It should be noted that equations F. 12a and F. 12b define a unique $\underline{\mathcal{E}}$ and $\underline{\mathcal{B}}$ from a given φ and $\underline{A}$. Interestingly enough the converse is not true. In particular, the transformation of φ and $\underline{A}$ into a new φ' and $\underline{A}'$ given by

$$\varphi' = \varphi - \frac{\partial}{\partial t} \chi \tag{F. 13a}$$

and

$$\underline{A}' = \underline{A} + \frac{\partial}{\partial \underline{x}} \chi \quad , \tag{F. 13b}$$

for any scalar function $\chi(\underline{x}, t)$, leaves $\underline{\mathcal{E}}$ and $\underline{\mathcal{B}}$ unchanged. Try it! Such a transformation is called a <u>gauge transformation</u>.

We can now solve for φ and $\underline{A}$, in terms of ρ_c and $\underline{j}_c$, using the Maxwell equations. In particular, from the divergence of $\underline{\mathcal{E}}$ equation, we have,

$$-\left(\frac{\partial}{\partial \underline{x}} \cdot \frac{\partial}{\partial \underline{x}}\right)\varphi - \frac{\partial}{\partial t}\left(\frac{\partial}{\partial \underline{x}} \cdot \underline{A}\right) = \frac{1}{\epsilon_0} \rho_c \quad . \tag{F. 14a}$$

The curl $\underline{\mathcal{E}}$ and divergence $\underline{\mathcal{B}}$ equations are identically satisfied, and the curl $\underline{\mathcal{B}}$ yields

$$\frac{\partial}{\partial \underline{x}} \times \left(\frac{\partial}{\partial \underline{x}} \times \underline{A}\right) = \frac{1}{c^2}\left(\frac{1}{\epsilon_0} \underline{j}_c - \frac{\partial}{\partial \underline{x}} \frac{\partial}{\partial t} \varphi - \frac{\partial^2}{\partial t^2} \underline{A}\right) \quad .$$

We can express the left hand side using the vector identity (see Appendix G)

$$\frac{\partial}{\partial \underline{x}} \times \left(\frac{\partial}{\partial \underline{x}} \times \underline{A}\right) \equiv \frac{\partial}{\partial \underline{x}}\left(\frac{\partial}{\partial \underline{x}} \cdot \underline{A}\right) - \left(\frac{\partial}{\partial \underline{x}} \cdot \frac{\partial}{\partial \underline{x}}\right) \underline{A} \quad ,$$

to obtain

$$-\left(\frac{\partial}{\partial \underline{x}} \cdot \frac{\partial}{\partial \underline{x}}\right) \underline{A} + \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \underline{A} = \frac{1}{\epsilon_0 c^2} \underline{j}_c - \frac{\partial}{\partial \underline{x}}\left(\frac{1}{c^2} \frac{\partial \varphi}{\partial t} + \frac{\partial}{\partial \underline{x}} \cdot \underline{A}\right) \tag{F. 14b}$$

We now note that the degree of freedom afforded to us by our choice of gauge (through the gauge transformation equations F. 13) allows us to set[#]

$$\frac{1}{c^2}\frac{\partial\varphi}{\partial t}+\frac{\partial}{\partial\underline{x}}\cdot\underline{A}=0 \quad . \qquad (F.15)$$

See problem F. 4. Substituting F. 15 in F. 14a and F. 14b then yields

$$-\left(\frac{\partial}{\partial\underline{x}}\cdot\frac{\partial}{\partial\underline{x}}\right)\varphi+\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\varphi=\frac{1}{\varepsilon_0}\rho_c \qquad (F.16a)$$

and

$$-\left(\frac{\partial}{\partial\underline{x}}\cdot\frac{\partial}{\partial\underline{x}}\right)\underline{A}+\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\underline{A}=\frac{1}{\varepsilon_0 c^2}\underline{j}_c \quad . \qquad (F.16b)$$

In other words the fields φ and $\underline{A}$ obey wave equations with source strengths ρ_c/ε_0 and $\underline{j}_c/\varepsilon_0 c^2$. These can be solved using a Green's function solution to yield

$$\varphi(\underline{x},t)=\frac{1}{4\pi\varepsilon_0}\int\rho_c(\underline{x}',t')\frac{d^3\underline{x}'}{|\underline{x}-\underline{x}'|} \quad , \qquad (F.17a)$$

and

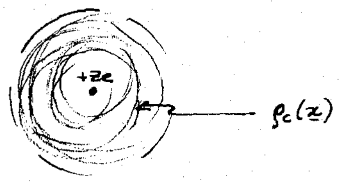
$$\underline{A}(\underline{x},t)=\frac{1}{4\pi\varepsilon_0 c^2}\int\underline{j}_c(\underline{x}',t')\frac{d^3\underline{x}'}{|\underline{x}-\underline{x}'|} \quad , \qquad (F.17b)$$

where t' is the retarded time

$$t'=t-\frac{1}{c}|\underline{x}-\underline{x}'| \quad . \qquad (F.17c)$$

F. 4. Electric fields in matter.

Consider a neutral atom sitting at the origin.



Its total charge (Z = atomic number),

$$Ze-\int\rho_c d^3\underline{x}=0$$

is of course zero (neutral atom). If the electrons are symmetrically distributed around the nucleus, its dipole moment

$$\underline{d}=-\int\underline{x}\,\rho_c(\underline{x})\,d^3\underline{x}=0 \qquad (F.18)$$

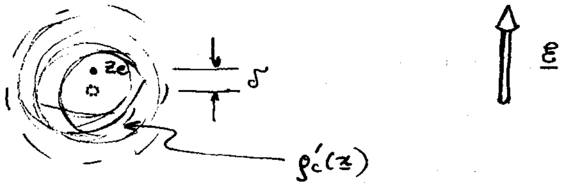
is also zero, since $\rho_c(\underline{x})=\rho_c(-\underline{x})$.

Imagine now that we turn on an electric field $\underline{\mathcal{E}}$ along, say, the z-direction. This will

---

[#] Equation F. 15 is called the Lorentz gauge.

result in a tendency for the nucleus (positively charged) to move along $\underline{\mathcal{E}}$, and in a tendency for the electrons to move along the opposite directions



resulting in an asymmetric distribution $\rho_c'(\underline{x})$ of electrons about the nucleus. Now the dipole moment is no longer zero. In particular, we will have a non-zero component along the z-direction (direction of $\underline{\mathcal{E}}$),

$$d_z = -\int z\, \rho_c'(\underline{x})\, d^3\underline{x} \quad ,$$

where, to compute the integral, we have shifted the origin to the new position of the nucleus.

To lowest order in the electric field strength, one finds that the induced dipole moment is proportional to the magnitude of $\underline{\mathcal{E}}$ and, for an originally symmetric distribution of charge, also along the direction of $\underline{\mathcal{E}}$. We can express this as a vector proportionality equation

$$\underline{d} = \varepsilon_0\, \alpha\, \underline{\mathcal{E}} \quad , \qquad (F.19)$$

where α is a constant, with dimensions of volume, called the polarizability of the atom. The dipole moment $\underline{d}$ has units of charge times displacement (distance) and we can write it as

$$\underline{d} = q_c \underline{\delta} \quad . \qquad (F.20)$$

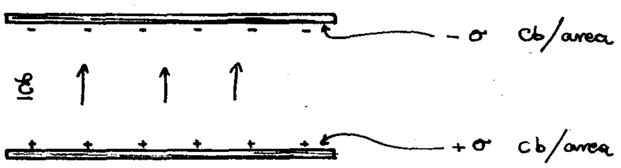
If $q_c = Ze$, the magnitude of the positive (or negative) charge, we see that $-\underline{\delta}$ is the average net displacement of the electrons from the nucleus.

Consider now a piece of material which need not be homogeneous and is exposed to an electric field which may not be uniform. Then the induced dipole moment will vary from one location to another. We may think of the local dipole moment per unit volume as a vector field, called the polarization field and given by

$$\underline{\mathcal{P}} = N\underline{d} = Nq_c\underline{\delta} \quad , \qquad (F.21)$$

where N is equal to the number of atoms per unit volume.

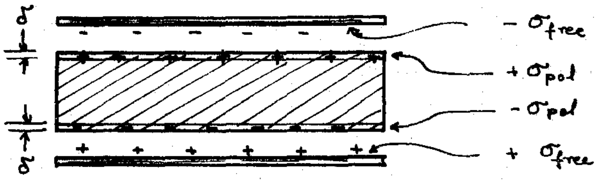
Consider next a parallel plate capacitor formed by two conducting plates charged to a surface charge of $\mp\sigma$ Coulombs per unit area on the top and bottom plate respectively.



This results in an electric field inside the capacitor, given by

$$\mathcal{E} = \frac{\sigma}{\epsilon_0} \quad .$$

If we now introduce a slab of material in between the two plates, it will be polarized in response to the electric field. This will result in a net displacement δ of the positive charges from the negative charges leaving a net positive charge on the upper surface and a net negative charge on the lower surface. This is sketched below.



The induced surface charges on the material are called polarization charges which we will denote by σ_{pol}. It can be seen that the electric field inside the material is now decreased and given by

$$\mathcal{E}_{inside} = \frac{\sigma_{free} - \sigma_{pol}}{\epsilon_0} \quad , \tag{F.22}$$

where the subscript 'free' is used to separate the two types of charge.

The polarization charge is now given by the total displaced charge per unit volume, times the displacement thickness δ, i. e.

$$\sigma_{pol} = (N q_c)\, \delta \quad ,$$

this, however, we recognize as the magnitude of the polarization field on the surface of the material, i. e.

$$\sigma_{pol} = (\mathcal{P})_{surface}$$

For a homogeneous material, however, the polarization field will be uniform and we can write

$$\sigma_{pol} = \mathcal{P} \quad .$$

Substituting in equation F. 22, we then have

$$\mathcal{E}_{inside} = \frac{\sigma_{free} - \mathcal{P}}{\epsilon_0} \quad . \tag{F.23}$$

Equation F. 23 motivates the definition of a new field, called the displacement field $\mathcal{D}$, defined by

$$\underline{\mathcal{D}} \equiv \varepsilon_0 \underline{\mathcal{E}} + \underline{\mathcal{P}} \quad . \qquad \text{(F. 24)}$$

Evidently, the displacement field inside the dielectric is given by

$$\mathcal{D}_{\text{inside}} = \sigma_{\text{free}} \quad .$$

In other words, the magnitude of the displacement field is equal to the electric field we would have had (times ε_0) if the dielectric was not there.

Using Gauss' theorem, one can show that the divergence of the displacement field is equal to the (free) charge density, i.e.

$$\frac{\partial}{\partial \underline{x}} \cdot \mathcal{D} = \rho_{\text{free}} \quad . \qquad \text{(F. 25)}$$

Now since the induced local dipole moment is proportional to $\underline{\mathcal{E}}$, we must also have that the polarization field is also proportional to the local $\underline{\mathcal{E}}$ and that we should be able to write a local proportionality equation,

$$\underline{\mathcal{P}} = \varepsilon_0 \chi_E \underline{\mathcal{E}} \quad . \qquad \text{(F. 26)}$$

The constant χ_E is dimensionless and is called the electric susceptibility.

It would appear that we could combine equations F. 19, F. 21 and F. 26 to compute χ_E. The situation, however, is a bit more complicated because the local electric field is altered by the neighboring polarization charges. The situation is even more complicated by the fact that in a crystal, for example, which is not isotropic, $\underline{\mathcal{P}}$ and $\underline{\mathcal{E}}$ will not even be parallel, i.e. χ_E is a tensor. In a gas, however, where we can neglect all these effects, we have

$$(\chi_E)_{\text{gas}} \simeq N\alpha \quad , \qquad \text{(F. 27)}$$

whereas in isotropic condensed matter (condensed noble gases for example) we have[#]

$$\mathcal{E}_{\text{local}} \simeq \mathcal{E} + \frac{1}{3} \frac{\mathcal{P}}{\varepsilon_0} \quad ,$$

so that,

$$\chi_E \simeq \frac{N\alpha}{1 - N\alpha/3} \quad , \qquad \text{(F. 28)}$$

known as the Clausius-Mossoti equation, which reduces to F. 27 for $N\alpha \ll 1$.

We can combine equations F. 24 and F. 26 to obtain

$$\underline{\mathcal{D}} = \varepsilon_0 (1 + \chi_E) \underline{\mathcal{E}}$$

or, combining the constants,

$$\underline{\mathcal{D}} = \varepsilon \underline{\mathcal{E}} \qquad \text{(F. 29)}$$

where the constant

$$\varepsilon \equiv \varepsilon_0 (1 + \chi_E) \quad , \qquad \text{(F. 30)}$$

[#] See for example Feynman lectures, v. II, chapter 11.

is called the permitivity. The ratio

$$\kappa = \frac{\epsilon}{\epsilon_o} = 1 + \chi_E \tag{F.31}$$

is called the dielectric constant and is the quantity frequently tabulated for dielectrics.

Before leaving this section, we should note that a time varying polarization field yields a current density given by (see equation)

$$\underline{j}_{pol} = \frac{\partial}{\partial t}\,\underline{\mathcal{P}} = N q_c \dot{\underline{\delta}} \quad , \tag{F.32}$$

called polarization current.

F. 5. Magnetic fields in matter.

The interaction of a magnetic field with an atom is more complicated than that of the electric field because it is the result of three mechanisms:

(i) interaction with the magnetic moment $\underline{\mu}_L$ associated with the orbital motion of the electrons,

(ii) interaction with the magnetic moment $\underline{\mu}_S$ associated with the electron spin,

and

(iii) interaction with the nuclear magnetic moment $\underline{\mu}_N$.

In a particular situation these interactions may oppose each other so that a macroscopic sample may tend to either weaken or strengthen an external field depending on the material and its state.

The interaction of the orbital angular momentum of the electrons with an external magnetic field can be understood as follows. If we assume that the external magnetic field is uniform over the extent of the atom, we can approximate the vector potential by

$$\underline{A}(\underline{x}, t) \simeq -\frac{1}{2}\,(\underline{x} \times \underline{\mathcal{B}}) \quad . \tag{F.33}$$

See problem F. 10. From the Hamiltonian for a charged particle in an electromagnetic field (problem 1. 7), we have

$$H = \frac{1}{2m}\,(\underline{p} - q_c\underline{A})\cdot(\underline{p} - q_c\underline{A}) + q_c\varphi \quad . \tag{F.34}$$

If we now choose the z-axis to lie along $\underline{\mathcal{B}}$, we have, in cylindrical coordinates (see appendix G)

$$\underline{A} = \hat{e}_\phi A_\phi = \hat{e}_\phi \frac{1}{2}\, r_\perp \mathcal{B}_z \quad , \tag{F.35}$$

and the Hamiltonian becomes

$$H = \frac{1}{2m}\,[p_\perp^2 + \frac{1}{r_\perp^2}\,(p_\phi - \frac{1}{2}\, q_c\, r_\perp^2\, \mathcal{B}_z)^2 + p_z^2] + q_c\varphi \tag{F.36}$$

where $p_\perp$, p_ϕ and p_z are the conjugate momenta to $r_\perp$, ϕ and z, classically given by

$$p_\perp = m\dot{r}_\perp \tag{F.37a}$$

$$p_\phi = m r_\perp^2 \dot{\phi} + \frac{1}{2} q_c r_\perp^2 \mathcal{B}_z \tag{F.37b}$$

$$p_z = m\dot{z} \ . \tag{F.37c}$$

There is a magnetic (dipole) moment associated with the orbital motion of the charge q_c, given by the circulating current times the orbit area, which is normal to the orbit, i.e.

$$\mu_z = \text{current} \cdot \text{area} = q_c \frac{\dot{\phi}}{2\pi} \cdot \pi r_\perp^2$$

or

$$\mu_z = \frac{1}{2} q_c r_\perp^2 \dot{\phi} \ . \tag{F.38}$$

Substituting for $r_\perp^2 \dot{\phi}$ from F.37b, we then have

$$\mu_z = \frac{q_c}{2m} p_\phi - \frac{q_c^2}{4m} r_\perp^2 \mathcal{B}_z \ . \tag{F.39}$$

Now since $H \neq H(\phi)$, we have that

$$\dot{p}_\phi = - \frac{\partial H}{\partial \phi} = 0$$

or p_ϕ is a constant of the motion. Therefore as we increase the magnetic field from zero to a final value $\mathcal{B}_z$, the magnetic moment <u>decreases</u> by an amount

$$- \delta\mu_z = \frac{q_c^2}{4m} r_\perp^2 \mathcal{B}_z \ , \tag{F.40}$$

since p_ϕ = constant.[#] Note that the change in μ_z is in the opposite direction to the change in $\mathcal{B}_z$ for both positive and negative charges.

Now for a spherically symmetric atom,

$$\langle r_\perp^2 \rangle = \langle x^2 + y^2 \rangle = \frac{2}{3} \langle r^2 \rangle \ ,$$

therefore we expect that for each (spherically symmetric) bound orbiting electron

$$\delta \underline{\mu}_L = - \frac{e^2}{6m_e} \langle r^2 \rangle \underline{\mathcal{B}} \ , \tag{F.41}$$

where $q_c = -e$ is the charge and $m = m_e$ is the mass of the electron.

If we now define a magnetization field $\underline{\mathfrak{M}}$ equal to the local magnetic dipole moment per unit volume, i.e.

$$\underline{\mathfrak{M}} = N\underline{\mu} \ , \tag{F.42}$$

where N is the number of magnetic dipoles per unit volume, we see that, as a result of this effect, an external magnetic field $\underline{\mathcal{B}}$ will result in a change in the local magnetization in the <u>opposite</u> direction to $\underline{\mathcal{B}}$. This behavior is called <u>diamagnetism</u>.

[#] note that $p_\phi = L_z = (\underline{x} \times \underline{p})_z$ = z-component of angular momentum. See also problem 1.5.

The spin angular momentum $\underline{S}$ of the electron is associated with a magnetic dipole moment

$$\underline{\mu}_S = - g_e(\frac{e}{2m_e}) \underline{S} \quad , \tag{F. 43}$$

where g_e is called the <u>electron spin g-factor</u> and found to be equal to (see problem F. 15)

$$g_e \simeq 2.00 \quad , \tag{F. 44}$$

and $\underline{S}$ is the spin angular momentum. The projection of $\underline{S}$ however, along a particular axis (say the z-axis) is restricted for an electron to the values

$$S_z = \pm \frac{1}{2} \hbar \quad , \tag{F. 45}$$

where $\hbar$ is Planck's constant divided by 2π.

If we ignore the possible finite size of the electron, this spin magnetic moment has a Hamiltonian in a magnetic field $\underline{\mathcal{B}}$ given by

$$H_S = - \underline{\mu}_S \cdot \underline{\mathcal{B}} \tag{F. 46}$$

Compare this with the Hamiltonian of problem F. 13 and the discussion in problem F. 15.

Consider now N_S magnetic spin dipoles per unit volume in thermal equilibrium at a temperature T. If we ignore mutual interaction effects and treat them independently, we must have that the probability of finding one parallel to $\underline{\mathcal{B}}$ (spin up) must be proportional to e^{-E/k_BT}, or

$$\text{prob } (\uparrow) \propto e^{+\mu_0 \mathcal{B}_z/k_BT} \quad , \tag{F. 47a}$$

where $\mu_0 = |\mu_z|$, whereas for spin down

$$\text{prob } (\downarrow) \propto e^{-\mu_0 \mathcal{B}_z/k_BT} \quad . \tag{F. 47b}$$

Combining these two, we have that the <u>fraction</u> per unit volume in the spin up state must be given by

$$\frac{N_\uparrow}{N_S} = \frac{e^{\mu_0 \mathcal{B}_z/k_BT}}{e^{\mu_0 \mathcal{B}_z/k_BT} + e^{-\mu_0 \mathcal{B}_z/k_BT}} \quad , \tag{F. 48a}$$

whereas the fraction per unit volume in the spin down state,[#]

$$\frac{N_\downarrow}{N_S} = \frac{e^{-\mu_0 \mathcal{B}_z/k_BT}}{e^{\mu_0 \mathcal{B}_z/k_BT} + e^{-\mu_0 \mathcal{B}_z/k_BT}} \quad . \tag{F. 48b}$$

Correspondingly, the magnetic dipole moment per unit volume (magnetization field) is given by

$$\mathfrak{M}_z = N_S \langle \mu_z \rangle = N_S \frac{e^{\mu_0 \mathcal{B}_z/k_BT} - e^{-\mu_0 \mathcal{B}_z/k_BT}}{e^{\mu_0 \mathcal{B}_z/k_BT} + e^{-\mu_0 \mathcal{B}_z/k_BT}} \mu_0 \quad , \tag{F. 49}$$

[#] Note that we have used the quantum mechanical result that for an electron these are the only two possibilities. A purely classical calculation would have to allow for a random dipole orientation.

where, for an electron (see equation F. 13)

$$\mu_0 = g_e\left(\frac{e}{2m_e}\right)\frac{\hbar}{2} \quad , \qquad \text{(F. 50)}$$

since $S_z = \pm \hbar/2$.

We recognize the temperature dependent part of equation F. 49 as the hyperbolic tangent, i. e.

$$\mathcal{M}_z = N_S\mu_0 \tanh\left(\frac{\mu_0\mathcal{B}_z}{k_BT}\right) \quad . \qquad \text{(F. 51)}$$

Now for weak magnetic fields, $\mu_0\mathcal{B}_z \ll k_BT$ (see problem F. 17) we have $\tanh x \sim x$, so that

$$\mathcal{M}_z \sim N_S \frac{\mu_0^2}{k_BT}\mathcal{B}_z \quad . \qquad \text{(F. 52)}$$

It can be seen that the direction of the magnetization is along $\mathcal{B}_z$, i. e. strengthens the magnetic field. This behavior of matter is called paramagnetism.[#] Note that as $\mathcal{B}$ increases, or T decreases, we reach a saturation magnetization ($\tanh x \to 1$) given by

$$(\mathcal{M}_z)_{sat} = N_S\mu_0 \quad . \qquad \text{(F. 53)}$$

The quantity μ_0, the absolute value of the electron spin moment projection is very nearly equal to the Bohr magneton

$$\mu_B \equiv \frac{e\hbar}{2m_e} = 9.27 \times 10^{-24} \text{ ampere}\cdot\text{meter}^2 \quad . \qquad \text{(F. 54)}$$

Matter displays both kinds of behavior and is found to be diamagnetic or paramagnetic depending on which of the two effects is stronger. One should also mention that some materials display a large scale magnetization whose behavior is called ferromagnetic. In these materials, the spin dipole moments can align each other and we end up with large scale net magnetization fields that extend to domain boundaries. In that case we are not aligning a single electron spin moment when we turn on an external magnetic field, but the magnetization field of a region in the material. The net effect is a response that can be several thousand times stronger than paramagnetism.

In an analogous fashion to the case of dielectrics in electric fields one defines an auxiliary field

$$\underline{\mathcal{H}} \equiv \underline{\mathcal{B}} - \frac{1}{\epsilon_0 c^2}\underline{\mathcal{M}} \quad . \qquad \text{(F. 55)}$$

Note that the curl of the magnetization field is a current density, i. e.

$$\frac{\partial}{\partial \underline{x}} \times \underline{\mathcal{M}} = \underline{j}_{mag} \, , \qquad \text{(F. 56)}$$

which can be identified as the current density that would be required in the magnetic dipoles

---

[#] Mnemonic rule: paramagnetism ⇒ spins parallel to $\mathcal{B}$.

to produce the magnetization field $\underline{\mathcal{M}}$. Correspondingly we can separate the total current in a material into three types:

$\underline{j}_{cond}$ = externally introduced conduction current

$\underline{j}_{mag} \equiv \frac{\partial}{\partial \underline{x}} \times \underline{\mathcal{M}}$ = magnetization current

$\underline{j}_{pol} \equiv \frac{\partial}{\partial t} \underline{\mathcal{P}}$ = polarization current

and correspondingly consider the total current as the sum, i.e.

$$\underline{j} = \underline{j}_{cond} + \underline{j}_{mag} + \underline{j}_{pol} \tag{F. 57}$$

If we introduce these quantities into Maxwell's equations we have

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{\mathcal{E}} = \frac{1}{\epsilon_0} (\rho_{free} + \rho_{pol}) \tag{F. 58a}$$

$$\frac{\partial}{\partial \underline{x}} \times \underline{\mathcal{E}} = - \frac{\partial}{\partial t} \underline{\mathcal{B}} \tag{F. 58b}$$

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{\mathcal{B}} = 0 \tag{F. 58c}$$

$$\frac{\partial}{\partial \underline{x}} \times \underline{\mathcal{B}} = \frac{1}{c^2} \left[\left(\frac{1}{\epsilon_0} \underline{j}_{cond} + \frac{\partial}{\partial \underline{x}} \times \underline{\mathcal{M}} + \frac{\partial}{\partial t} \underline{\mathcal{P}} \right) + \frac{\partial}{\partial t} \underline{\mathcal{E}} \right], \tag{F. 58d}$$

or, substituting the $\underline{\mathcal{D}}$ and $\underline{\mathcal{H}}$ fields,

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{\mathcal{D}} = \rho_{free} \tag{F. 59a}$$

$$\frac{\partial}{\partial \underline{x}} \times \underline{\mathcal{E}} = - \frac{\partial}{\partial t} \underline{\mathcal{B}} \tag{F. 59b}$$

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{\mathcal{B}} = 0 \tag{F. 59c}$$

$$\frac{\partial}{\partial \underline{x}} \times \underline{\mathcal{H}} = \underline{j}_{cond} + \frac{\partial}{\partial t} \underline{\mathcal{D}} , \tag{F. 59d}$$

which may look simpler but has of course swept the difficulty of the material response in the auxiliary fields $\underline{\mathcal{D}}$ and $\underline{\mathcal{H}}$.

Problems

F. 1 Assume that the electric and magnetic fields obey periodic boundary conditions at the edges of a box

$$|x| < L_x/2 \ , \quad |y| < L_y/2 \ , \quad |z| < L_z/2 \ ,$$

i. e.

$$\underline{\mathcal{E}}(-L_x/2, y, z) = \underline{\mathcal{E}}(L_x/2, y, z)$$

$$\underline{\mathcal{E}}(x, -L_y/2, z) = \underline{\mathcal{E}}(x, L_y/2, z)$$

$$\underline{\mathcal{E}}(x, y, -L_z/2) = \underline{\mathcal{E}}(x, y, L_z/2)$$

and similarly for $\underline{\mathcal{B}}$. Show that both fields can be expressed as Fourier series of the form

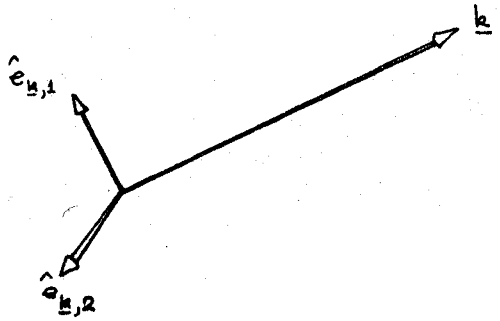
$$\underline{\mathcal{E}}(\underline{x}, t) = \sum_{\underline{k}} \left[\sum_{\ell=0}^{2} \mathcal{E}_{\underline{k}, \ell}(t) \ \hat{e}_{\underline{k}, \ell} \right] e^{i\underline{k} \cdot \underline{x}}$$

and

$$\underline{\mathcal{B}}(\underline{x}, t) = \sum_{\underline{k}} \left[\sum_{\ell=0}^{2} \mathcal{B}_{\underline{k}, \ell}(t) \ \hat{e}_{\underline{k}, \ell} \right] e^{i\underline{k} \cdot \underline{x}} \ ,$$

where $\hat{e}_{\underline{k},0}$ is a unit vector along the wavenumber $\underline{k}$ and $\hat{e}_{\underline{k},1}$, $\hat{e}_{\underline{k},2}$ are the two perpendicular unit vectors such that

$$\hat{e}_{\underline{k},1} \times \hat{e}_{\underline{k},2} = \hat{e}_{\underline{k},0} \quad (\text{along } \underline{k})$$



and where

$$\mathcal{E}_{\underline{k}, \ell}(t) = \frac{1}{\text{Volume}} \int [\hat{e}_{\underline{k}, \ell} \cdot \underline{\mathcal{E}}(\underline{x}, t)] \ e^{-i\underline{k} \cdot \underline{x}} \ d^3\underline{x}$$

and

$$\mathcal{B}_{\underline{k}, \ell}(t) = \frac{1}{\text{Volume}} \int [\hat{e}_{\underline{k}, \ell} \cdot \underline{\mathcal{B}}(\underline{x}, t)] \ e^{-i\underline{k} \cdot \underline{x}} \ d^3\underline{x}$$

Problems (continued)

Hint: Show that for $\underline{k}$ and $\underline{k}'$ whose components k_i, k'_i are given by

$$k_i,\ k'_i = n_i\left(\frac{2\pi}{L_i}\right)\ ; \quad n_i = 0, \pm 1, \pm 2, \ldots$$

$$\int_{\text{Volume}} e^{i(\underline{k}-\underline{k}')\cdot \underline{x}}\, d^3\underline{x} = (\text{Volume})\,\delta_{\underline{k},\underline{k}'}$$

F. 2 Assuming that the charge and current densities also obey periodic boundary conditions at the edges of the same box, show that

$$\rho_c(\underline{x}, t) = \sum_{\underline{k}} \rho_{\underline{k}}(t)\, e^{i\underline{k}\cdot\underline{x}}$$

and

$$\underline{j}_c(\underline{x}, t) = \sum_{\underline{k},\ell} [j_{\underline{k},\ell}(t)\ \hat{e}_{\underline{k},\ell}]\ e^{i\underline{k}\cdot\underline{x}}$$

and that the equation of charge conservation becomes

$$\dot{\rho}_{\underline{k}} + ik\, j_{\underline{k},0} = 0\ , \quad \text{for all admissible } \underline{k}$$

where $k^2 = k_x^2 + k_y^2 + k_z^2$. Note the form of the divergence in this representation and the fact that the charge conservation equation does not restrict the two transverse components $j_{\underline{k},1}$, $j_{\underline{k},2}$.

F. 3 Show that the Maxwell equations for the Fourier components of $\underline{\mathcal{E}}$ and $\underline{\mathcal{B}}$ reduce to algebraic equations, given by

$$ik\,\mathcal{E}_{\underline{k},0} = \frac{1}{\varepsilon_0}\rho_{\underline{k}}\ , \qquad ik\,\mathcal{B}_{\underline{k},0} = 0$$

$$-ik\,\mathcal{E}_{\underline{k},1} = \dot{\mathcal{B}}_{\underline{k},2}\ , \qquad ikc^2\,\mathcal{B}_{\underline{k},1} = \frac{1}{\varepsilon_0} j_{\underline{k},2} + \dot{\mathcal{E}}_{\underline{k},2}$$

$$ik\,\mathcal{E}_{\underline{k},2} = \dot{\mathcal{B}}_{\underline{k},1}\ , \qquad -ikc^2\,\mathcal{B}_{\underline{k},2} = \frac{1}{\varepsilon_0} j_{\underline{k},1} + \dot{\mathcal{E}}_{\underline{k},1}\ .$$

Note that these are six independent equations for six unknowns, not eight as the vector form of the Maxwell equations in $(\underline{x}, t)$ space might lead you to believe.

F. 4 Assume that a particular choice φ' and $\underline{A}'$ yielded

$$\frac{1}{c^2}\frac{\partial\varphi'}{\partial t} + \frac{\partial}{\partial\underline{x}}\cdot\underline{A}' = \psi\ .$$

We now transform these using equations F. 13. Show that

$$\frac{1}{c^2}\frac{\partial\varphi}{\partial t} + \frac{\partial}{\partial\underline{x}}\cdot\underline{A} = 0$$

<u>Problems</u> (continued)

if we chose χ such that

$$-\left(\frac{\partial}{\partial \underline{x}} \cdot \frac{\partial}{\partial \underline{x}}\right)\chi + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\chi = \psi \quad .$$

[Note that this is a wave equation for χ with a source ψ, whose solution is given by

$$\chi(\underline{x}, t) = \frac{1}{4\pi}\int \psi(\underline{x}', t')\frac{d^3\underline{x}'}{|\underline{x} - \underline{x}'|} \quad ,$$

where

$$t' = t - \frac{1}{c}\,|\underline{x} - \underline{x}'| \quad .$$

You don't have to show this!]

F. 5 If we separate the local charge density into externally introduced free charges and polarized bound charges, i. e.

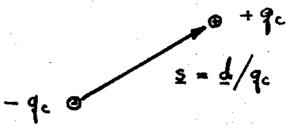
$$\rho_c(\underline{x}, t) = \rho_{free}(\underline{x}, t) + \rho_{pol}(\underline{x}, t) \quad ,$$

show that the divergence of the polarization field is given by,

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{\mathcal{P}} = -\rho_{pol} \quad ,$$

in general

A dipole can be idealized as a pair of charges $\pm q_c$, separated by a vector $\underline{s} = \underline{d}/q_c$, directed from the negative charge to the positive charge, in the limit of $\underline{s} \to 0$, keeping $\underline{d}$ constant.



F. 6 Show that an idealized stationary dipole generates an electrostatic potential given by

$$\varphi_1 = (\underline{d} \cdot \frac{\partial}{\partial \underline{x}})\,\varphi_0 \quad ,$$

where φ_0 is the electrostatic potential generated by a unit charge, i. e.

$$\varphi_0 = \frac{1}{4\pi\varepsilon_0 |\underline{x}|} \quad ,$$

and that therefore

$$\varphi_1(\underline{x}, t) = \frac{1}{4\pi\varepsilon_0}\frac{\underline{d} \cdot \underline{x}}{|\underline{x}|^3} \quad .$$

F. 7 Using the Lagrangian for charged particles in a constant electromagnetic field (equation 1.38), show that the Lagrangian for an idealized dipole in a constant electromagnetic field is given by

$$L(\underline{x}, \underline{\dot{x}}, \underline{d}, \underline{\dot{d}}; t) = \frac{1}{2} m |\underline{\dot{x}}|^2 - \underline{d} \cdot \frac{\partial}{\partial \underline{x}} \varphi + \underline{\dot{d}} \cdot \underline{A} + \underline{\dot{x}} \cdot (\underline{d} \cdot \frac{\partial}{\partial \underline{x}}) \underline{A}$$

and compute the conjugate momenta

$$\underline{p} \equiv \frac{\partial L}{\partial \underline{\dot{x}}} , \qquad \underline{\pi} = \frac{\partial L}{\partial \underline{\dot{d}}} .$$

Note that we are neglecting the effect *of* the dipole *on* φ and $\underline{A}$.

F. 8 Show that, in this approximation, the Hamiltonian for the dipole is given by

$$H = \frac{1}{2m} |\underline{p} - (\underline{d} \cdot \frac{\partial}{\partial \underline{x}}) \underline{A}|^2 + \underline{d} \cdot \frac{\partial}{\partial \underline{x}} \varphi ,$$

that the *energy* of a stationary dipole ($\underline{\dot{x}}=0$) is given by

$$E = - \underline{d} \cdot \underline{\mathcal{E}} ,$$

and that it feels a *torque* given by

$$\underline{\tau} = \underline{d} \times \underline{\mathcal{E}} .$$

F. 9 If $\underline{A} = 0$, derive the equations of motion for the center of mass of the dipole, i.e.

$$\underline{\dot{p}} = - \frac{\partial H}{\partial \underline{x}} , \qquad \underline{\dot{x}} = \frac{\partial H}{\partial \underline{p}} ,$$

and show that it feels a *force* given by

$$\underline{F} = (\underline{d} \cdot \frac{\partial}{\partial \underline{x}}) \underline{\mathcal{E}}$$

Hint: $\frac{\partial}{\partial \underline{x}} (\underline{u} \cdot \underline{v}) \equiv \underline{u} \times (\frac{\partial}{\partial \underline{x}} \times \underline{v}) + \underline{v} \times (\frac{\partial}{\partial \underline{x}} \times \underline{u}) + (\underline{u} \cdot \frac{\partial}{\partial \underline{x}}) \underline{v} + (\underline{v} \cdot \frac{\partial}{\partial \underline{x}}) \underline{u}$.

F. 10 Show that the vector potential of equation F. 33 yields the correct magnetic field, if $\underline{\mathcal{B}}$.

Hint: $\frac{\partial}{\partial \underline{x}} \times (\underline{u} \times \underline{v}) \equiv \underline{u} (\frac{\partial}{\partial \underline{x}} \cdot \underline{v}) - \underline{v} (\frac{\partial}{\partial \underline{x}} \cdot \underline{u}) + (\underline{v} \cdot \frac{\partial}{\partial \underline{x}}) \underline{u} - (\underline{u} \cdot \frac{\partial}{\partial \underline{x}}) \underline{v}$.

F. 11 Show that the Hamiltonian of equation F. 34 can be written, for a uniform magnetic field as

$$H = \frac{1}{2m} p^2 + q_c \varphi - \frac{q_c}{2m} \underline{L} \cdot \underline{\mathcal{B}} + \frac{q_c^2}{8m} [|\underline{x}|^2 \mathcal{B}^2 - (\underline{x} \cdot \underline{\mathcal{B}})^2] ,$$

where

$$\underline{L} \equiv \underline{x} \times \underline{p}$$

is the (orbital) angular momentum.

F. 12 From the definition of the orbital magnetic dipole moment (equation F. 39) show that, in general,

$$\underline{\mu}_L = \frac{q_c}{2m}\,\underline{L} - \frac{q_c^2}{4m}\,[\underline{x}\cdot\underline{x} - (\underline{x}\cdot\hat{e}_{\mathcal{B}})^2]\,\underline{\mathcal{B}} \;,$$

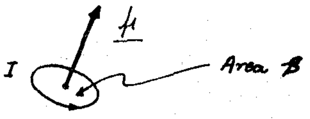
where $\underline{L} = \underline{x} \times \underline{p}$ is the (orbital) angular momentum, and $\hat{e}_{\mathcal{B}}$ is a unit vector in the direction of $\underline{\mathcal{B}}$.

Hint: Show that $p_\phi = L_z$ = z-component of angular momentum in equation F. 39.

F. 13 Show that if we increase a uniform magnetic field from zero to some final value $\underline{\mathcal{B}}$, the <u>change</u> in the Hamiltonian, as a result of the orbital motion of a charge q_c, is given by

$$H' = -\,\underline{\mu}_L\cdot\underline{\mathcal{B}} - \frac{q_c^2}{8m}\,[(\underline{x}\cdot\underline{x})\,\mathcal{B}^2 - (\underline{x}\cdot\underline{\mathcal{B}})^2] \;.$$

F. 14 A magnetic dipole $\underline{\mu}$ can be idealized as the product of the current in a small loop times the area of the loop, in the limit of zero area keeping the product constant.



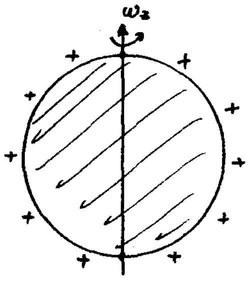
Show that the vector potential generated by such a idealized dipole is given by

$$\underline{A}_1 = \frac{1}{4\pi\epsilon_0 c^2}\cdot\frac{\underline{\mu}\times\underline{x}}{|\underline{x}|^3} \;.$$

F. 15 <u>Electron spin magnetic moment</u>. It should not be surprising that there is a need for a proportionality constant g_e to relate the dipole moment to the angular momentum. The magnetic dipole moment is related to the distribution of <u>current density</u>, whereas the angular momentum is a function of the distribution of mass. For example compute the g-factor for the magnetic moment μ_z

$$\mu_z = g\left(\frac{q_c}{2m}\right) L_z \;,$$

of a uniformly dense sphere of mass m whose charge, as a result of electrostatic repulsive forces, resides uniformly on the surface.



$4\pi a^2 \sigma_c = \rho_c$

$\frac{4\pi}{3} a^3 \rho_m = m$

Answer: $g = 5/3$.

Problems (continued)

F. 16 Derive the result of equation F. 40 using Maxwell's equations. Consider an orbiting charge at zero magnetic field and allow the field to rise to its final value. Integrate the resulting torque from the induced electric field as a consequence of the curl ($\underline{\mathcal{E}}$) equation.

F. 17 Compute the magnetic field strength (give your answer in gauss) such that

$$\frac{\mu_B \mathcal{B}}{k_B T} \sim 1$$

at room temperature, where μ_B is the Bohr magneton (equation F. 54).

Appendix G: Useful vector relations

G. 1. Vector identities

$$\underline{A}\cdot(\underline{B}\times\underline{C}) \equiv \underline{C}\cdot(\underline{A}\times\underline{B}) \equiv \underline{B}\cdot(\underline{C}\times\underline{A}) \equiv -\underline{A}\cdot(\underline{C}\times\underline{B}) \equiv -\underline{B}\cdot(\underline{A}\times\underline{C}) \equiv -\underline{C}\cdot(\underline{B}\times\underline{A}) \qquad (G.\ 1)$$

$$\underline{A}\times(\underline{B}\times\underline{C}) \equiv -(\underline{A}\cdot\underline{B})\underline{C} + (\underline{A}\cdot\underline{C})\underline{B} \qquad (G.\ 2)$$

$$\underline{A}\times(\underline{B}\times\underline{C}) + \underline{C}\times(\underline{A}\times\underline{B}) + \underline{B}\times(\underline{C}\times\underline{A}) \equiv 0 \qquad (G.\ 3)$$

$$\begin{aligned}(\underline{A}\times\underline{B})\cdot(\underline{C}\times\underline{D}) &\equiv A\cdot[\underline{B}\times(\underline{C}\times\underline{D})]\\ &= (A\cdot C)(B\cdot D) - (A\cdot D)(B\cdot C)\end{aligned} \qquad (G.\ 4)$$

G. 2. Differential identities $\left(\frac{\partial}{\partial\underline{x}} \equiv \hat{e}_x\frac{\partial}{\partial x} + \hat{e}_y\frac{\partial}{\partial y} + \hat{e}_z\frac{\partial}{\partial z}\right)$

$$\frac{\partial}{\partial\underline{x}}(\varphi\psi) \equiv \left(\frac{\partial\varphi}{\partial\underline{x}}\right)\psi + \varphi\frac{\psi}{\partial\underline{x}} \qquad (G.\ 5)$$

$$\frac{\partial}{\partial\underline{x}}\cdot(\psi\underline{A}) \equiv \frac{\partial\psi}{\partial\underline{x}}\cdot\underline{A} + \psi\frac{\partial}{\partial\underline{x}}\cdot\underline{A} \qquad (G.\ 6)$$

$$\frac{\partial}{\partial\underline{x}}\times(\psi\underline{A}) = \frac{\partial\psi}{\partial x}\times\underline{A} + \psi\frac{\partial}{\partial\underline{x}}\times\underline{A} \qquad (G.\ 7)$$

$$\frac{\partial}{\partial\underline{x}}\cdot(\underline{B}\times\underline{C}) = \left(\frac{\partial}{\partial\underline{x}}\times\underline{B}\right)\cdot\underline{C} - \underline{B}\cdot\left(\frac{\partial}{\partial\underline{x}}\times\underline{C}\right) \qquad (G.\ 8)$$

$$\frac{\partial}{\partial\underline{x}}\times(\underline{B}\times\underline{C}) = -\left(\frac{\partial}{\partial\underline{x}}\cdot\underline{B}\right)\underline{C} + \underline{B}\left(\frac{\partial}{\partial\underline{x}}\cdot\underline{C}\right) - \left(\underline{B}\cdot\frac{\partial}{\partial\underline{x}}\right)\underline{C} + \left(\underline{C}\cdot\frac{\partial}{\partial\underline{x}}\right)\underline{B} \qquad (G.\ 9)$$

$$\frac{\partial}{\partial\underline{x}}(\underline{B}\cdot\underline{C}) = -\left(\frac{\partial}{\partial\underline{x}}\times\underline{B}\right)\times\underline{C} + \underline{B}\times\left(\frac{\partial}{\partial\underline{x}}\times\underline{C}\right) + \left(\underline{B}\cdot\frac{\partial}{\partial\underline{x}}\right)\underline{C} + \left(\underline{C}\cdot\frac{\partial}{\partial\underline{x}}\right)\underline{B} \qquad (G.\ 10)$$

$$\left(\frac{\partial}{\partial\underline{x}}\cdot\frac{\partial}{\partial\underline{x}}\right)\underline{A} \equiv \frac{\partial}{\partial\underline{x}}\left(\frac{\partial}{\partial\underline{x}}\cdot\underline{A}\right) - \frac{\partial}{\partial\underline{x}}\times\left(\frac{\partial}{\partial\underline{x}}\times\underline{A}\right) \qquad (G.\ 11)$$

$$\frac{\partial}{\partial\underline{x}}\times\frac{\partial}{\partial\underline{x}}\psi \equiv 0 \qquad (G.\ 12)$$

$$\frac{\partial}{\partial\underline{x}}\cdot\left(\frac{\partial}{\partial\underline{x}}\times\underline{A}\right) \equiv 0 \qquad (G.\ 13)$$

Note also

$$\frac{\partial}{\partial\underline{x}}\cdot\underline{x} \equiv 3\ , \qquad \frac{\partial}{\partial\underline{x}}\times\underline{x} \equiv 0 \qquad \frac{\partial}{\partial\underline{x}}\frac{1}{|\underline{x}|} = \frac{\underline{x}}{|\underline{x}|^3} \qquad (G.\ 14)$$

G. 4. Integral identities

update

If $d^3\underline{x}$ is the volume element, $d\underline{S}$ is the unit outward normal to the surface element and $d\underline{\ell}$ is the counter-clockwise tangent to the contour $\mathcal{C}$ bounding the surface S we have;

Divergence theorem: $$\int_V d^3\underline{x}\,\left(\frac{\partial}{\partial\underline{x}}\cdot\underline{A}\right) \equiv \int_S d\underline{S}\cdot\underline{A} \tag{G. 15}$$

$$\int_V d^3\underline{x}\,\left(\frac{\partial\psi}{\partial\underline{x}}\right) \equiv \int_S d\underline{S}\;\psi \tag{G. 16}$$

$$\int_V d^3\underline{x}\,\left(\frac{\partial}{\partial\underline{x}}\times\underline{A}\right) \equiv \int_S d\underline{S}\times\underline{A} \tag{G. 17}$$

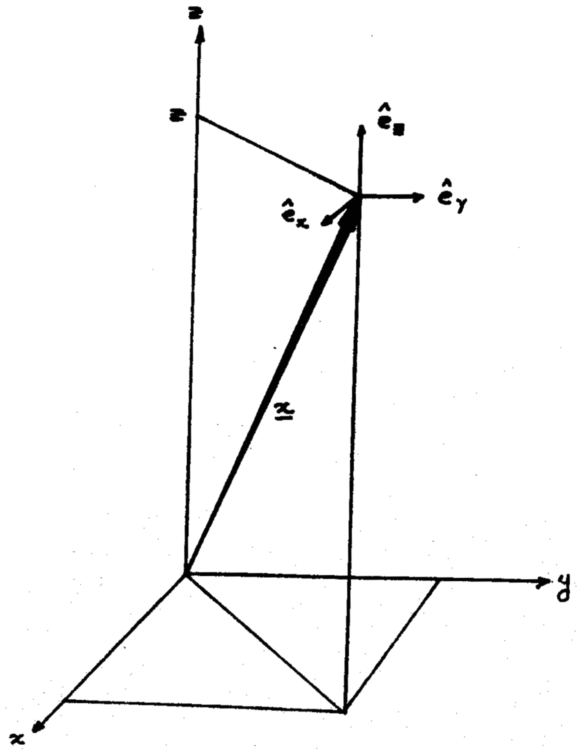
Green's 1st identity: $$\int_V d^3\underline{x}\,\left(\varphi\,\frac{\partial}{\partial\underline{x}}\cdot\frac{\partial\psi}{\partial\underline{x}} + \frac{\partial\varphi}{\partial\underline{x}}\cdot\frac{\partial\psi}{\partial\underline{x}}\right) \equiv \int_S d\underline{S}\cdot\varphi\,\frac{\partial\psi}{\partial\underline{x}} \tag{G. 18}$$

Green's theorem: $$\int_V d^3\underline{x}\,\left(\varphi\,\frac{\partial}{\partial\underline{x}}\cdot\frac{\partial\psi}{\partial\underline{x}} - \psi\,\frac{\partial}{\partial\underline{x}}\cdot\frac{\partial\varphi}{\partial\underline{x}}\right) \equiv \int_S d\underline{S}\cdot\left(\varphi\,\frac{\partial\psi}{\partial\underline{x}} - \psi\,\frac{\partial\varphi}{\partial\underline{x}}\right) \tag{G. 19}$$

Stokes's theorem: $$\int_S d\underline{S}\cdot\left(\frac{\partial}{\partial\underline{x}}\times\underline{A}\right) \equiv \oint_{\mathcal{C}} d\underline{\ell}\cdot\underline{A} \tag{G. 20}$$

$$\int_S d\underline{S}\times\left(\frac{\partial\psi}{\partial\underline{x}}\right) \equiv \oint_{\mathcal{C}} d\underline{\ell}\;\psi\ . \tag{G. 21}$$

G. 4. Cartesian coordinate system



$$d\hat{e}_x = 0$$

$$d\hat{e}_y = 0$$

$$d\hat{e}_z = 0$$

$$\underline{x} = \hat{e}_x x + \hat{e}_y y + \hat{e}_z z$$

$$d\underline{x} = \hat{e}_x dx + \hat{e}_y dy + \hat{e}_z dz$$

$$\frac{\partial}{\partial \underline{x}} = \hat{e}_x \frac{\partial}{\partial x} + \hat{e}_y \frac{\partial}{\partial y} + \hat{e}_z \frac{\partial}{\partial z}$$

Examples:

$$\text{grad}\ (\psi) = \frac{\partial}{\partial \underline{x}}\ \psi = \hat{e}_x \frac{\partial \psi}{\partial x} + \hat{e}_y \frac{\partial \psi}{\partial y} + \hat{e}_z \frac{\partial \psi}{\partial z}$$

$$\text{div}\ (\underline{A}) = \frac{\partial}{\partial \underline{x}} \cdot \underline{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

$$\text{curl}\ (\underline{A}) \equiv \frac{\partial}{\partial \underline{x}} \times \underline{A} = \det \begin{pmatrix} \hat{e}_x & \hat{e}_y & \hat{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{pmatrix}$$

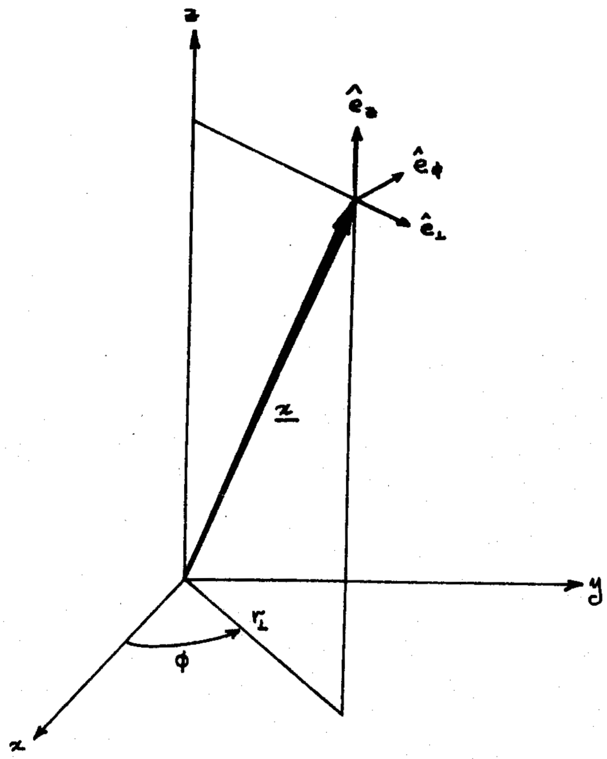
$$= \hat{e}_x \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)$$

$$+ \hat{e}_y \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)$$

$$+ \hat{e}_z \left(\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial y}\right)$$

$$\text{Lapl}\ (\psi) \equiv \left(\frac{\partial}{\partial \underline{x}} \cdot \frac{\partial}{\partial \underline{x}}\right) \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}$$

G. 5. Cylindrical coordinate system



$$d\hat{e}_\perp = \hat{e}_\varphi d_\varphi$$

$$d\hat{e}_\varphi = - \hat{e}_\perp dr_\perp$$

$$d\hat{e}_z = 0$$

$$\underline{x} = \hat{e}_\perp r_\perp + \hat{e}_z z$$

$$d\underline{x} = \hat{e}_\perp r_\perp + \hat{e}_\varphi r_\perp d\varphi + \hat{e}_z dz$$

$$\frac{\partial}{\partial \underline{x}} = \hat{e}_\perp \frac{\partial}{\partial r_\perp} + \hat{e}_\varphi \frac{1}{r_\perp} \frac{\partial}{\partial \varphi} + \hat{e}_z \frac{\partial}{\partial z}$$

Examples:

$$\frac{\partial}{\partial \underline{x}} \psi = \hat{e}_x \frac{\partial \psi}{\partial x} + \hat{e}_y \frac{\partial \psi}{\partial y} + \hat{e}_z \frac{\partial \psi}{\partial z}$$

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{A} = \frac{1}{r_\perp} \left[\frac{\partial}{\partial r_\perp} (r_\perp A_\perp) + \frac{\partial A_\varphi}{\partial \varphi}\right] + \frac{\partial A_z}{\partial z}$$

$$\frac{\partial}{\partial \underline{x}} \times \underline{A} = \frac{1}{r_\perp} \det \begin{pmatrix} \hat{e}_\perp & \hat{e}_\varphi r_\perp & \hat{e}_z \\ \frac{\partial}{\partial r_\perp} & \frac{\partial}{\partial \varphi} & \frac{\partial}{\partial z} \\ A_\perp & r_\perp A_\varphi & A_r \end{pmatrix}$$

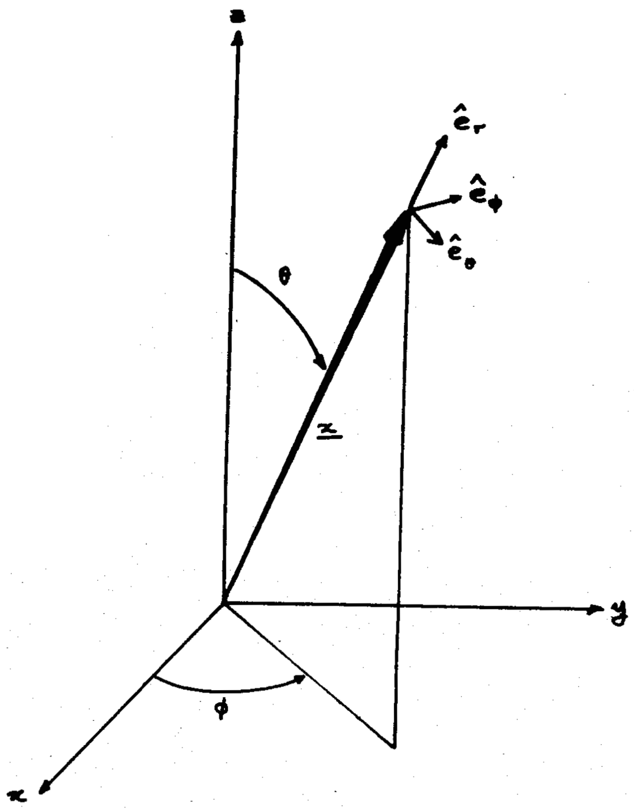
$$= \hat{e}_\perp \left(\frac{1}{r_\perp} \frac{\partial A_z}{\partial \varphi} - \frac{\partial A_\varphi}{\partial z}\right)$$

$$+ \hat{e}_\varphi \left(\frac{\partial A_\perp}{\partial z} - \frac{\partial A_z}{\partial r_\perp}\right)$$

$$+ \hat{e}_z \frac{1}{r_\perp} \left[\frac{\partial}{\partial r_\perp} (r_\perp A_\varphi) - \frac{\partial A_\perp}{\partial \varphi}\right]$$

$$\left(\frac{\partial}{\partial \underline{x}} \cdot \frac{\partial}{\partial \underline{x}}\right) \psi = \frac{1}{r_\perp} \frac{\partial}{\partial r_\perp} \left(r_\perp \frac{\partial \psi}{\partial r_\perp}\right) + \frac{1}{r_\perp^2} \frac{\partial^2 \psi}{\partial \varphi^2} + \frac{\partial^2 \psi}{\partial z^2}$$

G. 6. Spherical coordinate system



$$d\hat{e}_r = \hat{e}_\theta d\theta + \hat{e}_\varphi \sin\theta\, d\varphi$$

$$d\hat{e}_\theta = -\hat{e}_r d\theta + \hat{e}_\varphi \cos\theta\, d\varphi$$

$$d\hat{e}_\varphi = -\hat{e}_r \sin\theta\, d\varphi - \hat{e}_\theta \cos\theta\, d\varphi$$

$$\underline{x} = \hat{e}_r r$$

$$d\underline{x} = \hat{e}_r dr + \hat{e}_\theta r d\theta + \hat{e}_\varphi r \sin\theta\, d\varphi$$

$$\frac{\partial}{\partial \underline{x}} = \hat{e}_r \frac{\partial}{\partial r} + \hat{e}_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{e}_\theta \frac{1}{r\sin\theta} \frac{\partial}{\partial \varphi}$$

Examples:

$$\frac{\partial \psi}{\partial \underline{x}} = \hat{e}_r \frac{\partial \psi}{\partial r} + \hat{e}_\theta \frac{1}{r} \frac{\partial \psi}{\partial \theta} + \hat{e}_\varphi \frac{1}{r\sin\theta} \frac{\partial \psi}{\partial \varphi}$$

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{A} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 A_r) + \frac{1}{r\sin\theta} [\frac{\partial}{\partial \theta} (\sin\theta\, A_\theta) + \frac{\partial A_\varphi}{\partial \varphi}]$$

$$\frac{\partial}{\partial \underline{x}} \times \underline{A} = \frac{1}{r^2 \sin\theta} \det \begin{pmatrix} \hat{e}_r & \hat{e}_\theta r & \hat{e}_\varphi r \sin\theta \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} \\ A_r & rA_\theta & r\sin\theta\, A_\varphi \end{pmatrix}$$

$$= \hat{e}_r \frac{1}{r\sin\theta} [\frac{\partial}{\partial \theta} (\sin\theta\, A_\varphi) - \frac{\partial A_\theta}{\partial \varphi}]$$

$$+ \hat{e}_\theta [\frac{1}{r\sin\theta} \frac{\partial A_r}{\partial \varphi} - \frac{1}{r} \frac{\partial}{\partial r} (r A_\varphi)] + \hat{e}_\varphi \frac{1}{r} [\frac{\partial}{\partial r} (rA_\theta) - \frac{\partial A_r}{\partial \theta}]$$

$$(\frac{\partial}{\partial \underline{x}} \cdot \frac{\partial}{\partial \underline{x}}) \psi = \frac{1}{r^2} [\frac{\partial}{\partial r} (r^2 \frac{\partial \psi}{\partial r}) + \frac{1}{\sin\theta} \frac{\partial}{\partial \theta} (\sin\theta \frac{\partial \psi}{\partial \theta}) + \frac{1}{\sin^2\theta} \frac{\partial^2 \psi}{\partial \varphi^2}]$$

update

Appendix H. The spherical Bessel functions.

The spherical Bessel functions provide the non-trivial solutions to the equation

$$\mathcal{D}_\ell \, f_\ell(\rho) = f_\ell(\rho) \tag{H. 1a}$$

where

$$\mathcal{D}_\ell \equiv -\frac{1}{\rho^2}\frac{d}{d\rho}\left(\rho^2 \frac{d}{d\rho}\right) + \frac{\ell(\ell+1)}{\rho^2} \tag{H. 1b}$$

$$\equiv -\frac{d^2}{d\rho^2} - \frac{2}{\rho}\frac{d}{d\rho} + \frac{\ell(\ell+1)}{\rho^2} \tag{H. 1c}$$

We can eliminate the first derivative term by a transformation of the dependent variable. Making the substitution

$$g_\ell = \rho f_\ell \tag{H. 2}$$

results then in the equation

$$\left[\frac{d^2}{d\rho^2} + 1 - \frac{\ell(\ell+1)}{\rho^2}\right] g_\ell(\rho) = 0 \quad . \tag{H. 3}$$

For $\ell = 0$, we recognize the two solutions as the sine and cosine of ρ, and therefore since $f_0 = g_0/\rho$, we have

$$f_0(\rho) = \begin{cases} \frac{1}{\rho} \sin \rho \\ \frac{1}{\rho} \cos \rho \quad . \end{cases} \tag{H. 4}$$

These two solutions are called respectively the $\ell = 0$ regular spherical Bessel function (finite at the origin),

$$j_0(x) = \frac{1}{\rho} \sin \rho \quad , \tag{H. 5a}$$

and the $\ell = 0$ irregular spherical Bessel function, or spherical Neuman function (infinite at the origin),

$$n_0(x) = \frac{1}{\rho} \cos \rho \quad . \tag{H. 5b}$$

To solve the equation for $\ell \neq 0$, we note that the operator $\mathcal{D}_\ell$ can be factored into the product of two operators, i. e.

$$\mathcal{D}_\ell \equiv \left(-\frac{d}{d\rho} + \frac{\ell-1}{\rho}\right)\left(\frac{d}{d\rho} + \frac{\ell+1}{\rho}\right) \quad . \tag{H. 6}$$

This motivates the definition of two differential operators

$$F^+_\ell \equiv -\frac{d}{d\rho} + \frac{\ell}{\rho} \quad , \qquad F^-_\ell \equiv \frac{d}{d\rho} + \frac{\ell}{\rho} \quad . \tag{H. 7}$$

We then have, from equations H. 6 and H. 7, the identity

$$\mathcal{D}_\ell \equiv F^+_{\ell-1} \, F^-_{\ell+1} \quad . \tag{H. 8}$$

If we now compute the reverse product $F^-_{\ell+1} \, F^+_{\ell-1}$, we find by direct substitution that

$$F^-_{\ell+1} \, F^+_{\ell-1} \equiv \left(\frac{d}{d\rho} + \frac{\ell+1}{\rho}\right)\left(-\frac{d}{d\rho} + \frac{\ell-1}{\rho}\right)$$

$$= -\frac{d^2}{d\rho^2} - \frac{2}{\rho}\frac{d}{d\rho} + \frac{\ell(\ell-1)}{\rho^2}$$

or

$$F^{-}_{\ell+1}\, F^{+}_{\ell-1} \equiv \mathcal{D}_{\ell-1} \quad . \tag{H. 9}$$

Consider now the fact that

$$\mathcal{D}_{\ell-1}\, f_{\ell-1} = (F^{-}_{\ell+1}\, F^{+}_{\ell-1})\, f_{\ell-1} = f_{\ell-1} \quad ,$$

and operate on both sides of the equation with $F^{+}_{\ell-1}$. This yields

$$F^{+}_{\ell-1}\, (F^{-}_{\ell+1}\, F^{+}_{\ell-1})\, f_{\ell-1} = (F^{+}_{\ell-1}\, f_{\ell-1})$$

or

$$(F^{+}_{\ell-1}\, F^{-}_{\ell+1})\, (F^{+}_{\ell-1}\, f_{\ell-1}) = (F^{+}_{\ell-1}\, f_{\ell-1}) \quad .$$

Using equation H. 8, we have, however, that

$$\mathcal{D}_{\ell}\, (F^{+}_{\ell-1}\, f_{\ell-1}) = (F^{+}_{\ell-1}\, f_{\ell-1}) \quad ,$$

and therefore we see that if $f_{\ell-1}$ is an eigenfunction of $\mathcal{D}_{\ell-1}$, then $F^{+}_{\ell-1}\, f_{\ell-1}$ is an eigenfunction of $\mathcal{D}_{\ell}$. In other words,

$$F^{+}_{\ell-1}\, f_{\ell-1} = f_{\ell} \quad ,$$

or

$$F^{+}_{\ell}\, f_{\ell} = f_{\ell+1} \quad , \tag{H. 10}$$

within a constant. Using equations H. 9 and H. 10 we then also have,

$$F^{-}_{\ell+1}\, F^{+}_{\ell-1}\, f_{\ell-1} = f_{\ell-1}$$

or

$$F^{-}_{\ell+1}\, f_{\ell} = f_{\ell-1} \quad . \tag{H. 11}$$

Using equation H. 10 and the solutions for $\ell = 0$ we can now generate all the other solutions, in particular

$$j_{\ell+1}(\rho) = F^{+}_{\ell}\, j_{\ell}(\rho) \tag{H. 12a}$$

and

$$n_{\ell+1}(\rho) = F^{+}_{\ell}\, n_{\ell}(\rho) \quad . \tag{H. 12b}$$

The first few of these are given by

$$j_1(\rho) = \frac{1}{\rho^2} \sin \rho - \frac{1}{\rho} \cos \rho \quad , \tag{H. 13a}$$

$$j_2(\rho) = \left(\frac{3}{\rho^3} - \frac{1}{\rho}\right) \sin \rho - \frac{3}{\rho^2} \cos \rho \quad , \tag{H. 13b}$$

and

$$n_1(\rho) = \frac{1}{\rho^2} \cos \rho + \frac{1}{\rho} \sin \rho \quad , \tag{H. 14a}$$

$$n_2(\rho) = \left(\frac{3}{\rho^3} - \frac{1}{\rho}\right) \cos \rho + \frac{3}{\rho^2} \sin \rho \quad . \tag{H. 14b}$$

The functions $j_{\ell}(\rho)$ and $n_{\ell}(\rho)$ can also be expressed in terms of the cylindrical Bessel functions of half-integer order, through the equations

$$j_{\ell}(\rho) = \left(\frac{\pi}{2\rho}\right)^{1/2} J_{\ell+1/2}(\rho) \quad , \quad n_{\ell}(\rho) = (-1)^{\ell} \left(\frac{\pi}{2\rho}\right)^{1/2} J_{-(\ell+1/2)}(\rho) \quad . \tag{H. 15}$$

They possess an asymptotic behavior in the vicinity of the origin, given by $(\rho \leqslant \ell)$

$$j_\ell \sim \frac{\rho^\ell}{(2\ell+1)!!}\left[1 - \frac{\rho^2}{2(2\ell+3)} + \ldots\right] \tag{H. 16a}$$

$$n_\ell \sim \frac{(2\ell+1)!!}{(2\ell+1)}\,\rho^{-(\ell+1)}\left[1 + \frac{\rho^2}{2(2\ell-1)} + \ldots\right] , \tag{H. 16b}$$

and become oscillatory for large ρ. In particular, for $\rho \geqslant \ell(\ell+1)$,

$$\begin{pmatrix} j_\ell \\ n_\ell \end{pmatrix} \sim \frac{1}{\rho}\begin{pmatrix} \sin \\ \cos \end{pmatrix}(\rho - \ell\pi/2) \quad . \tag{H. 17}$$

For problems involving radial travelling wave solutions, it is often convenient to introduce the spherical Hankel functions, defined by

$$h_\ell^{(+)} = n_\ell + i j_\ell \tag{H. 18a}$$

and

$$h_\ell^{(-)} = n_\ell - i j_\ell \quad , \tag{H. 18b}$$

which, of course, are also solutions of the original equation. The first few of these are given by

$$\left.\begin{aligned} h_0^{(\pm)} &= \frac{1}{\rho}\, e^{\pm i\rho} \\ h_1^{(\pm)} &= (\frac{1}{\rho} \mp i)\,\frac{e^{\pm i\rho}}{\rho} \\ h_2^{(\pm)} &= (\frac{3}{\rho^2} \mp \frac{3i}{\rho} - 1)\,\frac{e^{\pm i\rho}}{\rho} \quad . \end{aligned}\right\} \tag{H. 19}$$

From the defining equations, it can be seen that

$$n_\ell = \frac{1}{\rho}\,(p_\ell \cos\rho - q_\ell \sin\rho) \tag{H. 20a}$$

and

$$j_\ell = \frac{1}{\rho}\,(q_\ell \cos\rho + p_\ell \sin\rho) \quad , \tag{H. 20b}$$

where p_ℓ and q_ℓ are polynomials of even or odd powers in ρ^{-1}, of degree ℓ and $\ell-1$ respectively. Substitution of these expressions in equation H. 18 then yields

$$h_\ell^{(\pm)}(\rho) = (p_\ell \pm i q_\ell)\,\frac{e^{\pm i\rho}}{\rho} \quad . \tag{H. 21}$$

It can be seen that for real ρ, $j_\ell(\rho)$ and $n_\ell(\rho)$ are real, whereas

$$h_\ell^{(-)}(\rho) = [h_\ell^{(+)}(\rho)]^* \quad . \tag{H. 22}$$

The asymptotic behavior of the spherical Hankel functions, for large ρ is given by

$$h_\ell^{(\pm)}(\rho) \sim [1 \pm i\,\frac{\ell(\ell+1)}{2\rho}]\,\frac{e^{\pm i(\rho - \ell\pi/2)}}{\rho} \quad . \tag{H. 23}$$

It should also be mentioned that these functions also solve the eigenvalue equation H. 1a with the opposite sign, if we replace ρ with $i\rho$ in the solutions, i. e.

$$\mathcal{D}_\ell\, f_\ell(i\rho) = -\, f_\ell(i\rho) \tag{H. 24}$$

or

$$\left[\frac{1}{\rho^2}\frac{d}{d\rho}\left(\rho^2 \frac{d}{d\rho}\right) - 1 - \frac{\ell(\ell+1)}{\rho^2}\right] f_\ell(i\rho) = 0 \quad . \tag{H. 24'}$$

A word of caution! The various arbitrary signs and phases for the spherical Bessel functions are not universal. The choices in these lectures were motivated by the goal of making the analogy between the solutions of the one dimensional problem [sin, cos & $e^{\pm ikx}$] and the radial problem [j_ℓ, n_ℓ & $h_\ell^{(\pm)}$] as explicit as possible. In most references, the spherical Neumann function $n_\ell(\rho)$ is defined with the opposite sign, while the spherical Hankel functions of the first and second kind are related to the $h_\ell^{(\pm)}$ through the equations

$$h_\ell^{(1)} = -i h_\ell^{(+)} \quad , \qquad h_\ell^{(2)} = i h_\ell^{(-)} \quad . \tag{H. 25}$$

Other useful relations

$$\frac{1}{\rho}(2\ell+1) f_\ell = f_{\ell-1} + f_{\ell+1} \tag{H. 26a}$$

$$(2\ell+1)\frac{d}{d\rho} f_\ell = \ell f_{\ell-1} - (\ell+1) f_{\ell+1} \tag{H. 26b}$$

$$\frac{d}{d\rho}(\rho^{\ell+1} f_\ell) = \rho^{\ell+1} f_{\ell-1} \tag{H. 26c}$$

$$\frac{d}{d\rho}(\rho^{-\ell} f_\ell) = -\rho^{-\ell} f_{\ell+1} \tag{H. 26d}$$

$$\left(\frac{d}{d\rho} + \frac{\ell+1}{\rho}\right) f_\ell \equiv F_{\ell+1}^- f_\ell = f_{\ell-1} \tag{H. 26e}$$

$$\left(-\frac{d}{d\rho} + \frac{\ell}{\rho}\right) f_\ell \equiv F_\ell^+ f_\ell = f_{\ell+1} \tag{H. 26f}$$

$$\frac{1}{\rho^\ell} f_\ell = \left(-\frac{1}{\rho}\frac{d}{d\rho}\right)^\ell f_0 \tag{H. 26g}$$

$$n_\ell j_{\ell+1} - n_{\ell+1} j_\ell = \frac{1}{\rho^2} \tag{H. 26h}$$

$$j_\ell n'_\ell - n_\ell j'_\ell = \frac{1}{\rho^2} \tag{H. 26i}$$

APPENDIX I. The auxiliary operator Q in the hydrogen polarizability calculation.

Substituting

$$H_0 = -\frac{\hbar^2}{2m_r}\nabla^2 - \frac{e^2}{4\pi\epsilon_0 r}$$

we have, if $Q = Q(\underline{x})$ only,

$$QH_0 - H_0Q \equiv [Q, H_0] = -\frac{\hbar^2}{2m_r}[Q, \nabla^2] .$$

Substituting in equation 8.58, we see that the operator Q must satisfy the differential equation

$$-\frac{\hbar^2}{2m_r}[Q\nabla^2\psi_0 - \nabla^2(Q\psi_0)] = z\psi_0 , \tag{I.1}$$

where ψ_0 is the ground state hydrogen atom wavefunction. By dimensional analysis we must have

$$Q = \frac{m_r a_0^3}{\hbar^2}\, g , \tag{I.2}$$

where g is a dimensionless function. Substituting I.2 into I.1 we then have

$$-\frac{a_0^2}{2}[g\nabla^2\psi_0 - \nabla^2(g\psi_0)] = \left(\frac{z}{a_0}\right)\psi_0 . \tag{I.3}$$

Now

$$\nabla^2(g\psi_0) = g\nabla^2\psi_0 + 2(\nabla g)\cdot(\nabla\psi_0) + \psi_0\nabla^2 g$$

so that I.3 can be written as

$$a_0^2[(\nabla g)\cdot(\nabla\psi_0) + \frac{1}{2}\psi_0\nabla^2 g] = \left(\frac{r}{a_0}\right)\cos\theta\,\psi_0 . \tag{I.4}$$

Substituting

$$\psi_0 \propto e^{-r/a_0}$$

we have

$$\nabla\psi_0 \propto -\frac{1}{a_0}e^{-r/a_0}\,\hat{e}_r$$

and therefore

$$\frac{a_0^2}{2}\nabla^2 g - a_0\frac{\partial g}{\partial r} = \left(\frac{r}{a_0}\right)\cos\theta . \tag{I.5}$$

By separation of variables or otherwise, we can see that

$$g = f(\rho)\cos\theta$$

where $\rho = r/a_0$ and

$$\frac{1}{2\rho^2}(\rho^2 f')' - f' = \rho ,$$

where a prime denotes differentiation with respect to ρ, or

$$\frac{1}{2}f'' + \left(\frac{1}{\rho} - 1\right)f' - \frac{1}{\rho^2}f = \rho . \tag{I.6}$$

The solution of equation I.6 which leaves I.1 finite as $\rho \to 0$ and $\rho \to \infty$ is given by

$$f(\rho) = -(\rho + \frac{1}{2}\rho^2) ,$$

so that

$$Q = \frac{m_r a_0^3}{\hbar^2} g = - \frac{m_r a_0^3}{\hbar^2} [1 + \frac{1}{2}(\frac{r}{a_0})] (\frac{r}{a_0}) \cos\theta$$

or

$$Q = - \frac{m_r a_0}{\hbar^2} (a_0 + \frac{1}{2} r) z ,$$

as required by equation 8.57.

Reference: Landau and Lifshitz, Quantum Mechanics, section 76, problem 4.

APPENDIX J: The Radiation Gauge.

Using the transformation equations,

$$\underline{A}'(\underline{x}, t) = \underline{A}(\underline{x}, t) + \frac{\partial}{\partial \underline{x}} \chi(\underline{x}, t) \tag{J. 1a}$$

$$\varphi'(\underline{x}, t) = \varphi(\underline{x}, t) - \frac{\partial}{\partial t} \chi(\underline{x}, t) \tag{J. 1b}$$

and the choice

$$\chi(\underline{x}, t) = \int_0^t \varphi(\underline{x}, t')\,dt' , \tag{J. 2}$$

we have

$$\underline{A}'(\underline{x}, t) = \underline{A}(\underline{x}, t) + \int_0^t \left[\frac{\partial}{\partial \underline{x}} \varphi(\underline{x}, t')\right] dt' \tag{J. 3a}$$

and

$$\varphi'(\underline{x}, t) = 0 . \tag{J. 3b}$$

We now transform the fields again, using

$$\chi'(\underline{x}, t) = -\frac{1}{4\pi}\int \left[\frac{\partial}{\partial \underline{x}} \cdot \underline{A}'(\underline{x}', t)\right] \frac{d^3\underline{x}'}{|\underline{x}' - \underline{x}|} . \tag{J. 4}$$

This results in a new vector potential given by

$$\underline{A}''(\underline{x}, t) = \underline{A}'(\underline{x}, t) + \frac{\partial}{\partial \underline{x}} \chi'(\underline{x}, t) , \tag{J. 5}$$

whose divergence is zero, since

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{A}''(\underline{x}, t) = \frac{\partial}{\partial \underline{x}} \cdot \underline{A}'(\underline{x}, t) + \left|\frac{\partial}{\partial \underline{x}}\right|^2 \chi'(\underline{x}, t)$$

and

$$\left|\frac{\partial}{\partial \underline{x}}\right|^2 \chi'(\underline{x}, t) = -\frac{\partial}{\partial \underline{x}} \cdot \underline{A}'(\underline{x}, t) , \tag{J. 6}$$

since equation J. 4 is the solution of equation J. 6 (inhomogeneous Laplace's equation) by construction.

The new scalar potential

$$\varphi''(\underline{x}, t) = -\frac{\partial}{\partial t} \chi'(\underline{x}, t)$$

is also zero since, from Gauss' law

$$\frac{\partial}{\partial \underline{x}} \cdot \underline{\mathcal{E}}(\underline{x}, t) = -\left|\frac{\partial}{\partial \underline{x}}\right|^2 \cancelto{0}{\varphi'(\underline{x}, t)} - \frac{\partial}{\partial t}\left[\frac{\partial}{\partial \underline{x}} \cdot \underline{A}'(\underline{x}, t)\right] = 0 ,$$

we have that $\chi'(\underline{x}, t)$ is not, in fact a function of time, i. e.,

$$\chi'(\underline{x}, t) = \chi'(\underline{x})$$

and therefore $\varphi''(\underline{x}, t) \equiv 0$.

Reference: J. D. Bjorken and S. D. Drell, Relativistic Quantum Fields (McGraw Hill, 1965), §14.2.