Quantum Notes

Kevin E. Schmidt
Department of Physics
Arizona State University
Tempe, Arizona 85287

March 31, 2015

This is a first stab at putting some quantum notes together. The text I use is Baym[1].

Contents

List of Figures

Chapter 1

Dirac Notation

1.1 Introduction

Dirac notation is a convenient way to describe Hilbert space and to transform between bases. However, since quantum mechanics has some complications, we will begin by rewriting standard vectors in Dirac notation. A vector can be written in a coordinate free notation as $\boldsymbol{r}$. This would tell you a position in space given an origin[1]. The definition $\boldsymbol{r}$ gives us a position whose specification is independent of the coordinate system. This position does not require a coordinate system to define it, however, to do calculations, it is usually necessary to define a coordinate system and describe the vector in those coordinates.

A quantum state is described by a ray in complex Hilbert space. Usually we pick a phase convention and normalization so that we describe a state by a particular unit vector in Hilbert space which we write in Dirac notation as $|\psi\rangle$.

1.2 Dirac Notation for coordinate vectors

We could use Dirac notation for our coordinates, and write the coordinate vector as $|r\rangle$. To define a coordinate system, we need a set of basis vectors which are the unit vectors along our coordinate axes. These are often written $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, $\hat{\boldsymbol{z}}$ or $\hat{\boldsymbol{e}}_1$ $\hat{\boldsymbol{e}}_2$, $\hat{\boldsymbol{e}}_3$. The coordinates of $\boldsymbol{r}$ are given by the dot products

$$
\begin{aligned}
x \text{ coordinate} &= \hat{\boldsymbol{x}} \cdot \boldsymbol{r} = \hat{\boldsymbol{e}}_1 \cdot \boldsymbol{r} \\
y \text{ coordinate} &= \hat{\boldsymbol{y}} \cdot \boldsymbol{r} = \hat{\boldsymbol{e}}_2 \cdot \boldsymbol{r} \\
z \text{ coordinate} &= \hat{\boldsymbol{z}} \cdot \boldsymbol{r} = \hat{\boldsymbol{e}}_3 \cdot \boldsymbol{r} \,.
\end{aligned}
\tag{1.1}
$$

Notice that coordinates are numbers. You can put them into a computer program. The vectors and unit vectors are concepts you keep in your mind when you write the program.

---

[1]The name vector, for other quantities, tells you that under coordinate transformations they transform like these positions.

In Dirac notation we could write the coordinate unit vectors and coordinates in an analogous way

$$\begin{aligned}
\hat{\boldsymbol{e}}_1 &\rightarrow |1\rangle \\
\hat{\boldsymbol{e}}_1 \cdot \boldsymbol{r} &\rightarrow \langle 1|r\rangle .
\end{aligned} \tag{1.2}$$

The Dirac bracket $\langle 1|r\rangle$ is like the dot product of the two vectors. If one of the vectors is a unit vector, the dot product gives the coordinate of the other vector along that unit vector. Our coordinate vectors form an orthonormal set

$$\hat{\boldsymbol{e}}_i \cdot \hat{\boldsymbol{e}}_j = \delta_{ij} \tag{1.3}$$

where δ_{ij} is the Kronecker delta which is equal to 1 if $i = j$ and is zero otherwise. In Dirac notation we would write

$$\langle i|j\rangle = \delta_{ij} . \tag{1.4}$$

Notice that the inner product of the basis vectors must be defined by you. Once defined, Dirac notation is often much clearer than standard vector notation (at least to me). For example, writing out a vector in terms of the coordinates and basis vectors becomes

$$\begin{aligned}
\boldsymbol{r} &= \hat{\boldsymbol{e}}_1(\hat{\boldsymbol{e}}_1 \cdot \boldsymbol{r}) + \hat{\boldsymbol{e}}_2(\hat{\boldsymbol{e}}_2 \cdot \boldsymbol{r}) + \hat{\boldsymbol{e}}_3(\hat{\boldsymbol{e}}_3 \cdot \boldsymbol{r}) \\
|r\rangle &= |1\rangle\langle 1|r\rangle + |2\rangle\langle 2|r\rangle + |3\rangle\langle 3|r\rangle .
\end{aligned} \tag{1.5}$$

These expressions show that any vector can be written as a linear combination of the basis vectors. In general we can write this as

$$|r\rangle = \sum_j |j\rangle\langle j|r\rangle . \tag{1.6}$$

Since this is true for any $\boldsymbol{r}$, we see that

$$\sum_j |j\rangle\langle j| = \text{Identity Operator} = 1 . \tag{1.7}$$

This is known as the completeness relation. The 1 on the right is really the identity operator. Sometimes the identity operator is written with a different symbol such as $\mathbb{I}$. This completeness relation is true for any orthonormal set of vectors that completely specifies the coordinate space. In quantum mechanics it will be true for any orthonormal set of basis states that completely describes all possible states of the physical system. One way to think of the completeness relation is that it calculates the part of a vector that is along $|j\rangle$ and puts it along $|j\rangle$.

In Dirac notation to get a number you need a matrix element like $\langle j|r\rangle$. Things like $|i\rangle$ and $|r\rangle$ are not numbers. Just as $\boldsymbol{r}$ and $\hat{\boldsymbol{x}}$ are not numbers.[2]

---

[2]We will see in a moment, when someone writes a wave function e.g. $\psi(\boldsymbol{r})$, that since given a position $\boldsymbol{r}$, the wave function has a value (i.e. a number) that the wave function is a matrix element (or a set of matrix elements, one for each $\boldsymbol{r}$ value). It is not a state vector. We will see that this wave function gives the coordinates of the state vector along a set of position state basis vectors.

Sometimes we are a little sloppy and write things like

$$|r\rangle = \begin{pmatrix} \langle 1|r\rangle \\ \langle 2|r\rangle \\ \langle 3|r\rangle \end{pmatrix} = \begin{pmatrix} x \text{ coordinate} \\ y \text{ coordinate} \\ z \text{ coordinate} \end{pmatrix} \tag{1.8}$$

which is fine as long as you realize that the column vector needs to have the same meaning as our other vector expressions – the corresponding coordinates must be multiplied by their basis vectors.

The advantages of Dirac Notation are that it is compact, and changing to a different basis is easy. Let's look at a coordinate rotation. Assume we know the coordinates of r in one frame, $\langle 1|r\rangle$, $\langle 2|r\rangle$, $\langle 3|r\rangle$. What are its coordinates in a frame rotated by an angle θ around the z axis? Let's call this rotated frame the primed frame with unit vectors $|1'\rangle$, $|2'\rangle$ and $|3'\rangle$. The coordinates in this frame are $\langle i'|r\rangle$. Inserting the complete set $1 = \sum_j |j\rangle\langle j|$ we can write

$$\langle i'|r\rangle = \langle i'| \left(\sum_j |j\rangle\langle j| \right) |r\rangle = \sum_j \langle i'|j\rangle\langle j|r\rangle\,. \tag{1.9}$$

In matrix notation, Eq. 1.9 is

$$\begin{pmatrix} \langle 1'|r\rangle \\ \langle 2'|r\rangle \\ \langle 3'|r\rangle \end{pmatrix} = \begin{pmatrix} \langle 1'|1\rangle & \langle 1'|2\rangle & \langle 1'|3\rangle \\ \langle 2'|1\rangle & \langle 2'|2\rangle & \langle 2'|3\rangle \\ \langle 3'|1\rangle & \langle 3'|2\rangle & \langle 3'|3\rangle \end{pmatrix} \begin{pmatrix} \langle 1|r\rangle \\ \langle 2|r\rangle \\ \langle 3|r\rangle \end{pmatrix}\,. \tag{1.10}$$

You should make sure you can go back and forth between these notations.

If we write Eq. 1.9 in the usual component notation we might write the vector components in the rotated coordinate system as r_i', the components in the original coordinate system as r_i, and the rotation matrix R_{ij}, so that

$$r_i' = \sum_j R_{ij} r_j\,. \tag{1.11}$$

Comparing with Eq. 1.9 shows that the rotation matrix is just $R_{ij} = \langle i'|j\rangle$. This matrix element is the dot product of the basis vectors in the rotated coordinate system with those in the original system.

Dirac notation doesn't do the calculation for you. You still have to work out these matrix elements. In figure 1.1 we have a new coordinate system defined by rotating the original coordinate system by an angle θ around the z axis. From the figure we can write

$$\begin{aligned} |1'\rangle &= \cos\theta|1\rangle + \sin\theta|2\rangle \\ |2'\rangle &= -\sin\theta|1\rangle + \cos\theta|2\rangle \\ |3'\rangle &= |3\rangle\,. \end{aligned} \tag{1.12}$$

The rotation matrix is

$$R = \begin{pmatrix} \langle 1'|1\rangle = \cos\theta & \langle 1'|2\rangle = \sin\theta & \langle 1'|3\rangle = 0 \\ \langle 2'|1\rangle = -\sin\theta & \langle 2'|2\rangle = \cos\theta & \langle 2'|3\rangle = 0 \\ \langle 3'|1\rangle = 0 & \langle 3'|2\rangle = 0 & \langle 3'|3\rangle = 1 \end{pmatrix} \tag{1.13}$$

3
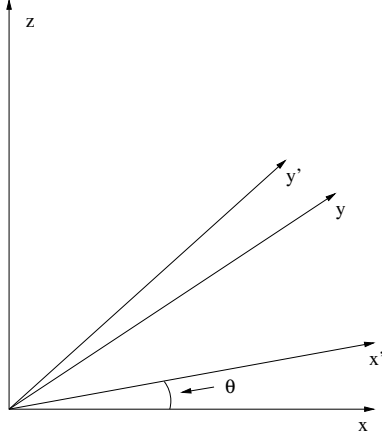
Figure 1.1: A new coordinate system defined by rotating the original coordinates around the z axis by an angle θ.

Here I have talked about rotating the coordinate axes. Notice that if I keep the same coordinate axes and rotate the vector in the opposite direction, I will have the same relationship between the components of the original and rotated vector as I did between the components in the original coordinate system and the rotated (in the opposite direction) coordinate system. We have exactly the same possibilities in quantum mechanics. We can transform the basis, or we can inversely transform our state: both cases give the same matrix elements, and the same mathematics. In quantum mechanics, you must keep track of which transformation you are doing.

1.3   Quantum Mechanics generalizations

The main differences between the coordinate system example and quantum mechanics are that the number of basis states (i.e. coordinate axes) usually is large or infinite, and the "coordinates" that is the matrix elements are complex. Further, there is a difference between the state vector $\langle\psi|$ and the state vector $|\psi\rangle$, although they both describe the same system, and we can easily relate one to the other.

For coordinate systems we get this difference and more if we use nonorthogonal coordinates. That is if the $\boldsymbol{e}_i$ vectors are independent but not orthogonal $\boldsymbol{e}_i \cdot \boldsymbol{e}_j \neq \delta_{ij}$. In that case to expand a general vector we can still write

$$\boldsymbol{r} = r_1\boldsymbol{e}_1 + r_2\boldsymbol{e}_2 + r_3\boldsymbol{e}_3 \,, \tag{1.14}$$

but $r_1 \neq \boldsymbol{e}_1 \cdot \boldsymbol{r}$. Instead to pull out the r_1 component, we need a new vector which when dotted with $\boldsymbol{e}_1$ gives one, and when dotted with $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ gives zero. I hope it is obvious that this vector which we will call $\boldsymbol{e}^1$ must be perpendicular to both $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$, so it is perpendicular to the plane formed by $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$, and has a length such that $\boldsymbol{e}^1 \cdot \boldsymbol{e}_1 = 1$. That

is

$$e^1 = \frac{e_2 \times e_3}{e_1 \cdot (e_2 \times e_3)} \tag{1.15}$$

and similarly for e^2 and e^3. We could then identify e^1 as $\langle 1|$. For coordinates we usually call the two types of vectors covariant and contravariant.

In quantum mechanics, the bra and ket states with the same label are more simply related. The state $\langle \psi|$ is the Hermitian adjoint of the state $|\psi\rangle$. We write this with a dagger notation

$$(|\psi\rangle)^\dagger \equiv \langle \psi|$$
$$((\langle \psi|\phi\rangle))^\dagger = \langle \phi|\psi\rangle = \langle \psi|\phi\rangle^* \tag{1.16}$$

where the asterisk or star means complex conjugate. Together, these show that

$$(\alpha|\phi\rangle + \beta|\psi\rangle)^\dagger = \langle \phi|\alpha^* + \langle \psi|\beta^* . \tag{1.17}$$

This formalism allows us to easily keep track of when we need to take complex conjugates when calculating physical properties.

Transformations between orthonormal bases are linear and unitary.[3] We can see this by writing

$$\begin{aligned}
\delta_{ij} &= \langle i|j\rangle = \sum_k \langle i|k'\rangle\langle k'|j\rangle \\
&= \sum_k (\langle k'|i\rangle)^* \langle k'|j\rangle = \sum_k R_{ki}^* R_{kj} ,
\end{aligned} \tag{1.18}$$

where we have inserted a complete set of states and written the two matrix elements in terms of our transformation matrix $R_{kj} = \langle k'|j\rangle$. The inverse of the matrix R is defined as

$$\delta_{ij} = \sum_k R_{ik}^{-1} R_{kj} \tag{1.19}$$

and comparing we have $R_{ik}^{-1} = R_{ki}^*$. That is, the inverse of R is its transpose conjugate also known as the Hermitian conjugate. This is the definition of a unitary matrix. For the orthogonal coordinate case, the matrix elements are real and the transpose is the inverse making coordinate transformation matrices orthogonal rather than the more general unitary matrices.

1.4   Dirac notation for operators

We mentioned above that rotating a coordinate system gives new coordinates for a vector which are equivalent to those obtained when rotating the vector in the original coordinates

---

[3]As we will see later, this is usually what we want, however symmetry transformations also can be antilinear and antiunitary. Physically, these latter coorespond to reversing the direction of time.

in the opposite sense. That is the same notation allows us to define linear operators that operate on vectors. Similar operators are needed in quantum mechanics. Examples could be operators that rotate or translate states, projection operators that pick out a part of a state (this is what we want to do to represent a measurement, where only the part of a state that corresponds to one or more basis states corresponds to a particular measurement value), time development operators, etc.

Some people like to use a hat notation in text books to distinguish operators from numbers. In most current papers, this notation is not used and whether an object is a number or operator is often inferred from context. Since I use the hat notation for standard unit vectors, I will just use a subscript "op" to indicate an object is an operator. We will drop this when it is clear from context.[4]

Our linear operators, operating on one state, give a linear combination of the basis states

$$O_{\mathrm{op}}|i\rangle = \sum_j O_{ji}|j\rangle \tag{1.20}$$

The coefficients O_{ji} are called the matrix elements of the operator

$$\langle k|O_{\mathrm{op}}|i\rangle = \langle k|\sum_j O_{ji}|j\rangle = \sum_j O_{ji}\langle k|j\rangle = O_{ki} \tag{1.21}$$

Since O_{op} is linear,[5] we have

$$O_{\mathrm{op}}\left(\alpha|i\rangle + \beta|j\rangle\right) = \sum_k |k\rangle\left(\alpha O_{ki} + \beta O_{kj}\right). \tag{1.22}$$

It is convenient to define the Hermitian conjugate or adjoint operator as

$$(O_{\mathrm{op}}|i\rangle)^\dagger \equiv \langle i|O_{\mathrm{op}}^\dagger. \tag{1.23}$$

From this you should verify that

$$\begin{aligned} (\alpha A_{\mathrm{op}} + \beta B_{\mathrm{op}})^\dagger &= \alpha^* A_{\mathrm{op}}^\dagger + \beta^* B_{\mathrm{op}}^\dagger \\ (A_{\mathrm{op}} B_{\mathrm{op}})^\dagger &= B_{\mathrm{op}}^\dagger A_{\mathrm{op}}^\dagger \\ \langle i|O_{\mathrm{op}}|j\rangle &= \langle j|O_{\mathrm{op}}^\dagger|i\rangle^*. \end{aligned} \tag{1.24}$$

The last result shows that the matrix elements of the operator are the transpose complex conjugate of the matrix elements of its adjoint operator. An operator can be defined in terms of the matrix elements

$$O_{\mathrm{op}} = \sum_{jk} |j\rangle\langle j|O|k\rangle\langle k| = \sum_{jk} |j\rangle O_{jk}\langle k| \tag{1.25}$$

---

[4]You will notice in electromagnetism that often the electric field is written $\boldsymbol{E}$ instead of including its space and time arguments $\boldsymbol{E}(\boldsymbol{r}, t)$. This is for much the same reason; writing out the arguments everywhere makes the equations hard to read and can obscure the physics. When the arguments are not clear from context, you need to write them out. Similarly, here, when it may not be clear that an object is an operator it needs to be indicated clearly.

[5]Again it is possible to have antilinear operators where the coefficients get complex conjugates, but only with time reversal.

1.5 Connecting Dirac notation to simple wave functions

Often quantum systems are described by wave functions. Let's look at writing wave functions for one particle. For example we can write the bound energy eigenstates of an electron in a hydrogen atom, as $\psi_k(\boldsymbol{r})$, where k is the state label. For the hydrogen bound states you would normally specify the quantum numbers n, ℓ, and m giving the radial excitation n, and the angular momentum ℓ and m.

First realize that a wave function like this evaluated at some position of the electron is a number. That is a wave function is a matrix element. In this case, we have hydrogenic states, $|\psi_k\rangle$, of the electron. In addition, we have position states of the electron. We write these states as $|\boldsymbol{r}\rangle$. If the electron is in a particular $|\boldsymbol{r}\rangle$ state, then it is at that particular position with probability amplitude 1, and at any other position with probability amplitude 0. The states are then orthogonal $\langle\boldsymbol{r}|\boldsymbol{r}'\rangle = 0$ if $\boldsymbol{r}' \neq \boldsymbol{r}$.[6] Since the possible position values are continuous, it is convenient to normalize these states so that the completeness relation is given by the integral over the states,

$$\int d^3r\, |\boldsymbol{r}\rangle\langle\boldsymbol{r}| = 1 \,. \tag{1.27}$$

If we write the state $|\boldsymbol{r}'\rangle$ using the completeness relation, we have

$$|\boldsymbol{r}'\rangle = \int d^3r\, |\boldsymbol{r}\rangle\langle\boldsymbol{r}|\boldsymbol{r}'\rangle \,. \tag{1.28}$$

Since we know that the matrix element is zero except when $\boldsymbol{r} = \boldsymbol{r}'$, this is

$$|\boldsymbol{r}'\rangle = \int d^3r\, |\boldsymbol{r}'\rangle\langle\boldsymbol{r}|\boldsymbol{r}'\rangle \tag{1.29}$$

or

$$1 = \int d^3r\, \langle\boldsymbol{r}|\boldsymbol{r}'\rangle \,. \tag{1.30}$$

The function that is zero when $\boldsymbol{r} \neq \boldsymbol{r}'$ and integrates to one is

$$\langle\boldsymbol{r}|\boldsymbol{r}'\rangle = \delta^3(\boldsymbol{r} - \boldsymbol{r}') = \delta(x - x')\delta(y - y')\delta(z - z') \tag{1.31}$$

the Dirac delta function. This or similar delta function normalizations are usually more convenient to use for continuous quantum numbers.

The usual wave function is the amplitude for finding a particle in the state $|\psi_k\rangle$ at position $\boldsymbol{r}$, that is the overlap of these two states,

$$\langle\boldsymbol{r}|\psi_k\rangle = \psi_k(\boldsymbol{r}) \,. \tag{1.32}$$

---

[6]An equivalent way to define these states is that they are the eigenstates of the position operator $\boldsymbol{r}_{\mathrm{op}}$, with eigenvalue $\boldsymbol{r}$,

$$\boldsymbol{r}_{\mathrm{op}}|\boldsymbol{r}\rangle = \boldsymbol{r}|\boldsymbol{r}\rangle \,. \tag{1.26}$$

Since they are eigenstates of an Hermitian operator they are orthogonal.

We might want to calculate the overlap $\langle \psi_1 | \psi_2 \rangle$. If we knew the wave functions, we could insert a complete set of $| \boldsymbol{r} \rangle$ states

$$\langle \psi_1 | \psi_2 \rangle = \int d^3 r \langle \psi_1 | \boldsymbol{r} \rangle \langle \boldsymbol{r} | \psi_2 \rangle = \int d^3 r \, \psi_1^*(\boldsymbol{r}) \psi_2(\boldsymbol{r}) \,. \tag{1.33}$$

This is the usual calculation with wave functions. However, Dirac notation shows immediately that you can calculate this overlap in any basis.

1.6 Problems

Problem 1.1. Below A is an operator on the Hilbert space containing $|\psi\rangle$, and α is a complex number. Of the expressions below (a) through (r) below identify which, if any, are nonsense, and, for those, explain why they are nonsense:

    (a) $A|\psi\rangle$,   (b) $\langle\psi|A$,   (c) $A\langle\psi|$,   (d) $|\psi\rangle A$,   (e) $|\psi\rangle\langle\psi|A$,   (f) $|\psi\rangle A\langle\psi|$,

    (g) $A^\dagger|\psi\rangle$,   (h) $\langle\psi|A^\dagger$,   (i) $A^\dagger\langle\psi|$,   (j) $|\psi\rangle A^\dagger$,   (k) $|\psi\rangle\langle\psi|A^\dagger$,   (l) $|\psi\rangle A^\dagger\langle\psi|$,

    (m) $\alpha|\psi\rangle$,   (n) $\langle\psi|\alpha$,   (o) $\alpha\langle\psi|$,   (p) $|\psi\rangle\alpha$,   (q) $|\psi\rangle\langle\psi|\alpha$,   (r) $|\psi\rangle\alpha\langle\psi|$,

Problem 1.2. A, B, C, D are operators and $|\psi\rangle$ is a ket in Hilbert space.

(a) Write $(A|\psi\rangle)^\dagger$ in terms of $A^\dagger$ and $\langle\psi|$.

(b) Write $(ABCD)^\dagger$ in terms of $A^\dagger$, $B^\dagger$, $C^\dagger$ and $D^\dagger$.

Chapter 2

Schrödinger equation and two-state systems

2.1 Schrödinger equation and picture

There are several equivalent ways of formulating quantum mechanics. We will begin in the Schrödinger picture or representation. In this picture the state of this system evolves in time. The equation of motion for the state is called the Schrödinger equation

$$H|\psi(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi(t)\rangle \tag{2.1}$$

where H is a linear operator called the Hamiltonian. Notice that this equation just says that, apart from some constants that are conventional, that the time derivative of a state is given by a linear operator on that state. Notice also that since the constant $\hbar$ has units of energy-time, that H has units of energy. In fact, in cases where there is an analogous classical system, the H operator will usually be derivable from the classical Hamiltonian by replacing the classical position and momentum variables with the corresponding quantum operators.

Since the normalization of the state has no physical consequences, we choose a normalization convention. Here we take the usual convention $\langle\psi(t)|\psi(t)\rangle = 1$. Since the derivative of a constant is zero,

$$
\begin{aligned}
0 &= \frac{\partial}{\partial t}\langle\psi(t)|\psi(t)\rangle = \left[\frac{\partial}{\partial t}\langle\psi(t)|\right]|\psi(t)\rangle + \langle\psi(t)|\left[\frac{\partial}{\partial t}|\psi(t)\rangle\right] \\
&= \frac{i}{\hbar}\langle\psi(t)|\left[H^\dagger - H\right]|\psi(t)\rangle
\end{aligned}
\tag{2.2}
$$

since this must be true for any state $|\psi(t)\rangle$, we must have

$$H^\dagger = H\,. \tag{2.3}$$

Such operators are called Hermitian, since they are equal to their Hermitian conjugate (or adjoint). The Hamiltonian is Hermitian. The matrix elements of a Hermitian operator in

any basis satisfy

$$\langle i|H|j\rangle = \langle j|H|i\rangle^*$$ (2.4)

so they are represented by Hermitian matrices, i.e. matrices that are equal to their complex conjugate transpose.

2.2 One-state systems are boring

Let's imagine that we have an isolated quantum system whose Hilbert space is spanned by one basis state. Let's call that basis state $|0\rangle$, with $\langle 0|0\rangle = 1$. Our Hamiltonian can only have one matrix element $\langle 0|H|0\rangle \equiv E_0$, which we have equated to the number E_0. Notice that Eq. 2.4 says that E_0 is real.

In fact,

$$H_{\text{op}}|0\rangle = E_0|0\rangle .$$ (2.5)

where we have written explicitly that H is an operator and E_0 is a (real) number. When operating on a state with an operator gives the same state back multiplied by a number, the state is an eigenstate of the operator, and the number is the eigenvalue. We will find these eigenstates and eigenvalues extremely useful in our study of quantum mechanics.

Since we have only one basis state, the only possible solution is

$$|\psi(t)\rangle = a(t)|0\rangle .$$ (2.6)

The normalization condition requires

$$1 = \langle \psi(t)|\psi(t)\rangle = \langle 0|a^*(t)a(t)|0\rangle = |a(t)|^2$$ (2.7)

which means that $a(t) = e^{i\phi(t)}$ with $\phi(t)$ real. Substituting into the Schrödinger equation we find

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}E_0 t}|\psi(0)\rangle = e^{-\frac{i}{\hbar}E_0 t}e^{i\phi(0)}|0\rangle .$$ (2.8)

That is the overall phase of the state changes with time, but nothing else. Since, this phase has no measurable effects, the only thing the system can do is stay forever in its one state, never changing, with no physical consequences. Obviously, the two-state system is the simplest interesting system.

2.3 More basis states

Of course, most isolated quantum systems that we want to study have a huge number of possible states. Even simple systems are far harder to solve than electromagnetic or fluid mechanical problems because of the large number of degrees of freedom. Let's take a simple example of looking at an oxygen atom. Oxygen has an atomic nucleus which, for most atomic physics applications you can approximate as a point particle with charge $8e$, and

a mass of about 16 atomic mass units. By making this approximation, we already have removed the proton and neutron degrees of freedom (or at a more elementary level the quark and gluon degrees of freedom), the spin degrees of freedom of the nucleus, its finite size, quadrupole moment, etc. Our approximate oxygen atom is then made up of 8 electrons, and the nucleus. We continue to approximate, working in Coulomb gauge and dropping the photon degrees of freedom, since they give small corrections that can be included later with perturbation theory, and taking a nonrelativistic model for the electrons, we write the interaction as instantaneous Coulomb's law between all pairs of electrons and the electrons and nucleus. Electrons, are spin 1/2 particles, but the nonrelativistic Coulomb interaction does not flip spins, so we can assign spins to the electrons if we like. If we do all this, then to specify a wave function for particular spin assignments requires a function that depends on the coordinates of the 8 electrons and the nucleus. That is, the wave function is a function like

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_8, \boldsymbol{R}, t) \tag{2.9}$$

There are 28 degrees of freedom, and each of the cartesian coordinates can take any value. The Schröedinger equation that describes the time development of the wave function is a 28 dimensional partial differential equation. Let's imagine that you realize that none of the methods you learned in differential equation class will work here. Instead you decide to let a computer solve this. One way would be to make a grid. Let's say you decide that you can get by with a 100 point grid in each dimension. The number of grid points for our 28 degrees of freedom is then 10^{56}. To get an idea of what this number corresponds to, the mass of the earth is 6×10^{27} grams. The mass of the hydrogen atom is about 1 over Avogadro's number in grams, so the number of electrons in the earth is less than 36×10^{50}. If we could store a complex number on one electron, we would still need 10^4 earths to just store this wave function.

Quantum mechanics is hard. Without clever approximations, we have no hope of solving any complicated quantum systems on classical computers.

One way of learning how quantum mechanics works is to look at the opposite extreme of a two-state system.

Sometimes only two states out of the many or infinite number of states for a physical system are important for a particular process. For example, if a system is cooled to a low enough temperature it will be in its lowest energy or ground state. Applying a weak constant electric field to a system that can be approximated by a charged particle moving in a harmonic oscillator potential. Only the ground state and first excited states are important, and the measurements can be described by the results of a two-state system.

Other times, there are clear two-state degrees of freedom. The spin of spin one-half particles can be described by two states (for example spin-up or spin-down), and the polarization of a photon (which is essentially its spin) can be described as a two-state system.

2.4   Photon polarization

A classical electromagnetic wave can be thought of as a state of the electromagnetic field consisting of many photons. Like the field, the photons can carry momentum and energy. They also have a polarization, which corresponds to their spin, and a state with many similar photons will have an electric field polarization that can be readily measured.

Since the free Maxwell's equations are linear and invariant under space and time translations (they contain only derivatives) we can write solutions which have space and time dependence of $e^{i\boldsymbol{k}\cdot\boldsymbol{r}-i\omega t}$ with $\boldsymbol{k}$ and ω real. Obviously these solutions are unphysical since they are complex, and physical fields are real. Since Maxwell's equations are linear and real, we can superpose this mathematical solution with its complex conjugate to obtain a real physical solution, or equivalently take the real part. Taking

$$
\begin{aligned}
\boldsymbol{E}_c(\boldsymbol{r},t) &= \boldsymbol{E}_0 e^{i\boldsymbol{k}\cdot\boldsymbol{r}-i\omega t} \\
\boldsymbol{B}_c(\boldsymbol{r},t) &= \boldsymbol{B}_0 e^{i\boldsymbol{k}\cdot\boldsymbol{r}-i\omega t}
\end{aligned}
\tag{2.10}
$$

where $\boldsymbol{E}_0$ and $\boldsymbol{B}_0$ are constant complex vectors, and substituting into the free Maxwell's equations

$$
\begin{aligned}
\boldsymbol{\nabla} \times \boldsymbol{E}_c &= -\frac{1}{c}\frac{\partial}{\partial t}\boldsymbol{B}_c \; [c]_{\mathrm{SI}} \\
\boldsymbol{\nabla} \times \boldsymbol{B}_c \; [c]_{\mathrm{SI}} &= \frac{1}{c}\frac{\partial}{\partial t}\boldsymbol{E}_c
\end{aligned}
\tag{2.11}
$$

we have

$$
\begin{aligned}
\boldsymbol{k} \times \boldsymbol{E}_0 &= \frac{\omega}{c}\boldsymbol{B}_0 \; [c]_{\mathrm{SI}} \\
\boldsymbol{k} \times \boldsymbol{B}_0 \; [c]_{\mathrm{SI}} &= -\frac{\omega}{c}\boldsymbol{E}_0 \, .
\end{aligned}
\tag{2.12}
$$

Crossing $\boldsymbol{k}$ with the first equation and substituting the second gives

$$
\boldsymbol{k} \times (\boldsymbol{k} \times \boldsymbol{E}_0) = -\frac{\omega^2}{c^2}\boldsymbol{E}_0.
\tag{2.13}
$$

Either working out the components or otherwise, we have

$$
\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = \boldsymbol{B}(\boldsymbol{A}\cdot\boldsymbol{C}) - \boldsymbol{C}(\boldsymbol{A}\cdot\boldsymbol{B})
\tag{2.14}
$$

or

$$
-k^2\left[\boldsymbol{E}_0 - \hat{\boldsymbol{k}}(\hat{\boldsymbol{k}}\cdot\boldsymbol{E}_0)\right] = -\frac{\omega}{c^2}\boldsymbol{E}_0
\tag{2.15}
$$

we see that $\omega = kc$, and $\hat{\boldsymbol{k}}\cdot\boldsymbol{E}_0 = 0$. Substituting back into one of the original equations shows that $\hat{\boldsymbol{k}} \times \boldsymbol{E}_0 = \boldsymbol{B}_0 \; [c]_{\mathrm{SI}}$. So the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields are perpendicular to each other and $\boldsymbol{E} \times \boldsymbol{B}$ points in the direction of propagation.

As noted above, we have mathematical solutions with $\boldsymbol{E}_0$ (and therefore $\boldsymbol{B}_0$) complex, but plugging them into Maxwell's equations as above, and taking the complex conjugate is equal to plugging the complex conjugates of the fields into Maxwell's equations. Therefore the complex conjugate fields are also solutions. Therefore the real (or imaginary) parts of the solutions above are solutions which are real.

We can show that any solution of the free Maxwell equations can be written as a linear combination of the solutions with

$$\begin{aligned} \boldsymbol{E}(\boldsymbol{r},t) &= \operatorname{Re}\boldsymbol{E}_0 e^{i\boldsymbol{k}\cdot\boldsymbol{r}-i\omega t} \\ \boldsymbol{B}(\boldsymbol{r},t)\,[c]_{\mathrm{SI}} &= \boldsymbol{k}\times\boldsymbol{E}(\boldsymbol{r},t) \end{aligned} \tag{2.16}$$

and $\boldsymbol{E}_0$ complex.

Let's take one of these solutions and choose the coordinate system so that $\hat{\boldsymbol{z}} = \hat{\boldsymbol{k}}$. In that case, $\boldsymbol{E}_0 = E_{0x}\hat{\boldsymbol{x}} + E_{0y}\hat{\boldsymbol{y}}$, since $\boldsymbol{E}_0$ has no component along $\boldsymbol{k}$. We can write

$$\boldsymbol{E}_0 = |E_0|\hat{\boldsymbol{E}}_0 \tag{2.17}$$

where

$$|E_0| = \sqrt{E_{0x}^* E_{0x} + E_{0y}^* E_{0y}} = \sqrt{|E_{0x}|^2 + |E_{0y}|^2}\,. \tag{2.18}$$

$\hat{\boldsymbol{E}}_0$ is a normalized vector with complex coordinates. It is called the polarization of the plane wave.

Let's see what the physical meaning of the components of $\hat{\boldsymbol{E}}_0$ are with some examples before writing the general case. A very simple case is $\hat{\boldsymbol{E}}_0 = \hat{\boldsymbol{x}}$. In this case

$$\boldsymbol{E}(\boldsymbol{r},t) = |E_0|\hat{\boldsymbol{x}}\cos\left[\frac{\omega}{c}(z-ct)\right] \tag{2.19}$$

If we measure at $z = 0$ (or any other fixed position), we see that the field oscillates harmonically in the x direction

$$\boldsymbol{E}(0,t) = |E_0|\hat{\boldsymbol{x}}\cos\left(\omega t\right)\,. \tag{2.20}$$

This is called an x polarized beam.

Similarly, if we take $\hat{\boldsymbol{E}}_0 = \hat{\boldsymbol{y}}$, we will have a y polarized beam. Let's take a real linear combination of these two

$$\hat{\boldsymbol{E}}_0 = \cos\theta\hat{\boldsymbol{x}} + \sin\theta\hat{\boldsymbol{y}} \tag{2.21}$$

Plugging into our defining equation for the physical field, you will see that this is again a harmonically oscillating field in the direction of $\hat{\boldsymbol{E}}_0$. All of these choices have produced this same form. These are called linear polarizations, since the field always points along (or opposite) the same direction. Notice that we could have obtained this form by rotating the x polarized beam through an angle θ.

Since we can have complex $\boldsymbol{E}_0$, let's next look at multiplying the $\hat{\boldsymbol{y}}$ term above by i,

$$\hat{\boldsymbol{E}}_0 = \cos\theta\hat{\boldsymbol{x}} + i\sin\theta\hat{\boldsymbol{y}} \tag{2.22}$$

15

The physical field at $z = 0$ will be

$$\boldsymbol{E}(0,t) = |\boldsymbol{E}_0| \left[\hat{\boldsymbol{x}} \cos\theta \cos(\omega t) + \hat{\boldsymbol{y}} \sin\theta \sin(\omega t) \right] \qquad (2.23)$$

In this case, the two components oscillate out of phase. If we trace out the tip of $\boldsymbol{E}(0,t)$ as a function of time, it will be an ellipse with one axis along $\hat{\boldsymbol{x}}$ and the other along $\hat{\boldsymbol{y}}$. If we take $\theta = \pi/4$, the two axes will be equal, and we trace out a circle. So right circular polarization is

$$\hat{\boldsymbol{E}}_r = \frac{\hat{\boldsymbol{x}} + i\hat{\boldsymbol{y}}}{\sqrt{2}} \qquad (2.24)$$

Going through the same excersize multiplying y by $-i$, you get left circular polarization is

$$\hat{\boldsymbol{E}}_l = \frac{\hat{\boldsymbol{x}} - i\hat{\boldsymbol{y}}}{\sqrt{2}} \; . \qquad (2.25)$$

Notice that $\hat{\boldsymbol{E}}_r \cdot \hat{\boldsymbol{E}}_l \neq 0$. However, $\hat{\boldsymbol{E}}_r^* \cdot \hat{\boldsymbol{E}}_l = 0$. That is, just as we found the length by $\boldsymbol{E}^* \cdot \boldsymbol{E}$, our rule for finding orthogonal polarizations (which comes from calculating the time average in classical electromagnetism) requires that we take the complex conjugate just like in quantum mechanics. It is convenient to use Dirac notation. So we will call $|x\rangle$ x polarized light, $|y\rangle$, y polarized light, $2^{-1/2}(|x\rangle + i|y\rangle)$ right circularly polarized, $2^{-1/2}(|x\rangle - i|y\rangle)$ left circularly polarized, etc.

At this point, lets make the connection to the photon polarization. If we have a beam of unpolarized light and we send it through an ideal polaroid oriented along x, half the beam energy will be stopped by the polaroid, and the other half will emerge as x polarized light. If we now put the beam of x polarized light through a polarizer oriented along the direction in the x-y plane an angle θ from the x axis, the electric field component along that direction, will make it through. That is, if we define the polarization direction of the polaroid to be

$$|\theta\rangle = |x\rangle \cos\theta + |y\rangle \sin\theta \qquad (2.26)$$

the electric field magnitude will be reduced by

$$\langle\theta|x\rangle = \cos\theta \qquad (2.27)$$

and the emerging beam energy (which is proportional to the square of the field) will be given by $|\langle\theta|x\rangle|^2 = \cos^2\theta$ times the incoming energy.

We now assume, correctly, that the original beam can be thought of as a large number of individual photons. After the first x polarizer, all of the photons that emerge are x polarized photons. While we cannot measure all of the properties of an electromagnetic field because of the uncertainty relations, we can measure the photon polarization (and this can be verified experimentally). Now, we take our x polarized photons and look at how many make it through our polaroid tilted at an angle θ. In quantum mechanics, all of the x polarized photons are equivalent. We cannot have part of a photon making it through the polaroid. We instead infer that there is a probability of each x polarized photon making it through

the polaroid. In order for this to agree with the classical result above, this probability has to be $\cos^2 \theta$. Further, we know classically that the emerging beam will be polarized along the direction of the polaroid. For quantum mechanics to agree with this, we require that the polarization state of the photons that make it through are $|\theta\rangle$.

This is an example of a quantum measurement (at least for the photons that make it through) using the Copenhagen interpretation. The probability of making it through (i.e. having polarization along $|\theta\rangle$) for an incoming polarization state $|\psi\rangle$, is $|\langle\theta|\psi\rangle|^2$ and the the measured photons' state has "collapsed" to $|\theta\rangle$ corresponding to that polarization. The use of the behavior of the classical limit to figure out what the quantum probabilities must be, is an example of the correspondence principle. The correspondence principle is unnecessary once you have the correct quantum mechanical equations and rules, but was a great help historically in discovering those equations and rules.

Let's look at the mathematics of the operation of a polarizer in several "languages". We start with an initial photon state polarized along $|\psi\rangle$. For our experiments, we will have many identically prepared copies of this photon.

Our single photon's polarization quantum state can be specified, using, for example, the $|x\rangle$ and $|y\rangle$ polarization bases,

$$|\psi\rangle = \frac{\alpha}{|\alpha|^2 + |\beta|^2}|x\rangle + \frac{\beta}{|\alpha|^2 + |\beta|^2}|y\rangle \tag{2.28}$$

we could also write a "wave function"

$$\psi(p) = \begin{cases} \langle x|\psi\rangle = \frac{\alpha}{|\alpha|^2+|\beta|^2} & p = 1 \\ \langle y|\psi\rangle = \frac{\beta}{|\alpha|^2+|\beta|^2} & p = 2 \end{cases} \tag{2.29}$$

where the value of $\psi(p)$ for the two polarizations 1 and 2 gives the amplitude for being in the $|x\rangle$ and $|y\rangle$ states. We can also write a column vector representation

$$\begin{pmatrix} \frac{\alpha}{|\alpha|^2+|\beta|^2} \\ \frac{\beta}{|\alpha|^2+|\beta|^2} \end{pmatrix} \tag{2.30}$$

to represent the state.

The probability of passing through the x polarizer is given by

$$P_x = |\langle x|\psi\rangle|^2 = \langle\psi|x\rangle\langle x|\psi\rangle = \frac{|\alpha|^2}{|\alpha^2| + |\beta|^2} \tag{2.31}$$

Notice that $|x\rangle\langle x|$ is the operator that projects out the component along $|x\rangle$.

We can also calculate this using the wave function form as

$$P_x = |\psi(1)|^2 = \sum_{p=1}^{2} \psi^*(p)\delta_{p,1}\psi(p) = \frac{|\alpha|^2}{|\alpha^2| + |\beta|^2} \tag{2.32}$$

17

where $\delta_{p,1}$ gives 1 for $p = 1$, i.e. for polarization along x, and 0 for $p = 2$, i.e. polarization along y. It is the x polarization projection operator for the wavefunction.

Finally, we can use the matrix form,

$$P_x = \left| \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\alpha}{|\alpha|^2+|\beta|^2} \\ \frac{\beta}{|\alpha|^2+|\beta|^2} \end{pmatrix} \right|^2 = \begin{pmatrix} \frac{\alpha^*}{|\alpha|^2+|\beta|^2} & \frac{\beta^*}{|\alpha|^2+|\beta|^2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{\alpha}{|\alpha|^2+|\beta|^2} \\ \frac{\beta}{|\alpha|^2+|\beta|^2} \end{pmatrix} = \frac{|\alpha|^2}{|\alpha^2| + |\beta|^2}$$

(2.33)

Each of these three ways of writing the mathematics gives the same result. Each form (and especially their generalizations to more degrees of freedom as well as continuous degrees of freedom) has places where they can simplify the notation and the calculations.

polarization analyzer Calcite - extraordinary-ordinary - can use to separate in space two polarizations, and recombine them.

Two-slit like experiment – polarization interference.

Turpentine - optically active - rotates polarization - find polarizations that don't change (up to an overall phase) -eigenvectors/values.

Problems – set of N polarizers from x photons to y polarizer. Probability $= cos(pi/(2*N))^{(2N)}$. Large N goes to $exp(-Pi^2/4N)$. Plot. Other baymish problems.

Chapter 3

Measurements

You should have already studied the basic mathematics and physics of measurements in quantum mechanics. Here are the rules we will follow for the idealized quantum measurement:

1. Physical observables have a corresponding Hermitian operator O. It has normalized eigenstates and eigenvalues $O|n\rangle = \lambda_n|n\rangle$.

2. For a system in state $|\psi\rangle$, a measurement of the observable quantity gives an eigenvalue λ of the operator O with probability

$$
\begin{aligned}
P(\lambda) &= \sum_{\text{states with eigenvalue } \lambda_m = \lambda} |\langle m|\psi\rangle|^2 \\
&= \sum_{\text{states with eigenvalue } \lambda_m = \lambda} \langle \psi|m\rangle\langle m|\psi\rangle,
\end{aligned}
\tag{3.1}
$$

where I assume that the states $|m\rangle$ are normalized to 1. For other normalizations make sure that the probability of being in some state is unity.

3. Subsequent to the measurement, the results for any following measurement will be the same as if the state is now $|\psi'\rangle$ where

$$
|\psi'\rangle = \frac{\sum_{\text{states with eigenvalue } \lambda_m = \lambda} |m\rangle\langle m|\psi\rangle}{\left[\sum_{\text{states with eigenvalue } \lambda_m = \lambda} \langle \psi|m\rangle\langle m|\psi\rangle\right]^{1/2}}
\tag{3.2}
$$

4. A corollary to the above is that the average of many measurements of an ensemble of systems all prepared identically in the state $|\psi\rangle$ is

$$
\sum_{\text{possible } \lambda \text{ values}} P(\lambda)\lambda = \sum_m \lambda_m \langle \psi|m\rangle\langle m|\psi\rangle = \langle \psi|O|m\rangle\langle m|\psi\rangle = \langle \psi|O|\psi\rangle.
\tag{3.3}
$$

This is the expectation value.

5. Another corollary is that if, in item 2, we define a projection operator $P_O(\lambda)$ onto the eigenstates of O with eigenvalue λ,

$$P_O(\lambda) = \sum_{\text{states with eigenvalue } \lambda_m = \lambda} |m\rangle\langle m| \qquad (3.4)$$

then the probability of measuring the value λ for the operator O is the expectation value of $P_O(\lambda)$

$$P(\lambda) = \langle\psi|P_O(\lambda)|\psi\rangle. \qquad (3.5)$$

Any real measurement of a quantum system is just an application of some particular Hamiltonian to that state. A typical measurement would be like those in Baym Chapter 1, like passing a photon through a polarizer. Instead of dealing with the Hamiltonian and states of the polarizer and the photon, we simplify and say that the states corresponding to the photon being absorbed and exciting the polarizer will have zero probability of ever making a transition back to having a photon again with this new state component being able to interfere with the photon that passes through the polarizer. The portion of Hilbert space corresponding to the absorbed photon is then dropped from our calculation. The Copenhagen interpretation says that the wave function collapses after the measurement. The many-worlds interpretation says the other portions of the wave function correspond to one of the other worlds. Since the difference between these two is not measurable, both interpretations give the same results for all experiments. There are various other interpretations that also yield identical results.

Sometimes people speak of "dephasing" of the state. This means that we have somehow lost the phase information between different parts of the state. An example of how this happens would be to set up the standard two slit interference experiment where an electron passing through both slits gives a set of interference fringes on a detecting screen behind the slits. Now imagine that instead of the slits we had two identical bending magnets that guided the electrons down two equivalent long paths and then back to where the slits were. In a cartoon theoretical description, we could imagine having perfectly identical apparatus for both paths, so the two parts of the electron state vector would maintain their same relative phase. When we add the two components at the screen and square the result we would get an interference pattern as before. The sum of the two amplitudes has a term where the amplitude from the left path is squared, a term where the amplitude from the right path is squared, and two cross terms where the electron has traveled along both paths that lead to the interference fringes. However in a real experiment, the magnetic fields would likely fluctuate, the vacuum would not be perfect and the electron would experience different interactions along the two different paths. The cross term in square of the sum of the amplitudes would then have a different phase. When we repeat the experiment many times, each experiment would be slightly different and we would get many different phases for the cross terms. The fringe positions would then average out, and would not be detectable. In this case, the result is the same as if we had measured which slit the electron went through, even though we do not have that information.

3.1 Mixed states

Our analysis above assumes that we have a system in a quantum state $|\psi\rangle$. In order to measure the probabilities, we need to prepare multiple copies of the system each in the same state $|\psi\rangle$. That is we need an ensemble of identically prepared systems.

Often, instead, we have an ensemble of systems in a distribution of states. For example, if we prepare identical systems in contact with a heat bath at temperature T, statistical mechanics tells us that the probability of the system being in an energy eigenstate $|n\rangle$ with eigenvalue E_n is proportional to the Boltzmann factor $\exp(-E_n/k_B T)$. Therefore the probability of obtaining a particular measurement will be the probability that the system is in a particular quantum state multiplied by the probability of getting that measurement in the quantum state.

Therefore the probability of measuring the value of O being λ becomes

$$P(\lambda) = \sum_i p_i \langle \psi_i | P_O(\lambda) | \psi_i \rangle, \tag{3.6}$$

where p_i is the probability of the copy of the system we are measuring being in the state $|\psi_i\rangle$.

We can think of a pure state as the special case where one state has probability 1 and the rest have probability 0.

Take a moment and make sure that you understand that there is an enormous difference between the two cases:

1. For a pure state which is a linear combination of $|\psi_1\rangle$ and $|\psi_2\rangle$

$$|\psi\rangle = a_1 |\psi_1\rangle + a_2 |\psi_2\rangle \tag{3.7}$$

   the probabiliy of measuring the value of O to be λ would be

$$
\begin{aligned}
P_{\text{pure}}(\lambda) &= \left[a_1^* \langle \psi_1 | + a_2^* \langle \psi_2 | \right] P_O(\lambda) \left[a_1 |\psi_1\rangle + a_2 |\psi_2\rangle \right] \\
&= |a_1|^2 \langle \psi_1 | P_O(\lambda) | \psi_1 \rangle + |a_2|^2 \langle \psi_2 | P_O(\lambda) | \psi_2 \rangle + 2\mathrm{Re} a_1 a_2^* \langle \psi_2 | P_O(\lambda) | \psi_1 \rangle
\end{aligned}
\tag{3.8}
$$

2. For a mixed state system where the probability of being in $|\psi_1\rangle$ is p_1 and the probability of being in $|\psi_2\rangle$ is p_2. the probabiliy of measuring the value of O to be λ would be

$$P_{\text{mixed}}(\lambda) = p_1 \langle \psi_1 | P_O(\lambda) | \psi_1 \rangle + p_2 \langle \psi_2 | P_O(\lambda) | \psi_2 \rangle \tag{3.9}$$

If $|a_1|^2 = p_1$, $|a_2|^2 = p_2$, we see that the pure state contains the interference term $2\mathrm{Re} a_1 a_2^* \langle \psi_2 | P_O(\lambda) | \psi_1 \rangle$ while the mixed state does not. The answers will be completely different.

The difference is analogous to the two-slit experiment. If we have identically prepared particles then when a particle passes through the two slits, it will be in a superposition of the

states where it passes through slit 1 or slit 2. We get the usual two-slit interference pattern from the effect of the interference term. If instead we measure whether the particle passed through slit 1 or slit 2, then after this measurement, we know that half the particles will be in state $|\psi_1\rangle$ and half in $|\psi_2\rangle$. The pattern will just be the superposition of the patterns we would see from each slit separately with no interference pattern.

3.2   Density matrices

Since we can use the mixed state formalism for both mixed and pure states, we write

$$P(\lambda) = \sum_i p_i \langle \psi_i | P_O(\lambda) | \psi_i \rangle \tag{3.10}$$

and now insert a complete set of states $|j\rangle$ and rearrange

$$
\begin{aligned}
P(\lambda) &= \sum_j \sum_i p_i \langle \psi_i | j \rangle \langle j | P_O(\lambda) | \psi_i \rangle \\
&= \sum_j \langle j | P_O(\lambda) | \psi_i \rangle p_i \langle \psi_i | j \rangle \,.
\end{aligned}
\tag{3.11}
$$

By rearranging we have written the trace, which is invariant under basis change since we didn't specify what complete set we used

$$P(\lambda) = \mathrm{tr} P_O \rho \tag{3.12}$$

where the density matrix ρ is defined by

$$\rho = \sum_i |\psi_i\rangle p_i \langle \psi_i | \,. \tag{3.13}$$

Obviously the density matrix for a pure state is

$$\rho_{\text{pure}} = |\psi\rangle\langle\psi| \,. \tag{3.14}$$

Repeating these arguments, we see that the expectation value of any operator O is

$$\langle O \rangle = \mathrm{tr} O \rho \tag{3.15}$$

and since $\sum p_i = 1$ and $\langle \psi_i | \psi_i \rangle = 1$, $\mathrm{tr}\rho = 1$.

Since the results of any measurement of a pure or mixed state system can be obtained from the density matrix, it contains all possible information about the system.

3.3 Eigenvectors and eigenvalues of the density matrix

Often the eigenvectors and values of the density matrix can give useful information about what we can learn from measurements.

Since the density matrix is Hermitian by construction, we know its eigenvalues are real. Further, they are greater than or equal to zero. We can show this by first noting that, the expectation value of ρ for an eigenstate of ρ is the eigenvalue, $\langle\phi_n|\rho|\phi_n\rangle = \lambda_n$. The expectation value for an arbitrary state $|\phi\rangle$, is greater than or equal to zero since

$$\langle\phi|\rho|\phi\rangle \;\; = \;\; \sum_i p_i\langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle = \sum_i p_i\,|\langle\phi|\psi_i\rangle|^2 \geq 0 \tag{3.16}$$

and therefore all the eigenvalues of ρ are greater than or equal to zero.

Inserting the complete set of eigenvectors of ρ gives the usual expansion

$$\rho = \sum_n \rho|\psi_n\rangle\langle\psi_n| = \sum_n \lambda_n|\psi_n\rangle\langle\psi_n|\,, \tag{3.17}$$

with $\lambda_n \geq 0$. Since the trace of ρ is 1, $\sum_n \lambda_n = 1$. This looks like our mixed state form. However, there is a big difference. Here the $|\psi_n\rangle$ form a complete orthogonal set. There is no limit on the number of states in the general mixed state form, and the states do not need to be orthogonal. For example, we can form a beam of photons where half the photons are polarized along x and half are right-circularly polarized. These two states are not orthogonal, but give a well-defined density matrix,

$$\rho = \frac{1}{2}|x\rangle\langle x| + \frac{1}{2}\frac{|x\rangle + i|y\rangle}{\sqrt{2}}\frac{\langle x| - i\langle y|}{\sqrt{2}} \tag{3.18}$$

The corresponding 2x2 matrix

$$\begin{pmatrix} \langle x|\rho|x\rangle & \langle x|\rho|y\rangle \\ \langle y|\rho|x\rangle & \langle y|\rho|y\rangle \end{pmatrix} = \begin{pmatrix} \frac{3}{4} & \frac{-i}{4} \\ \frac{i}{4} & \frac{1}{4} \end{pmatrix} \tag{3.19}$$

has eigenvalues $\lambda_\pm = \frac{1}{2} \pm \frac{\sqrt{2}}{4}$ roughly 0.854 and 0.146, and eigenvectors

$$\frac{1}{\sqrt{\left(\frac{1\pm\sqrt{2}}{4}\right)^2 + \frac{1}{16}}}\left[\frac{1\pm\sqrt{2}}{4}|x\rangle + \frac{i}{4}|y\rangle\right]\,. \tag{3.20}$$

This shows that this mixed state could equally well have been prepared with these two orthogonal polarizations.

3.4 Reduced density matrix

Density matrices often appear when we look at just part of a system. If you imagine the entire universe being just one quantum state, then all of our density matrices are derived this way.

For example if we have an N-particle system in a pure state $|\Psi\rangle$, we could write a wave function as

$$\langle \boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N | \Psi \rangle = \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N) \tag{3.21}$$

where I have assumed that the particles have no spin. You could include spin degrees of freedom as well as positions in a more general case. The density matrix for this pure state is $\rho = |\Psi\rangle\langle\Psi|$. If we have an operator that operates (and therefore measures) just particle 1, O_1, we would write its expectation value as

$$\langle O_1 \rangle = \mathrm{tr} O_1 \rho \,. \tag{3.22}$$

Performing the trace using the position eigenstates, we have

$$\int d^3 r_1 d^3 r_2 ... d^3 r_N \langle \boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N | O_1 | \Psi \rangle \langle \Psi | \boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N \rangle \tag{3.23}$$

The most general form for the matrix elements of O_1 in the position eigenstate basis is

$$O_1 = \int d^3 r_1 d^3 r'_1 d^3 r_2 ... d^3 r_N |\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N \rangle O_1(\boldsymbol{r}_1, \boldsymbol{r}'_1) \langle \boldsymbol{r}'_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N | \tag{3.24}$$

since it can only depend on particle 1 and cannot change the position of the other particles. Plugging in this form, we get

$$\mathrm{tr} O_1 \rho = \int d^3 r_1 d^3 r'_1 O_1(\boldsymbol{r}_1, \boldsymbol{r}'_1) \rho_1(\boldsymbol{r}'_1, \boldsymbol{r}_1) = \mathrm{tr} O_1^{(r)} \rho_1^{(r)} \tag{3.25}$$

where

$$
\begin{aligned}
\rho_1^{(r)} &= \int d^3 r_1 d^3 r'_1 |\boldsymbol{r}'_1\rangle \rho_1(\boldsymbol{r}'_1, \boldsymbol{r}_1) \langle \boldsymbol{r}_1| \\
O_1^{(r)} &= \int d^3 r_1 d^3 r'_1 |\boldsymbol{r}'_1\rangle O(\boldsymbol{r}'_1, \boldsymbol{r}_1) \langle \boldsymbol{r}_1| \\
\rho_1(\boldsymbol{r}'_1, \boldsymbol{r}_1) &= \int d^3 r_2 ... d^3 r_N \Psi^*(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N) \Psi(\boldsymbol{r}'_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N) \,.
\end{aligned} \tag{3.26}
$$

That is, we can define a reduced 1-body density matrix $\rho_1^{(r)}$ which contains all of the information we can possibly obtain by measuring just particle 1. We can find the eigenvectors and eigenvalues of this reduced density matrix. In the r basis, these will look like 1-particle orbitals, and the eigenvalues will be their occupation probabilities. These eigenfunctions of the 1-body density matrix are called the natural orbitals.

Weak scattering probes often can be approximated by assuming that the weak scatterer interacts with just one particle during the scattering. In that case, the response of the system is determined by the one-body density matrix. In other words, in that case, we can use the probe to measure properties of the one-body density matrix such as the natural orbital occupation numbers.

Other bases work the same way. To obtain the reduced density matrix we write a basis which is the outer product of states spanning our desired states (these were the position states for particle 1 in our example above, $|\boldsymbol{r}_1\rangle$) with states spanning the other degrees of freedom of our Hilbert space (these are the $|\boldsymbol{r}_2, ..., \boldsymbol{r}_N\rangle$ states in our example). We then trace over the second set leaving our reduced matrix density matrix.

3.5 Problems

Problem 3.1. Let's examine the two-slit experiment using the reduced density matrix. If the electron is isolated (and we view the slits as just boundary conditions rather than additional interacting particles) we can write the state of the electron after passing through the slits and before being detected on the screen as

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left[|\psi_1\rangle + |\psi_2\rangle \right] \tag{3.27}$$

where $|\psi_1\rangle$ is the state produced from the paths through slit one, and $|\psi_2\rangle$ is the state produced from the paths through slit two. The density matrix for the electron would be

$$\rho = |\psi\rangle\langle\psi| \tag{3.28}$$

Now let's include other parts of our laboratory, e.g. a measuring apparatus. If the electron still does not interact, we can write

$$|\Psi_{\text{total}}\rangle = |\psi\rangle \otimes |\Psi\rangle \tag{3.29}$$

where $|\Psi\rangle$ is a state in the Hilbert space for the rest of the universe.

a. Show that the reduced density matrix produced by tracing over the part of Hilbert space representing everything but the electron gives a reduced density matrix in agreement with the single particle density matrix Eq. 3.28.

b. Now assume that interactions change $|\Psi\rangle$ for paths where the electron goes through slit 1 to $|\Psi'\rangle$ orthogonal to $|\Psi\rangle$, but doesn't change $|\psi_1\rangle$ substantially. Calculate the reduced density matrix for the electron and show that it is the same as we would get if we describe an ensemble of electrons using a mixed state density matrix with the probabilities each one half of preparing the electron in state $|\psi_1\rangle$ or in $|\psi_2\rangle$.

c. Notice that this reduction with interactions that change the state of the rest of the universe occurs whether we would describe the result as a measurement or not. It certainly occurs when we make a measurement since to have something called a measurement we must put the rest of the universe in a different quantum state (usually very different). However, a random collision from the background gas in the chamber where we are making the two-slit interference experiment will also have the same effect without giving us any information. In a real two-slit interference experiment this would give us an experimental noise "background." Assume that such interactions in a well designed experiment change the state to

$$|\Psi_{\text{total}}\rangle = \frac{1}{\sqrt{2}} \left\{ |\psi_1\rangle \otimes [\cos\theta_1 |\Psi\rangle + \sin\theta_1 |\Psi'\rangle] + |\psi_2\rangle \otimes [\cos\theta_2 |\Psi\rangle + \sin\theta_2 |\Psi''\rangle] \right\} . \tag{3.30}$$

where $|\Psi\rangle$, $|\Psi'\rangle$, $|\Psi''\rangle$ are all orthogonal, and θ_1 and θ_2 are small. Calculate the reduced density matrix. Calculate the natural orbitals and their occupation numbers. Show that the reduced density matrix is equivalent to that of a single electron prepared in a mixed state constructed from $|\psi_1\rangle$, $|\psi_2\rangle$ and $|\psi\rangle$, and calculate their probabilities. Assume for convenience that the states $|\psi_1\rangle$ and $|\psi_2\rangle$ here represent wave packets near the slits and therefore are still separated and orthogonal (they will later overlap at the detector to give interference).

Problem 3.2. A particle moving in the x-y plane is constrained to the region $0 < x < L$, and $0 < y < x$, and is in the state whose wave function is

$$\langle xy|\phi\rangle = \phi(x,y) = \frac{2}{L}\left[\sin\left(\frac{\pi x}{L}\right)\sin\left(\frac{2\pi y}{L}\right) - \sin\left(\frac{2\pi x}{L}\right)\sin\left(\frac{\pi y}{L}\right)\right]. \tag{3.31}$$

If we have a probe that can only measure along x, we can calculate a reduced density matrix, $\rho_r(x, x')$, by integrating out y. After performing this integral, choose a basis of orthogonal functions, truncate your basis in a physically reasonable way, and calculate the following parts numerically using any convenient numerical tools. If you prefer, feel free to numerically integrate to get the reduced density matrix numerically.

a. Calculate the occupation numbers for the natural orbitals with occupation numbers greater than 0.01. Your occupation numbers should have at least 4 place accuracy. Verify this by increasing the size of your truncated basis.

b. Plot the $\sqrt{L}$ times the natural orbitals versus x/L for those natural orbitals corresponding to the occupation numbers of part b.

Problem 3.3. Spin $\frac{1}{2}$ is represented by a state vector with two components exactly like the photon polarization discussed in Baym. The standard convention is to write the spin operator in the $|\uparrow\rangle$, $|\downarrow\rangle$ basis as the matrix

$$\boldsymbol{S} = \frac{\hbar}{2}\left(\hat{\boldsymbol{x}}\sigma_x + \hat{\boldsymbol{y}}\sigma_y + \hat{\boldsymbol{z}}\sigma_z\right) \tag{3.32}$$

where the Pauli matrices are

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.33}$$

a. Find the eigenvectors and eigenvalues of these three matrices. Write the eigenvectors of the corresponding operators in the form $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$.

b. A density matrix for the spin of a mixed state of a beam of spin $\frac{1}{2}$ particles can be written as
$$\rho = \sum_j p_j|\psi_j\rangle\langle\psi_j| \tag{3.34}$$

27

where p_j is the probability of finding a particle in the spin state $|\psi_j\rangle$, and the sum runs over all of ways the particles can be prepared. Just as in the photon case, three measurements are needed to determine the density matrix. For example a Stern-Gerlach experiment can measure the probability of having a component $\hbar/2$ of spin along any single direction. Show that measuring the spin component along the three directions x, y, and z in the standard $|\uparrow\rangle$ $|\downarrow\rangle$ representation gives exactly the same mathematics as measuring the photon polarization along x, $45°$ to x, and right circular polarization, in the x-y polarization basis.

Problem 3.4. In the Einstein-Podolsky-Rosen or EPR gendanken experiment, you imagine that a particle decays emitting two back-to-back photons in a net angular momentum zero state. We can represent the polarization of such a state as

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left[|xx\rangle + |yy\rangle \right] \tag{3.35}$$

where the first x or y represents the polarization of photon 1, and the second represents the polarization of photon two. (The choice of relative sign is a phase convention.) The EPR paradox is that after the photons have propagated a great distance apart, measuring the polarization of photon 2 to be along x, according to the Copenhagen convention, means that photon 1 is immediately changed into an x state. This "instantaneous action at a distance" bothered Einstein, Podolsky and Rosen. The purpose of this problem is to show that no measurement on photon 1 can detect whether any measurement has been made on photon 2.

a. Calculate the density matrix for an ensemble of such systems in the pure state $|\psi\rangle$, and trace over the states of photon 2 to produce the reduced density matrix for photon 1.

b. Photon 2 is now measured to find the probability of it having polarization $|\psi_a\rangle$ or the orthogonal polarization $|\psi_b\rangle$. Calculate the density matrix for the ensemble after this measurement, and then trace over the states of photon 2 to produce the reduced density matrix for photon 1.

The two reduced density matrices of parts a and b should be identical showing that the measurement of photon 2 does not change the results for photon 1 and there is no measurable weird "instantaneous action at a distance."

Chapter 4

Symmetries in quantum Mechanics

4.1 Definition of a symmetry in Quantum Mechanics [1]

Often a quantum system has symmetries. One way to think about symmetries is imagine two observers related by a symmetry transformation (examples are observers separated by a translation, or being rotated which respect to each other). They must measure the same probabilities for corresponding physical events. For example in our photon polarization studies, we saw that if $|\psi\rangle$ is the wave vector of an x polarized photon traveling along z, the probability of passing through a polaroid aligned along an angle θ to the x axis with the corresponding polarization along θ described by the state $|\theta\rangle$ is $p = |\langle\theta|\psi\rangle|^2 = \cos^2\theta$. An observer using a rotated (primed) frame must compute the same probability with the transformed states $p = |\langle\theta'|\psi'\rangle|^2$ if the transformation is a symmetry of the system.

4.2 Symmetry Representations

Wigner[2] showed that because the probabilities are the same under a symmetry transformation, symmetries in quantum mechanics can be represented by operators, T, that are either linear and unitary

$$
\begin{aligned}
T\left[\alpha_1|\Psi_1\rangle + \alpha_2|\Psi_2\rangle\right] &= \alpha_1 T|\Psi_1\rangle + \alpha_2 T|\Psi_2\rangle \\
\langle\Psi|\Phi\rangle &= \langle\Psi|T^\dagger T|\Phi\rangle
\end{aligned}
\tag{4.1}
$$

or antilinear and antiunitary

$$
\begin{aligned}
T\left[\alpha_1|\Psi_1\rangle + \alpha_2|\Psi_2\rangle\right] &= \alpha_1^* T|\Psi_1\rangle + \alpha_2^* T|\Psi_2\rangle \\
\langle\Psi|\Phi\rangle &= \langle\Psi|T^\dagger T|\Phi\rangle^* .
\end{aligned}
\tag{4.2}
$$

---

[1] I follow the development of S. Weinberg, *The Quantum Theory of Fields*, (Cambridge University Press, Cambridge, 1995).

[2] E.P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*, Pure and Applied Physics, Volume 5, (Academic, New York, 1959).

To demonstrate Wigner's result, let's take two observers Alice and Bob, denoted by A and B, whose frames of reference (i.e. sets of quantum states) are related by a symmetry transformation. This means that for equivalent measurements Alice and Bob will measure the same probabilities. Alice can define a complete set of orthonormal basis states $|n^A\rangle$. The matrix element magnitude squared $|\langle m^A|n^A\rangle|^2 = \delta_{mn}$ is the probability of the system being in the state $|m^A\rangle$ and measured in the state $|n^A\rangle$. Bob can construct a set of corresponding basis states $|n^B\rangle$ to define the same system. The transformation T is defined to take Alice's state and change it into Bob's $T|n^A\rangle = |n^B\rangle$. Bob must measure the same probabilities, so $|\langle m^B|n^B\rangle|^2 = \delta_{mn}$ and Bob's states are orthonormal. Since both Alice and Bob must be able to describe any physical state, their sets of states are complete.

The physical system is described by Alice in the state $|\alpha^A\rangle$ which she expands as a linear combination of her basis states

$$|\alpha^A\rangle = \sum_n \alpha_n^A |n^A\rangle. \tag{4.3}$$

Bob can similarly describe the system with a transformed state

$$T|\alpha^A\rangle = |\alpha^B\rangle = \sum_n \alpha_n^B |n^B\rangle. \tag{4.4}$$

They each can measure the probability for their state to be in their corresponding $|m\rangle$ states and these must give the same probabilities

$$|\langle m^A|\alpha^A\rangle|^2 = |\langle m^B|\alpha^B\rangle|^2 \tag{4.5}$$

which shows that $|\alpha_m^A| = |\alpha_m^B|$.

To investigate the possibilities for the phases of the coefficients, we define a second set of states related by the symmetry transformation

$$|\beta^A\rangle = \sum_n \beta_n^A |n^A\rangle.$$

$$T|\beta^A\rangle = |\beta^B\rangle = \sum_n \beta_n^B |n^B\rangle. \tag{4.6}$$

Writing the coefficients as $\alpha_n^A = |\alpha_n^A| e^{i\phi_{\alpha n}^A}$ etc., equating the probabilities

$$|\langle \beta^A|\alpha^A\rangle|^2 = |\langle \beta^B|\alpha^B\rangle|^2 \tag{4.7}$$

and then equating the magnitudes of the corresponding coefficients of the magnitudes of $|\alpha_n|$ $|\beta_m|$, gives the set of requirements

$$\cos(\phi_{\alpha n}^A - \phi_{\alpha m}^A - \phi_{\beta n}^A + \phi_{\beta m}^A) = \cos(\phi_{\alpha n}^B - \phi_{\alpha m}^B - \phi_{\beta n}^B + \phi_{\beta m}^B). \tag{4.8}$$

The cosines will be equal if their arguments are the same modulo 2π or if they have opposite signs modulo 2π. Since how we transform the state $|\alpha^A\rangle$ cannot depend on the state $|\beta^A\rangle$, either we must pick the phase differences to be the same in the transformed state or the

phase differences to have opposite signs. Phase differences of 2π have no physical effect, and can be ignored. Eq. 4.8 can only be satisfied if the same choice is made for all terms. The transformation can give an overall phase change to the state, but this has no physical effect. We therefore can represent all symmetry transformations as either unitary linear or antiunitary antilinear operators. In the former case, matrix elements are invariant under the transformation, and in the latter matrix elements are changed to their complex conjugates under transformations.

Taking the complex conjugate of the x-representation of the Schrödinger equation with a real Hamiltonian then changing variables to $t' = -t$ gives the Schrödinger equation back with $\psi(-t')$ as a solution. Taking the complex conjugate in this case corresponds to time reversal. Antiunitary transformations in quantum mechanics all involve time reversal.

4.3   Continuous Symmetries

The identity is a trivial symmetry, where $T = 1$ and T is unitary. Therefore any symmetry that can become the identity by continously changing a parameter (for example, proper rotations or translations) has a unitary representation.

We now use our knowledge of how translations and rotations work in space to work out some properties of the corresponding unitary operators. A rotation and translation of the coordinates is given by

$$x'_k = \sum_{j=1}^{3} R_{kj} x_j + a_k \tag{4.9}$$

where R_{kj} is a 3×3 orthogonal rotation matrix and a_k are the coordinates of the translation. We write the corresponding transformation operator in quantum mechanics as $T(R, \boldsymbol{a})$. Since R is orthogonal, its inverse is its transpose

$$\sum_{k=1}^{3} R_{jk} R_{k\ell}^{-1} = \sum_{k=1}^{3} R_{jk} R_{\ell k} = \delta_{j\ell} . \tag{4.10}$$

If we write a second transformation

$$
\begin{aligned}
x_k{}'' &= \sum_{j=1}^{3} R'_{kj} x'_j + a'_k = \sum_{j=1,\ell}^{3} R'_{kj} \left[R_{j\ell} x_\ell + a_j \right] + a'_k \\
&= \sum_{\ell=1}^{3} \left[\sum_{j=1}^{3} R'_{kj} R_{j\ell} \right] x_\ell + R'_{kj} a_j + a'_k .
\end{aligned} \tag{4.11}
$$

This shows that the composition rule for the unitary transformations is

$$T(R', \boldsymbol{a}') T(R, \boldsymbol{a}) = T(R'R, R'\boldsymbol{a} + \boldsymbol{a}') . \tag{4.12}$$

In principle there can be an additional phase factor on the right, but this can always be removed, and I will not consider it here. Since $T(R^{-1}R, 0)$ is the identity transformation, a corollary is

$$T^{-1}(R, \boldsymbol{a}) = T(R^{-1}, -R^{-1}\boldsymbol{a}) \tag{4.13}$$

Since finite rotations and translations can be built up by repeatedly rotating and translating by an infinitesimal amount, most of the properties of continous symmetry operators can be deduced by looking at infinitesimal transformations. An infinitesimal rotation matrix can be written as

$$R_{jk}(\alpha) = \delta_{jk} + \sum_{\ell=1}^{3} \epsilon_{jk\ell}\alpha_\ell \tag{4.14}$$

where α_ℓ is infinitesimal and $\epsilon_{jk\ell}$ is the Levi-Civita symbol[3]. This can be seen by looking at rotations of θ around the three axes. For a z rotation, we get $\cos\theta$ for the x and y diagonals, and $\pm\sin(\theta)$ for the x and y off diagonals. Taking θ infinitesimal, the linear terms in θ are the two off diagonals of opposite sign.

Since any unitary operator can be written as $U = e^{iO}$ with O hermitian, the infinitesimal unitary transformation becomes

$$T(R(\alpha), \boldsymbol{a}) = 1 + \frac{i}{\hbar}\sum_{j=1}^{3}\alpha_j L_j + \frac{i}{\hbar}\sum_{j=1}^{3}a_j P_j \tag{4.15}$$

with α_j and a_j infinitesimals. The hermitian operators P_j and L_j will turn out to be the momentum and angular momentum operators, but to begin with we only know they are the hermitian generators of rotations and translations.

We can use the composition rule to see how the operators L_j and P_j transform under rotations. The matrix element

$$\langle n|O|m\rangle = \langle n|T^\dagger TOT^\dagger T|m\rangle = \langle n'|(TOT^\dagger)|m'\rangle \tag{4.16}$$

shows that the transformed operators are $TOT^\dagger = TOT^{-1}$. Transforming the infinitesimal T and using Eq. 4.13 gives

$$\begin{aligned}
T(R', a')T(R(\alpha), \boldsymbol{a})T^{-1}(R', a') &= T(R', a')T(R(\alpha), \boldsymbol{a})T(R'^{-1}, -R'^{-1}\boldsymbol{a}') \\
&= T(R'R(\alpha), R'\boldsymbol{a} + \boldsymbol{a}')T(R'^{-1}, -R'^{-1}\boldsymbol{a}') \\
&= T(R'R(\alpha)R'^{-1}, -R'R(\alpha)R'^{-1}\boldsymbol{a}' + R'\boldsymbol{a} + \boldsymbol{a}').
\end{aligned} \tag{4.17}$$

The tranformed infinitesimal transformation will therefore be

$$T(R(\beta), \boldsymbol{b}) = T(R'R(\alpha)R'^{-1}, -R'R(\alpha)R'^{-1}\boldsymbol{a}' + R'\boldsymbol{a} + \boldsymbol{a}'). \tag{4.18}$$

---

[3]That is $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$, $\epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1$, and all other subscript combinations are zero.

or

$$1 + \frac{i}{\hbar} \sum_{j=1}^{3} \beta_j L_j + \frac{i}{\hbar} \sum_{j=1}^{3} b_j P_j = T(R', a') \left[1 + \frac{i}{\hbar} \sum_{j=1}^{3} \alpha_j L_j + \frac{i}{\hbar} \sum_{j=1}^{3} a_j P_j \right] T^{-1}(R', a')$$

(4.19)

with[4]

$$\sum_{p=1}^{3} \epsilon_{nmp} \beta_p = \sum_{j,k,\ell=1}^{3} R'_{nj} \epsilon_{jk\ell} R'_{mk} \alpha_\ell$$

$$b_n = - \sum_{m,j,k,\ell=1}^{3} R'_{nj} \epsilon_{jk\ell} R'_{mk} \alpha_\ell a'_m + \sum_{m=1}^{3} R'_{nm} a_m .$$

(4.20)

Notice that the transformation of α and a are exactly the usual classical ones under rotation and translation.

We now want to work out the required commutation relations for the generators. Taking primed transformation to also be infinitesimal, with rotation angle α'_n, and using Eq. 4.14, the terms that give commutators will have two infinitesimals. Since operators commute with themselves, we can ignore terms that contain an infinitesimal squared. Keeping just these terms, we have

$$\beta_j = \alpha_j + \sum_{k,n=1}^{3} \epsilon_{jkn} \alpha_k \alpha'_n + ...$$

$$b_n = a_n + \sum_{m,p,k=1}^{3} \epsilon_{nmp} \alpha'_p a_m + ...$$

(4.21)

and substituting this into

$$1 + \frac{i}{\hbar} \sum_{j=1}^{3} \beta_j L_j + \frac{i}{\hbar} \sum_{j=1}^{3} b_j P_j = \left[1 + \frac{i}{\hbar} \sum_{j=1}^{3} \alpha'_j L_j + \frac{i}{\hbar} \sum_{j=1}^{3} a'_j P_j \right]$$

$$\cdot \left[1 + \frac{i}{\hbar} \sum_{j=1}^{3} \alpha_j L_j + \frac{i}{\hbar} \sum_{j=1}^{3} a_j P_j \right]$$

$$\cdot \left[1 - \frac{i}{\hbar} \sum_{j=1}^{3} \alpha'_j L_j - \frac{i}{\hbar} \sum_{j=1}^{3} a'_j P_j \right] + ...$$

(4.22)

---

[4]The first expression in Eq. 4.20 shows that $\sum_{\ell=1}^{3} \epsilon_{jk\ell} \alpha_\ell$ transforms like a second rank tensor. However it is antisymmetric, so just like the case of cross products, α_ℓ transforms like a vector under proper rotations i.e. those with determinant 1. The appendix of this chapter has a derivation.

and equating corresponding terms gives the commutation relations

$$
\begin{aligned}
i\hbar L_j \epsilon_{jkn} &= [L_k, L_n] \\
i\hbar P_j \epsilon_{jkn} &= [L_k, P_n] \\
0 &= [P_k, P_n] \,.
\end{aligned}
\tag{4.23}
$$

The important point here is that these drop out naturally from just the requirement that quantum mechanical rotations and translations must combine in the same way that spatial rotations and translations combine.

4.4 Translations and Momentum

We now want to identify the generators of translations with the canonical momentum conjugate to the variable that is translated. We call p_x the generator of translations of a particle along x. The position operator is therefore x and it has eigenstates corresponding to a definite position $|x\rangle$. For the infinitesimal translation operator we can write[5]

$$
\begin{aligned}
1 + \frac{i}{\hbar} p_x a &= \int dx |x - a\rangle\langle x| = \int dx |x\rangle\langle x + a| \\
&= \int dx |x\rangle \left[\langle x| + a \frac{\partial}{\partial x} \langle x| \right]
\end{aligned}
\tag{4.24}
$$

so that

$$
p_x = \int dx |x\rangle \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x| \,.
\tag{4.25}
$$

Sometimes this is written

$$
p_x = \int dx dx' |x'\rangle \delta(x - x') \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x| \,.
\tag{4.26}
$$

We can verify that these expressions for the momentum give the usual wave function description by calculating the matrix element $\langle \psi | p_x | \phi \rangle$,

$$
\langle \psi | p_x | \phi \rangle = \int dx \langle \psi | x \rangle \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x | \phi \rangle = \int dx \psi^*(x) \frac{\hbar}{i} \frac{\partial}{\partial x} \phi(x) \,.
\tag{4.27}
$$

Using the x basis, it is easy to calculate the commutation relation

$$
\begin{aligned}
x p_x &= \int dx |x\rangle x \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x| \\
p_x x &= \int dx |x\rangle \frac{\hbar}{i} \frac{\partial}{\partial x} x \langle x| \\
&= \int dx |x\rangle x \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x| + \int dx \frac{\hbar}{i} |x\rangle\langle x|
\end{aligned}
\tag{4.28}
$$

_____

[5]Remember we defined our translation operator as a coordinate transformation. Therefore a wave function $\psi(x)$ gets transformed to $\psi(x + a)$. If $\psi(x)$ is peaked around $x = 0$, $\psi(x + a)$ will be peaked around $x = -a$. This accounts for the $|x - a\rangle\langle x|$ in Eq.4.24.

and the commutator is

$$[x, p_x] = i\hbar \underbrace{\int dx |x\rangle\langle x|}_{\text{The identity operator}} = i\hbar. \tag{4.29}$$

Alternatively, we can calculate the commutator of the translation operator with x,

$$\left[x, \int dx |x - a\rangle\langle x|\right] = -a \int dx |x - a\rangle\langle x| \tag{4.30}$$

For an infinitesimal translation a goes to zero, and this becomes

$$\left[x, 1 + \frac{i}{\hbar} p_x a\right] = -a + O(a^2) \tag{4.31}$$

and for $a \to 0$, $[x, p] = i\hbar$ as before.

We can use our commutator to show that

$$\left[p_x, 1 - \frac{i}{\hbar} x\alpha\right] = -\alpha \tag{4.32}$$

so that $1 - \frac{i}{\hbar} x\alpha$ is the infinitesimal translation operator for the momentum.

Applying these infinitesimal translations repeatedly, they go over to the exponentials

$$\lim_{N \to \infty} \left(1 + \frac{i}{\hbar} p_x \frac{a}{N}\right)^N = e^{\frac{i}{\hbar} p_x a}$$

$$\lim_{N \to \infty} \left(1 - \frac{i}{\hbar} x \frac{\alpha}{N}\right)^N = e^{-\frac{i}{\hbar} x\alpha} \tag{4.33}$$

which can be used to shift the position or the momentum of a state.

We have normalized our $|x\rangle$ states such that

$$\langle x|x'\rangle = \delta(x - x')$$

$$\int dx |x\rangle\langle x| = 1. \tag{4.34}$$

Calling $|p\rangle$ momentum eigenstates such that $p_x|p\rangle = p|p\rangle$ with p a number (or c-number), we can write

$$\langle x|p_x|p\rangle = \int dx' \langle x|x'\rangle \frac{\hbar}{i} \frac{\partial}{\partial x'} \langle x'|p\rangle = p\langle x|p\rangle. \tag{4.35}$$

The $\langle x|x'\rangle$ gives a delta function which can be integrated immediately to give

$$\frac{\hbar}{i} \frac{\partial}{\partial x} \langle x|p\rangle = p\langle x|p\rangle. \tag{4.36}$$

We can solve the differential equation to give

$$\langle x|p\rangle \propto e^{\frac{i}{\hbar} px}. \tag{4.37}$$

35

To find the eigenvalues and normalization we can define the boundaries of our system and then take the limit that they go to infinity. The simplest method is to assume periodic boundary conditions in a box of side L, so our x coordinate runs from $-L/2$ to $L/2$ with L becoming huge. If we require periodicity, then $\langle -L/2|p\rangle = \langle L/2|p\rangle$, and the possible values of momentum are

$$p_n = \frac{2\pi\hbar n}{L} \tag{4.38}$$

with n integer.

If we normalize the momentum states to 1, we have

$$1 = \langle p_n|p_n\rangle = \int dx \langle p_n|x\rangle\langle x|p_n\rangle \tag{4.39}$$

which gives

$$\langle x|p_n\rangle = \frac{1}{\sqrt{L}} e^{\frac{i}{\hbar}p_n x} . \tag{4.40}$$

The completeness relation is

$$1 = \sum_n |p_n\rangle\langle p_n| = \int dn |p_n\rangle\langle p_n| = \frac{L}{2\pi\hbar} \int dp |p\rangle\langle p| , \tag{4.41}$$

where in the limit that $L \to \infty$, the spacing of the p_n goes to zero, so we can change $\sum_n \to \int dn$, and use $dn = L dp_n/2\pi\hbar$. Since p_n is a dummy variable, we can change its name to p since the n subscript no longer indicates integer values of n.

An alternative normalization is

$$\langle p_n|p_n\rangle = L \tag{4.42}$$

in which case

$$1 = \int \frac{dp}{2\pi\hbar} |p\rangle\langle p|$$
$$\langle x|p\rangle = e^{\frac{i}{\hbar}px} . \tag{4.43}$$

Notice that this normalization, in the limit $L \to \infty$, looks like a delta function. That is $\langle p|p'\rangle$ is zero except when $p = p'$, where it goes to infinity. We can verify this by calculating the limit $L \to \infty$,

$$\langle p_n|p_m\rangle = L\delta_{nm} \tag{4.44}$$

so that

$$\sum_n \langle p_n|p_m\rangle = L \tag{4.45}$$

Writing $p_n = p$ and $p_m = p'$, and converting to an integral,

$$\frac{L}{2\pi\hbar} \int dp \langle p|p'\rangle = L , \tag{4.46}$$

36

canceling the L on both side, we see that

$$\langle p|p' \rangle = 2\pi\hbar\delta(p - p').$$ (4.47)

With this normalization choice, the L factors have dropped out. This is probably the most popular momentum state normalization.

For either normalization all L must drop out of any physical result. Some people use a large box of side L at the beginning of their calculations to define their basis, and then simply set $L = 1$ since they know it will cancel in any correct calculation.

4.5 Time reversal reprise

Let's look at time reversal symmetry in more depth. In classical mechanics with conservative forces and in classical electrodynamics, the fundamental equations of motion are not changed if the time variable is reversed. For example for the Kepler motion of the earth around the sun, ignoring the other planets (and smaller objects), the equation of motion of the earth and the sun are

$$
\begin{aligned}
M_e \frac{d^2 \boldsymbol{r}_e(t)}{dt^2} &= \boldsymbol{\nabla}_e \frac{GM_e M_s}{|\boldsymbol{r}_e(t) - \boldsymbol{r}_s(t)|} \\
M_s \frac{d^2 \boldsymbol{r}_s(t)}{dt^2} &= \boldsymbol{\nabla}_s \frac{GM_e M_s}{|\boldsymbol{r}_e(t) - \boldsymbol{r}_s(t)|}
\end{aligned}
$$ (4.48)

where the subscripts e and s stand for the earth and the sun. The time reversed equations are given by substituting $t_r = -t$,

$$
\begin{aligned}
M_e \frac{d^2 \boldsymbol{r}_e(-t_r)}{dt_r^2} &= \boldsymbol{\nabla}_e \frac{GM_e M_s}{|\boldsymbol{r}_e(-t_r) - \boldsymbol{r}_s(-t_r)|} \\
M_s \frac{d^2 \boldsymbol{r}_s(t)}{dt_r^2} &= \boldsymbol{\nabla}_s \frac{GM_e M_s}{|\boldsymbol{r}_e(-t_r) - \boldsymbol{r}_s(-t_r)|}
\end{aligned}
$$ (4.49)

and we see that $\boldsymbol{r}_e(-t_r)$ and $\boldsymbol{r}_s(-t_r)$ satisfies the same equations of motion as $\boldsymbol{r}_e(t)$ and $\boldsymbol{r}_s(t)$. Therefore, the reversed trajectories $\boldsymbol{r}_e(-t)$, and $\boldsymbol{r}_s(-t)$ for t running forward are perfectly valid solutions to the equations of motion.

Frictional forces are not time reversal invariant. So a damped harmonic oscillator has the equation of motion

$$m\frac{d^2 x(t)}{dt^2} = -kx(t) - \gamma\frac{dx(t)}{dt}$$ (4.50)

where the frictional force here is taken to be proportional to the speed and in the opposite direction. Making the same substitution as before, the equation of motion in terms of the reversed time $t_r = -t$ is

$$m\frac{d^2 x(-t_r)}{dt^2} = -kx(-t_r) + \gamma\frac{dx(-t_r)}{dt_r}$$ (4.51)

here the sign of the frictional term has changed, so $x(-t)$ will not be a solution of the equation of motion when $x(t)$ is.

We know now that the frictional forces are macroscopic manifestations of the statistical properties of many particles, and the underlying classical equations of motion when all of the other degrees of freedom are included are time reversal invariant.

However, let me be clear, most physicists are reasonably confident that the fundamental microscopic laws of physics are *not* time reversal invariant. Experimentally the the symmetries P and CP where C is charge conjugation (essentially changing particles into their antiparticles) and P is parity (changing the coordinate axis to their negatives) are violated by the weak interaction. All local, Lorentz invariant field theories (which includes the standard model and nearly all others) are invariant under the combination of CPT where T is time reversal. So if CP is violated, T must also be violated so that CPT is a good symmetry.

We won't study the weak interaction here. All of the other interactions do seem to be time reversal invariant.

Let's look at what we expect from a quantum time-reversal transformation. If we work in a coordinate basis $|\boldsymbol{r}\rangle$ which are eigenstates of the position operator $\boldsymbol{r}_{\text{op}}$, then under time reversal, we would not expect the position eigenstates to change. We would expect the transformed momentum eigenstates to have the opposite momemtum. Momentum eigenstates would be constructed as we just showed by taking the linear combinations,

$$|\boldsymbol{p}\rangle = \int d^3r |\boldsymbol{r}\rangle\langle\boldsymbol{r}|\boldsymbol{p}\rangle = \int d^3r e^{\frac{i}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}}|\boldsymbol{r}\rangle\,. \tag{4.52}$$

Obviously if we have a unitary transformation that does not change $\boldsymbol{r}$, it will not change the momentum eigenstates either. However, an antiunitary transform will since it takes the complex conjugate of the coefficients. The transformed momentum states are, where the T operator represents that antiunitary, antilinear operator where the $|\boldsymbol{r}\rangle$ states are unchanged

$$T|\boldsymbol{p}\rangle = \int d^3r e^{-\frac{i}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}}|\boldsymbol{r}\rangle = |-\boldsymbol{p}\rangle\,. \tag{4.53}$$

The antilinear, antiunitary transformation does just what we want.

If you look at classical electromagnetism you will see that the magnetic field must change sign under time reversal. You probably already know that the magnetic field couples to the spins of particles, so that the Hamiltonian has a term proportional to $\boldsymbol{S}\cdot\boldsymbol{B}$, where $\boldsymbol{S}$ is the spin. This should tell you that we have to look more carefully at what time reversal means for systems with spin. We will see later that the time reversal operator flips the spins (like you would expect from an angular momentum), and perhaps, more surprisingly, can change the sign depending on the spin.

Appendix

Here we rotate an antisymmetric second rank tensor A, where $A_{jk} = \epsilon_{jkn}\alpha_n$, by brute force calculation of the matrix product RAR^{-1}. Writing this in matrix notation

$$\begin{pmatrix} 0 & \beta_3 & -\beta_2 \\ -\beta_3 & 0 & \beta_1 \\ \beta_2 & -\beta_1 & 0 \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} \begin{pmatrix} 0 & \alpha_3 & -\alpha_2 \\ -\alpha_3 & 0 & \alpha_1 \\ \alpha_2 & -\alpha_1 & 0 \end{pmatrix} \begin{pmatrix} R_{11} & R_{21} & R_{31} \\ R_{12} & R_{22} & R_{32} \\ R_{13} & R_{23} & R_{33} \end{pmatrix}$$

$$(4.54)$$

and multiplying the matrices gives

$$\begin{aligned} \beta_1 &= \alpha_1 \left[R_{22}R_{33} - R_{23}R_{32}\right] + \alpha_2 \left[R_{33}R_{31} - R_{21}R_{33}\right] + \alpha_3 \left[R_{21}R_{32} - R_{22}R_{31}\right] \\ &= \text{Det(R)} \left(\alpha_1 R_{11}^{-1} + \alpha_2 R_{21}^{-1} + \alpha_3 R_{31}^{-1}\right) \\ \beta_2 &= \alpha_1 \left[R_{13}R_{32} - R_{22}R_{33}\right] + \alpha_2 \left[R_{11}R_{33} - R_{13}R_{31}\right] + \alpha_3 \left[R_{12}R_{31} - R_{11}R_{32}\right] \\ &= \text{Det(R)} \left(\alpha_1 R_{12}^{-1} + \alpha_2 R_{22}^{-1} + \alpha_3 R_{32}^{-1}\right) \\ \beta_3 &= \alpha_1 \left[R_{11}R_{22} - R_{12}R_{21}\right] + \alpha_2 \left[R_{11}R_{22} - R_{12}R_{21}\right] + \alpha_3 \left[R_{11}R_{22} - R_{12}R_{21}\right] \\ &= \text{Det(R)} \left(\alpha_1 R_{13}^{-1} + \alpha_2 R_{23}^{-1} + \alpha_3 R_{33}^{-1}\right) \end{aligned} \qquad (4.55)$$

The quantities multiplying the α_j are cofactors of the elements of the matrix R. The cofactor matrix is the determinant times the inverse transpose[6]. The inverse transpose of an orthogonal matrix is the matrix, and the determinant of a proper rotation matrix is one. If the transformation contains a spatial inversion the determinant gives an additional negative sign. For proper rotations

$$\beta_j = \sum_{k=1}^{3} R_{jk}\alpha_k \qquad (4.56)$$

---

[6]You can verify this by noting that multiplying the elements of any row of a matrix by their cofactors and summing gives the determinant. Multiplying the elements of a row by the cofactors of a different row gives the determinant of a matrix which has that row twice and therefore has determinant zero. Therefore the transpose of the matrix of cofactors is the inverse times the determinant.

4.6 Problems

Problem 4.1. A Hamiltonian for one particle of mass m moving in one dimension in Galilean frame A takes the form

$$H_A = \frac{p^2}{2m} + V(x), \tag{4.57}$$

and the Schrödinger equation is

$$H_A|\psi_A(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi_A(t)\rangle \tag{4.58}$$

An observer in a frame B moving along x with a velocity $-v$ relative to frame A could use the time dependent Hamiltonian

$$H_B = \frac{p^2}{2m} + V(x - vt), \tag{4.59}$$

and the Schrödinger equation is

$$H_B|\psi_B(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi_B(t)\rangle. \tag{4.60}$$

(a) Find a time-dependent unitary transformation $T(v,t)$ that transforms between the two frames, i.e. such that $T(v,t)|\psi_A(t)\rangle = |\psi_B(t)\rangle$.

(b) Use your transformation from part (a) to transform a solution $\psi_A(x,t)$ to the Schrödinger differential equation

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\psi_A(x,t) = -\frac{\hbar}{i}\frac{\partial}{\partial t}\psi_A(x,t). \tag{4.61}$$

into a solution to the Schrödinger differential equation

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x - vt)\right]\psi_B(x,t) = -\frac{\hbar}{i}\frac{\partial}{\partial t}\psi_B(x,t). \tag{4.62}$$

(c) Verify by direct substitution substitution that your expression for $\psi_B(x,t)$ is a solution to Eq. 4.62.

Chapter 5

Natural Units

5.1 Simplifying calculations with natural units

We teach freshman physics students to check the units of their calculations as a way of finding mistakes. This is clearly helpful when the first step in a calculation is to substitute numerical values for the parameters as many freshman do.

For more complicated problems, where we carry around symbolic parameters, this becomes much less useful. Generally a mistake in units just means we dropped a factor of $\hbar$ or m somewhere. The difficulty is that we often end up with expressions that do not add to our understanding or calculation, they just take a lot longer to write down.

5.2 Harmonic oscillator example

For example, the harmonic oscillator has a Hamiltonian

$$
\begin{aligned}
H &= \frac{p_{\mathrm{op}}^2}{2m} + \frac{1}{2}m\omega^2 x_{\mathrm{op}}^2 \\
&= \int_{-\infty}^{\infty} dx |x\rangle \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}m\omega^2 x^2 \right] \langle x| .
\end{aligned}
\tag{5.1}
$$

We can solve for the eigenfunctions $\langle x|n\rangle$ and the eigenvalues by solving the eigenvalue differential equation. We will be carrying around a lot of factors of $\hbar$, m and ω.

Typically we have units of length, mass, and time.[1]

---

[1]SI adds a unit of current so that its units are meter, kilogram, second and ampere. CGS uses centimeter, gram, and second. The unit of charge is defined by $|F| = \left|\frac{q_1 q_2}{r_{12}^2}\right|$ so that two unit charges separated by 1 centimeter feel a force of one gram-cm/sec^2. The CGS or gaussian unit of charge therefore has dimensions $\sqrt{\mathrm{gram} - \mathrm{cm}^3/\mathrm{sec}^2}$. SI introduces an extra conversion factor (not unlike the english system that introduces gallons to measure volume rather than using feet^3) so that in SI the magnitude of the force between two charges at rest is $|F| = \left|\frac{\mu_0 c^2}{4\pi}\frac{q_1 q_2}{r^2}\right|$. The charges then have units of Ampere-second, and to get the units right, the conversion factor μ_0, must have dimensions of $\mathrm{Kgram} - \mathrm{m}/\mathrm{Ampere}^2 - \mathrm{sec}^2$.

However, realize that we could measure all masses using whatever particular mass m we have in the lab as the unit mass. In that case our mass has numerical value 1 in units of m. Similarly, we can measure time in units of the value of ω^{-1} of our oscillator. $\hbar$ is the only other parameter in the harmonic oscillator, and we can use it to form a length unit $\sqrt{\frac{\hbar}{m\omega}}$. Using these units, the numerical values of $\hbar$, m and ω are all equal to 1. Sometimes we just say that we take $\hbar = m = \omega = 1$.

In this set of units, there is only one harmonic oscillator Hamiltonian,

$$
\begin{aligned}
H &= \frac{p_{\text{op}}^2}{2} + \frac{1}{2}x_{\text{op}}^2 \\
&= \int_{-\infty}^{\infty} dx \, |x\rangle \left[-\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}x^2 \right] \langle x| \, .
\end{aligned}
\tag{5.2}
$$

The ground state wave function in these units would be

$$
\langle x|0\rangle = \psi_0(x) = \frac{1}{\pi^{1/4}} e^{-\frac{x^2}{2}} \, .
\tag{5.3}
$$

If we want to put back the factors of $\hbar$, m, and ω, we simply need to know the units of the various quantities. Since x is a length, it is measured in units of $\sqrt{\frac{\hbar}{m\omega}}$. Therefore if we measure x in arbitrary units, we need to convert back and replace $x \to x\sqrt{\frac{m\omega}{\hbar}}$. Similarly, the wavefunction has units of length$^{-1/2}$, so we need to multiply it by $\left(\frac{m\omega}{\hbar}\right)^{1/4}$. Therefore to convert from natural units to arbitrary units, we simply write

$$
\psi_0(x) = \left(\frac{m\omega}{\hbar}\right)^{1/4} \frac{1}{\pi^{1/4}} e^{-\frac{m\omega x^2}{2\hbar}} \, ,
\tag{5.4}
$$

which is the standard result.

5.3 Alternative view

An alternative way of getting to the same place is to define new variables. So for example, we could define

$$
\begin{aligned}
x &= \bar{x}\sqrt{\frac{\hbar}{m\omega}} \\
p &= \bar{p}\sqrt{\hbar m\omega} \\
\psi_n(x) &= \bar{\psi}_n(\bar{x}) \left(\frac{m\omega}{\hbar}\right)^{1/4} \\
E_n &= \hbar\omega \bar{E}_n
\end{aligned}
\tag{5.5}
$$

and similarly for any other variables of interest. Writing everything in terms of the barred variables is then identical to using natural units. The main difference is that when you substitute to get the expressions in terms of the barred variables, you will end up cancelling all of the $\hbar$, m and ω factors rather than just setting them to 1.

5.4   Additional parameters

A typical Hamiltonian could have other scales. For example an anharmonic oscillator might have a Hamiltonian

$$H = \frac{p_{op}^2}{2m} + \frac{1}{2}m\omega^2 x_{op}^2 + \alpha x_{op}^4 \,. \tag{5.6}$$

In this case we have the parameters $\hbar$, m, ω, α. The units of α must be Energy/length^4. If we choose our natural units as before, we would write this as

$$H = \frac{p_{op}^2}{2} + \frac{1}{2}x_{op}^2 + \alpha x_{op}^4 \,. \tag{5.7}$$

and the numerial value of α in our calculations would be the value of the dimensionless combination $\alpha\hbar/m^2\omega^3$ when we return to arbitrary units. This dimensionless parameter is the only interesting parameter in the problem; for large values the system looks like an x^4 potential, while for small values it looks like a harmonic oscillator.

5.5   Natural units for some common situations

As shown above, for the harmonics oscillator we usually use $\hbar = m = \omega = 1$.

For the infinite square well we usually use $\hbar = m = L = 1$.

For atomic and molecular physics, atomic units are often used. Converting from CGS these correspond to $\hbar = m = e = 1$. While for SI, they correspond to $\hbar = m = e^2 \frac{\mu_0 c^2}{4\pi} = 1$.

For relativistic quantum mechanics often $\hbar = c = 1$ are used.

Please note that the fine structure constant

$$\begin{aligned}
\alpha &= \frac{e^2}{\hbar c} \quad \text{CGS} \\
\alpha &= \frac{e^2}{4\pi\hbar c} \quad \text{Heaviside} - \text{Lorentz} \\
\alpha &= \frac{e^2 \mu_0 c}{4\pi\hbar} \quad \text{SI}
\end{aligned} \tag{5.8}$$

is unitless and has a numerical value of about $1/137.036$. Therefore you cannot use e, $\hbar$, and c equal to 1 as natural units. For example, in atomic units, c has a value of about 137.036.

5.6   Dimensional analysis and natural units

If you choose the natural units so that the parameters that you pick to be 1 are the important ones of your system, the natural units will set the order of magnitude for your problem, and will be automatically sensibly sized. For example for atomic units, the unit of energy is

$$\alpha^2 m c^2 = \frac{e^4 m}{\hbar^2} \simeq 27.2 \text{ eV} \tag{5.9}$$

while the unit of length is the Bohr radius

$$\frac{\hbar^2}{me^2} = a_0 \simeq 0.511 \text{ A} . \tag{5.10}$$

These are roughly the energy and length scales we expect for hydrogen orbitals. We will calculate the polarizability of the hydrogen atom in its ground state later. This is the ratio of the induced dipole moment to the electric field. A dipole moment is a charge times a distance, while an electric field is a charge divided by a distance squared. Therefore the ratio has units of distance cubed. Without detailed calculation, we would then expect the polarizability to have an order of magnitude around a_0^3. In fact, the ground-state polarizibility of hydrogen is $4.5a_0^3$.

Often if you can figure out the dominant physical reason for a property, you can use the parameters that that physical reason implies to form a quantity that has the correct dimensions which then tells you the order of magnitude of a quantity.

5.7 A White Dwarf and Neutron Star Example

The collapse of a star to a white dwarf or neurton star is one example of the power of applying dimensional analysis to get rough orders of magnitude once the basic physics is understood.

A star that exhausts its nuclear fuel can end its life as a white dwarf made mostly of iron or carbon. However, there is an upper mass limit for white dwarfs. If the mass is beyond the Chandrasekar limit (about 2 solar masses) the star collapses in a supernova to become a neutron star or a black hole. The basis physical processes and scales are:

1. As the star cools it must be supported against its gravity by the pressure from the electron fermi gas.

2. The equation of state of a fermi gas softens when the particle velocities become relativistic.

At zero temperature the energy per particle of a noninteracting fermi gas is proportional to its fermi energy. The pressure will be proportional to the negative derivative of the energy with respect to volume. Since only one fermion can occupy a state, noninteracting fermions fill the lowest energy orbitals. These orbitals can be characterized by their momentum eigenvalue or equivalently their wave vector. The wave vector for free particles in a box of side L is proportional to $1/L$. Therefore it $k_f \propto n^{1/3}$, where n is the number density. The noninteracting fermi energy is the energy of a fermion with the fermi momentum, p_f,

$$E_f(p_f) = \sqrt{p_f^2 c^2 + m^2 c^4} . \tag{5.11}$$

Using the chain rule, the pressure behaves like

$$P \sim \frac{\partial E_f(p_f)}{\partial p_f} \sim \frac{x}{\sqrt{x^2 + 1}} \tag{5.12}$$

where $x = p_f^2/m^2c^2$. At small x this is linear in x, while at large x it is constant. The softening of the equation of state occurs at roughly $p_f = mc$. The value of $1/k_f$ is roughly the distance between the electrons, which therefore must be $r_c = \hbar/mc$ which is the Compton radius.

As an aside the classical radius of the electron, that is the length scale where the electrostatic and relativistic energies are the same, is $r_0 = e^2/mc^2$. Notice since this is classical, there are no $\hbar$ factors. The fine structure constant $\alpha = e^2/\hbar c$ is unitless. Dividing the classical radius by α gives the compton radius $\hbar/mc$ which has no electron charge factor. It is the scale where quantum mechanics and relativity are of equal importance. Dividing by a fine structure constant again gives the Bohr radius $\hbar^2/e^2m$. Here the speed of light factors have canceled, so this is the length scale where nonrelativistic quantum mechanics and electrostatics are of equal importance. As expected from this analysis, the Hartree (twice a Rydberg) is simply $\alpha^2 mc^2$.

The gravitional energy is roughly

$$E_G \sim -\frac{GM^2}{R} \tag{5.13}$$

where M is the mass of the star and R is the radius. The fermi energy will go like the inverse size of the star for a relativistic gas, just as the gravitational energy does. If the gravitational energy dominates, the star will collaps. For a nonrelativistic gas, the total kinetic energy will be proportional to R^{-2} and this will stabilize the star.

The electron fermi energy at the compton radius will be roughly its mass. We assume one or two nucleon masses, m_n, per electron. The number of electrons is roughly R^3/r_c^3. The kinetic and potential energies will be roughly equal magnitude for

$$\frac{GM^2}{R} = \frac{R^3}{r_c^3}mc^2 = \frac{R^3}{r_c^4}\hbar c \tag{5.14}$$

and the total mass of the star must be roughly the number electrons times the nucleon mass

$$M = \frac{R^3}{r_c^3}m_n \, . \tag{5.15}$$

Solving the two equations we have

$$
\begin{aligned}
R &= \frac{r_c M_p}{m_n} \\
M &= \frac{M_p^3}{m_n^2}
\end{aligned}
\tag{5.16}
$$

where the Planck mass M_p is

$$M_p = \sqrt{\frac{\hbar c}{G}} \, . \tag{5.17}$$

We can plug in some numbers and see if these dimensional arguments make sense. We have the Planck mass, the atomic mass unit, and the compton radius for the electron

$$
\begin{aligned}
M_p &= 2.18 \times 10^{-5} \text{ g} \\
m_u &= 1.66 \times 10^{-24} \text{ g} \\
r_c &= 3.86 \times 10^{-11} \text{ cm}
\end{aligned}
\tag{5.18}
$$

which gives

$$
\begin{aligned}
R &\sim 5 \times 10^{8} \text{ cm} \\
M &\sim 4 \times 10^{33} \text{ g}
\end{aligned}
\tag{5.19}
$$

or in terms of the solar radius and mass

$$
\begin{aligned}
R_\odot &= 6.95 \times 10^{10} \text{ cm} \\
M_\odot &= 1.98 \times 10^{33} \text{ g}
\end{aligned}
\tag{5.20}
$$

$$
\begin{aligned}
R &\sim 7 \times 10^{-3} R_\odot \\
M &\sim 2 \times M_\odot .
\end{aligned}
\tag{5.21}
$$

Notice that if we apply the same reasoning to the collapse of a neutron star, we would simply replace the electron compton radius with the neutron compton radius in Eq. 5.16. Since the mass scale has no r_c factor, the mass limit for a white dwarf and a neutron star will be roughly the same, around two solar masses, which is verified from observations. The critical radius of the neutron star will be smaller than that of the white dwarf by roughly the ratio of the electron to nucleon mass or about 2000 times smaller. So R for a neutron star will be roughly 30 km, which is again the correct observational scale.

Of course, real electrons and nuclei or real neutrons interact. Therefore, to get more than just orders of magnitude, detailed calculations with realistic interactions must be performed. However, if we have correctly identified the main physics, these detailed calculations will modify the prefactors making them functions of the dimensionless parameters of the system, but will not change the overall behavior.

Chapter 6

Eigenvectors and eigenvalues

6.1 Eigenstates of Hermitian and Unitary Operators

Often we want to work in the basis of eigenstates of an operator. In quantum mechanics we will deal mostly with either Hermitian or unitary operators. A Hermitian operator is one where $H^\dagger = H$, while for a unitary operator $U^\dagger = U^{-1}$.

Given eigenvectors and eigenvalues for a Hermitian operator H, we have

$$H|\psi_n\rangle = \lambda_n|\psi_n\rangle. \tag{6.1}$$

and taking the adjoint, we have

$$\langle\psi_n|H^\dagger = \langle\psi_n|H = \langle\psi_n|\lambda_n^*. \tag{6.2}$$

Evaluating $\langle\psi_m|H|\psi_n\rangle$ gives

$$\langle\psi_m|H|\psi_n\rangle = \lambda_n\langle\psi_m|\psi_n\rangle = \lambda_m^*\langle\psi_m|\psi_n\rangle. \tag{6.3}$$

or

$$(\lambda_m - \lambda_n^*)\langle\psi_n|\psi_m\rangle = 0. \tag{6.4}$$

Taking $m = n$, the overlap of $|\psi_n\rangle$ with itself is non zero, so λ_n must be real. For $\lambda_m \neq \lambda_n$, the states are orthogonal. If $\lambda_m = \lambda_n$ for different m and n, we have a *degeneracy*. In this case, any linear combinations of $|\psi_m\rangle$ and $|\psi_n\rangle$ are also eigenvectors with the same eigenvalues. We can therefore use, for example, Gram-Schmidt orthogonalization to find orthogonal eigenvectors.[1] The result is that the eigenvalues of a Hermitian matrix are real, and the eigenvectors can always be chosen to be orthogonal. Eigenvectors of Hermitian operators with nondegenerate eigenvalues are always orthogonal.

---

[1]Gram-Schmidt orthogonalization produces an orthogonal set from a linearly independent but nonorthogonal set of states, $|1\rangle$, $|2\rangle$,...,$|N\rangle$, by simply subtracting off the components along the previously chosen states. That is we take $|1'\rangle = \mathcal{N}_1|1\rangle$, $|2'\rangle = \mathcal{N}_2\left[|2\rangle - |1'\rangle\langle 1'|2\rangle\right]$, ..., $|j'\rangle = \mathcal{N}_j\left[|j\rangle - \sum_{i=1}^{j-1}|j'\rangle\langle j'|j\rangle\right]$, where the $\mathcal{N}_j$ are chosen so that $\langle j'|j'\rangle = 1$. You can choose any ordering of the states since the initial ordering is arbitrary.

A similar argument goes through for Unitary operators

$$U|\phi_n\rangle = \alpha_n|\phi_n\rangle. \tag{6.5}$$

and taking the adjoint, we have

$$\langle\phi_n|U^\dagger = \langle\phi_n|U^{-1} = \langle\phi_n|\alpha_n^*. \tag{6.6}$$

We now multiply this equation from the right with U and rearrange to give

$$\langle\phi_n|U = \langle\phi_n|\frac{1}{\alpha_n^*} \tag{6.7}$$

The matrix element of $\langle\phi_n|U|\phi_m\rangle$ then gives the equation

$$\left(\alpha_m - \frac{1}{\alpha_n^*}\right)\langle\phi_n|\phi_m\rangle = 0. \tag{6.8}$$

For $m = n$, we have $|\alpha_n| = 1$, so the eigenvalues of a unitary operator are phase factors. Just as for Hermitian operators, the left eigenvectors are the adjoint of the right eigenvectors, and the eigenvectors of different eigenvalues are orthogonal. Eigenvectors with the same eigenvalues can always be orthogonalized.

6.2 Diagonalization

Usually, an operator is defined by its matrix elements in some particular basis. That is we can write

$$O = \sum_{ij}|i\rangle\langle i|O|j\rangle\langle j| = \sum_{ij}|i\rangle O_{ij}\langle j| \tag{6.9}$$

with

$$O_{ij} \equiv \langle i|O|j\rangle. \tag{6.10}$$

Starting with the eigenvalue equation, $O|\psi_n\rangle = \lambda_n|\psi_n\rangle$, inserting a complete set of basis functions $|j\rangle$, and multiplying on the left by the basis state $\langle i|$, this becomes

$$\sum_j\langle i|O|j\rangle\langle j|\psi_n\rangle = \lambda_n\langle i|\psi_n\rangle. \tag{6.11}$$

If we define $\psi_i^{(n)} \equiv \langle i|\psi_n\rangle$, we can write this in matrix notation as

$$\sum_j O_{ij}\psi_j^{(n)} = \lambda_n\psi_i^{(n)} \tag{6.12}$$

or

$$\begin{pmatrix} O_{11} & O_{12} & ... & O_{1N} \\ O_{21} & O_{22} & ... & O_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ O_{N1} & O_{N2} & ... & O_{NN} \end{pmatrix} \begin{pmatrix} \psi_1^{(n)} \\ \psi_2^{(n)} \\ \vdots \\ \psi_N^{(n)} \end{pmatrix} = \lambda_n \begin{pmatrix} \psi_1^{(n)} \\ \psi_2^{(n)} \\ \vdots \\ \psi_N^{(n)} \end{pmatrix}. \tag{6.13}$$

We can use any matrix diagonalization method to find the eigenvectors and eigenvalues of the operator. That is,

$$|\psi_n\rangle = \sum_j |j\rangle\langle j|\psi_n\rangle = \sum_j |j\rangle\psi_j^{(n)}, \tag{6.14}$$

the calculation of the eigenvectors of the matrix gives us the basis transformation.

Notice that if we write the operator in its eigenbasis,

$$O = \sum_{mn} |\psi_m\rangle\langle\psi_m|O|\psi_n\rangle\langle\psi_n|. \tag{6.15}$$

the matrix

$$\mathcal{O}_{mn} \equiv \langle\psi_m|O|\psi_n\rangle = \lambda_n\delta_{nm} \tag{6.16}$$

is diagonal, and we can also write

$$\mathcal{O} = \sum_n |\psi_n\rangle\lambda_n\langle\psi_n|. \tag{6.17}$$

Changing to the eigenbasis has diagonalized the matrix of the matrix elements. Explicitly, we can write

$$\mathcal{O}_{mn} = \langle\psi_m|O|\psi_n\rangle = \sum_{ij}\langle\psi_m|i\rangle\langle i|O|j\rangle\langle j|\psi_n\rangle = \sum_{ij}\psi_i^{(m)*}O_{ij}\psi_j^{(n)} \tag{6.18}$$

or

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{pmatrix} =$$

$$\begin{pmatrix} \psi_1^{(1)*} & \psi_2^{(1)*} & \dots & \psi_N^{(1)*} \\ \psi_1^{(2)*} & \psi_2^{(2)*} & \dots & \psi_N^{(2)*} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1^{(N)*} & \psi_2^{(N)*} & \dots & \psi_N^{(N)*} \end{pmatrix} \begin{pmatrix} O_{11} & O_{12} & \dots & O_{1N} \\ O_{21} & O_{22} & \dots & O_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ O_{N1} & O_{N2} & \dots & O_{NN} \end{pmatrix} \begin{pmatrix} \psi_1^{(1)} & \psi_1^{(2)} & \dots & \psi_1^{(N)} \\ \psi_2^{(1)} & \psi_2^{(2)} & \dots & \psi_2^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N^{(1)} & \psi_N^{(2)} & \dots & \psi_N^{(N)} \end{pmatrix} \tag{6.19}$$

which shows explicitly that the eigenvectors of the matrix give a transformation matrix that diagonalizes the operator. Similarly we can write

$$O_{ij} = \langle i|O|j\rangle = \sum_n \langle i|O|\psi_n\rangle\langle\psi_n|j\rangle = \sum_n \langle i|\psi_n\rangle\lambda_n\langle\psi_n|j\rangle = \sum_n \psi_i^{(n)}\lambda_n\psi_j^{(n)*} \tag{6.20}$$

49

or

$$
\begin{pmatrix}
O_{11} & O_{12} & ... & O_{1N} \\
O_{21} & O_{22} & ... & O_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
O_{N1} & O_{N2} & ... & O_{NN}
\end{pmatrix} =
$$

$$
\begin{pmatrix}
\psi_1^{(1)} & \psi_1^{(2)} & ... & \psi_1^{(N)} \\
\psi_2^{(1)} & \psi_2^{(2)} & ... & \psi_2^{(N)} \\
\vdots & \vdots & \ddots & \vdots \\
\psi_N^{(1)} & \psi_N^{(2)} & ... & \psi_N^{(N)}
\end{pmatrix}
\begin{pmatrix}
\lambda_1 & 0 & ... & 0 \\
0 & \lambda_2 & ... & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & ... & \lambda_N
\end{pmatrix}
\begin{pmatrix}
\psi_1^{(1)*} & \psi_2^{(1)*} & ... & \psi_N^{(1)*} \\
\psi_1^{(2)*} & \psi_2^{(2)*} & ... & \psi_N^{(2)*} \\
\vdots & \vdots & \ddots & \vdots \\
\psi_1^{(N)*} & \psi_2^{(N)*} & ... & \psi_N^{(N)*}
\end{pmatrix}
$$

$$(6.21)$$

6.3   Ordering and Phases

The ordering of basis states and the choice of their phases has no effect on the results. Therefore you can choose these to simplify your problem. For example, multiplying $\psi_j^{(1)}$ by the same phase for all j, obviously will not change the result in Eq. 6.20, since the only term affected is $\psi_i^{(1)}\lambda_1\psi_j^{(1)*}$ and the phase factors cancel. Remember that negative one is an example of a phase, so changing the sign of $\psi_j^{(1)}$ will not change the result. Similarly interchanging $\psi_j^{(1)}$ and $\psi_j^{(2)}$ along with λ_1 and λ_2 in Eq. 6.20 just changes the order of the terms in the summation which again cannot change the result. Since Eq. 6.21 is just the matrix representation of Eq. 6.20, reordering or picking arbitrary phases for the states does not change the result.

Be careful! You can get in trouble if you don't use the same ordering and phase convention everywhere. Therefore if you obtain matrix elements from other people, make sure you understand the phase convention.

6.4   Block diagonal matrices

Block diagonal matrices have a standard form like in the example below

$$
\begin{pmatrix}
\times & \times & \times & 0 & 0 & 0 \\
\times & \times & \times & 0 & 0 & 0 \\
\times & \times & \times & 0 & 0 & 0 \\
0 & 0 & 0 & \times & \times & 0 \\
0 & 0 & 0 & \times & \times & 0 \\
0 & 0 & 0 & 0 & 0 & \times
\end{pmatrix}
\tag{6.22}
$$

where an $\times$ indicates an arbitrary nonzero entry (not all equal!). In this case, the corresponding operator has off-diagonal matrix elements between (i.e. couples) states 1, 2, and

3, but these states do not couple to 4, 5, or 6. Similarly states 4 and 5 couple only to each other and states 6 couples only to itself. Such a matrix can be diagonalized by diagonalizing separately each of the sub blocks. For example, given the matrix

$$\begin{pmatrix} a & b & 0 & 0 & 0 \\ b & a & 0 & 0 & 0 \\ 0 & 0 & c & d & 0 \\ 0 & 0 & d & c & 0 \\ 0 & 0 & 0 & 0 & e \end{pmatrix} \tag{6.23}$$

with a, b c, d and e real, we would find the eigenvectors and eigenvalues of the upper left 2×2 matrix

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix} \tag{6.24}$$

to be

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{6.25}$$

with eigenvalues $a + b$ and $a - b$ respectively. The corresponding eigenvectors of the full matrix are then

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{6.26}$$

with the same eigenvalues. You can check this by multiplying the original matrix times these eigenvectors.. Similarly, the other three eigenvectors are

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{6.27}$$

with eigenvalues $c+d$, $c-d$, and e. It is much easier to diagonalize the smaller matrix blocks than to try to diagonalize the original matrix.

Here's another example matrix form

$$\begin{pmatrix} \times & 0 & 0 & 0 & \times & \times \\ 0 & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & 0 & 0 & 0 \\ 0 & 0 & 0 & \times & 0 & 0 \\ \times & 0 & 0 & 0 & \times & \times \\ \times & 0 & 0 & 0 & \times & \times \end{pmatrix}. \tag{6.28}$$

At first glance this does not look block diagonal. But if you study a bit more you will see that states 1, 5, and 6 couple only to each other, states 2 and 3 couple only to each other, and state 4 couples only to itself. This is just the previous example with the states reordered $1, 2, 3, 4, 5, 6 \rightarrow 1, 5, 6, 2, 3, 4$. Therefore we will also call this block diagonal, since you can choose the matrix representation to be block diagonal in the standard form.

You do not need to actually reorder the states to find the eigenvectors and eigenvalues for the blocks. For example, the matrix

$$\begin{pmatrix} a & 0 & 0 & 0 & b \\ 0 & c & 0 & d & 0 \\ 0 & 0 & e & 0 & 0 \\ 0 & d & 0 & c & 0 \\ b & 0 & 0 & 0 & a \end{pmatrix} \tag{6.29}$$

is just a reordering of the one shown in Eq. 6.23. Since state 1 only couples to 5, just as in the explicit block form, we can "strike out" the rows and columns for states 2, 3, and 4, and the resulting matrix is that of Eq. 6.24. We can do the same thing for states 2 and 4 and the uncoupled state 3. The eigenvectors are then

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{6.30}$$

with eigenvalues $a + b$, $a - b$, $c + d$, $c - d$ and e respectively. Again, you can check this by direct multiplication of the original matrix times these eigenvectors to see how it works.

6.5   Functions of Operators

The eigenvalues and eigenvectors are often useful to calculate functions of operators. A typical example is solving the Schrödinger equation

$$H|\psi(t)\rangle = -\frac{\hbar}{i} \frac{\partial}{\partial t} |\psi(t)\rangle , \tag{6.31}$$

with H independent of time. The formal solution is

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar} H t} |\psi(0)\rangle \tag{6.32}$$

where $|\psi(0)\rangle$ is the state of the system at time $t = 0$. To go further we need to have an expression for the exponential of an operator. A function of an operator can be defined as its power series if the series converges. That is if

$$f(x) = \sum_k \alpha_k x^k \tag{6.33}$$

we define

$$f(O) = \sum_k \alpha_k O^k \tag{6.34}$$

since we know how to repeatedly operate on a state we can calculate each of the terms and add them. If the series converges, we have a method of calculating a function of the operator. An alternative method is to operate on eigenvectors of O. If we apply the power series to an eigenvector, we get the power series with O replaced by its eigenvalue multiplied by the eigenvector. Performing the sum, gives

$$f(O)|\psi_n\rangle = f(\lambda_n)|\psi_n\rangle. \tag{6.35}$$

We will take this to be the definition of a function a Hermitian operator. The result will agree with the power series whenever it converges. The advantage is that it is no longer necessary to have a convergent power series.

Using the completeness relation we can write

$$f(O) = \sum_n f(O)|\psi_n\rangle\langle\psi_n| = \sum_n |\psi_n\rangle f(\lambda_n)\langle\psi_n|. \tag{6.36}$$

The matrix elements of $f(O)$ in the original basis are then

$$f(O)_{ij} = \langle i|f(O)|j\rangle = \sum_n \langle i|\psi_n\rangle f(\lambda_n)\langle\psi_n|j\rangle = \sum_n \psi_i^{(n)} f(\lambda_n)\psi_j^{(n)*} \tag{6.37}$$

or

$$\begin{pmatrix} f(O)_{11} & f(O)_{12} & \ldots & f(O)_{1N} \\ f(O)_{21} & f(O)_{22} & \ldots & f(O)_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ f(O)_{N1} & f(O)_{N2} & \ldots & f(O)_{NN} \end{pmatrix} =$$

$$\begin{pmatrix} \psi_1^{(1)} & \psi_1^{(2)} & \ldots & \psi_1^{(N)} \\ \psi_2^{(1)} & \psi_2^{(2)} & \ldots & \psi_2^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N^{(1)} & \psi_N^{(2)} & \ldots & \psi_N^{(N)} \end{pmatrix} \begin{pmatrix} f(\lambda_1) & 0 & \ldots & 0 \\ 0 & f(\lambda_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & f(\lambda_N) \end{pmatrix} \begin{pmatrix} \psi_1^{(1)*} & \psi_2^{(1)*} & \ldots & \psi_N^{(1)*} \\ \psi_1^{(2)*} & \psi_2^{(2)*} & \ldots & \psi_N^{(2)*} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1^{(N)*} & \psi_2^{(N)*} & \ldots & \psi_N^{(N)*} \end{pmatrix}$$

$$\tag{6.38}$$

In principle, we can therefore solve our time-dependent Schrödinger equation by diagonalizing H, that is by finding its eigenvalues and eigenvectors. We then transform (with the matrix of eigenvectors) to the eigenvector basis. In this basis we multiply each term by the corresponding phase factor $e^{-\frac{i}{\hbar}E_n t}$ where E_n is the eigenvalue of H. We then transform back to the original basis.

6.6 Analytic Diagonalization

Analytic methods typically start by finding the eigenvalues of the matrix. The eigenvalue equations are N coupled linear *homogeneous* equations.

$$[O_{ij} - \lambda \delta_{ij}] \psi_i = 0 \tag{6.39}$$

Since there is no source term, for most λ values, only $\psi_i = 0$ is a solution. The only way for there to be nonzero ψ_i solutions is if there are fewer than N independent equations, that is the determinant of the matrix $O_{ij} - \lambda \delta_{ij}$ is zero. The determinant is a polynomial of order N, the size of the matrix, so there are N eigenvalues. Since only up to quartic equations can be solved analytically in general, except in special cases, only up to four by four matrices can be analytically diagonalized. Once we have found the eigenvalues, we can plug them into the eigenvalue equation and solve for the eigenvectors. For the nondegenerate case there are $N - 1$ equations, so one of the nonzero values is set by our normalization condition and phase convention.

If there are no degeneracies, a quick way to write down the eigenvectors is to use the cofactors of the original matrix. Recall that if we write the determinant of A, the factor multiplying A_{ij} in the determinant expression is the cofactor. The determinant can be calculated by the sum of the elements in any row times their cofactors or any column times their cofactors. Furthermore if we sum the product of the elements of a row times the cofactors of a different row, we get the determinant of a matrix with two rows identical which has determinant zero. From these statements, we can see that multiplying the matrix times the transpose of the matrix of cofactors gives a diagonal matrix with the diagonal elements all equal to the determinant. Therefore, the Cofactor matrix C_{ij} is the determinant times the transpose of the inverse of the matrix. We can write the inverse matrix as

$$(O - \lambda I)_{ij}^{-1} = \sum_n \langle i|\psi_n\rangle \frac{1}{\lambda_n - \lambda} \langle \psi_n|j\rangle = \sum_n \psi_i^{(n)} \frac{1}{\lambda_n - \lambda} \psi_j^{(n)*} \tag{6.40}$$

and since the determinant is the product of eigenvalues, the cofactor matrix of $O - \lambda I$ is

$$C_{ij}(\lambda) = \sum_n \psi_j^{(n)} \psi_i^{(n)*} \prod_{m \neq n} (\lambda_m - \lambda) . \tag{6.41}$$

We therefore have

$$\psi_j^{(k)} \propto C_{ij}(\lambda_k) \tag{6.42}$$

as long as there is no degeneracy and $\psi_i^{(k)*} \neq 0$.

Let's apply this to a simple two by two matrix

$$\begin{pmatrix} 0 & \alpha \\ \alpha & 0 \end{pmatrix} . \tag{6.43}$$

The determinant is

$$\text{Det} \begin{pmatrix} -\lambda & \alpha \\ \alpha & -\lambda \end{pmatrix} = \lambda^2 - \alpha^2 = 0 \tag{6.44}$$

so that $\lambda_\pm = \pm\alpha$ are the two eigenvalues. The cofactors of the first row give

$$\psi \propto \begin{pmatrix} -\lambda \\ -\alpha \end{pmatrix} \tag{6.45}$$

Plugging in for $\lambda_\pm$, normalizing and picking the phase so the first element is positive, gives the eigenvectors

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{6.46}$$

6.7 Putting it together

Let's use the results to look at a spin-half particle initially with z component $\hbar/2$. We take the basis to be spin up and down in the z direction $|\uparrow\rangle$ and $|\downarrow\rangle$. A magnetic field B is applied along the x direction. What is the state vector as a function of t? What is the probability of measuring the particle as spin down after a time t? The Hamiltonian for the spin is

$$H = -\mu\boldsymbol{\sigma} \cdot \boldsymbol{B} \tag{6.47}$$

where μ is the magnetic moment.

The matrix elements of the spin operator $\boldsymbol{\sigma}$ in the up-down basis are given by the Pauli matrices

$$\begin{pmatrix} \langle\uparrow|\boldsymbol{\sigma}|\uparrow\rangle & \langle\uparrow|\boldsymbol{\sigma}|\downarrow\rangle \\ \langle\downarrow|\boldsymbol{\sigma}|\uparrow\rangle & \langle\downarrow|\boldsymbol{\sigma}|\downarrow\rangle \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{z}} & \hat{\boldsymbol{x}} - i\hat{\boldsymbol{y}} \\ \hat{\boldsymbol{x}} + i\hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \end{pmatrix} = \hat{\boldsymbol{x}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \hat{\boldsymbol{y}}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \hat{\boldsymbol{z}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
$$= \hat{\boldsymbol{x}}\sigma_x + \hat{\boldsymbol{y}}\sigma_y + \hat{\boldsymbol{z}}\sigma_z. \tag{6.48}$$

The Hamiltonian matrix elements in the up-down basis, for a field along x, are

$$\begin{pmatrix} \langle\uparrow|H|\uparrow\rangle & \langle\uparrow|H|\downarrow\rangle \\ \langle\downarrow|H|\uparrow\rangle & \langle\downarrow|H|\downarrow\rangle \end{pmatrix} = -\mu\sigma_x B = \begin{pmatrix} 0 & -\mu B \\ -\mu B & 0 \end{pmatrix} \tag{6.49}$$

The eigenvalues and vectors were calculated above,

$$|1\rangle = \frac{1}{\sqrt{2}}[|\uparrow\rangle + |\downarrow\rangle], \quad E_1 = -\mu B$$

$$|2\rangle = \frac{1}{\sqrt{2}}[|\uparrow\rangle - |\downarrow\rangle], \quad E_2 = \mu B. \tag{6.50}$$

We write

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}Ht}|\psi(0)\rangle$$
$$= \sum_{n=1,2} |n\rangle e^{-\frac{i}{\hbar}E_n t}\langle n|\psi(0)\rangle$$
$$= \sum_{n=1,2}\sum_{s=\uparrow,\downarrow} |n\rangle e^{-\frac{i}{\hbar}E_n t}\langle n|s\rangle\langle s|\psi(0)\rangle \tag{6.51}$$

Taking the matrix elements with $\langle s'|$ which is either $\langle \uparrow |$ and $\langle \downarrow |$, we have

$$\langle s'|\psi(t)\rangle = \sum_{n=1,2}\sum_{s=\uparrow,\downarrow} \langle s'|n\rangle e^{-\frac{i}{\hbar}E_n t}\langle n|s\rangle\langle s|\psi(0)\rangle \qquad (6.52)$$

which in matrix notation becomes

$$\begin{aligned}
\left(\begin{array}{c} \langle \uparrow |\psi(t)\rangle \\ \langle \downarrow |\psi(t)\rangle \end{array}\right) &= \left(\begin{array}{cc} \langle \uparrow |1\rangle & \langle \uparrow |2\rangle \\ \langle \downarrow |1\rangle & \langle \downarrow |2\rangle \end{array}\right)\left(\begin{array}{cc} e^{-\frac{i}{\hbar}E_1 t} & 0 \\ 0 & e^{-\frac{i}{\hbar}E_2 t} \end{array}\right)\left(\begin{array}{cc} \langle 1| \uparrow\rangle & \langle 1| \downarrow\rangle \\ \langle 2| \uparrow\rangle & \langle 2| \downarrow\rangle \end{array}\right)\left(\begin{array}{c} \langle \uparrow |\psi(0)\rangle \\ \langle \downarrow |\psi(0)\rangle \end{array}\right) \\
&= \frac{1}{\sqrt{2}}\left(\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right)\left(\begin{array}{cc} e^{\frac{i}{\hbar}\mu Bt} & 0 \\ 0 & e^{-\frac{i}{\hbar}\mu Bt} \end{array}\right)\frac{1}{\sqrt{2}}\left(\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right)\left(\begin{array}{c} \langle \uparrow |\psi(0)\rangle \\ \langle \downarrow |\psi(0)\rangle \end{array}\right) \\
&= \left(\begin{array}{cc} \cos(\omega t) & i\sin(\omega t) \\ i\sin(\omega t) & \cos(\omega t) \end{array}\right)\left(\begin{array}{c} \langle \uparrow |\psi(0)\rangle \\ \langle \downarrow |\psi(0)\rangle \end{array}\right) \qquad (6.53)
\end{aligned}$$

where $\omega = \mu B/\hbar$. For $|\psi(0)\rangle = |\uparrow\rangle$, we have

$$\begin{aligned}
|\psi(t)\rangle &= |\uparrow\rangle\langle \uparrow |\psi(t)\rangle + |\downarrow\rangle\langle \downarrow |\psi(t)\rangle \\
&= \cos(\omega t)|\uparrow\rangle + i\sin(\omega t)|\downarrow\rangle \qquad (6.54)
\end{aligned}$$

and the probability of measuring the particle with down spin at time t is

$$P_\downarrow(t) = |\langle \downarrow |\psi(t)\rangle|^2 = \sin^2(\omega t) = \frac{1}{2}\left[1 - \cos\left(\frac{2\mu Bt}{\hbar}\right)\right] = \frac{1}{2}\left[1 - \cos\left(\frac{E_2 - E_1}{\hbar}t\right)\right]. \quad (6.55)$$

6.8 Numerical Diagonalization

Fairly large matrices can be diagonalized numerically. For a general real symmetric matrix or hermitian matrix $N = 1000$ will take a minute or two on a single processor few GHz machine to calculate all eigenvalues and eigenvectors. The algorithms scale with time as N^3, so $N = 10,000$ will take a day or so. Typically we are most interested in a few low lying eigenvalues and eigenvectors and these can be calculated iteratively, with order M matrix vector multiplies for M eigenvalues and eigenvectors. If the matrix is sparse, the matrix vector multiplies can be of order N operations, and it is possible to find the lowest few eigenvalues and vectors for matrices with $N > 10^6$.

To give us the flavor of how the numerical methods work, I will describe the Jacobi method very briefly. This method is several times slower than the Householder method generally used, but it has the advantage of being easy to understand. Let's look for the transformation that diagonalizes a two by two hermitian matrix. Since we can choose the phase of our eigenvectors, we can choose one of their elements to be real. For the first eigenvector we then choose this element to be $\cos\theta$. The magnitude of the second element must then be the magnitude of $\sin\theta$, and it can have any phase. The other eigenvector can now be formed from Gram-Schmidt orthogonalization so that the general form for our unitary transformation is

$$\left(\begin{array}{cc} \cos\theta & e^{i\phi}\sin\theta \\ -e^{-i\phi}\sin\theta & \cos\theta \end{array}\right). \qquad (6.56)$$

Tranforming a general Hermitian two by two matrix to diagonal form

$$
\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & e^{i\phi}\sin\theta \\ -e^{-i\phi}\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} O_{11} & O_{12} \\ O_{12}^* & O_{22} \end{pmatrix} \begin{pmatrix} \cos\theta & -e^{i\phi}\sin\theta \\ e^{-i\phi}\sin\theta & \cos\theta \end{pmatrix}. \quad (6.57)
$$

Looking at either of the off diagonal terms that must be zero, we get the requirement

$$
0 = (O_{22} - O_{11})\sin 2\theta + O_{12}e^{-i\phi}(1 + \cos 2\theta) - O_{12}^* e^{i\phi}(1 - \cos 2\theta). \quad (6.58)
$$

Since the diagonal terms are real, we can set the imaginary part to zero by picking ϕ so that $O_{12}e^{-i\phi}$ is real. To keep the transformation as close to the identity as possible we pick the smallest magnitude ϕ. Solving for θ we have

$$
\tan 2\theta = -\frac{2O_{12}e^{-i\phi}}{O_{22} - O_{11}} \quad (6.59)
$$

again we take the smallest magnitude θ.

For the two by two system, this solves for the eigenvectors. The eigenvalues are calculated from Eq. 6.57.

The Jacobi method generalizes this transformation to zero a single off diagonal element of the matrix. For example for a three by three matrix, to zero the off diagonal elements, we use the transformation matrices

$$
\begin{pmatrix} \cos\theta_{12} & e^{i\phi_{12}}\sin\theta_{12} & 0 \\ -e^{-i\phi_{12}}\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \cos\theta_{13} & 0 & e^{i\phi_{13}}\sin\theta_{13} \\ 0 & 1 & 0 \\ -e^{-i\phi_{13}}\sin\theta_{13} & 0 & \cos\theta_{13} \end{pmatrix},
$$

$$
\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & e^{i\phi_{23}}\sin\theta_{23} \\ 0 & -e^{-i\phi_{23}}\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \quad (6.60)
$$

where the ϕ angles are chosen so that $O_{ij}e^{-i\phi_{ij}}$ is real and

$$
\tan 2\theta_{ij} = -\frac{2O_{ij}e^{-i\phi_{ij}}}{O_{ii} - O_{jj}}. \quad (6.61)
$$

The problem is that zeroing one of the off diagonal elements will change the other off diagonal elements in those rows and columns. However the magnitude of the sum of the squares of the off-diagonal elements always decreases if the process is repeated[2], and the method converges.

---

[2]The sum of the magnitudes squared of the off diagonal elements of a Hermitian matrix is

$$
\sum_i \sum_{j\neq i} A_{ij}A_{ij}^* = \sum_i \sum_{j\neq i} A_{ij}A_{ji} = \mathrm{tr}A^2 - \sum_i A_{ii}A_{ii}. \quad (6.62)
$$

For the transformed matrix, the trace of A^2 does not change. For a Jacobi step, only two diagonal elements change, and they change exactly like they would for diagonalizing a 2 by 2 matrix. Since in that case, the off diagonal elements go to zero, the sum of the squares of the two changed diagonal elements must be greater than it was before the transformation. Therefore the magnitude of the negative term in Eq. 6.62 always increases, and the sum of the magnitude squares of the off diagonal terms always decreases.

The Jacobi method therefore sweeps through all the off-diagonal elements repeatedly until they are small enough that they can be ignored. At the same time, we can multiply the transformation matrices together as we use them since this gives the complete transformation matrix we have used. When the original matrix has been transformed to be diagonal, its diagonal elements are the eigenvalues and the transformation matrix gives the eigenvectors.

6.9 Using Symmetries to Simplify Diagonalization

If two operators commute,

$$[A, B] = AB - BA = 0 \tag{6.63}$$

they can be simultaneously diagonalized, that is, they can share the same eigenvectors. To see this look at the eigenstates of A,

$$A|\psi_n^A\rangle = \lambda_n^A|\psi_n^A\rangle . \tag{6.64}$$

Multiplying by B we have

$$BA|\psi_n^A\rangle = \lambda_n^A B|\psi_n^A\rangle . \tag{6.65}$$

If the operators commute, this becomes

$$A\left[B|\psi_n^A\rangle\right] = \lambda_n^A \left[B|\psi_n^A\rangle\right] \tag{6.66}$$

and we see that $B|\psi_n^A\rangle$ is an eigenstate of A with eigenvalue λ_n^A. If λ_n^A is not degenerate, then there is only one eigenvector of A with that eigenvalue, so $B|\psi_n^A\rangle$ must be proportional to $|\psi_n^A\rangle$ or

$$B|\psi_n^A\rangle = \lambda_n^B|\psi_n^A\rangle \tag{6.67}$$

and $|\psi_n^A\rangle$ is an eigenvector of B. If λ_n^A is degenerate, then $B|\psi_n^A\rangle$ must be a linear combination of those degenerate states. If we write out the matrix representation of B in the eigenstates of A, it will be block diagonal with B only coupling degenerate states of A together. Diagonalizing the block diagonal B will only mix the states of A that were degenerate, so they will remain eigenvectors of A. Therefore, in all cases, we can simultaneously diagonalize A and B. To reiterate, if A and B commute, we can diagonalize A, finding its eigenvectors and eigenvalues. Writing the matrix elements of B in this eigenbasis of A gives a block diagonal representation of B. Diagonalizing each of the blocks, i.e. finding their eigenvectors and eigenvalues gives simultaneous eigenvectors of A and B.

The power of this method is that if a symmetry operation like rotations, translations, or parity commutes with our Hamiltonian, we can first diagonalize the symmetry operator. This is often much easier. The Hamiltonian in the symmetry eigenstates will then be block diagonal with the elements in each block having the same symmetry operator eigenvalue. The block diagonal matrix is then easier to diagonalize.

Also notice that if one of the operators is unitary, then if $UO - OU = 0$, we can multiply by U^{-1} to get $U^{-1}OU = O$. This shows that a symmetry operator commutes with O if O is invariant under the symmetry transformation. Often you can immediately see that a

Hamiltonian or other operator is invariant under a symmetry transformation and you don't need to go through the detailed calculation of the commutator to see that it must be zero.

Here's an example. If we had a 4 states in a line and imagine a parity operation so that 1 goes to 4, 2 goes to 3, 3 goes to 2, and 4 goes to 1. In matrix representation this parity operator is

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{6.68}$$

The parity operator is interesting in that it is both unitary and Hermitian, as you can check.

A general Hamiltonian would be

$$\begin{pmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{12}^* & H_{22} & H_{23} & H_{24} \\ H_{13}^* & H_{23}^* & H_{33} & H_{34} \\ H_{14}^* & H_{24}^* & H_{34}^* & H_{44} \end{pmatrix}. \tag{6.69}$$

For the commutator to be zero we must have

$$HP - PH = 0 = \begin{pmatrix} H_{14} & H_{13} & H_{12} & H_{11} \\ H_{24} & H_{23} & H_{22} & H_{12}^* \\ H_{34} & H_{33} & H_{23}^* & H_{13}^* \\ H_{44} & H_{34}^* & H_{24}^* & H_{14}^* \end{pmatrix} - \begin{pmatrix} H_{14}^* & H_{24}^* & H_{34}^* & H_{44} \\ H_{13}^* & H_{23}^* & H_{33} & H_{34} \\ H_{12}^* & H_{22} & H_{23} & H_{24} \\ H_{11} & H_{12} & H_{13} & H_{14} \end{pmatrix}. \tag{6.70}$$

So that the most general form that is invariant under parity $H = P^{-1}HP$ (but notice that $P^{-1} = P$), has H_{14} and H_{23} real and

$$\begin{pmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{12}^* & H_{22} & H_{23} & H_{13}^* \\ H_{13}^* & H_{23} & H_{22} & H_{12}^* \\ H_{14} & H_{13} & H_{12} & H_{11} \end{pmatrix}. \tag{6.71}$$

Since H and P commute we diagonalize the simpler operator P first.

If we reorder the states, P is block diagonal since state 1 only couples to state 4, and 2 couples to state 3. Both sub matrices are of the form of Eq. 6.43, and the four eigenvectors are

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}. \tag{6.72}$$

The first two states have eigenvalues 1 and the second have eigenvalues -1, that is they are the even and odd parity states. If we now transform the Hamiltonian to this basis, it will be block diagonal with two, two by two, blocks. Labeling the parity states with p_k and p_m, the column vectors above are $\langle i|p_m \rangle$. To get H in the standard block diagonal form, we want to

keep the eigenvectors with the same eigenvalues adjacent. So we order the parity states as above. We can write

$$\langle p_k|H|p_m\rangle = \sum_{i,j}\langle p_k|j\rangle\langle j|H|i\rangle\langle i|p_m\rangle \tag{6.73}$$

which is

$$
\frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix}
\begin{pmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{12}^* & H_{22} & H_{23} & H_{13}^* \\ H_{13}^* & H_{23} & H_{22} & H_{12}^* \\ H_{14} & H_{13} & H_{12} & H_{11} \end{pmatrix}
\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}
$$

$$
= \begin{pmatrix} H_{11}+H_{14} & H_{12}+H_{13} & 0 & 0 \\ H_{12}^*+H_{13}^* & H_{22}+H_{23} & 0 & 0 \\ 0 & 0 & H_{11}-H_{14} & H_{12}-H_{13} \\ 0 & 0 & H_{12}^*-H_{13}^* & H_{22}-H_{23} \end{pmatrix}. \tag{6.74}
$$

We can now easily find the eigenvectors and eigenvalues of these two, two by two blocks. These eigenvectors are the eigenvectors in the parity operator eigenbasis. That is we will solve the equation

$$H|\psi_n\rangle = E_n|\psi_n\rangle \tag{6.75}$$

by writing

$$\sum_m \langle p_k|H|p_m\rangle\langle p_m|\psi_n\rangle = E_n\langle p_k|\psi_n\rangle \tag{6.76}$$

The resulting eigenvectors of this matrix are $\langle p_k|\psi_n\rangle$. To convert to the original basis, we need to transform back with

$$\langle i|\psi_n\rangle = \sum_k \langle i|p_k\rangle\langle p_k|\psi_n\rangle. \tag{6.77}$$

6.10 Defective Matrices

Rarely in physics do we come across defective matrices. These are N by N matrices that do not have N eigenvectors. Hermitian and Unitary matrices are never defective, so quantum mechanics is nearly immune to this disease. However, here is an example of a defective matrix

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \tag{6.78}$$

We can find its eigenvalues

$$\mathrm{Det}\begin{pmatrix} -\lambda & 1 \\ 0 & -\lambda \end{pmatrix} = \lambda^2 = 0 \tag{6.79}$$

so their are two degenerate eigenvalues. Defective matrices always have degenerate eigenvalues. Let's look for the right eigenvectors with eigenvalue zero,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0. \tag{6.80}$$

Multiplying out, we get the equation $\beta = 0$. The only eigenvector is then $\alpha = 1$, $\beta = 0$. Also notice that if the zero elements are infinitesimally different than zero, the matrix will no longer be defective. In the few cases in physics where defective matrices appear, you can take a limit from the non defective case to see what is going on.

6.11 Problems

Problem 6.1. Show that an operator that is both Hermitian and unitary has eigenvalues that can only be ± 1. Show that its square is the identity operator.

Problem 6.2. A particle of mass m is in a 1-dimensional infinite square well potential so that $-\frac{L}{2} < x < \frac{L}{2}$. The even parity energy eigenstates of the square well are written as $|n^{(e)}\rangle$ and the odd parity energy eigenstates are $|n^{(o)}\rangle$. Below, parts a. and b. are to point you in the right direction. Parts c. and d. are the heart of the problem.

a. Write down the normalized eigenfunctions $\langle x|n^{(e)}\rangle$ and $\langle x|n^{(o)}\rangle$ and the corresponding energy eigenvalues.

b. The particle at $t = 0$ is in the state $|\psi(0)\rangle$

$$|\psi(0)\rangle = \int_{-L/2}^{L/2} dx|x\rangle \frac{1}{(2\pi\sigma^2)^{1/4}} e^{-\frac{x^2}{4\sigma^2} + ikx} \tag{6.81}$$

where you can assume that $L \gg \sigma$. Calculate the coefficients $\langle n^{(e)}|\psi(0)\rangle$ and $\langle n^{(o)}|\psi(0)\rangle$.

c. Use your coefficients to make a set of plots of the probability density $P(x)$ of finding a particle at x as a function of time. Plot $P(x)L$ versus x/L for the values of $\tau = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$ where

$$t = \frac{mL}{\hbar k}\tau, \tag{6.82}$$

and $\sigma = 0.02L$, $k = \frac{100}{L}$.

d. Discuss qualitatively, the peak position, the width and the oscillations of the 5 plots. Compare the peak position to the position of a classical particle with velocity $\hbar k/m$.

e. Repeat the calculations of part c with τ values of 127.1, 127.3, 127.5. Why has the wave packet regenerated?

Problem 6.3. The Hamiltonian for a particle of mass m separates in x, y and z and the y and z states do not change so we can study this physical particle as a particle of mass m in a 1-dimensional potential

$$H = \frac{p_x^2}{2m} + v(x) \tag{6.83}$$

where

$$v(x) = \begin{cases} \lambda\delta(x) & |x| < \frac{L}{2} \\ \infty & |x| > \frac{L}{2} \end{cases}. \tag{6.84}$$

Since this Hamiltonian commutes with the parity operator $P = \int_{-L/2}^{L/2} |x\rangle\langle -x|$, we can find eigenvectors of P and H, $|n^{(e)}\rangle$, with even parity (parity eigenvalue 1), and eigenvectors of P and H, $|n^{(o)}\rangle$, with odd parity (parity eigenvalue -1).

Below, parts a. and b. are to point you in the right direction. Parts c. and d. are the heart of the problem.

a. (10 percent) Write down the normalized eigenfunctions $\langle x|n^{(e)}\rangle$ and $\langle x|n^{(o)}\rangle$ in terms of $k_n^{(e)}$, and $k_n^{(o)}$ where $E_n^{(e),(o)} = \frac{\hbar^2 [k_n^{(e),(o)}]^2}{2m}$ and give equations which define $k_n^{(o)}$ and $k_n^{(e)}$.

b. (10 percent) The particle at $t = 0$ is in the state $|\psi(0)\rangle$

$$|\psi(0)\rangle = \int_{-L/2}^{L/2} dx|x\rangle \frac{1}{(2\pi\sigma^2)^{1/4}} e^{-\frac{(x+L/4)^2}{4\sigma^2} + ikx} \tag{6.85}$$

where you can assume that $L \gg \sigma$. Calculate the coefficients $\langle n^{(e)}|\psi(0)\rangle$ and $\langle n^{(o)}|\psi(0)\rangle$.

c. (40 percent) Use your coefficients to make a set of plots of the probability density $P(x)$ of finding a particle at x as a function of time. Plot $P(x)L$ versus x/L for the values of $\tau = 0, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$ where

$$t = \frac{mL}{\hbar k}\tau, \tag{6.86}$$

and $\sigma = 0.04L$, $k = \frac{100}{L}$, and $\frac{mL\lambda}{\hbar^2} = 50$.

d. (40 percent) Discuss qualitatively, the peak positions, the width and the oscillations of the 6 plots. Compare the peak positions to the position of a classical particle with velocity $\hbar k/m$.

Problem 6.4. A benzene molecule has 6 carbon atoms and 6 hydrogen atoms arranged in a ring. A picture borrowed from Wikipedia is shown below:
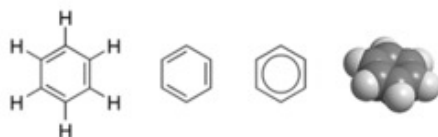
A carbon atom has 6 electrons, and from our knowledge of hydrogenic orbitals, we expect 2 electrons to be in $n = 1$, 1s like orbitals, the other 4 electrons are in the $n = 2$, 2s and 2p like orbitals. These are degenerate in hydrogen, and close in energy in carbon so that in the molecule, the coulomb interaction of the other atoms causes them to be mixed together strongly. They are hybridized into what are called sp$_2$ orbitals that are essentially mixtures of the 2p$_x$, 2p$_y$ and 2s atomic orbitals, leaving the 2p$_z$ orbital which is called a π orbital. The 3 sp$_2$ hybrid orbitals point outward with angles of 120° between them. These then mix together with their nearest neighbors to form the backbone chemical bonds that hold the molecule together in the hexagonal form shown.

There are 42 electrons total. Twelve of these are taken up by the carbon inner core 1s orbitals. Three of each carbon's sp$_2$ orbitals and one hydrogenic 1s orbital from each hydrogen form the backbone with 24 more electrons. The remaining 6 electrons are in the π orbitals. These are weakly interacting and weakly bonding so they are the orbitals that participate the most when the benzene is probed, for example, by putting it into external

fields. We will assume the electrons in the π orbitals are noninteracting with each other, and the matrix elements for a single electon, in the basis of the 6 π orbitals can be written as

$$H = \begin{pmatrix} \epsilon_0 & -t & 0 & 0 & 0 & -t \\ -t & \epsilon_0 & -t & 0 & 0 & 0 \\ 0 & -t & \epsilon_0 & -t & 0 & 0 \\ 0 & 0 & -t & \epsilon_0 & -t & 0 \\ 0 & 0 & 0 & -t & \epsilon_0 & -t \\ -t & 0 & 0 & 0 & -t & \epsilon_0 \end{pmatrix}. \tag{6.87}$$

This is often called a Hückel model. The Hamiltonian has an on-site energy, ϵ_0, and a hopping term $-t$ which, from the time dependent Schrödinger equation, we see is proportional to the amplitude per unit time for an electron to move to an adjacent carbon atom.

a. Write a matrix T representating the unitary operator that corresponds to relabeling the states $1 \to 2$, $2 \to 3$, $3 \to 4$, $4 \to 5$, $5 \to 6$, $6 \to 1$.

b. Show that the T matrix commutes with the H matrix.

c. Calculate analytically the eigenvalues and eigenvectors of T.

d. Using your result for part b and c write down the eigenvectors of H immediately, and then calculate the eigenvalues of H.

e. Using the Pauli exclusion principle and the fact that electrons are fermions with 2 spin states, show that the ground state of the 6 π electrons has energy $6\epsilon_0 - 8t$ and the first excited state has energy $6\epsilon_0 - 6t$ with degeneracy 16.

Problem 6.5. A particle is trapped in a box with constant cross section. The Hamiltonian separates, and the cross sectional states remain the same throughout, so we can equivalently solve just a one-dimensional Schrödinger equation. Initially there is a barrier across the center of the box, so that the one-dimensional Hamiltonian is

$$H(\lambda) = \frac{p^2}{2m} + V_0(x) \tag{6.88}$$

where

$$V_0(x) = \begin{cases} \lambda\delta(x) & |x| < L/2 \\ \infty & |x| > L/2 \end{cases}, \tag{6.89}$$

with $\lambda \to \infty$. The system is prepared by placing the particle in the region $0 < x < L/2$, and the particle is allowed to couple to the rest of the universe and decays into the lowest energy state contained in $0 < x < L/2$. Sometime after that, at time $t = 0$, the barrier is removed. That is λ becomes zero instantly.

a. Show that the state of the system for time $t > 0$ can be written as

$$|\psi(t)\rangle = \sum_n e^{-\frac{i}{\hbar}E_n t}|n\rangle\langle n|\psi(0)\rangle \tag{6.90}$$

where E_n and $|n\rangle$ are the eigenvalues and eigenstates of $H(\lambda = 0)$. Give the values of E_n and $\langle n|\psi(0)\rangle$.

b. At time $t = \tau 4mL^2/(\pi\hbar)$, the barrier between the two equal halves of the box is reinstated, (i.e. $\lambda \to \infty$ again). Calculate the probability as a function of τ, that the particle will be measured in the regions $0 < x < L/2$.[3]

Plot this probability as a function of τ, for $0 < \tau < 1$. What is the relationship of this probability for $\tau > 1$ to its value in the range $0 < \tau < 1$? What is the value of the probability for $\tau = 1/2$?

Problem 6.6. Polyactelyne is a hydrocarbon chain with equal numbers of hydrogen and carbon atoms. Under certain circumstances it becomes an organic conductor (Heeger, Mac-Diarmid and Shirakawa won the 2000 Nobel prize in chemistry for this work). The chain has two basic forms, cis and trans depending on the dihedral angles of the bonds. Physical chains can combine both forms as well as have other conformations. We will look at the trans form shown in fig. 6.1.
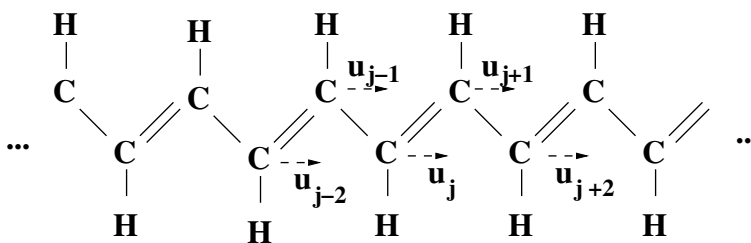
Figure 6.1: A schematic of the structure of trans-polyacetylene. The u_j correspond to the vibrational coordinate for the CH group to move along the chain.

To understand the basic physics, we first think about building up an approximate many-electron wave function from single electron orbitals. The carbon atoms have 6 electrons, 2 of these are strongly bound in the $1s$ orbitals and can be neglected in understanding the bonding. The remaining 4 carbon electrons in the atom are in the $2s$ and $2p$ orbitals. These are degenerate for hydrogen, and their energies are close for carbon, so they mix together strongly when bonding. In this case, if we think of the atoms as being in the x-y plane, the $2p_x$, $2p_y$ and $2s$ orbitals "hybridize" to give sp-2 molecular orbitals that point out from the atom at 120° angles. These form bonds with the adjacent carbon sp-2 orbitals and the hydrogen $1s$ orbital. This gives (counting spin up and spin down) 4 molecular orbitals per carbon atom which form the backbone that holds the molecule together. The remaining

---

[3]My result for this is

$$p(\tau) = \frac{1}{2} + \frac{32}{\pi^2} \sum_{n=1}^{\infty} \cos\left([(2n-1)^2 - 4]2\pi\tau\right) \frac{1}{[(2n-1)^2 - 4]^2}. \tag{6.91}$$

65

electron (we have not yet counted the hydrogen's original electron) goes into a state formed from the $2p_z$ carbon orbital which gives a π bond. Since the $2p_z$ orbitals point out of the plane of the molecule, they give weaker bonds than the backbone (or Σ) bonds.

We will first look at a model for the one-electron π-bonded Hamiltonian. If the carbon atoms were far apart, the $2p_z$ orbitals would be single-electron orbital eigenstates. As the carbons atoms are brought closer, these orbitals overlap, and the Hamiltonian will couple them together. The nearest neighbor overlaps will be the largest, so we ignore the others. Sometimes these chains form rings, and the analysis is slightly easier for them, so lets label the carbon atom $2p_z$ orbitals by integers, $j = 0, 1, 2, ..., N - 1$, with N even, and connect them into a ring so that $j = 0$ is the same as $j = N$. Our single electron Hamiltonian for equally spaced carbon atoms would be

$$H^{(1)} = -t \sum_{j=0}^{N-1} [|j\rangle\langle j + 1| + |j + 1\rangle\langle j|] \tag{6.92}$$

The negative sign is conventional. This kind of Hamiltonian is called a tight-binding model in physics and a Hückel model in chemistry.

a. Show that the translation operator $T_1 = \sum_{j=0}^{N-1} |j + 1\rangle\langle j|$ commutes with $H^{(1)}$.

b. Find the eigenvalues and eigenvectors of T_1.

c. Are any of the eigenvalues of T_1 degenerate? What does that tell you about the relationship between the eigenvectors of T_1 and the eigenvectors of $H^{(1)}$?

d. Use the result of part c. to calculate the eigenvalues of $H^{(1)}$. For the case where the number N becomes large, calculate the ground-state energy per particle by adding up the energies of the lowest N states (recall electrons are fermions). Remember each orbital can have spin up or spin down.

Problem 6.7. Polyacetylene undergoes a Peierls transition. That is, the chain dimerizes. In the schematic of the trans form, shown in figure 6.1, assume that u_{j-2}, u_j, u_{j+2} etc. are small and positive and u_{j-1}, u_{j+1} etc. are small and negative. The double bonds shown will become shorter than the single bonds shown. There is then additional neat physics. For example if we start with an equally spaced chain, the carbon atoms with odd number labels could all move to the left to dimerize, or they could all move to the right. If different parts of the chain dimerize differently there is a region of mismatch. This turns out to be a soliton excitation (for example a chain with an odd number of carbon atoms will necessary have at least one soliton).

Let's investigate the simplest dimerization. With a ring of an even number N of carbon atoms, moving the even labeled atoms to the right by u and the odd labeled atoms by $-u$ will make the overlap with their nearer neighbor stronger and with their other neighbor weaker.

If we keep up to just the linear term in the change of the overlap contribution, the effective π Hamiltonian becomes

$$H^{(2)} = -\sum_j (t + (-1)^j 2\alpha u) \left[|j\rangle\langle j+1| + |j+1\rangle\langle j| \right] \tag{6.93}$$

a. Show that the translation operator $T_2 = \sum_{j=0}^{N-1} |j+2\rangle\langle j|$ commutes with $H^{(2)}$.

b. Show that T_2 can be written in block diagonal form. Find the eigenvalues and eigenvectors of the blocks.

c. Are any of the eigenvalues of T_2 degenerate? What does that tell you about the relationship between the eigenvectors of T_2 and the eigenvectors of $H^{(2)}$?

d. Use the result of part c. to calculate the eigenvalues of $H^{(2)}$. For the case where the number N becomes large, calculate the ground-state energy per particle by adding up the energies of the lowest N states. My result is:

$$E_0(u) = -\frac{4t}{\pi} E\left(1 - \frac{4\alpha^2 u^2}{t^2}\right). \tag{6.94}$$

where

$$E(m) = \int_0^{\pi/2} d\phi \sqrt{1 - m\sin^2\phi} \tag{6.95}$$

is the complete elliptic integral of the second kind.[4]

Problem 6.8. We would expect the backbone to give an energy that is minimized when the bond lengths are all equal. The overall zero of the energy is unimportant so in a normal mode formalism we would take the backbone energy to be

$$\sum_j \frac{1}{2} K(u_{j+1} - u_j)^2. \tag{6.96}$$

The energy per particle for our case is then $E_b(u) = 2Ku^2$ Detailed electronic structure calculations or fits to experiment give values of the parameters. Use $t = 2.5$ eV, $K = 21$ eV/A^2, $\alpha = 4.1$ eV/A.

a. Plot the energy $E_b(u) + E_0(u)$ as a function of u to show that, with the parameters given, the solution $u = 0$ is a local maximum (this is the Peierls instability).

b. Find the value of u from your plot that gives a minimum in the energy.

Problem 6.9. The equations of motion for each of the following physical problems have similar structure:

$$\frac{\partial^n X_i(t)}{\partial t^n} = -\alpha \sum_j A_{ij} X_j(t) \tag{6.97}$$

---

[4]Some references use $E(k) = \int_0^{\pi/2} d\phi \sqrt{1 - k^2 \sin^2\phi}$. Be careful.

where α is a constant, A_{ij} is a Hermitian or real symmetric matrix, n is 1 or 2 depending on whether the equations of motion are first or second order in time, and $X_i(t)$ are the desired quantities.

i. The 1-dimensional Schrödinger equation on a grid is illustrated in Baym chapter 3. For the "infinite square well" on a grid of spacing Δx, the Hamiltonian is

$$H = \frac{\hbar^2}{2m\Delta x^2} \left[2\sum_{j=1}^{N} |x_j\rangle\langle x_j| - \sum_{j=1}^{N-1} (|x_j\rangle\langle x_{j+1}| + |x_{j+1}\rangle\langle x_j|) \right] . \qquad (6.98)$$

You are given the complex numbers $\psi_i(t=0)$ that describe an initial state, $|\psi(t=0)\rangle = \sum_{i=1}^{N} \psi_i(0)|x_i\rangle$. You need to calculate $|\psi(t)\rangle = \sum_{i=1}^{N} \psi_i(t)|x_i\rangle$ by solving for the $\psi_i(t)$.

ii. N masses of mass m are constrained to move only along x. Each is connected by a spring of spring constant k to its two nearest neighbors. Masses 1 and N which have only one neighbor, are also connected by similar springs to stationary walls. You are given their initial displacements from their equilibrium positions $x_i(t=0)$ and their initial velocities $v_i(t=0)$. You need to calculate $x_i(t)$.

iii. N equal capacitors of capacitance C are arranged in a line and have one terminal connected to ground. The other terminal of each capacitor is connected to the corresponding terminal of the next capacitor by an inductor of inductance L. The capacitors 1 and N at the ends also have inductors from their ungrounded terminals to ground. You are given the inital voltage $V_i(t=0)$ across and the initial current $I_i(t=0)$ through each capacitor. You need to calculate $V_i(t)$.

a. Show that the three problems with appropriate choice for α lead to identical matrices A_{ij}.

b. Find the eigenvalues and eigenvectors of the A matrix analytically[5]

c. Show how you express the solutions for the desired quantities in terms of the initial conditions, the eigenvalues and eigenvectors.[6]

---

[5]One way is to look at the periodic system with the ends tied together instead of the boundary conditions given. The resulting Hamiltonian is invariant under relabeling $i \to i+1$, which is like a translation, and also invariant under reflection around any point. You can first diagonalize the translation operator, then find the combinations that give parity eigenstates, and note that the odd parity states of a roughly twice as large system give the boundary conditions needed.

[6]For example, my result for the mass and spring problem is

$$x_k(t) = \frac{2}{N+1} \sum_{j=1}^{N} \sum_{n=1}^{N} \sin\left(\frac{kn\pi}{N+1}\right) \sin\left(\frac{jn\pi}{N+1}\right) \left[\cos(\omega_n t)x_j(0) + \frac{1}{\omega_n}\sin(\omega_n t)v_j(0) \right]$$

$$\omega_n = \sqrt{\frac{k}{m}} \left[2\sin\left(\frac{n\pi}{2(N+1)}\right) \right] . \qquad (6.99)$$

Problem 6.10. A particle moves in two dimensions, x, and y[7] Take the particle to have mass m, and it moves in the potential given by

$$V(x,y) = \begin{cases} 0 & 0 < x < L \text{ and } 0 < y < L \\ \infty & \text{otherwise} \end{cases} \tag{6.100}$$

i.e. it moves freely in a square of side L.

a. Write down the energy eigenvalues and eigenfunctions of the particle in the basis that separates in x and y.

b. Show that the Hamiltonian commutes with the operator, $R_{x=y}$, corresponding to reflection through the line $y = x$.

c. Calculate the matrix elements of $R_{x=y}$ in the basis of part a. to find simultaneous eigenfunctions of $R_{x=y}$ and the Hamiltonian.

d. Use your results of part c. to find *all* the eigenfunctions for a particle trapped in the triangular cavity

$$V(x,y) = \begin{cases} 0 & 0 < x < L \text{ and } 0 < y < x \\ \infty & \text{otherwise} \end{cases} \tag{6.101}$$

Be sure to make a convincing argument that you have not left out any states.

e. You probably studied wave guides in undergraduate electromagnetism. Relate how the techniques used here can greatly simplify the solution of problem 8.5 part a in J.D. Jackson, *Classical Electrodynamics*. Recognizing connections like this across fields will make you a better physicist and can greatly help you in your research. Do not turn this part in.

---

[7]The physical system corresponds to motion in 3-dimensions, but the motion along z either separates or can be integrated out to give this 2-dimensional problem. For example this might represent the motion of an electron along the plane of the junction of $GaAs$ and $AlGaAs$. The electron feels a strong trapping potential perpendicular to the junction. At the temperatures of interest, the electron will decay to its ground state in the strong trapping potential and never be excited. It's state will then be a product of this bound z state and the state for its motion along x and y. We are dealing with just this latter part.

Chapter 7

Diagonalization Example

7.1 Diagonalization to find the eigenfunctions of a differential equation

As an example let's imagine that we want to calculate the time dependence of a quantum state with a bounded harmonic oscillator hamiltonian given by

$$H = \frac{p^2}{2m} + v_0(x) + v(x) = T + v(x) = T + V \tag{7.1}$$

with

$$v_0(x) = \begin{cases} 0 & |x| < \frac{L}{2} \\ \infty & |x| > \frac{L}{2} \end{cases} \tag{7.2}$$

the potential for an infinite square well, and $v(x)$ a harmonic oscillator potential

$$v(x) = \frac{1}{2}m\omega^2 x^2 \,. \tag{7.3}$$

Since T restricts the states to those with $|x| < L/2$, we can restrict $v(x)$ to the same range.

In order to calculate the time dependence, we decide to expand in the eigenstates of H. This problem does not have an analytic solution. Therefore we must calculate these eigenstates numerically. One method would be to integrate the differential equation with, for example, the shooting method using Numerov's method to get the eigenfunctions in the x basis. Here we will use diagonalization in an alternative basis.

First, let's look to see if symmetries can simplify the problem. The potential is symmetric under the interchange $x \leftrightarrow -x$, so we should expect that both T and V commute with the parity operator

$$P = \int_{-L/2}^{L/2} dx \, |x\rangle\langle -x| \,. \tag{7.4}$$

Let's demonstrate this in detail.

Writing V in the x representation gives

$$V = \int_{-L/2}^{L/2} dx |x\rangle v(x)\langle x| \tag{7.5}$$

where x is now a c-number, we have

$$PV - VP = \int_{-L/2}^{L/2} dx \int_{-L/2}^{L/2} dx' |x\rangle \underbrace{\langle -x|x'\rangle}_{\delta(x+x')} v(x')\langle x'| - \int_{-L/2}^{L/2} dx \int_{-L/2}^{L/2} dx' |x\rangle v(x) \underbrace{\langle x|x'\rangle}_{\delta(x-x')}\langle -x'|$$

$$= \int_{-L/2}^{L/2} dx |x\rangle \left[v(-x) - v(x)\right]\langle -x| = 0. \tag{7.6}$$

Working in the x representation for p is simplest. There replacing $v(x)$ above with $-d^2/dx^2$ shows that P also commutes with T.

Since P and T commute they can be simultaneously diagonalized. We can solve the differential equation for the x-representation for the eigenvectors and eigenvalues of T.

The parity operator is hermitian and its own inverse, by inspection. Therefore squaring it must gives the identity. We can check that

$$P^2 = \int_{-L/2}^{L/2} dx \int_{-L/2}^{L/2} dx' |x\rangle \underbrace{\langle -x|x'\rangle}_{\delta(x+x')}\langle -x'| = \int_{-L/2}^{L/2} dx |x\rangle\langle x| = 1 \tag{7.7}$$

so its only possible eigenvalues are ± 1. In fact any operator that is both Hermitian and unitary must have eigenvalues that are both real and are phase factors. This immediately tells you that the eigenvalues can only be ± 1. The eigenvectors of P in the x representation are therefore the even and odd functions in x. The T eigenfunction equation becomes

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\langle x|\psi\rangle = E_n\langle x|\psi\rangle. \tag{7.8}$$

The even solutions are cosines and the odd are sines. Enforcing the boundary conditions gives

$$\langle x|n^{(e)}\rangle = \sqrt{\frac{2}{L}}\cos\left(\frac{[2n-1]\pi x}{L}\right)$$

$$\langle x|n^{(o)}\rangle = \sqrt{\frac{2}{L}}\sin\left(\frac{2\pi nx}{L}\right) \tag{7.9}$$

where normalization is given by setting the integral of the squares to 1.

Notice that the eigenstates of T are nondegenerate. Therefore any method of finding the eigenstates of T will necessarily give parity eigenstates. Since P commutes with V, states with different parity will not be coupled by V, and the resulting representation will be block diagonal.

For our particular case, the eigenvalues are

$$
\lambda_n^{(e)} = \frac{\hbar^2}{2mL^2}\pi^2(2n-1)^2 = \hbar\omega\alpha^{-1}\frac{\pi^2(2n-1)^2}{2}
$$

$$
\lambda_n^{(o)} = \frac{\hbar^2}{2mL^2}4\pi^2 n^2 = \hbar\omega\alpha^{-1}2\pi^2 n^2 \tag{7.10}
$$

where $\alpha = m\omega L^2/\hbar$.

The matrix elements of V are

$$
\langle m^{(e)}|V|n^{(e)}\rangle = m\omega^2\frac{2}{L}\int_{-L/2}^{L/2}dx\, x^2\langle m^{(e)}|x\rangle\langle x|n^e\rangle
$$

$$
= \hbar\omega\alpha\int_{-1/2}^{1/2}dx\, x^2\cos([2n+1]\pi x)\cos([2m+1]\pi x)
$$

$$
\langle m^{(o)}|V|n^{(o)}\rangle = m\omega^2\frac{2}{L}\int_{-L/2}^{L/2}dx\, x^2\langle m^{(o)}|x\rangle\langle x|n^o\rangle
$$

$$
= \hbar\omega\alpha\int_{-1/2}^{1/2}dx\, x^2\sin(2n\pi x)\sin(2m\pi x)\,. \tag{7.11}
$$

You can do these integrals in many ways, for example writing

$$
\int dx\, x^2\sin(\alpha x)\sin(\beta x) = -\frac{d^2}{d\alpha^2}\int dx\,\sin(\alpha x)\sin(\beta x)
$$

$$
= -\frac{1}{2}\frac{d^2}{d\alpha^2}\int dx\,\{\cos[(\alpha+\beta)x] - \cos[(\alpha-\beta)x]\} \tag{7.12}
$$

The last function can be trivially integrated and taking the α derivatives is straightforward. The limit where $\alpha = \beta$ needs to be done carefully and gives the extra term with δ_{mn} in the expressions. The cosine integral can be done similarly.

The result is

$$
\langle m^e|H|n^{(e)}\rangle = \hbar\omega\left\{\delta_{mn}\left[\alpha^{-1}\frac{\pi^2(2n-1)^2}{2} + \alpha\left(\frac{1}{24} - \frac{1}{4\pi^2(2n-1)^2}\right)\right]\right.
$$

$$
\left.+(1-\delta_{mn})\alpha\frac{(-1)^{m+n}(2n-1)(2m-1)}{4\pi^2(m+n-1)^2(n-m)^2}\right\}
$$

$$
\langle m^o|H|n^{(o)}\rangle = \hbar\omega\left\{\delta_{mn}\left[\alpha^{-1}2n^2\pi^2 + \alpha\left(\frac{1}{24} - \frac{1}{16\pi^2 n^2}\right)\right]\right.
$$

$$
\left.+(1-\delta_{mn})\alpha\frac{(-1)^{m+n}mn}{\pi^2(n^2-m^2)^2}\right\} \tag{7.13}
$$

Note that in the case where the integrals are not analytic, we would perform the integrations numerically to get the matrix elements of H.

73

Now that we have expressions for the matrix elements we can use any numerical tools we like to diagonalize the matrix.

We need to give a numerical value to α. For a particular case, we take take $L/2 = 10\sqrt{\hbar/m\omega}$, and we then need to evaluate the appropriate block of the matrix, truncate it so that we accurately calculate all of the states that contribute to our initial state, but have a finite matrix.

To demonstrate the results, here I will calculate the ground state and the 50th excited state of the bounded harmonic oscillator. I decide that 6 place accuracy is sufficient. To demonstrate the accuracy of the result, let's plot the eigenfunctions and compare them to those obtained by integrating the differential equation.

It is convenient to use natural units where $\hbar = m = \omega = 1$.

The 50th excited state is even, so we just need the matrix given by the even states. The eigenstates of H are then

$$|\psi_m\rangle = \sum_n |n^{(e)}\rangle\langle n^{(e)}|\psi_m\rangle \tag{7.14}$$

and converting to the x representation to make a plot gives

$$\psi_m(x) = \langle x|\psi_m\rangle = \sum_n \langle x|n^{(e)}\rangle\langle n^{(e)}|\psi_m\rangle. \tag{7.15}$$

I found 40 even square well states was more than adequate to get the 50th excited state with energy 50.9459 $\hbar\omega$ in agreement with the result from the applet integration. Adding more states does not change the results.

The plot of the state is shown in the figure. The result is indistinguishable from the applet plot.

Below I give two ways to solve this problem. I will first show my solution using maple. I wrote this maple script since using it or a similar package will probably be the easiest way if you have no previous programming experience.

```
with(linalg):
with(plots):

seth := proc(n,alpha)
   local h,i,j;
   h := matrix(n,n,0);
   for i from 1 by 1 to n do
      for j from 1 by 1 to n do
         if (i = j) then
            h[i,j] := evalf(Pi^2*(2*i-1)^2/(2*alpha)
                    +alpha*(1/24-1/(4*Pi^2*(2*i-1)^2)))
         else
            h[i,j] := evalf(alpha*(-1)^(i+j)*(2*i-1)*(2*j-1)/
                    (4*Pi^2*(i+j-1)^2*(i-j)^2))
         end if;
```
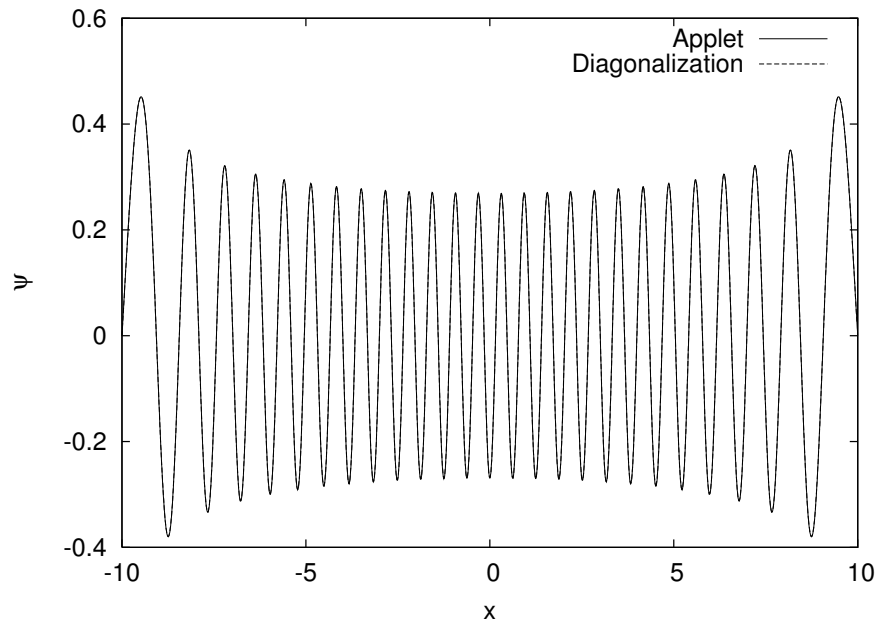
Figure 7.1: The fiftieth excited state from the diagonalization using 40 even square well basis states compared to the applet result. The results are indistinguishable on the scale of the plot.

```
      end do;
   end do;
   return h;
end:

evalvec := proc(vec,n,x,L)
   local j,psi;
   psi := 0;
   for j from 1 by 1 to n do
      psi := psi+evalf(sqrt(2/L)*cos(Pi*x*(2*j-1)/L))*vec[j]
   end do;
   return psi;
end:

L := 20:
n := 40:
h := seth(n,L^2):
vals := evalf(Eigenvals(h,vecs)):
print(vals[1],vals[26]);
```

```
psi0 := x -> evalvec(subvector(vecs,1..n,1),n,x,L):
psi50 := x -> -evalvec(subvector(vecs,1..n,26),n,x,L):
plot([psi0(x),exp(-x^2/2)/Pi^(1/4)],x=-L/2..L/2,axes=BOXED
    ,style=[line,point],labels=[x,psi_0]);

napplet := 5000:
fd := fopen("applet50.out",READ):
for i from 1 by 1 to 8 do
    readline(fd);
end do:
psinorm := 0:
applet := array(1..napplet,1..2):
for i from 1 by 1 to napplet do
    line := readline(fd);
    data := sscanf(line,"%lg %lg");
    applet[i,1] := op(1,data);
    applet[i,2] := op(2,data);
    psinorm := psinorm+applet[i,2]^2:
end do:
psinorm := 1.0/sqrt(psinorm*L/napplet):
for i from 1 by 1 to napplet do
    applet[i,2] := applet[i,2]*psinorm;
end do:

p1 := plot(convert(applet,listlist),x=-L/2..L/2,axes=BOXED):
p2 := plot(psi50(x),x=-L/2..L/2,axes=BOXED,style=point,labels=[x,psi_50]
    ,color=green):
display({p1,p2});
```

to set up the matrix, diagonalize it and plot the ground and fiftieth excited states. I included the analytic ground state, and read in and normalized the applet fiftieth excited state to show that they are identical

If you already know a programming language which has linear algebra libraries available, it will probably be easier to just write a little code. Since the Intel fortran compiler I use comes with LAPACK libraries, I wrote the following code to do the same thing as above:

```
program hdiag2
use diagonalize
implicit none
integer, parameter :: i4=selected_int_kind(9)
integer, parameter :: r8=selected_real_kind(15,9)
real(kind=r8) :: pi,el2,alpha,dx,x,psi,fun
integer(kind=i4) :: n,iparity,i,j,k,is,js,ngrid
real(kind=r8), allocatable :: h(:,:),vector(:,:),value(:)
```

```fortran
pi=4.0_r8*atan(1.0_r8)
read (5,*) n        ! number of states
read (5,*) el2      ! L/2 in units of sqrt(hbar/m w)
read (5,*) iparity  ! parity +1 or -1
read (5,*) ngrid    ! grid points for wave functions
allocate(h(n,n),vector(n,n),value(n))
alpha=4.0_r8*el2**2
select case (iparity)
   case (1) ! even parity
   is=1
   do i=1,n
      is=-is
      h(i,i)=0.5_r8*pi**2*(2*i-1)**2/alpha &
         +alpha*(1.0_r8/24.0_r8-1.0/(4.0_r8*pi**2*(2*i-1)**2))
      js=is
      do j=i+1,n
         js=-js
         h(i,j)=is*js*alpha*(2*i-1)*(2*j-1) &
            /(4.0_r8*pi**2*(i+j-1)**2*(i-j)**2)
         h(j,i)=h(i,j)
      enddo
   enddo
   case (-1) ! odd parity
   is=1
   do i=1,n
      is=-is
      h(i,i)=pi**2*2.0_r8*i**2/alpha &
         +alpha*(1.0_r8/24.0_r8-1.0/(16.0_r8*pi**2*i**2))
      js=is
      do j=i+1,n
         js=-js
         h(i,j)=alpha*is*js*i*j/(pi**2*(i**2-j**2)**2)
         h(j,i)=h(i,j)
      enddo
   enddo
   case default
      write (6,'(''Illegal iparity value'',i10)') iparity
      stop
end select
call eigen(h,n,vector,value)
do i=1,n
   write (6,'(''Eigenvalue = '',1p,e20.10)') value(i)
```

```
            write (6,'(''Eigenvector = '')')
            dx=el2*2.0_r8/(ngrid+1)
            do j=0,ngrid
                fun=0.0_r8
                x=j*dx-el2
                do k=1,n
                    select case (iparity)
                    case (1)
                        psi=cos((2*k-1)*pi*x/(2.0_r8*el2))/sqrt(el2)
                    case (-1)
                        psi=sin(k*pi*x/el2)/sqrt(el2)
                    end select
                    fun=fun-vector(k,i)*psi
                enddo
            write (6,'(1p,2e20.10)') x,fun
            enddo
        enddo
    end program hdiag2
```

I wrote this quick and dirty interface to LAPACK which overloads the eigen routine to call
either the real symmetric or hermitian LAPACK routines:

```
module diagonalize
    implicit none
    integer, private, parameter :: i4=selected_int_kind(9)
    integer, private, parameter :: r8=selected_real_kind(15,9)

interface eigen
    module procedure eigenh,eigenrs
end interface

contains
    subroutine eigenh(h,n,vector,value)
!
! call lapack routine for hermitian matrix
!
    integer(kind=i4) :: n,info
    complex(kind=r8) :: h(n,n),vector(n,n),work((n+3)*n)
    real(kind=r8) :: value(n),rwork(3*n)
    vector=h
    call zheev('v','u',n,vector,n,value,work,size(work),rwork,info)
    if (info.ne.0) then
        write (6,'('' error in zheev '',i10)') info
        stop
```

```
      endif
   end subroutine eigenh

   subroutine eigenrs(s,n,vector,value)
   integer(kind=i4) :: n,info
   real(kind=r8) :: s(n,n),vector(n,n),value(n),work((n+3)*n)
!
! call lapack routine for real symmetric matrix
!
   vector=s
   call dsyev('v','u',n,vector,n,value,work,size(work),info)
   if (info.ne.0) then
      write (6,'('' error in dsyev '',i10)') info
      stop
   endif
   end subroutine eigenrs
end module diagonalize
```

The input data was

```
40          #number of states
10          # L/2
1           # parity
1000        # grid
```

To normalize the output of the applet I used the awk script

```
#!/bin/sh
awk '
BEGIN {
   L=20
   i=0
   norm=0
}
{
   if ($0 ~ /#/ ) next
   x[i] = $1
   psi[i] = $2
   norm += psi[i]*psi[i]
   i++
}

END {
   norm=1.0/sqrt(norm*L/i)
```

```
    for (j=0;j<i;j++) print x[j],psi[j]*norm
}' -
```

and for the plot I used gnuplot with the commands

```
#set term post eps enhanced "Helvetica" 20
#set output 'p3f1.eps'
#set ylabel "{/Symbol=20 y}"
set ylabel "psi"

set yrange [-.4:.6]
set xlabel "x"
plot "<normer < applet50.out" title "Applet" w l, \
'temp.out' title "Diagonalization" w p
```

7.2   Another Example

Let's imagine that we have a system in our lab that we allow to come into thermal equilibrium at room temperature, and we know that the excited state energy levels will be separated from the ground state by several electron volts. Since room temperature corresponds to $k_b T \simeq \frac{1}{40}$eV, we would expect our system to be in its ground state with a high probability.

   We then turn on and off abruptly, for a time t, an electric field that is sensibly close to uniform across our system. We ask what is the probability that the system is in one of its excited states? Or we might ask, what is the probability that the system is still in its ground state after? We can ask many such questions. Let's look at the answer to the first two questions.

   To be specific, let's take our initial system to be modeled by a 1-dimensional infinite square well Hamiltonian. That is

$$H_0 = \frac{p^2}{2m} + v_0(x) \tag{7.16}$$

with

$$v_0(x) = \left\{ \begin{array}{ll} 0 & |x| < \frac{L}{2} \\ \infty & |x| > \frac{L}{2} \end{array} \right. . \tag{7.17}$$

I will write the eigenfunctions of H_0 as

$$\langle x|n \rangle = \left\{ \begin{array}{ll} \sqrt{\frac{2}{L}} \cos\left(\frac{n\pi x}{2L}\right) & n \text{ odd} \\ \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{2L}\right) & n \text{ even} \end{array} \right. \tag{7.18}$$

When the electric field is turned on, the Hamiltonian changes to

$$H = H_0 + V \tag{7.19}$$

where
$$V = -q\mathcal{E}x \tag{7.20}$$

for a constant electric field $\mathcal{E}$ in the x direction and a particle of charge q.

The Schrödinger equation
$$H(t)|\psi(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi(t)\rangle, \tag{7.21}$$

shows that if $H(t)$ changes discontinuously at time t, the time derivative of $|\psi(t)\rangle$ changes discontinuously, but unless $H(t)$ contains a delta function in time, integrating both sides infinitesimally across the discontinuity shows that $|\psi(t)\rangle$ is continuous. That is,
$$\lim_{\epsilon\to 0}\int_{t_0-\epsilon}^{t_0+\epsilon} dt\, H(t)|\psi(t)\rangle = 0 \tag{7.22}$$

unless $H(t)$ contains a delta function at $t = t_0$. Since it doesn't,
$$\begin{aligned} 0 &= \lim_{\epsilon\to 0}\int_{t_0-\epsilon}^{t_0+\epsilon} dt\, H(t)|\psi(t)\rangle = \lim_{\epsilon\to 0}\frac{\hbar}{i}\int_{t_0-\epsilon}^{t_0+\epsilon} dt\,\frac{\partial}{\partial t}|\psi(t)\rangle \\ &= \lim_{\epsilon\to 0}|\psi(t+\epsilon)\rangle - |\psi(t-\epsilon)\rangle \end{aligned} \tag{7.23}$$

so that $|\psi(t)\rangle$ is continuous.

Therefore the probability amplitude of finding the particle in the eigenstate $|n\rangle$ of H_0 after the electric field is turned on for time t is
$$\langle n|\psi(t)\rangle = \langle n|e^{-\frac{i}{\hbar}Ht}|1\rangle \tag{7.24}$$

since $|\psi(0)\rangle = |1\rangle$, i.e. the particle is initially in the ground state of H_0.

We evaluate the exponential by inserting a complete set of eigenstates of H,
$$H|\Psi_k\rangle = E_k|\Psi_k\rangle \tag{7.25}$$

so that
$$\langle n|\psi(t)\rangle = \sum_k \langle n|e^{-\frac{i}{\hbar}E_k t}|\Psi_k\rangle\langle\Psi_k|1\rangle = \sum_k e^{-\frac{i}{\hbar}E_k t}\langle n|\Psi_k\rangle\langle\Psi_k|1\rangle. \tag{7.26}$$

One way to find the needed matrix elements is to write
$$\sum_m \langle n|H|m\rangle\langle m|\Psi_k\rangle = E_k\langle n|\Psi_k\rangle \tag{7.27}$$

This is an eigenvalue/vector problem. We need the matrix elements of H, and these are readily calculated for our case to be
$$\langle n|H|m\rangle = \delta_{nm}\frac{\hbar^2 n^2\pi^2}{2mL^2} - q\mathcal{E}\int_{-L/2}^{L/2} dx\,\langle n|x\rangle x\langle x|m\rangle. \tag{7.28}$$

81

Substituting the expressions for the infinite square well eigenfunctions, the integral is straight-forward and gives

$$\langle n|H|m\rangle = \delta_{nm}\frac{\hbar^2 n^2\pi^2}{2mL^2} + \frac{8q\mathcal{E}L}{\pi^2}\begin{cases} (-1)^{(n+m-1)/2}\frac{nm}{(n^2-m^2)^2} & n+m \text{ odd} \\ 0 & n+m \text{ even}\end{cases}\tag{7.29}$$

We truncate the matrix at a high enough energy that we know it can be safely ignored (we can check by recalculating with a different cut off), and calculate its eigenvalues and eigenvectors numerically as we did in the previous example. This will give E_k and $\langle n|\Psi_k\rangle$. We then substitute these values into Eq. 7.26, and perform the sum. The final answer for the probability is the magnitude squared of the resulting amplitude.

For the case where the final state is $|2\rangle$, we can define the transition probability $P_{1\to2}$, for the particle to be in the first excited state of the infinite square well after the electric field is turned off when it was in the ground state of the infinite square well before the electric field was turned on.

To get an idea of the results, we can make plots of $P_{1\to2}$ as a function of $\tau = \frac{\hbar}{mL^2}t$, for $0 \le \tau \le 1$ for each of the four cases of $\tau = 0.1, 1.0, 10.0, 100.0$.

Later we will use time dependent perturbation theory to calculate quantities like this when the field is small. The lowest order result result is

$$P_{1\to2}^{\text{LO}} = \frac{q^2\mathcal{E}^2 m^2 L^6}{\hbar^4}\frac{2^{12}}{3^6\pi^8}\sin^2\left(\frac{3\hbar\pi^2}{4mL^2}t\right).\tag{7.30}$$

I used the maple code

```
with(linalg):
with(plots):

seth := proc(n,efield)
   local h,i,j;
   h := matrix(n,n,0);
   for i from 1 by 1 to n do
      for j from 1 by 1 to n do
         if (i = j) then
            h[i,j] := evalf(Pi^2*i^2/2)
         else
            if (((i+j) mod 2) = 1) then
               h[i,j] := evalf((-1)^((i+j-1)/2)*i*j/((i^2-j^2)^2))
                  *evalf(8.*efield/Pi^2)
            end if;
         end if;
      end do;
   end do;
   return h;
```

```
end:

evalp := proc(vec1,vec2,vals,n,t)
   local amp,p;
   amp := 0;
   for j from 1 by 1 to n do
       amp := amp + evalf(vec1[j]*vec2[j]*exp(-I*vals[j]*t));
   end do;
   p := evalf(amp*conjugate(amp));
   return p;
end:

efield := 100.0;
n := 40:
h := seth(n,efield):
iinitial := 1;
ifinal := 2;

vals := evalf(Eigenvals(h,vecs)):
print(vals[1],vals[16]);
vec1 := subvector(vecs,iinitial,1..n):
vec2 := subvector(vecs,ifinal,1..n):

prob := t -> evalp(vec1,vec2,vals,n,t);
plotsetup(ps,plotoutput='efield.ps',
   plotoptions='color,portrait,width=8in,height=5in,noborder,shrinkby=0.5');

plot([prob(x),efield^2*2^12*sin(Pi^2*3/4*x)^2/(3^6*Pi^8)],x=0..1,axes=BOXED
   ,style=line,labels=[t,prob]);
```

to make the plots shown.
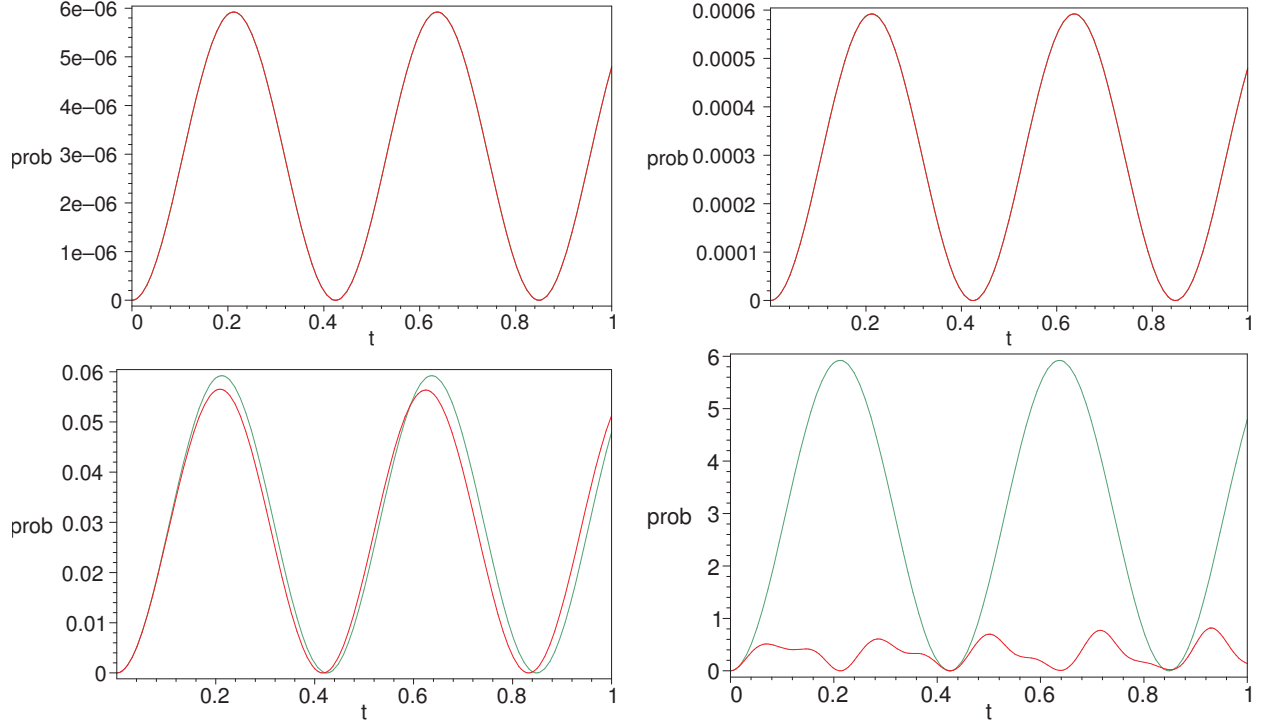
Figure 7.2: The calculated probabilities for the values of $\frac{q\mathcal{E}mL^3}{\hbar^2} = 0.1, 1.0, 10.0, 100.0$ in english reading order. The green curve is the lowest order prediction. Notice it agrees well for the two smaller fields and gets progressively worse (even violating the requirement that the probability must be less than 1) for larger field values.

Chapter 8

Differential equations for eigenvalues and eigenvectors

8.1 Introduction

If the energy eigenvalue equation separates into a one dimensional second order differential equation, it is usually just integrated numerically. That is for a Hamiltonian

$$
\begin{aligned}
H &= \frac{p_{x\ \mathrm{op}}^2}{2m} + v(x_{\mathrm{op}}) \\
&= \int dx |x\rangle \left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + v(x) \right] \langle x|
\end{aligned}
\tag{8.1}
$$

We can solve for its eigenvalues and eigenvectors

$$
H|\psi\rangle = E|\psi\rangle
\tag{8.2}
$$

or

$$
\begin{aligned}
\langle x| \int dx' |x'\rangle \left[-\frac{\hbar^2}{2m}\frac{d^2}{dx'^2} + v(x') \right] \langle x'|\psi\rangle &= \langle x|E|\psi\rangle \\
\int dx' \underbrace{\langle x|x'\rangle}_{=\delta(x-x')} \left[-\frac{\hbar^2}{2m}\frac{d^2}{dx'^2} + v(x') \right] \underbrace{\langle x'|\psi\rangle}_{=\psi(x')} &= E \underbrace{\langle x|\psi\rangle}_{=\psi(x)}
\end{aligned}
\tag{8.3}
$$

so that

$$
\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + v(x) \right] \psi(x) = E\psi(x)
\tag{8.4}
$$

where we write the wave function, $\langle x|\psi\rangle = \psi(x)$.

We can rewrite this as either a single second order or two coupled first order ordinary differential equations:

$$
\frac{d^2\psi(x)}{dx^2} = \frac{2m}{\hbar^2}(v(x) - E)\psi(x)
\tag{8.5}
$$

or

$$\frac{d\psi'(x)}{dx} = \frac{2m}{\hbar^2}(v(x) - E)\psi(x)$$
$$\frac{d\psi}{dx} = \psi'(x). \tag{8.6}$$

At this point, if E is known, these are in the form needed by standard ordinary differential equation solvers. Typical choices are the Numerov method for the second order equation, and Runge-Kutta or many others for the coupled first order equations.

Cases where E is known before hand are the continuum solutions we need, for example, in scattering problems. Bound states of finite systems and other examples have a discrete spectrum where only certain eigenvalues E are allowed (other values give wave functions that cannot be normalized or violate the boundary conditions). For these latter cases often a shooting method is employed. That is, a guess for E is made, the differential equation is integrated with that E, and the solution examined to see how E needs to be changed to get to the correct solution.

8.2   Crude integration

To understand how we would proceed, let's look at a special case – calculating numerically the eigenstates and eigenvalues of the infinite square well that runs from $-L/2 < x < L/2$. We must find the solutions of the differential equation that match the boundary conditions that $\psi(x = \pm L/2) = 0$.

The crudest integration method is Euler's method. This just uses the definition of the derivative

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{8.7}$$

and instead of taking the limit, it approximates the derivative for a small but finite Δx as,

$$\frac{df(x)}{dx} = \frac{f(x + \Delta x) - f(x)}{\Delta x} + ... \tag{8.8}$$

where ... indicates higher order terms in Δx. Rearranging we simply get the first term in the Taylor series

$$f(x + \Delta x) = f(x) + \frac{df(x)}{dx}\Delta x + ... \tag{8.9}$$

In Euler's method we substitute the differential equation for the derivatives

$$\psi'(x + \Delta x) = \psi'(x) + \frac{2m}{\hbar^2}[v(x) - E]\psi(x)\Delta x + ...$$
$$\psi(x + \Delta x) = \psi(x) + \psi'(x)\Delta x + ... \tag{8.10}$$

and we drop the ... terms.

In order to start the integration, we need $\psi(x)$ and $\psi'(x)$ at some point. For the infinite square well we have $\psi(-L/2) = 0$. If we multiply our wave function by a constant, we multiply the slope by the same constant. Therefore, we can take any value we like for the value of $\psi'(-L/2)$. Once we have a solution, we will need to properly normalize it. A reasonable value might be $L^{-3/2}$.

Looking at Euler's equations, we see that we now know everything except E. We therefore need to guess an E value to be able to integrate the differential equation. A not crazy starting value might be something like $\hbar^2/mL^2$ which has the right units. Given this value, we could now iterate the equations to get $\psi(x + n\Delta x)$. Obviously, we would like to choose Δx so that we eventually land on $x = L/2$. Therefore we take $\Delta x = L/N$, where N is an integer. Writing $\psi(x + i\Delta x) = \psi_i$, and $\Delta x \psi'(x + i\Delta x) = \psi_i'$, Euler's method for the infinite square well where $v(x) = 0$ gives the equations

$$\begin{aligned} \psi_{i+1}' &= \psi_i' - \lambda \psi_i \\ \psi_{i+1} &= \psi_i + \psi_i' \end{aligned} \tag{8.11}$$

with $\lambda = \frac{2mE\Delta x^2}{\hbar^2}$.

For the shooting method, we guess any reasonable λ, and iterate the equations until we get to ψ_N. We count how many times the ψ_j changes sign since this tells us how many nodes the wave function has, and the nth excited state has n nodes. If the number of nodes is too large, the chosen energy is too high. If the number of nodes is too small or if the number of nodes is correct, but $\psi_N \neq 0$, the chosen energy is too low. A typical way to proceed is to start with some value of energy and double it until it is too high. The last energy is then too high and the previous energy too low. If the first energy chosen is too high, we can pick a negative energy and double it until it is too low. Once we have an energy that is too high and one that is too low, we can "binary chop" by taking their average, integrating to see if it is too high or too low, and replacing the appropriate bound. Each iteration then improves the energy by one binary digit. IEEE double precision has a 53 bit mantissa, so somewhat more than 50 iterations will guarantee convergence to machine accuracy. There are, of course, ways to speed things up considerably, but the crude method described will get the job done.

Notice that adding the potential requires no additional work, it just needs to be evaluated at the grid points and included in the ψ' equation.

8.3 Numerical integration errors

We can actually solve these equations analytically for $v = 0$ (but realize this would be a silly way to get the correct result since we can solve the correct differential equation much more easily) to see what a computer program would do. Subtracting the equation for ψ_{i+1} from the equation for ψ_i, we get

$$\psi_{i+1} - \psi_i = \psi_i - \psi_{i-1} + (\psi_i' - \psi_{i-1}') = \psi_i - \psi_{i-1} - \lambda \psi_{i-1} \tag{8.12}$$

or

$$\psi_{i+1} = 2\psi_i - \psi_{i-1} - \lambda\psi_{i-1} \,. \tag{8.13}$$

Since this is translationally invariant, we expect possible solutions of the form α^j. Substituting we get the equation $\alpha^2 - 2\alpha - (1 + \lambda) = 0$, with solutions $\alpha = 1 \pm i\sqrt{\lambda}$. Taking the linear combination that gives $\psi_0 = 0$, we have

$$\psi_j = (1 + i\sqrt{\lambda})^j - (1 - i\sqrt{\lambda})^j \,. \tag{8.14}$$

Writing $1 + i\sqrt{\lambda} = \sqrt{1 + \lambda}\,e^{i\tan^{-1}\sqrt{\lambda}} = e^{\alpha + i\beta}$ and enforcing the boundary condition that $\psi_N = 0$ gives the requirement $e^{2i\beta N} = 1$, or $\beta = \pi n/N$, which becomes

$$\beta = \tan^{-1}\sqrt{\lambda} = \frac{\pi n}{N} \,. \tag{8.15}$$

The energy eigenvalues are then

$$E_n = \frac{\hbar^2}{2mL^2} N^2 \tan^2\left(\frac{\pi n}{N}\right) \tag{8.16}$$

and the wave functions on the grid points are

$$\psi_j^{(n)} = N_n e^{j\frac{\ln(1+\lambda)}{2}} \sin\left(\frac{\pi n j}{N}\right) \tag{8.17}$$

where N_n is a normalization. For $n \ll N$, we can expand to get

$$
\begin{aligned}
E_n &= \frac{\hbar^2\pi^2 n^2}{2mL^2}\left[1 + \frac{2\pi^2 n^2}{3N^2} + \dots\right] \\
\psi_j^{(n)} &= N_n \sin\left(\frac{\pi n j}{N}\right)\left[1 + \frac{j\pi^2 n^2}{2N^2} + \dots\right]
\end{aligned}
\tag{8.18}
$$

which shows that the first term of each expression gives the correct answer, and the higher order terms are the numerical errors from the Euler integration method.

8.4   Numerov's method

The especially large wave function errors can be greatly lessened by using an improved integration method. Popular choices are the fourth order Runge-Kutta method for the pair of first order differential equations instead of Euler's method or Numerov's method for the second order equation. Numerov's method is especially easy to derive.

Eq. 8.5 has the form

$$\frac{d^2 f(x)}{dx^2} = g(x)f(x) + h(x) \tag{8.19}$$

where $f(x) = \psi(x)$, $g(x) = \frac{2m}{\hbar^2}[v(x) - E]$, and $h(x) = 0$. Numerov expands $f(x \pm \Delta x)$ in a Taylor series

$$f(x \pm \Delta x) = f(x) \pm f'(x)\Delta x + \frac{1}{2}f''(x)(\Delta x)^2 \pm \frac{1}{6}f'''(x)(\Delta x)^3 + \frac{1}{24}f''''(x)(\Delta x)^4 + O(\Delta x^5)$$
(8.20)

Adding the two expressions and realizing that all of the odd orders cancel, we have

$$f(x + \Delta x) + f(x - \Delta x) = 2f(x) + f''(x)(\Delta x)^2 + \frac{1}{12}f''''(x)(\Delta x)^4 + O(\Delta x^6) \qquad (8.21)$$

Repeating for the expansion of $f''(x \pm \Delta x)$ we have

$$f''(x + \Delta x) + f''(x - \Delta x) = 2f''(x) + f''''(x)(\Delta x)^2 + O(\Delta x^4) \qquad (8.22)$$

Substituting this last expression for the $f''''(x)$ gives

$$\begin{aligned} f(x + \Delta x) + f(x - \Delta x) &= 2f(x) + f''(x)(\Delta x)^2 \\ &\quad + \frac{1}{12}\left[f''(x + \Delta x) - 2f''(x) + f''(x - \Delta x)\right](\Delta x)^2 + O(\Delta x^6) \end{aligned}$$
(8.23)

Substituting the differential equation gives the three-term recursion relation

$$\begin{aligned} & f(x + \Delta x)\left[1 - \frac{(\Delta x)^2}{12}g(x + \Delta x)\right] + f(x - \Delta x)\left[1 - \frac{(\Delta x)^2}{12}g(x - \Delta x)\right] \\ = \; & f(x)\left[2 + \frac{5(\Delta x)^2}{6}g(x)\right] + \frac{(\Delta x)^2}{12}\left[h(x + \Delta x) - 2h(x) + h(x - \Delta x)\right] \qquad (8.24) \end{aligned}$$

Instead of the function and its derivative at the starting point, we need the function at two adjacent points. If the function is zero at some point, we can then view the value at the adjacent point as a normalization choice and start with any convenient value. The shooting method works just as in Euler's method, but the numerical integration will be much more accurate for a given choice of Δx.

The applet on the class web page uses Numerov's method.

First order integration methods like Runge-Kutta are especially useful for discontinuous potentials. This is because they give the wave function and its derivative at the grid points. The integration is performed up to the point where the potential is discontinuous. The discontinuity does not change the function or its derivative, so the integration can then continue from the discontinuous point on without change.

Second order methods like Numerov break down at discontinuities since the Taylor expansions used are no longer well defined. Generally this means that if they are used for discontinuous potentials, the numerical errors will be of lower order (i.e. larger) than the expansions, which assume continuity, would lead you to believe.

Chapter 9

Simple wave packets

A wave packet is a state which has a reasonably well defined position and momentum. Here we will look at a packet in one dimension along x. The generalization to more dimensions is straightforward.

9.1   General wave packet

A general wave packet $|\psi\rangle$ can be described by inserting a complete set of states $|x\rangle$, and giving the wave function, $\langle x|\psi\rangle$,

$$|\psi\rangle = \int dx |x\rangle\langle x|\psi\rangle. \tag{9.1}$$

Similarly in momentum space

$$|\psi\rangle = \int \frac{dp_x}{2\pi\hbar} |p_x\rangle\langle p_x|\psi\rangle. \tag{9.2}$$

We can convert back and forth by inserting a complete set of states

$$\langle p_x|\psi\rangle = \int dx \langle p_x|x\rangle\langle x|\psi\rangle = \int dx e^{-\frac{i}{\hbar}p_x x}\langle x|\psi\rangle. \tag{9.3}$$

The expectation value of the momentum and position are

$$\langle p_x\rangle \equiv \langle\psi|p_{x\;\mathrm{op}}|\psi\rangle$$
$$\langle x\rangle \equiv \langle\psi|x_{\mathrm{op}}|\psi\rangle. \tag{9.4}$$

For these to be reasonably well defined, we want the fluctuations in the possible measured momentum and position to be small. The fluctuations of a hermitian operator are given by its variance,

$$\Delta O^2 = \langle\psi|O^2|\psi\rangle - \langle\psi|O|\psi\rangle^2 \tag{9.5}$$

91

9.2 Minimum uncertainty wave packet

The uncertainty in the momentum and position is written as $\Delta x \Delta p_x$. Let's calculate the minimum uncertainty wave packet with $\langle x \rangle = \langle p_x \rangle = 0$. We can then minimize

$$\langle p_x^2 \rangle \langle x^2 \rangle \tag{9.6}$$

subject to the constraints that $|\psi\rangle$ is normalized, and the expectation values of x and p_x are zero. Varying ψ we get the equation for the minimum uncertainty packet $|\psi_0\rangle$ to be

$$\langle x^2 \rangle p_x^2 |\psi_0\rangle + \langle p_x^2 \rangle x^2 |\psi_0\rangle - \lambda_x x |\psi - \lambda_{p_x} p_x |\psi_0\rangle - \lambda |\psi_0\rangle = 0 \,. \tag{9.7}$$

For fixed values of λ_x and λ_{p_x} this is in the form of the time independent Schrödinger equation for the harmonic oscillator. For $\lambda_x = \lambda_{p_x} = 0$, the expectation values of x and p_x in the eigenstates of the harmonic oscillator are zero. Any other values of these constants will shift the expectation values of x and p_x. Therefore to maintain the constraint, we must choose $\lambda_x = \lambda_{p_x} = 0$. The uncertainty is the proportional to the product of the expectation values of the kinetic and potential energies. These will be minimized for the ground state. The minimum uncertainty wave packet with momentum and position expectation values zero therefore has the form of the ground state of the harmonic oscillator. Recalling that the harmonic oscillator ground state is a gaussian, whose square is normalized,

$$\langle x | \psi_0 \rangle = \frac{1}{(2\pi\sigma^2)^{1/4}} \exp\left(-\frac{x^2}{4\sigma^2}\right) \,. \tag{9.8}$$

Converting to momentum space

$$
\begin{aligned}
\langle p | \psi_0 \rangle &= \int dx \langle p | x \rangle \langle x | \psi_0 \rangle = \int dx \frac{1}{(2\pi\sigma^2)^{1/4}} \exp\left(-\frac{x^2}{4\sigma^2} - \frac{i}{\hbar} p x\right) \,. \\
&= \int dx \frac{1}{(2\pi\sigma^2)^{1/4}} \exp\left(-\left[\frac{x}{2\sigma} - i\frac{p\sigma}{\hbar}\right]^2 - \frac{p^2\sigma^2}{\hbar^2}\right) \,. \\
&= (8\pi\sigma^2)^{1/4} \exp\left(-\frac{p^2\sigma^2}{\hbar^2}\right) \,.
\end{aligned} \tag{9.9}
$$

Plugging in we have

$$
\begin{aligned}
\Delta x^2 &= \sigma^2 \\
\Delta p_x^2 &= \frac{\hbar^2}{4\sigma^2}
\end{aligned} \tag{9.10}
$$

with the uncertainty

$$\Delta p_x \Delta x = \frac{\hbar}{2} \,. \tag{9.11}$$

Notice that translating the position does not change either Δx or Δp_x. This is easily verified since the translation operator commutes with p_x (p_x is its generator), and since

Δx^2 can be written as $(x - \langle x \rangle)^2$, translating changes both x and $\langle x \rangle$ by the same amount. Similarly translating the momentum does not change either Δx or Δp_x. Therefore the minimum uncertainty wave packets for other expectation values of x and p_x can be obtained by simply translating our wave packet in position and momentum. There cannot be a packet with nonzero $\langle p_x \rangle$ or $\langle x \rangle$ with less uncertainty or it could have been translated back to have zero expectation values and would give a smaller result than the smallest possible result calculated.

We therefore have the uncertainty principle result

$$\Delta p_x \Delta x \geq \frac{\hbar}{2} \tag{9.12}$$

The minimum uncertainty wave packet for arbitrary $\langle x \rangle$ and $\langle p_x \rangle$ is then

$$|\psi\rangle = e^{\frac{i}{\hbar} x \langle p_x \rangle} e^{-\frac{i}{\hbar} p_x \langle x \rangle} |\psi_0\rangle \tag{9.13}$$

where interchanging the order of the two translations would introduce only an unimportant phase difference.

Evaluating, the translated minimum uncertainty wave packet, we have

$$\langle x|\psi\rangle = \frac{1}{(2\pi\sigma^2)^{1/4}} \exp\left(-\frac{(x - \langle x \rangle)^2}{4\sigma^2} + \frac{i}{\hbar} x \langle p_x \rangle \right) . \tag{9.14}$$

By either Fourier transforming this or translating the minimum uncertainty wave packet in momentum space, we get

$$\langle p|\psi\rangle = (8\pi\sigma^2)^{1/4} \exp\left[-\frac{(p - \langle p_x \rangle)^2 \sigma^2}{\hbar^2} - \frac{i}{\hbar} (p - \langle p_x \rangle) \langle x \rangle \right] . \tag{9.15}$$

9.3 Free particle gaussian wave packet motion

The Hamiltonian for the free particle is

$$H = \frac{p_{\text{op}}^2}{2m} \tag{9.16}$$

which obviously commutes with the momentum so that the momentum eigenstates (which are nondegenerate) are also eigenstates of H. We can propagate the free particle wave packet using

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar} \frac{p_{\text{op}}^2}{2m} t} |\psi(0)\rangle . \tag{9.17}$$

This is simplest using the momentum representation, so that

$$
\begin{aligned}
\langle p|\psi(t)\rangle &= \langle p|e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}|\psi(0)\rangle \\
&= (8\pi\sigma^2)^{1/4}\exp\left[-\frac{(p-\langle p_x\rangle)^2\sigma^2}{\hbar^2}-\frac{i}{\hbar}(p-\langle p_x\rangle)\langle x\rangle-\frac{i}{\hbar}\frac{p^2}{2m}t\right] \\
&= (8\pi\sigma^2)^{1/4}\exp\left[-(p-\langle p_x\rangle)^2\underbrace{\left(\frac{\sigma^2}{\hbar^2}-\frac{it}{2m\hbar}\right)}_{\text{increasing }\Delta x}-\frac{i}{\hbar}p\underbrace{\left(\langle x\rangle+\frac{\langle p_x\rangle t}{m}\right)}_{\text{packet motion}}+\frac{i}{\hbar}\langle p_x\rangle\underbrace{\left(\langle x\rangle+\frac{\langle p_x\rangle}{2m}t\right)}_{\text{unimportant overall phase}}\right].
\end{aligned}
$$

$$(9.18)$$

Up to an unimportant phase factor, we can now write the result in position representation by making the substitutions for the width $\sigma^2\to\sigma^2-it\hbar/2m$, and position $\langle x\rangle\to\langle x\rangle+\langle p\rangle t/m$ in Eq. 9.14. Since H commutes with p, the width of the packet in momentum space, Δp, does not change, but the width of the packet in position space increases,

$$\Delta x^2(t)=\Delta x^2(0)+\frac{\hbar^2 t^2}{4m^2\Delta x^2(0)}=\Delta x^2(0)+\frac{\Delta p^2 t^2}{m^2}\,. \tag{9.19}$$

The width change can be understood since the higher momentum components propagate faster and the lower momentum components propagate slower. The change of the spatial width would roughly go like $t\Delta p/m$ at large times.

9.4 General wave packet motion

In many cases, we do not have to characterize a wave packet completely. For example, for scattering experiments, we usually only need to have the beam of particles well enough collimated so that the momentum and energy are well enough defined so that the scattering for the range of the energies and momenta in the beam are the same within measurement errors. The spatial extent of a wave packet for a particle needs to satisfy the uncertainty principle constraint, but can often be much larger without changing the experimental results.

Usually, in these cases, we can ignore the wave packet spreading and we just want to know the trajectory of the wave packet. In that case we can simplify the analysis by making Taylor series expansions around the momentum peak (or position peak) of the packet. For the few cases where these approximations are not adequate, the initial wave packet must be measured so that it can be propagated accurately as we did for the gaussian wave packet above.

If we are given a free particle wave packet, $|\psi(p_0,x_0,t=0)\rangle$, that is strongly peaked around a momentum $p=p_0$ and broadly peaked around the position x_0, at time $t=0$, we can propagate it forward in time as usual with

$$
\begin{aligned}
|\psi(p_0,x_0,t)\rangle &= e^{-\frac{i}{\hbar}\epsilon(p_{\text{op}})t}|\psi(p_0,x_0,0)\rangle \\
&= \int\frac{dp}{2\pi\hbar}e^{-\frac{i}{\hbar}\epsilon(p)t}|p\rangle\langle p|\psi(p_0,x_0,0)\rangle\,,
\end{aligned}
\tag{9.20}
$$

where, here, $\epsilon(p) = \frac{p^2}{2m}$, but our result will be correct for any $\epsilon(p)$. Since $\langle p|\psi(p_0, x_0, 0)\rangle$ is strongly peaked around $p = p_0$, we can Taylor series expand and write[1]

$$\epsilon(p) = \epsilon(p_0) + \left.\frac{d\epsilon(p)}{dp}\right|_{p=p_0}(p - p_0) + \dots . \tag{9.21}$$

From the classical Hamilton equations we know that the classical velocity is

$$v = \dot{x} = \frac{dH}{dp} = \frac{d\epsilon(p)}{dp} . \tag{9.22}$$

We therefore write the corresponding derivative of our quantum kinetic energy function as v_0, $v_0 = d\epsilon(p_0)/dp_0$. Expanding, we can approximate

$$
\begin{aligned}
|\psi(p_0, x_0, t)\rangle &= e^{-\frac{i}{\hbar}[\epsilon(p_0) - v_0 p_0]t} \int \frac{dp}{2\pi\hbar} e^{-\frac{i}{\hbar}pv_0 t}|p\rangle\langle p|\psi(p_0, x_0, 0)\rangle \\
&= e^{-\frac{i}{\hbar}[\epsilon(p_0) - v_0 p_0]t} e^{-\frac{i}{\hbar}p_{\mathrm{op}}v_0 t}|\psi(p_0, x_0, 0)\rangle \\
&= e^{-\frac{i}{\hbar}[\epsilon(p_0) - v_0 p_0]t}|\psi(p_0, x_0 + v_0 t, 0)\rangle ,
\end{aligned}
\tag{9.23}
$$

where in the last line we realized that $e^{-\frac{i}{\hbar}p_{\mathrm{op}}v_0 t}$ is the translation operator for a translation $v_0 t$. We see that the packet moves with the velocity v_0. Unless the packet splits (like it would in a two-slit interference experiment), the overall phase factor is unimportant. The packet simply moves along, without changing shape, at velocity v_0. Notice that this same kind of analysis is used in optics and in wave guides and the corresponding $d\epsilon(p_0)/dp_0$ is the group velocity. The group velocity is always defined as the velocity of a well defined wave packet.

Keeping the next term in the Taylor series expansion would give a term proportional to p^2 and, as we saw above, this would give a term that causes the spatial width of the packet to change. Since Δp is small this width change will be small. In any case, for most scattering experiments the spatial width is large (because the momentum uncertainty is small) and the additional change from these higher order terms can be ignored. Within these approximations we do not need to further characterize the packet. This is the main reason that scattering experiments are done with well characterized beam energies and directions. Once these are well determined, the details of the size and shape of the wave packet become unimportant for understanding its motion.

9.5 Free particle propagator

The free particle propagator is $e^{-\frac{i}{\hbar}\frac{p_{\mathrm{op}}^2}{2m}t}$. It's position space matrix elements are

$$\langle x|e^{-\frac{i}{\hbar}\frac{p_{\mathrm{op}}^2}{2m}t}|x'\rangle = \int_{-\infty}^{\infty} \frac{dp}{2\pi\hbar}\langle x|p\rangle e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}\langle p|x'\rangle = \int_{-\infty}^{\infty} \frac{dp}{2\pi\hbar} e^{\frac{i}{\hbar}\left[p\frac{(x-x')}{t} - \frac{p^2}{2m}\right]t} . \tag{9.24}$$

---

[1] The first neglected term is $\frac{1}{2}\frac{d^2\epsilon(p_0)}{dp_0^2}(p - p_0)^2$, which will be small compared to the term kept if $|p - p_0| \ll |[\frac{1}{2}\frac{d^2\epsilon(p_0)}{dp_0^2}]^{-1}\frac{d\epsilon(p_0)}{dp_0}|$. For $\epsilon(p) = \frac{p^2}{2m}$ this is $|p - p_0| \ll |2p_0|$. This says the range is much less than value of the momentum, which is what we mean by strongly peaked.

Trying to perform the integral, we note that since the magnitude of the integrand is constant, the integral is not well defined for large $|p|$.

Here we use our knowledge of physics. The only way that very large $|p|$ values will contribute to a propagated state is if that state contained components with those large $|p|$ values. However, these correspond to very high energies. We know that our theories are not valid at very high energies because there are no experiments there, and our propagated states will not have extremely large $|p|$ components – if they did, we would have experiments and could push the limit higher. There will therefore be a natural cut off of the high momenta and we can include an appropriate convergence factor at high momentum.

This just means that we perform the gaussian integral as usual and our result is

$$\langle x|e^{-\frac{i}{\hbar}\frac{p_{op}^2}{2m}t}|x'\rangle = \sqrt{\frac{m}{it2\pi\hbar}} \exp\left(\frac{i}{\hbar}\left[\frac{1}{2}m\frac{(x-x')^2}{t^2}\right]t\right) . \qquad (9.25)$$

Appendix - Variations in Hilbert space

It is often useful to formulate quantum mechanical methods in terms of variational principles. This leads to equations like the ones we used for the minimum uncertainty wave packet. These variations are done exactly like those you have seen in classical mechanics. Here I will go through, in more detail, how this works in Hilbert space.

To set the notation, let's go through the equivalent calculation for a function $f(x)$, where we want to find the set of points where the function is stationary, that is it does not change with small (i.e. infinitesimal) changes to x. If x_0 is such a point then we can write

$$\delta f(x)|_{x=x_0} = 0 \qquad (9.26)$$

which is often just written $\delta f = 0$ or $\delta f(x) = 0$. This means

$$\delta f(x)|_{x=x_0} = f(x_0 + \delta x) - f(x_0) = 0 \qquad (9.27)$$

where δx is an infinitesimal. Expanding $f(x)$ in a Taylor series around x_0, we have

$$f(x_0 + \delta x) = f(x_0) + \delta x \frac{\partial f(x_0)}{\partial x_0} + ... \qquad (9.28)$$

So that keeping lowest order nonvanishing term in the infinitesimal

$$\delta x \frac{\partial f(x_0)}{\partial x_0} = 0. \qquad (9.29)$$

Since δx is small but not zero, we require

$$\frac{\partial f(x_0)}{\partial x_0} = 0. \qquad (9.30)$$

96

So we solve for all of the points where the derivative of $f(x)$ is zero. If we were looking for the minimum of $f(x)$, we would then substitute all of these values back into $f(x)$ and choose the one that gives the smallest value.

Let's go through the same calculation where we want to find the states in Hilbert space where the expectation value of a hermitian operator O is stationary. Let's call one of these states $|\psi_0\rangle$. The requirement is

$$\delta\langle\psi|O|\psi\rangle|_{|\psi\rangle=|\psi_0\rangle} = 0\,, \tag{9.31}$$

which is often written $\delta\langle\psi|O|\psi\rangle = 0$.

This means that the expectation value should not change for arbitrary infinitesimal variations of the state $|\psi_0\rangle$. This is usually written as

$$|\psi\rangle = |\psi_0\rangle + |\delta\psi\rangle\,. \tag{9.32}$$

Here $|\delta\psi\rangle$ is any infinitesimal complex coefficient times any possible state in Hilbert space. You can of course incorporate the coefficient into the state as long as you realize that the state $|\delta\psi\rangle$ is not normalized and can take any phase. To be able to use regular calculus, we can define $|\delta\psi\rangle$ by taking any complete set of states $|n\rangle$ (so that they span all of Hilbert space) and write

$$|\delta\psi\rangle = \sum_n \delta a_n |n\rangle \tag{9.33}$$

where δa_n are arbitrary infinitesimal complex coefficients.

Substituting into our expectation value, the variation becomes

$$(\langle\psi_0| + \langle\delta\psi|)\, O\, (|\psi_0\rangle + |\delta\psi\rangle) - \langle\psi_0|O|\psi_0\rangle = \langle\delta\psi|O|\psi_0\rangle + \langle\psi_0|O|\delta\psi\rangle = 0\,. \tag{9.34}$$

Using our expansion above, this becomes

$$0 = \sum_n [\delta a_n^*\langle n|O|\psi_0\rangle + \delta a_n\langle\psi_0|O|n\rangle] = \sum_n [\delta a_n^*\langle n|O|\psi_0\rangle + \delta a_n\langle n|O|\psi_0\rangle^*] \tag{9.35}$$

where we use the hermiticity of O to set $O^\dagger = O$ in the last expression. Since this must be zero for all possible choices for the δa_n, both the coefficient of the real and imaginary parts of each δa_n must separately vanish. We then get

$$\begin{aligned} \mathrm{Re}\langle n|O|\psi_0\rangle &= 0 \\ \mathrm{Im}\langle n|O|\psi_0\rangle &= 0 \end{aligned} \tag{9.36}$$

so that $\langle n|O|\psi_0\rangle = 0$. We can then use completeness to write

$$O|\psi_0\rangle = \sum_n |n\rangle\langle n|O\psi_0\rangle = 0\,. \tag{9.37}$$

Having gone through this in detail, we can now take some shortcuts. Notice that for hermitian operators our result is the same as if we had just varied the bra of the Dirac bracket. We can therefore do these kinds of variations by writing

$$\delta\langle\psi|O|\psi\rangle = 0 = \langle\delta\psi|O|\psi_0\rangle \tag{9.38}$$

so that $O|\psi_0\rangle = 0$ immediately.

So as an example, we can set the variation of the expectation value of the energy to zero subject to the constraint that the state is normalized. That is

$$\delta\left[\langle\psi|H|\psi\rangle - \lambda\langle\psi|\psi\rangle\right] = 0 \tag{9.39}$$

where the λ is a Lagrange multiplier enforcing the constraint that $\langle\psi|\psi\rangle = $ constant. Using our shortcut, we vary just $\langle\psi|$, and get

$$H|\psi_0\rangle = \lambda|\psi_0\rangle \tag{9.40}$$

so we see that the eigenstates of H are the states which have a stationary energy expectation values. The Lagrange multiplier is the energy eigenvalue. Obviously the ground state has the lowest energy and will minimize the energy expectation value. This is the Rayleigh-Ritz variational principle.

The Lagrange multiplier method is usually simplest. The alternative is to substitute the constraint before variation. We could have looked at the variation of the normalized energy expectation

$$\delta\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} = 0\,. \tag{9.41}$$

Again taking our shortcut we have

$$
\begin{aligned}
0 &= \frac{\left(\langle\psi_0| + \langle\delta\psi|\right)H|\psi_0\rangle}{\left(\langle\psi_0| + \langle\delta\psi_0|\right)|\psi_0\rangle} - \frac{\langle\psi_0|H|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} \\
&= \frac{\langle\delta\psi|H|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle} - \frac{\langle\psi_0|H|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle}\frac{\langle\delta\psi|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle}
\end{aligned} \tag{9.42}
$$

and writing the expectation value $\langle\psi_0|H|\psi_0\rangle/\langle\psi_0|\psi_0\rangle = \lambda$ we get

$$H|\psi_0\rangle = \lambda|\psi_0\rangle \tag{9.43}$$

as before. Notice that using the Lagrange multiplier is much simpler.

Chapter 10

Generalized uncertainty principle

10.1 Generalized Uncertainty Principle

The generalized uncertainty principle is usually written for two Hermitian operators A and B, as

$$\Delta A^2 \Delta B^2 \geq \frac{1}{4} \langle i[A, B] \rangle^2 \tag{10.1}$$

where as before

$$
\begin{aligned}
\langle O \rangle &= \langle \psi | O | \psi \rangle \\
\Delta O^2 &= \langle O^2 \rangle - \langle O \rangle^2 = \langle (O - \langle O \rangle)^2 \rangle .
\end{aligned} \tag{10.2}
$$

It is less often written as

$$\Delta A^2 \Delta B^2 \geq \frac{1}{4} \langle \{ A - \langle A \rangle, B - \langle B \rangle \} \rangle^2 + \frac{1}{4} \langle i[A, B] \rangle^2 \tag{10.3}$$

where $\{ A, B \} = AB + BA$ is the anticommutator. We will derive these expressions below.

For a single operator, we can vary $|\psi\rangle$ to minimize its variance subject to the constraint that $|\psi\rangle$ remains normalized,

$$\delta \left[\langle \psi | O^2 | \psi \rangle - \langle \psi | O | \psi \rangle^2 - \lambda \langle \psi | \psi \rangle \right] = 0 \tag{10.4}$$

or

$$O^2 |\psi\rangle = \lambda' |\psi\rangle , \tag{10.5}$$

where the values multiplying $|\psi\rangle$ have all been combined into the new eigenvalue λ'. The eigenvectors of O^2 are the same as the eigenvectors of O, and the variance of O is zero for an eigenstate since every measurement on an ensemble of identical eigenstates of O gives the same eigenvalue.

We immediately see that the generalized uncertainty principle in some sense only says something useful when A and B are unbounded operators (like position, momentum, angular momentum which have unbounded eigenvalues). If one of the operators is bounded, we can

choose an eigenstate of the other, its variance will be zero, the variance of the bounded operator is bounded, so the left side is zero. If you compute the right side, you will find that it is also zero. So this "minimum uncertainty wave packet" isn't really a wave packet at all, it is an eigenstate of one of the operators. For any particular state, $|\psi\rangle$, the generalized uncertainty principles will still be true, but there doesn't seem to be any big advantage to calculating the right-hand side versus the left-hand side. That is, if you want to know the variances in a particular state, it is just as easy to calculate them directly as to calculate a bound from calculating the expectation value of their commutator.

The real value of this uncertainty principle is for the unbounded case, and especially where the commutator is proportional to the identity. That is, for canonical variables. In that case the right-hand side of the usual expression is a constant, and we get uncertainty principles like those for position and momentum.

10.2   Derivation of the generalized uncertainty principle

We know that the variance of an operator in an eigenstate is zero, and we know that two Hermitian operators that commute share the same eigenstates, so for commuting operators, we can make both variances zero. Therefore, we would like to relate the product of the variances to the commutator. That is we want to relate expectation values that have just A or just B to those that have AB or BA. The Schwarz inequality does this.

Let's write

$$
\begin{aligned}
a &= A - \langle A \rangle \\
b &= B - \langle B \rangle .
\end{aligned}
\tag{10.6}
$$

For any Hermitian operators a and b, we can write

$$
\begin{aligned}
|\psi_a\rangle &= a|\psi\rangle \\
|\psi_b\rangle &= b|\psi\rangle
\end{aligned}
\tag{10.7}
$$

so that the left-hand side of the uncertainty relation is

$$
\begin{aligned}
\Delta A^2 \Delta B^2 &= \langle\psi|a^2|\psi\rangle\langle\psi|b^2|\psi\rangle = \langle\psi_a|\psi_a\rangle\langle\psi_b|\psi_b\rangle \\
&\geq |\langle\psi_a|\psi_b\rangle|^2
\end{aligned}
\tag{10.8}
$$

where the last line is the Schwarz inequality in Hilbert space. It is the analog of the inequality for standard vectors that their dot product squared is less than the product of their lengths squared. Expanding, we can now write

$$
\Delta A^2 \Delta B^2 \geq |\langle\psi_a|\psi_b\rangle|^2 = |\langle\psi|\tfrac{1}{2}\{a,b\} + \tfrac{1}{2}[a,b]|\psi\rangle|^2 .
\tag{10.9}
$$

Notice that

$$\{A, B\}^\dagger = (AB + BA)^\dagger = B^\dagger A^\dagger + A^\dagger B^\dagger = BA + AB = \{A, B\}$$
$$[A, B]^\dagger = (AB - BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB = -[A, B] \qquad (10.10)$$

so the anticommutator is Hermitian and will have real eigenvalues and real expectation value. The commutator is antihermitian, that is i times the commutator will be hermitian, and the commutator must have imaginary eigenvalues and imaginary expectation value. We can therefore write

$$\Delta A^2 \Delta B^2 \geq \frac{1}{4}\langle\psi|\{a, b\}|\psi\rangle^2 + \frac{1}{4}\langle\psi|i[a, b]|\psi\rangle^2$$
$$= \frac{1}{4}\langle\psi|\{A - \langle A\rangle, B - \langle B\rangle\}|\psi\rangle^2 + \frac{1}{4}\langle\psi|i[A, B]|\psi\rangle^2 . \qquad (10.11)$$

We have dropped the subtracted averages in the commutator since they give zero. Keeping both terms gives the second generalized uncertainty principle, while dropping the anticommutator gives the usual generalized uncertainty principle.

You need to be careful around the limits where things go to infinity. For the position-momentum case, you can get into trouble if you choose, for example, a position eigenstate. The position uncertainty is zero. If you evaluate the right-hand side by first evaluating the commutator, you will get the standard $\frac{1}{4}\hbar^2$ result. If however you evaluate by operating on the position state with $x - \langle x\rangle$, you would get zero. The problem, of course, is that the momentum variance is infinite, and subtracting two infinities (as you do in the commutator) can give any value you like if you are not careful. One proper way is to take a limit. That is, in fact, what we did when evaluating the minimum uncertainty wave packet. In that case, you find that for any states with bounded expectation values of Δp^2 and Δx^2, all of the results are finite and well behaved.

10.3   Minimum uncertainty wave packet

Notice that the Schwarz inequality gives the same uncertainty principle result for $\Delta p^2 \Delta x^2$ as our full variation. Because of this, you can find the minimum uncertainty wave packet, by enforcing the equalities. The Schwarz inequality becomes an equality only if the two vectors are in the same direction, i.e. they are multiples of each other,

$$a|\psi\rangle = \gamma b\psi\rangle \qquad (10.12)$$

where γ is any complex number. For the anticommutator term to be dropped it must be zero. Substituting the result above,

$$0 = \langle\psi|ab + ba|\psi\rangle = \langle\psi|\gamma^* bb + bb\gamma|\psi\rangle = (\gamma^* + \gamma)\langle\psi|b^2|\psi\rangle . \qquad (10.13)$$

The anticommutator term will be zero only if γ is imaginary.

We get for position and momentum and real β,

$$0 = \left[\frac{d}{dx} + \beta x - \frac{i}{\hbar}\langle p\rangle - \beta\langle x\rangle\right]\psi(x) \tag{10.14}$$

which can be integrated

$$\ln\psi = C - \frac{1}{2}\beta x^2 + \frac{i}{\hbar}\langle p\rangle x + \beta\langle x\rangle x \tag{10.15}$$

where C is a normalization constant. Rewriting,

$$\psi(x) = \frac{1}{\sqrt{2\pi\beta^{-1}}}e^{-\frac{(x-\langle x\rangle)^2}{2\beta^{-1}}}e^{\frac{i}{\hbar}\langle p\rangle x} \tag{10.16}$$

which agrees with our direct evaluation. However, there are some decisions that were made that might not be obvious. For example, why did we require the anticommutator to be zero? We could have required the commutator to be zero, which would have lead to an imaginary β. Doing the same integral would show you that the $\psi(x)$ is not normalizable. You might wonder if there is some other combination where both the commutator and anticommutator terms are nonzero, but with a sum that is smaller. Of course, the way to answer all of these questions is not to rely on the Schwarz inequality, but rather to do the full variation that we did to find the minimum uncertainty wave packet. In that case, there can be no doubt that it is the packet with the product of $\Delta p^2 \Delta x^2$ minimized.

10.4　Problems

Problem 10.1. Show that for two Hermitian Operators A and B and a real positive constant β (which has units to make all the units consistent),

$$\Delta A^2 + \beta^2 \Delta B^2 \geq \beta \left| \langle i[A, B] \rangle \right| . \tag{10.17}$$

where

$$\begin{aligned}
\langle O \rangle &= \langle \psi | O | \psi \rangle \\
\Delta O^2 &= \langle (O - \langle O \rangle)^2 \rangle .
\end{aligned} \tag{10.18}$$

Chapter 11

Path integral

11.1 Introduction

The path integral formalism of quantum mechanics is simply another way of looking at the mathematics of quantum mechanics. It has some nice physical interpretations and can help in understanding how to perform some calculations. A number of numerical techniques are simpler to understand in terms of path integrals. We will concentrate here on a Hamiltonian $H(p, x)$ but also allow it to be time dependent so that it can be $H(p, x, t)$.

11.2 Path integral derivation

For a general Hamiltonian of the form $H(p, x, t)$ we first specify that the Hamiltonian is written in a normal order. For example we require that all momenta are to the left of all positions. If someone writes down a Hamiltonian where this is not true, we simply commute the terms to define $H(p, x, t)$. So if I wrote down

$$H = \frac{(m\omega x + ip)(m\omega x - ip)}{2m} \tag{11.1}$$

you would first expand

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \frac{\omega ipx}{2} - \frac{\omega ixp}{2} \tag{11.2}$$

and then write $xp = [x, p] + px = i\hbar + px$, which gives the normal ordered form

$$H(p, x, t) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 - \frac{1}{2}\hbar\omega \,. \tag{11.3}$$

The short time propagator moves a state forward by an infinitesimal time Δt. Rather than label the states x and x', it is convenient to label them as $x(t + \Delta t)$ and $x(t)$, but here the t and $t + \Delta t$ mean the same thing as the x and x' they just specify that the x basis

with different dummy labels, and similarly for $p(t)$. The propagator in position space for an infinitesimal time is

$$\langle x(t+\Delta t)|e^{-\frac{i}{\hbar}H(p_{\text{op}},x_{\text{op}},t)\Delta t}|x(t)\rangle . \tag{11.4}$$

To order Δt we can write

$$
\begin{aligned}
\langle p(t)|e^{-\frac{i}{\hbar}H(p_{\text{op}},x_{\text{op}},t)\Delta t}|x(t)\rangle &= \langle p(t)|\left[1 - \frac{i}{\hbar}H(p_{\text{op}},x_{\text{op}},t)\Delta t + ...\right]|x(t)\rangle \\
&= \langle p(t)|\left[1 - \frac{i}{\hbar}H(p(t),x(t),t)\Delta t + ...\right]|x(t)\rangle \\
&= \langle p(t)|e^{-\frac{i}{\hbar}H(p(t),x(t),t)\Delta t+\cdots}|x(t)\rangle \\
&= e^{-\frac{i}{\hbar}H(p(t),x(t),t)\Delta t+\cdots}\langle p(t)|x(t)\rangle
\end{aligned} \tag{11.5}
$$

where we use our normal ordering assumption to be able to operate to the left with the momenta operators and replace them with their eigenvalues, and to the right with position operators and replace them with their eigenvalues.

The propagator, correct to order Δt, becomes

$$
\begin{aligned}
\langle x(t+\Delta t)|e^{-\frac{i}{\hbar}H(p_{\text{op}},x_{\text{op}},t)\Delta t}|x(t)\rangle &= \int_{-\infty}^{\infty} \frac{dp(t)}{2\pi\hbar}\langle x(t+\Delta t)|p(t)\rangle e^{-\frac{i}{\hbar}H(p(t),x(t),t)\Delta t}\langle p(t)|x(t)\rangle \\
&= \int_{-\infty}^{\infty} \frac{dp(t)}{2\pi\hbar}e^{\frac{i}{\hbar}\left[p(t)\frac{(x(t+\Delta t)-x(t))}{\Delta t}-H(p(t),x(t),t)\right]\Delta t} .
\end{aligned} \tag{11.6}
$$

In classical physics the Hamiltonian and Lagrangian are related by a Legendre transformation with

$$\mathcal{L} = p\dot{x} - H \tag{11.7}$$

If we identify $[x(t+\Delta t) - x(t)]/\Delta t$ as the x time derivative, then we see that the quantity in the exponent is $\frac{i}{\hbar}\int_{t}^{t+\Delta t} dt'\mathcal{L}$. The time integral of the Lagrangian is called the action.

Combining many time slices gives the propagator from $x(t_0)$ to $x(t)$

$$\prod_{t'=t_0}^{t}\int \frac{dx(t')dp(t')}{2\pi\hbar}e^{\frac{i}{\hbar}\left[p(t')\frac{(x(t'+\Delta t)-x(t'))}{\Delta t}-H(p(t'),x(t'),t')\right]\Delta t} \tag{11.8}$$

which is the path integral. Often the momentum integrals can be done analytically, and the usual path integral then corresponds to doing all of the intermediate $x(t')$ integrations. Viewing this as an enormous nested set of integrals, any single point from each integrand corresponds to a position of the particle at that time slice. Connecting these points gives a path, and doing all of the integrations corresponds to summing over all possible paths between the initial point $x(t_0)$ and the final point $x(t)$. Each path is weighted by the integral of the action along the path.

Taking advantage of the Lagrangian form, the path integral propagator can be abbreviated as

$$\int \mathcal{D}x \; e^{\frac{i}{\hbar} \int dt' \mathcal{L}(x(t'), \dot{x}(t'), t')} \; . \tag{11.9}$$

However, for any numerical calculation, these symbols just mean the limit of taking $\Delta t \rightarrow 0$ of Eq. 11.8.

Since the path integral was derived by starting with the Schrödinger equation propagator, the path integral is simply another way of thinking about the Schrödinger equation.

The path integral can be used as a convenient way to take the classical limit. The classical limit is when the action changes much more than $\hbar$ for small changes of the position and momenta. In that case, the path integral is completely dominated by the trajectories where these changes are zero i.e. where the action is extremized (usually minimized). That path, where the variation of the action is zero, is just the classical trajectory.

Chapter 12

Scattering in a slab geometry

12.1   Separation of the Hamiltonian

The experiment we are examining is one where a beam of particles scatters from a potential that depends only on one direction.

The Hamiltonian we use to describe this is

$$H = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} + v(x). \tag{12.1}$$

We can write this as

$$H = H_x + H_y + H_z \tag{12.2}$$

where

$$
\begin{aligned}
H_x &= \frac{p_x^2}{2m} + v(x) \\
H_y &= \frac{p_y^2}{2m} \\
H_z &= \frac{p_z^2}{2m}.
\end{aligned}
\tag{12.3}
$$

Notice that all of these operators commute, so we can simultaneously diagonalize all of them. But more than that, H_x, H_y, and H_z operate on different parts of the Hilbert space. For example, if we use the $|\boldsymbol{r}\rangle = |x, y, z\rangle$ basis, then H_x does not change the y or z values and similarly for H_y and H_z. We can always write the Hilbert space as an outer product $|x, y, z\rangle = |x\rangle \otimes |y\rangle \otimes |z\rangle$ which is usually written as $|x\rangle|y\rangle|z\rangle$ with the separate kets indicating an outer product. For these cases, H_x only operates on the $|x\rangle$ part of the outer product, H_y on the $|y\rangle$ part, H_z on the $|z\rangle$ part. We can therefore first find the eigenvectors and eigenvalues of H_x, H_y and H_z in each of their separate spaces

$$
\begin{aligned}
H_x |\psi_n^{(x)}\rangle &= E_{n_x}^{(x)} |\psi_n^{(x)}\rangle \\
H_y |\psi_n^{(x)}\rangle &= E_{n_x}^{(y)} |\psi_n^{(y)}\rangle \\
H_z |\psi_n^{(x)}\rangle &= E_{n_x}^{(z)} |\psi_n^{(z)}\rangle
\end{aligned}
\tag{12.4}
$$

and the eigenstates of H are

$$|\psi_{n_x}^{(x)}\rangle|\psi_{n_y}^{(y)}\rangle|\psi_{n_z}^{(z)}\rangle \tag{12.5}$$

with eigenvalue $E_{n_x,n_y,n_z} = E_{n_x}^{(x)} + E_{n_x}^{(y)} + E_{n_x}^{(z)}$.

12.2   A simple example

One of the simplest examples is the potential step,

$$H_x = \frac{p_x^2}{2m} + V\theta(x) \tag{12.6}$$

where

$$\theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases} . \tag{12.7}$$

We take the boundary condition that the only incoming waves for positive k_x come from the left. The solution is in the form

$$\langle x|\psi_{k_x}^{(x)}\rangle = \begin{cases} e^{ik_x x} + r(k_x)e^{-ik_x x} & x < 0 \\ t(k_x)e^{ik'_x x} & x > 0 \end{cases} \tag{12.8}$$

with

$$E_{k_x}^{(x)} = \frac{\hbar^2 k_x^2}{2m} = \frac{\hbar^2 k'_x{}^2}{2m} + V \tag{12.9}$$

where k'_x is real and has the same sign as k_x if $\frac{\hbar^2 k_x^2}{2m} > V$. If $\frac{\hbar^2 k_x^2}{2m} < V$, then k'_x is positive imaginary so that the wave function goes to zero for $x \to \infty$ and the wave function is integrable. Since there are no delta functions in the potential, the wave function and its derivative are continuous. Equating the function and its logarithmic derivative gives

$$\begin{aligned} 1 + r(k_x) &= t(k_x) \\ \frac{ik_x[1 - r(k_x)]}{1 + r(k_x)} &= ik'_x \end{aligned} \tag{12.10}$$

Solving, this becomes

$$\begin{aligned} r(k_x) &= \frac{k_x - k'_x}{k'_x + k_x} \\ t(k_x) &= \frac{2k_x}{k'_x + k_x} . \end{aligned} \tag{12.11}$$

We now need to relate this solution to the physical scattering.

12.3 Wave packet motion and scattering

The Hamiltonian is invariant under translations along y and z. This means that the momentum eigenstates in these directions (which are nondegenerate) must give the eigenstates of H for those parts. Equivalently H_y and H_z are translationally invariant which tells us that $|\psi_n^{(y)}\rangle$ and $|\psi_n^{(z)}\rangle$ can be taken to be the momentum eigenstates, $|k_y\rangle$, and $|k_z\rangle$. To get the H_x eigenstates we must solve

$$H_x|\psi^{(x)}\rangle = E^{(x)}|\psi^{(x)}\rangle . \tag{12.12}$$

This is a one dimensional Hamiltonian and is the only nontrivial part of finding the eigenstates. Therefore we concentrate on it.

For a general scattering problem, an experimentalist must measure and tell us the initial wave packet of the particle. We then solve for a complete set of H_x eigenstates. We combine these with the H_y and H_z momentum eigenstates to get eigenstates of H. Since we know the initial wave packet, we can find the overlap of the initial packet with our H eigenstates. We then write the initial packet as a linear combination of H eigenstates. To move the packet forward in time we multiply each of the coefficients of the H eigenstates by their $e^{-\frac{i}{\hbar}Et}$ value. Since we know the result, we know that this wave packet will collide with the potential slab, and after interaction, will split into a transmitted and a reflected part. The full quantum solution can tell us everything about the resulting packet at any time. Usually, we don't care about this. In fact, the main reason that we do scattering is not to test if quantum mechanics works, but to find out what $v(x)$ is. For example in x-ray or electron diffraction, the scattering is used to determine the potential which gives the charge density of the sample. We therefore want to make the experiment as easy as possible to carry out.

We saw in our analysis of free-particle wave packets that having a packet sharply peaked in momentum and broadly peaked in position made the analysis of its motion simple without having to know the details of the packet.

For scattering problems, we take the particle to be initially in a wave packet peaked far from the potential.

Since any bound states of the potential will be localized in the region of the potential, we begin by assuming that the particle is initially in a packet so far away from the potential that it does not contain any bound states. The calculation is easy enough if you include these states, the hard part is for the experimentalist to characterize the initial packet. The theoretical work needs to tell experimentalists what measurements must be done to characterize the experiment. In addition, it is very helpful if there are particular cases where very little characterization needs to be done.

We begin by imagining that the target is removed. A particle in the beam is then described by a wave packet. Let's call this wave packet $\psi_0(x, y, z, t)$, and take our coordinate system such that the packet position will be peaked around the origin at time $t = 0$. It is strongly peaked around $\boldsymbol{k}_0 = k_{x0}\hat{\boldsymbol{x}} + k_{y0}\hat{\boldsymbol{y}} + k_{z0}\hat{\boldsymbol{z}}$. We can write the motion of this free

particle wave packet by expanding as before

$$\begin{aligned}
\langle \boldsymbol{r}|\psi_0(t)\rangle &= \int \frac{d^3k}{(2\pi)^3}\langle \boldsymbol{r}|\boldsymbol{k}\rangle e^{-\frac{i}{\hbar}E_k t}\langle \boldsymbol{k}|\psi_0(0)\rangle \\
&= \int \frac{d^3k}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} e^{-\frac{i}{\hbar}\left[E_{k_0}+\boldsymbol{\nabla}_{k_0}E_{k_0}\cdot(\boldsymbol{k}-\boldsymbol{k}_0)\right]t}\langle \boldsymbol{k}|\psi_0(0)\rangle \\
&= e^{-\frac{i}{\hbar}\left[E_{k_0}-\boldsymbol{\nabla}_{k_0}E_{k_0}\cdot\boldsymbol{k}_0\right]t} \int \frac{d^3k}{(2\pi)^3} e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{v}_0 t)}\langle \boldsymbol{k}|\psi_0(0)\rangle \\
&= e^{i\phi(t)}\psi_0(x-v_{x0}t, y-v_{y0}t, z-v_{z0}t, 0)\,, \tag{12.13}
\end{aligned}$$

where

$$\boldsymbol{v}_0 = \frac{1}{\hbar}\boldsymbol{\nabla}_{k_0}E_{k_0} \equiv v_{x0}\hat{\boldsymbol{x}} + v_{y0}\hat{\boldsymbol{y}} + v_{z_0}\hat{\boldsymbol{z}}\,, \tag{12.14}$$

and $e^{i\phi(t)}$ is a phase factor that will be unimportant as long as we do not have two packets interfering. I have repeatedly dropped higher order terms that correspond to the packet width changing etc. since these are small if the packet is sharply peaked in momentum.

12.4   Packet motion in our example

Let's now calculate the motion of the packet in our example. We write[1]

$$\begin{aligned}
\langle \boldsymbol{r}|\psi(t)\rangle &= \int \frac{d^3k}{(2\pi)^3}\langle xyz|\psi_{\boldsymbol{k}}\rangle e^{-\frac{i}{\hbar}E_k t}\langle \psi_{\boldsymbol{k}}|\psi(0)\rangle \\
&= \int \frac{d^3k}{(2\pi)^3}\left[\theta(-x)e^{ikx} + \theta(-x)r(k_x)e^{-ikx} + \theta(x)t(k_x)e^{ik'_x x}\right] e^{ik_y y}e^{ik_z z}e^{-\frac{i}{\hbar}E_k t}\langle \psi_{\boldsymbol{k}}|\psi(0)\rangle
\end{aligned}$$
$$\tag{12.15}$$

We now expand as before. We keep the lowest order terms in the expansion that affect the result. The energy expansion is the same as before, the zeroth order term in the expansion of $\boldsymbol{k}$ in the phases gives an overall phase with no physical effect, so we keep the linear terms in the phases. For the magnitudes, the zeroth order is the lowest order term, so we stop there. That means we write the reflection and transmission coefficients as magnitudes times

---

[1]The second line is not generally correct. It relies on having the packet strongly peaked around $\boldsymbol{k}_0$ with k_{x0} positive. For negative values of k_x, the eigenfunction should have incoming and reflected waves for positive x and an outgoing wave for negative x. It is not the solution shown with k_x negative. This will be clear when we use the Lippmann-Schwinger equation to obtain scattering solutions in three-dimensions.

phases and expand

$$
\begin{aligned}
r(k_x) &= |r(k_x)|e^{i\phi_r(k_x)} = |r(k_{x0})|e^{i\phi_r(k_{x0})+i\frac{d\phi_r(k_{x0})}{dk_{x0}}(k_x-k_{x0})} \\
t(k_x) &= |t(k_x)|e^{i\phi_t(k_x)} = |t(k_{x0})|e^{i\phi_t(k_{x0})+i\frac{d\phi_t(k_{x0})}{dk_{x0}}(k_x-k_{x0})} \\
\frac{dk'_x}{dk_x} &= \frac{\frac{dE^{(x)}}{dk_x}}{\frac{dE^{(x)}}{dk'_x}} = \frac{v_x}{v'_x} \\
e^{ik'_x x} &= e^{ik'_{x0}x+i\frac{dk'_{x0}}{dk}(k_x-k_{x0})x} = e^{ik'_{x0}x-i\frac{v_{x0}}{v'_{x0}}k_{x0}x}\,e^{i\frac{v_{x0}}{v'_{x0}}k_x x} .
\end{aligned}
\tag{12.16}
$$

The energy expansion is identical to the free particle wave packet expansion, we get

$$
\begin{aligned}
\psi(x,y,z,t) &= e^{i\chi_i(\boldsymbol{r},t)}\theta(-x)\int \frac{d^3k}{(2\pi)^3}e^{ik_x(x-v_{x0}t)}e^{ik_y(y-v_{y0}t)}e^{ik_z(z-v_{z0}t)}\langle \boldsymbol{k}|\psi(0)\rangle \\
&\quad +e^{i\chi_r(\boldsymbol{r},t)}\theta(-x)|r(k_{x0})|\int \frac{d^3k}{(2\pi)^3}e^{ik_x[-x-v_{x0}t+\frac{d\phi_r(k_{x0})}{dk_{x0}}]}e^{ik_y(y-v_{y0}t)}e^{ik_z(z-v_{z0}t)}\langle \boldsymbol{k}|\psi(0)\rangle \\
&\quad +e^{i\chi_t(\boldsymbol{r},t)}\theta(x)|t(k_{x0})|\int \frac{d^3k}{(2\pi)^3}e^{ik_x\left[\frac{v_{x0}}{v'_{x0}}(x-v'_{x0}t)+\frac{d\phi_t(k_{x0})}{dk_{x0}}\right]}e^{ik_y(y-v_{y0}t)}e^{ik_z(z-v_{z0}t)}\langle \boldsymbol{k}|\psi(0)\rangle
\end{aligned}
\tag{12.17}
$$

where

$$
\begin{aligned}
\chi_i(\boldsymbol{r},t) &= -\frac{1}{\hbar}\left[E_{k_0}-\boldsymbol{\nabla}_{k_0}E_{k_0}\cdot \boldsymbol{k}_0\right]t \\
\chi_r(\boldsymbol{r},t) &= \chi_i(\boldsymbol{r},t)+\phi_r(k_{x0})-\frac{d\phi_r(k_{x0})}{dk_{x0}}k_{x0} \\
\chi_t(\boldsymbol{r},t) &= \chi_i(\boldsymbol{r},t)+\phi_t(k_{x0})-\frac{d\phi_t(k_{x0})}{dk_{x0}}k_{x0}+k'_{x0}x-\frac{v_{x0}}{v'_{x0}}k_{x0}x .
\end{aligned}
\tag{12.18}
$$

Taking the coefficients $\langle \psi_{\boldsymbol{k}}|\psi(0)\rangle$ to be the same as the free particle packet $\langle \boldsymbol{k}|\psi(0)\rangle$, we see that the incident packet is identical and the full solution for the wave packet is

$$
\begin{aligned}
\psi(x,y,z,t) &= e^{i\chi_i(\boldsymbol{r},t)}\theta(-x)\psi_0(x-v_{x0}t,y-v_{y0}t,z-v_{z0}t,0) \\
&\quad +e^{i\chi_r(\boldsymbol{r},t)}\theta(-x)|r(k_{x0})|\psi_0\left(-x-v_{x0}t+\frac{d\phi_r(k_{x0})}{dk_{x0}},y-v_{y0}t,z-v_{z0}t,0\right) \\
&\quad +e^{i\chi_t(\boldsymbol{r},t)}\theta(x)|t(k_{x0})|\psi_0\left(\frac{v_{x0}}{v'_{x0}}(x-v'_{x0}t)+\frac{d\phi_t(k_{x0})}{dk_{x0}},y-v_{y0}t,z-v_{z0}t,0\right) .
\end{aligned}
\tag{12.19}
$$

We choose the spatial origin so that $\psi_0(x,y,z,t)$ is broadly peaked around it at $t=0$. We see that the first term corresponds to the incident packet, it is broadly peaked around

$$
\begin{aligned}
x_{\text{peak}} &= v_{x0}t \\
y_{\text{peak}} &= v_{y0}t \\
z_{\text{peak}} &= v_{z0}t
\end{aligned}
\tag{12.20}
$$

113

and always zero for $x > 0$. For large positive t, this packet will be zero since the ψ_0 part will be peaked around large positive x and this is multiplied by $\theta(-x)$ which is zero for positive x.

The reflected packet is the second term. It is also always zero for $x > 0$ and is broadly peaked around

$$
\begin{aligned}
x_{\text{peak}} &= -v_{x0}t + \frac{d\phi_r(k_{x0})}{dk_{x0}} \\
y_{\text{peak}} &= v_{y0}t \\
z_{\text{peak}} &= v_{z0}t \, .
\end{aligned}
\tag{12.21}
$$

It will be zero for large negative t since the ψ_0 term will be peaked around a large positive x value and it is multiplied by $\theta(-x)$ which is zero for positive x.

The transmitted packet is the last term. It is zero for $x < 0$ and is broadly peaked around

$$
\begin{aligned}
x_{\text{peak}} &= v'_{x0}t - \frac{v'_{x0}}{v_{x0}}\frac{d\phi_t(k_{x0})}{dk_{x0}} \\
y_{\text{peak}} &= v_{y0}t \\
z_{\text{peak}} &= v_{z0}t \, .
\end{aligned}
\tag{12.22}
$$

It will be zero for large negative t since the ψ_0 term will be peaked around a large negative x value and it is multiplied by $\theta(x)$ which is zero for negative x.

We see that at large negative times, there is only the incident packet – exactly like the experimental case. We normalize the wave function to one for convenience. This means that at large negative times where we have only the incident packet that the integral of $|\psi_0|^2$ is one. At large positive times, we have only the reflected and transmitted packets. They have no overlap since one is multiplied by $\theta(-x)$ and the other by $\theta(x)$. For large positive times we can calculate the probability of finding the particle in the reflected wave packet as

$$
R = |r(k_{x0})|^2 \int d^3r \left| \psi_0 \left(-x - v_{x0}t + \frac{d\phi_r(k_{x0})}{dk_{x0}}, y - v_{y0}t, z - v_{z0}t, 0 \right) \right|^2 = |r(k_{x0})|^2 \, .
\tag{12.23}
$$

To perform the integral we simply change variables to the arguments of ψ_0. The Jacobian is 1 and we get the standard result that the reflection probability R is $|r(k_{x0})|^2$.

The transmission probability is

$$
T = |t(k_{x0})|^2 \int d^3r \left| \psi_0 \left(\frac{v_{x0}}{v'_{x0}}(x - v'_{x0}t) + \frac{d\phi_t(k_{x0})}{dk_{x0}}, y - v_{y0}t, z - v_{z0}t, 0 \right) \right|^2 = \frac{v'_{x0}}{v_{x0}}|t(k_{x0})|^2
\tag{12.24}
$$

where we again change variables to the arguments. The ratio of the velocities factor is from the Jacobian.

If we plug in for our simple example, we get

$$
\begin{aligned}
R &= \left| \frac{k_{x0} - k'_{x0}}{k_{x0} + k'_{x0}} \right|^2 = \frac{k_{x0}^2 - 2k_{x0}k'_{x0} + k'^2_{x0}}{k_{x0}^2 + 2k_{x0}k'_{x0} + k'^2_{x0}} \\
T &= \frac{k'_{x0}}{k_{x0}} \frac{4k_{x0}^2}{k_{x0}^2 + 2k_{x0}k'_{x0} + k'^2_{x0}} = \frac{4k_{x0}k'_{x0}}{k_{x0}^2 + 2k_{x0}k'_{x0} + k'^2_{x0}} .
\end{aligned}
\tag{12.25}
$$

Notice that $T + R = 1$ as expected. Often we want to write these in terms of more physically measured variables. Make sure you understand how to write these in terms of the incident energy and angle.

Notice that for our simple example that r and t are both real if the particle is able to be transmitted. However, if the barrier is high enough that there is no transmitted wave, the reflection coefficient becomes complex and the reflected packet is displaced.

12.5   General scattering

The analysis for a general scattering problem is nearly identical to that of the simple example. The only difference is that the eigenstates of the Hamiltonian will only become linear combinations of plane waves at some distance from the target where the target potential has become a constant (usually zero). For large positive and negative times these regions are the only ones that contribute, and the analysis above goes through.

12.6   Time reversal and reciprocity

For spin independent, time reversal invariant, Hamiltonians, the complex conjugate of an eigenstate of the Hamiltonian with energy E is also an eigenstate, but with different boundary conditions. We can use this to relate the reflection and transmission coefficients from the left and right. For the general case where the potential goes to a constant (but possibly different value) for large $|x|$, the solution for an incoming wave from the left is

$$
\psi(|x| \to \infty) = \begin{cases} e^{ikx} + r_L e^{-ikx} & x \to \infty \\ t_L e^{ik'x} & x \to \infty \end{cases}
\tag{12.26}
$$

Taking the complex conjugate gives the time reversed solution which is also a solution to $H\psi = E\psi$,

$$
\psi^*(|x| \to \infty) = \begin{cases} e^{-ikx} + r_L^* e^{ikx} & x \to -\infty \\ t_L^* e^{-ik'x} & x \to \infty \end{cases}
\tag{12.27}
$$

The time reversed solution corresponds to two incoming packets miraculously combining to give just one outgoing packet toward the left. We can construct the physical solution where a particle is incident from the right by taking a linear combination of ψ and ψ^* that has no

incoming component from the left, and normalizing the incoming plane wave from the right to one,

$$-\frac{r_L^* \psi(x) - \psi^*(x)}{t_L^*} = \begin{cases} t_L e^{-ikx} & x \to -\infty \\ -\frac{r_L^* t_L}{t_L^*} e^{ik'x} + e^{-ik'x} & x \to \infty \end{cases} . \tag{12.28}$$

We see that time reversal tells us that

$$t_R = t_L$$
$$r_R = -\frac{r_L^* t_L}{t_L^*} . \tag{12.29}$$

12.7 Parity

If the potential is symmetric about the origin, the Hamiltonian commutes with parity, and operating on a scattering eigenstate gives another eigenstate, so $\psi(-x)$ is also an eigenstate. In this case $k' = k$ so that

$$\psi(-x) = \begin{cases} t_L e^{-ikx} & x \to -\infty \\ e^{-ikx} + r_L e^{ikx} & x \to \infty \end{cases} \tag{12.30}$$

which is immediately the state incoming from the right. For a symmetric potential we have

$$r_R = r_L$$
$$t_R = t_L \tag{12.31}$$

and substituting into Eq. 12.29

$$\text{Re } r_L t_L^* = 0 . \tag{12.32}$$

12.8 Translated Hamiltonian

If we have a solution for the scattering from a potential $v(x)$,

$$\psi(|x| \to \infty) = \begin{cases} e^{ikx} + r_L e^{-ikx} & x \to \infty \\ t_L e^{ik'x} & x \to \infty \end{cases} \tag{12.33}$$

the solution for a potential $v(x - b)$ translated along x by b, is

$$\psi(x - b) = \begin{cases} e^{ik(x-b)} + r_L e^{-ik(x-b)} & x \to \infty \\ t_L e^{ik'(x-b)} & x \to \infty \end{cases} \tag{12.34}$$

and factoring out e^{-ikb} gives

$$\psi(x - b) = e^{-ikb} \begin{cases} e^{ikx} + r_L e^{ik2b} e^{-ikx} & x \to \infty \\ t_L e^{i(k-k')b} e^{ik'x} & x \to \infty \end{cases} \tag{12.35}$$

so that

$$r_L(b) = r_L(0) e^{ik2b}$$
$$t_L(b) = t_L(0) e^{i(k-k')b} \tag{12.36}$$

116

12.9 Scattering matrix

The Scattering or S-matrix is often used in understanding more complicated scattering. To make the analysis simpler, we will now specialize to the usual case where the potential is zero outside the target. In that case $k' = k$. The general solution with arbitrary boundary conditions of the time-independent Schrödinger equation has the form

$$\psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & x \to -\infty \\ Ce^{ikx} + De^{-ikx} & x \to \infty \end{cases}. \tag{12.37}$$

Solving the time-independent Schrödinger equation gives two linear relationships between the coefficients. The S-matrix relates the coefficients of the incoming waves to those of the outgoing waves

$$\begin{pmatrix} B \\ C \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A \\ D \end{pmatrix}. \tag{12.38}$$

The case where $A = 1$, and $D = 0$, is the usual case where the incoming wave is from the left, and therefore $B = r_L$, and $C = t_L$. Plugging in shows that $S_{11} = r_L$, $S_{21} = t_L$. Similarly the case where $D = 1$ and $A = 0$, is the case where the incoming wave is from the right, and plugging in shows that $S_{22} = r_R$, and $S_{12} = t_R$.

The S-matrix is unitary. We can see this by applying probability conservation. The continuity equation you derived, $\nabla \cdot j + \frac{\partial \rho}{\partial t} = 0$, for the probability current density becomes, for an eigenstate of H, $\nabla \cdot j = 0$. Applying this just to the H_x eigenstates gives $dj_x/dx = 0$, or j_x is a constant, independent of x. The current density for $x \to -\infty$ is $|A|^2 - |B|^2$ multiplied by $\hbar k/m$, and the current density for $x \to \infty$ is $|C|^2 - |D|^2$ multiplied by $\hbar k/m$. Equating these shows that $|A|^2 + |D|^2 = |B|^2 + |C|^2$. That is the total incoming current is equal to the total outgoing current. We can write this as

$$\begin{pmatrix} A^* & D^* \end{pmatrix} \begin{pmatrix} A \\ D \end{pmatrix} = \begin{pmatrix} B^* & C^* \end{pmatrix} \begin{pmatrix} B \\ C \end{pmatrix} = \begin{pmatrix} A^* & D^* \end{pmatrix} S^\dagger S \begin{pmatrix} A \\ D \end{pmatrix} \tag{12.39}$$

where S represents the 2×2 S-matrix of Eq. 12.38. Since this must be true for any choice of A and D, we require $S^\dagger S = 1$, or S is unitary. Multiplying S times $S^\dagger$ and setting it to the identity gives the four equations

$$\begin{aligned} S_{11}S_{11}^* + S_{12}S_{12}^* &= 1 \\ S_{11}S_{21}^* + S_{12}S_{22}^* &= 0 \\ S_{21}S_{11}^* + S_{22}S_{12}^* &= 0 \\ S_{21}S_{21}^* + S_{22}S_{22}^* &= 1 \end{aligned} \tag{12.40}$$

The second and third equations are complex conjugates, so they give one condition. Substituting into the other equations gives the requirements

$$\begin{aligned} 1 &= |S_{11}|^2 + |S_{21}|^2 \\ |S_{11}| &= |S_{22}| \\ |S_{12}| &= |S_{21}| \\ 0 &= S_{11}S_{21}^* + S_{12}S_{22}^*. \end{aligned} \tag{12.41}$$

This will be true for any scattering system with just two channels.

For our scattering, we know the potential is real and the system is time reversal invariant. We already studied this above. Taking the complex conjugate of the solution interchanges the incoming and outgoing waves. Therefore we can equate the S matrix with the complex conjugate of its inverse. Since the inverse is the complex conjugate transpose, time reversal invariance means the S-matrix is equal to its transpose, so it must be symmetric. The additional condition is that $S_{12} = S_{21}$. This is the $t_L = t_R$ condition (for $k = k'$) that we derived directly above. The condition that $r_R = -r_L^* t_L / t_R^*$ is the unitary condition $S_{11} S_{21}^* + S_{12} S_{22}^* = 0$ with $S_{12} = S_{21}$.

For the special potentials that are invariant under parity, we can take $x \to -x$ which interchanges 1 and 2, so we have $S_{11} = S_{22}$, and $S_{12} = S_{21}$. Since these are already time reversal invariant, $S_{11} = S_{22}$ is the only additional constraint.

12.10   Phase shifts

You may have seen that, many times, scattering experiments can be described by reporting phase shifts. We see here that scattering can be described by the unitary S-matrix. If we diagonalize that S-matrix, we get its eigenvectors and eigenvalues. The eigenvalues can be written as phase factors $e^{i2\delta}$, and the δ values are called the phase shifts. In general we need to calculate the eigenvectors too.

For some physical systems, the eigenvectors are given by symmetries of the system, and are independent of the details of the interaction. For example, for spherically symmetric potentials in three dimensions (i.e. central potentials), the eigenvectors are the usual angular momentum eigenfunctions, and we just need to give the phase shifts for each angular momentum value to completely specify the S-matrix and therefore the scattering.

The equivalent of a central potential for a 1-dimensional problem, is a potential that only depends on $|x|$. Such a potential is invariant under $x \to -x$, so the Hamiltonian commutes with the parity operator. As we saw above, this means that the S-matrix has the form

$$S = \begin{pmatrix} S_{11} & S_{21} \\ S_{21} & S_{11} \end{pmatrix}. \tag{12.42}$$

The eigenvectors of any matrix of this form are $\frac{1}{\sqrt{2}}(1, \pm 1)$, so we can fully specify the matrix by giving its phase shifts. Writing S in terms of its eigenvalues $e^{i\delta_+}$ and $e^{i\delta_-}$ and eigenvectors

$$S = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} e^{i2\delta_+} & 0 \\ 0 & e^{i2\delta_-} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^{i2\delta_+} + e^{i2\delta_-} & e^{i2\delta_+} - e^{i2\delta_-} \\ e^{i2\delta_+} - e^{i2\delta_-} & e^{i2\delta_+} + e^{i2\delta_-} \end{pmatrix}$$
$$\tag{12.43}$$

The phase shifts get their name since these combinations of incoming waves get multiplied by the phase factor to become the combination of outgoing waves. Here we see that since parity is conserved, equal amplitude incoming waves from the left and write are an even parity solution. The outgoing waves must also be even parity so they also have equal amplitudes on the left and right. The only effect of the potential is to change the phase. A similar

arguments goes through for the odd parity combinations. For central potentials, exactly the same argument goes through for the different angular momentum components. An incoming component with a particular ℓ, m value must give the same magnitude outgoing component. Only the phase can be modified by the potential giving the different phase shifts in different angular momentum channels.

12.11   Translation of the S-matrix

Applying the result of section 12.8 the potential $v(x - b)$ gives an S-matrix

$$S(b) = \begin{pmatrix} S_{11}e^{i2kb} & S_{12} \\ S_{21} & S_{22}e^{-i2kb} \end{pmatrix}, \qquad (12.44)$$

in terms of the untranslated S-matrix components.

12.12   Combining non overlapping potentials

Sometimes scattering occurs from a layered structure where the layers repeat. For example we might scatter from the potential $v(x + b/2) + v(x - b/2)$, where if the potential $v(x)$ is centered around the origin, $v(x \pm b/2)$ is centered around $x = \mp b/2$. If $v(x)$ goes to zero fast enough so that the potentials do not overlap, we can solve the scattering from each separately and combine the results. We wish to solve with the boundary conditions

$$\psi(x) = \begin{cases} e^{ikx} + re^{-ikx} & x \to -\infty \\ Ae^{ikx} + Be^{-ikx} & x \text{ between } v(x) \text{ and } v(x - b) \\ te^{ikx} & x \to \infty \end{cases} . \qquad (12.45)$$

In terms of the S matrix for $v(x)$, we can write down the necessary conditions. The only incoming wave on the potential $v(x - b/2)$ is from the A term, so we have

$$AS_{21} = t \qquad (12.46)$$
$$AS_{11}e^{ikb} = B \qquad (12.47)$$
$$\qquad (12.48)$$

For the potential $v(x + b/2)$ we have incoming terms from B and the 1 term. The S-matrix gives

$$r = S_{11}e^{-ikb} + BS_{12}$$
$$A = BS_{22}e^{ikb} + S_{21} . \qquad (12.49)$$

This gives 4 equations and 4 unknowns. Solving, we have

$$t = \frac{S_{21}^2}{1 - S_{11}S_{22}e^{2ikb}}$$
$$r = S_{11}e^{-ikb}\left[1 + \frac{S_{21}S_{12}e^{2ikb}}{1 - S_{11}S_{22}e^{2ikb}}\right] \qquad (12.50)$$

Sometimes an interative solution is more illuminating. We start with an incoming wave of unit amplitude from the left. Matching the boundary conditions at the potential at $x = -b/2$, we get a reflected wave of amplitude $S_{11}e^{-ikb}$, and a wave of amplitude S_{21} traveling to the right toward the potential at $x = b/2$. At this point, we have matched the boundary conditions for the potential at $x = -b/2$, but not $x = b/2$. Matching the boundary condition for this wave at the potential at $x = b/2$, we get a transmitted wave of amplitude S_{21}^2, and a reflected wave of amplitude $S_{21}S_{11}e^{ikb}$. At this point, we have matched the boundary conditions for the potential at $x = b/2$, but not $x = -b/2$. The wave traveling to the left between the potentials gives a reflected contribution of $S_{21}S_{11}e^{ikb}S_{12}$, and a wave traveling to the right with amplitude $S_{21}S_{11}e^{ikb}S_{12}S_{22}e^{ikb}$. Matching at the potential at $x = b/2$, this gives a transmitted contribution of $S_{21}S_{11}e^{ikb}S_{12}S_{22}e^{ikb}S_{21}$, etc. Continuing the series we get

$$
\begin{aligned}
x &= S_{11}S_{22}e^{2ikb} \\
t &= S_{21}^2\left[1 + x + x^2 + x^3...\right] = \frac{S_{21}^2}{1-x} = \frac{S_{21}^2}{1 - S_{11}S_{22}e^{2ikb}} \\
r &= S_{11}e^{-ikb} + S_{21}S_{11}e^{ikb}S_{12}\left[1 + x + x^2 + x^3...\right] = S_{11}e^{-ikb}\left[1 + \frac{S_{21}S_{12}e^{2ikb}}{1 - S_{11}S_{22}e^{2ikb}}\right],
\end{aligned}
\tag{12.51}
$$

which agrees with our direct solution.

For more complicated structures, other matrices that relate the waves on each side of a potential can be used. The transfer matrix is one choice. It relates the wave function and its derivavive at x to the wave function and its derivative at x'. The transfer matrices for the different potentials can then be multiplied together to get the total transfer matrix.

12.13 Resonant Form

Unitarity requires

$$
S_{11}S_{22} = -\frac{S_{12}|S_{22}|^2}{S_{21}^*} = -(1 - |S_{21}|^2)\frac{S_{12}}{S_{21}^*}
\tag{12.52}
$$

and with time reversal invariance $S_{12} = S_{21}$, so writing $S_{21} = |S_{21}|e^{i\phi_{21}}$, the transmission coefficient becomes

$$
t = \frac{|S_{21}|^2 e^{2i\phi_{21}}}{1 + (1 - |S_{21}|^2)e^{2i(kb+\phi_{12})}}
\tag{12.53}
$$

which shows that $|t| = 1$ if $e^{2i(kb+\phi_{12})} = -1$. If at some energy E_0, we have this relationship, we can expand $1/t \equiv t^{-1}$ around E_0, as

$$
t^{-1}(E) = t^{-1}(E_0) + \frac{dt^{-1}(E_0)}{dE_0}(E - E_0) + ...
\tag{12.54}
$$

and if we can drop the higher order terms, the transmission coefficient takes the resonant form

$$t(E) = t(E_0)\frac{i\Gamma/2}{E - E_0 + i\Gamma/2} \tag{12.55}$$

with

$$i\frac{\Gamma}{2} = \left[t(E_0)\frac{dt^{-1}(E_0)}{dE_0}\right]^{-1}. \tag{12.56}$$

Let's look at the case where $v(x) = \lambda\delta(x)$. Writing

$$\psi(x) = \begin{cases} e^{ikx} + S_{11}e^{-ikx} & x < 0 \\ S_{21}e^{ikx} & x > 0 \end{cases}, \tag{12.57}$$

integrating across the origin of the time independent Schrödinger equation we get the boundary conditions

$$ikS_{21} - ik(1 - S_{11}) = \frac{2m\lambda}{\hbar^2}S_{21}$$
$$1 + S_{11} = S_{21} \tag{12.58}$$

with solution

$$S_{11} = -\frac{i\frac{m\lambda}{\hbar^2 k}}{1 + i\frac{m\lambda}{\hbar^2 k}}$$
$$S_{21} = \frac{1}{1 + i\frac{m\lambda}{\hbar^2 k}} = \frac{1}{1 + i\lambda\sqrt{\frac{m}{2\hbar^2 E}}} \tag{12.59}$$

Notice that when $\lambda \to \infty$, that $S_{21} \to -i/\lambda\sqrt{\frac{m}{2\hbar^2 E}}$ so that $\phi_{21} \to -\pi/2$, and the transmission resonances for a pair of delta function potentials separated by b occur when $k = k_n = n\pi/b$, that is at the energies of the bound state of the almost infinite square well formed by the two very repulsive delta functions.

Therefore the resonances occur at $E_0 = \hbar^2 k_n^2/2m$. In this limit of strong repulsive delta function potentials, we can simplify our calculations and write

$$|S_{21}|^2 \to \frac{2\hbar^2 E}{m\lambda^2}$$
$$e^{2i\phi_{21}} \to -1 \tag{12.60}$$

Notice that for large λ, $|S_{21}|^2$ is small and changes slowly. In this limit, the dominant term in the derivative of Eq. 12.56 is the derivative of the e^{2ikb} term. Keeping just this part, we get

$$\frac{\Gamma}{2} = \frac{|S_{21}|^2\hbar^2 k_n}{2mb}. \tag{12.61}$$

Physically, the resonance can be thought of as the particle trapped in the almost bound state of the well between the delta functions. If we make a semiclassical picture, we could imagine

that the particle is bouncing back and forth between the delta functions, and $\Gamma/\hbar$ represents the rate with which it escapes. We could estimate that rate as the rate at which the particle hits a delta function which would be $v/b = \hbar k_n/mb$ times the probability of transmission through the delta function which is $|S_{21}|^2$. Multiplying these together we see we get Γ, so the semiclassical picture is reasonable.

12.14 Problems

Problem 12.1. A particle of mass m is incident on the one-dimensional potential barrier

$$v(x) = \begin{cases} 0 & x < 0 \\ \lambda x & x > 0 \end{cases} ,$$ (12.62)

where $\lambda > 0$ is a strength parameter.

a. Calculate the reflection coefficient as a function of $\epsilon = \left(\frac{2m}{\lambda^2 \hbar^2}\right)^{1/3} E_x$ where $E_x = \frac{p_x^2}{2m}$, and p_x is the x component of the incident momentum. My result is

$$r = \exp\left[2i \tan^{-1}\left(\frac{\text{Ai}'(-\epsilon)}{\sqrt{\epsilon}\text{Ai}(-\epsilon)}\right)\right]$$ (12.63)

where Ai is the Airy function[2], the solution of $\text{Ai}''(z) - z\text{Ai}(z) = 0$ that goes to zero for $z \to \infty$.

b. The time delay t_{delay} for a reflected packet is the additional time it takes the reflected packet to be peaked around some large negative x value compared to a packet that reflected off an infinitely repulsive potential at $x = 0$. Define $\tau = \left(\frac{\lambda^2}{2m\hbar}\right)^{1/3} t_{\text{delay}}$. Show that τ diverges like $\epsilon^{-1/2}$ for $\epsilon \to 0$. Plot τ as a function of ϵ for the range $0.1 < \epsilon < 10$. Plot on the same graph $\tau_{\text{classical}}$, the τ value of the time delay for a classical particle of the same mass bouncing off this same potential.

---

[2]See for example, M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, (National Bureau of Standards, New York, 1964), Section 10.4. This is also available on line at http://www.math.sfu.ca/~cbm/aands/

Chapter 13

Inelastic scattering in a slab geometry

13.1   An example problem

Let's assume that we have a target particle in a bound state around the origin. This could be a model of an electron in a hydrogen atom bound to a proton. Alternatively, it could be an atom bound by an electromagnetic trap. The point is that our target doesn't just look like a simple potential, but rather has quantum structure of its own.

Often the reason to scatter from a target is to learn about the structure of the target, not about the interaction of the projectile and the target. For example, we might want to use a scattering experiment to determine the excitation energies of the target. Neutron scattering in solids is often used in just this way to determine the phonon dispersion curves (that is the phonon energy as a function of wave vector, $\omega(\boldsymbol{k})$).

As a simple model of inelastic scattering, we can study the two-particle Hamiltonian, in the usual x-representation,

$$-\frac{\hbar^2}{2m_1}\frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m_2}\frac{\partial^2}{\partial x_2^2} + v(x_2) + v_{12}(x_1, x_2) \tag{13.1}$$

In Dirac notation, the Hamiltonian operator is,

$$
\begin{aligned}
H &= \int dx_1 dx_2 |x_1, x_2\rangle \left[-\frac{\hbar^2}{2m_1}\frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m_2}\frac{\partial^2}{\partial x_2^2} + v(x_2) + v_{12}(x_1, x_2) \right] \langle x_1, x_2| \\
&= H_1 + H_2 + v_{12}, \tag{13.2}
\end{aligned}
$$

where in the last line I define the three terms of the Hamiltonian.

The experiment we will describe starts with the target particle in its ground state and a projectile particle traveling toward it with a well defined energy E. The interaction term v_{12} is assumed to be small except when the projectile and target are close together.

I take a basis for solving this problem to be the outer product of the usual x basis for particle 1 and the eigenstates of the Hamiltonian for particle 2. That is,

$$\left[-\frac{\hbar^2}{2m_2}\frac{\partial^2}{\partial x_2^2} + v(x_2) \right] \psi_n(x_2) = \epsilon_n \psi_n(x_2), \tag{13.3}$$

with $\psi_n(x_2) = \langle x_2|n\rangle$. A general wave function can then be written as,

$$|\Psi\rangle = \sum_n \int dx_1 a_n(x_1)|x_1\rangle|n\rangle, \qquad (13.4)$$

where $|x_1\rangle|n\rangle$ is equivalent to $|x_1, n\rangle$, but stresses that the basis states are outer product states. $a_n(x)$ is therefore $\langle x_1, n|\Psi\rangle$, which is the amplitude for finding particle 1 at position x_1, and particle two in energy eigenstate n.

We obtain differential equation forms of the Schroedinger equation

$$H|\Psi\rangle = E_T|\Psi\rangle, \qquad (13.5)$$

by dotting it on the left with a basis function $\langle x, m|$. The matrix elements we need are,

$$
\begin{aligned}
\langle x, m|E_T|\Psi\rangle &= \sum_n \int dx_1 dx' \langle x|x_1\rangle\langle m|n\rangle E_T \langle x_1|x'\rangle a_m(x') \\
&= E_T a_m(x) \qquad (13.6) \\
\langle x, m|H_1|\Psi\rangle &= \sum_n \int dx_1 dx' \langle x|x_1\rangle\langle m|n\rangle \left(-\frac{\hbar^2}{2m_1}\frac{\partial^2}{\partial x_1^2}\right) \langle x_1|x'\rangle a_n(x') \\
&= -\frac{\hbar^2}{2m_1}\frac{\partial^2}{\partial x^2} a_m(x) \qquad (13.7) \\
\langle x, m|H_2|\Psi\rangle &= \sum_n \int dx_1 dx' \langle x|x_1\rangle\langle m|n\rangle \epsilon_n \langle x_1|x'\rangle a_n(x') \\
&= \epsilon_m a_m(x) \qquad (13.8) \\
\langle x, m|v_{12}|\Psi\rangle &= \sum_n \int dx_1 dx_2 dx' \langle x|x_1\rangle\langle m|x_2\rangle v_{12}(x_1, x_2)\langle x_2|n\rangle\langle x_1|x'\rangle a_n(x') \\
&= \sum_n a_n(x) \int dx_2 \psi_m^*(x_2) v_{12}(x, x_2)\psi_n(x_2) \\
&\equiv \sum_n a_n(x)\alpha_{mn}(x) \qquad (13.9)
\end{aligned}
$$

where the last line defines $\alpha_{mn}(x)$ in terms of the matrix elements of v_{12}.

Combining all these gives us a set of coupled differential equations that we need to solve,

$$-\frac{\hbar^2}{2m_1}\frac{\partial^2}{\partial x^2} a_m(x) + \sum_n \alpha_{mn}(x)a_n(x) = (E_T - \epsilon_m)a_m(x). \qquad (13.10)$$

A physical interaction v_{12} will go to zero when the two particles are well separated. We are looking for a wave function in this region whose only incoming amplitude is particle 1 coming from the left with particle 2 in its ground state. Particle 1 can be outgoing in either direction, and in principle if particle 2 has unbound solutions, we can also have outgoing

126

solutions to particle 2. That means that the solution for our wave function must satisfy the boundary conditons

$$
\begin{aligned}
a_n(x \to -\infty) &= \delta_{n,0} \exp(ik_n x) + r_n \exp(-ik_n x) \\
a_n(x \to \infty) &= t_n \exp(ik_n x)
\end{aligned}
\tag{13.11}
$$

with $\hbar k_n = \sqrt{2m_1(E_T - \epsilon_n)}$ when $E_T - \epsilon_n > 0$, and

$$
\begin{aligned}
a_n(x \to -\infty) &= r_n \exp(\kappa_n x) \\
a_n(x \to \infty) &= t_n \exp(-\kappa_n x)
\end{aligned}
\tag{13.12}
$$

with $\hbar \kappa_n = \sqrt{-2m_1(E_T - \epsilon_n)}$ when $E_T - \epsilon_n < 0$.

Solving a set of many coupled differential equations like this is not difficult numerically. To make an analytic solution that displays the physics, we need to simplify things a bit. I assume that the physical energies of the projectile particle 1 and the interaction v_{12} are such that our solution will have $a_n(x)$ small if $n \neq 0, 1$. It is easy to see how to generalize this to more states.

Further I take an interaction that has no "diagonal" term, and always causes the state of particle 2 to change. This is not so far fetched since a reasonable potential $v_{12}(x_1, x_2) = f(x_1 - x_2) - f(x_1)$ where the target consists of two "oppositely charged" particles, particle 2 and a heavy particle at the origin. The projective interacts with each with opposite sign. If the size of the bound state is much smaller than the range of v, this can be expanded as $f(x_1 - x_2) - f(x_1) \sim f'(x_1)x_2$, and for a symmetric potentials $v(x)$, the states have good parity and this interaction will have a zero matrix elements between states of the same parity.

Further, I take an interaction where the range over which it is not zero is much smaller than a deBroglie wave length of particle 1, and assume that the bound state wave functions of particle 2 are well localized at the origin. In that case we can approximate α_{mn} by a δ function which has the same integral. This reduces our $\alpha_{mn}(x)$ to

$$
\alpha_{10}(x) = \alpha_{01}(x) = \lambda \delta(x)
\tag{13.13}
$$

and all other subscript values give zero.

Defining $E = E_T - \epsilon_0$ and $\epsilon = \epsilon_1 - \epsilon_0$ the differential equations are

$$
\begin{aligned}
-\frac{\hbar^2}{2m_1} \frac{\partial^2}{\partial x^2} a_0(x) + \lambda \delta(x) a_1(x) &= E a_0(x), \\
-\frac{\hbar^2}{2m_1} \frac{\partial^2}{\partial x^2} a_1(x) + \lambda \delta(x) a_0(x) &= (E - \epsilon) a_0(x),
\end{aligned}
\tag{13.14}
$$

I define the four quantities that I want to calculate:

- $R_0 =$ the probability of the projectile being reflected leaving the target in its ground state.

- $T_0 =$ the probability of the projectile being transmitted leaving the target in its ground state.

- R_1 = the probability of the projectile being reflected and the target being excited to its first excited state.

- T_1 = the probability of the projectile being transmitted and the target being excited to its first excited state.

The solutions with the delta function interaction are exactly of the form of Eqs. 13.11 and 13.12. For the below threshold case where $E < \epsilon$,

$$
a_0(x) = \begin{cases} \exp(ik_0 x) + r_0 \exp(-ik_0 x) & x < 0 \\ t_0 \exp(-ik_0 x) & x > 0 \end{cases}
$$
$$
a_1(x) = \begin{cases} r_1 \exp(\kappa_1 x) & x < 0 \\ t_1 \exp(-\kappa_1 x) & x > 0 \end{cases}
$$
$$(13.15)$$

Using continuity of the wave function and integrating the differential equation across the delta function gives the boundary conditions,

$$
\begin{aligned}
1 + r_0 &= t_0 \\
r_1 &= t_1 \\
ik_0(1 - r_0) - ik_0 t_0 &= -\frac{2m\lambda}{\hbar^2} t_1 \\
\kappa_1(r_1 + t_1) &= -\frac{2m\lambda}{\hbar^2} t_0
\end{aligned}
$$
$$(13.16)$$

It is convenient to define the dimensionless variable

$$
s \equiv \frac{m\lambda^2}{2\hbar^2 \sqrt{|E(E - \epsilon)|}}
$$
$$(13.17)$$

The solution for $E < \epsilon$ gives

$$
\begin{aligned}
t_0 &= \frac{1}{1 - is} \\
r_0 &= \frac{-is}{1 - is}.
\end{aligned}
$$
$$(13.18)$$

Since the amplitudes go to zero at large $|x|$ values, we get the physical result that if the projectile does not have enough energy, the target cannot be left in an excited state. We get

$$
\begin{aligned}
R_0 &= |r_0|^2 = \frac{s^2}{1+s^2} & E < \epsilon \\
T_0 &= |t_0|^2 = \frac{1}{1+s^2} & E < \epsilon \\
R_1 &= 0 & E < \epsilon \\
T_1 &= 0 & E < \epsilon
\end{aligned}
$$
$$(13.19)$$
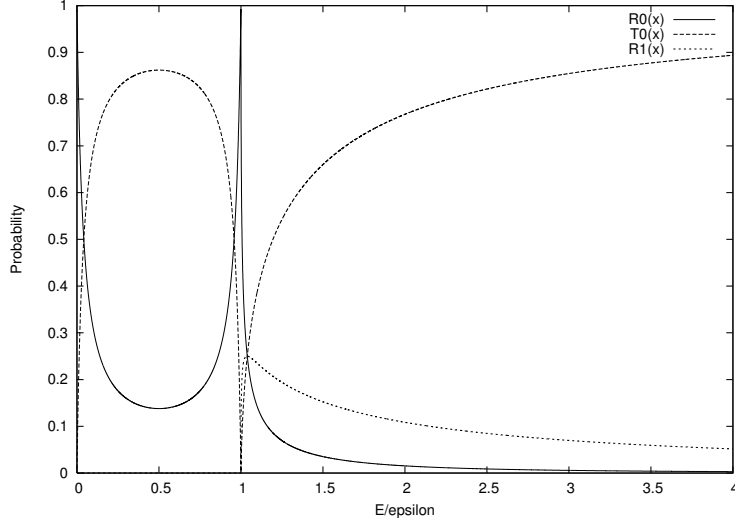
Figure 13.1: The scattering probabilities as a function of the the energy E/ϵ with $m\lambda^2/(2\hbar^2\epsilon) = .2$

The case where $E > \epsilon$ is nearly identical, and it is easily verified that the differential equation solutions and boundary conditions at the origin are the same with κ_1 replaced by $-ik_1$. The solutions are

$$
\begin{aligned}
t_0 &= \frac{1}{1+s} \\
r_0 &= \frac{s}{1+s} \\
r_1 &= = t_1 = \frac{\lambda}{2ik_1}\frac{1}{1+s}.
\end{aligned}
\tag{13.20}
$$

Note that since incident particle changes energy when the target is excited, we need to include the same flux factors that we need to determine the transmission through a step. The result is

$$
\begin{aligned}
R_0 &= |r_0|^2 = \frac{s^2}{(1+s)^2} & E > \epsilon \\
T_0 &= |t_0|^2 = \frac{1}{(1+s)^2} & E > \epsilon \\
R_1 = T_1 &= \frac{k_1}{k_0}|t_1|^2 = \frac{s}{(1+s)^2} & E > \epsilon
\end{aligned}
\tag{13.21}
$$
.

Notice that the sum of the four probabilities adds to 1 in all cases.

In figures 13.1 to 13.3 below, I show R_0, $R_1 = T_1$, and T_0, as a function of the energy E/ϵ. The three figures are for different strengths of the delta function compared to the excitation energy of the target.
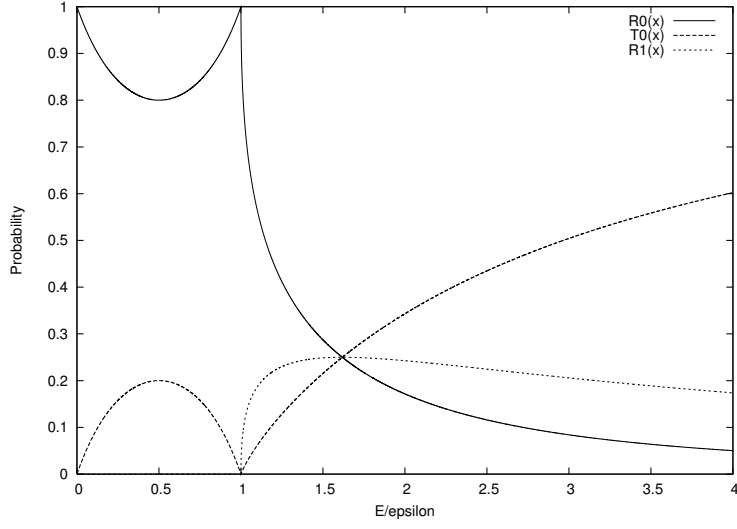
Figure 13.2: The scattering probabilities as a function of the the energy E/ϵ with $m\lambda^2/(2\hbar^2\epsilon) = 1$.
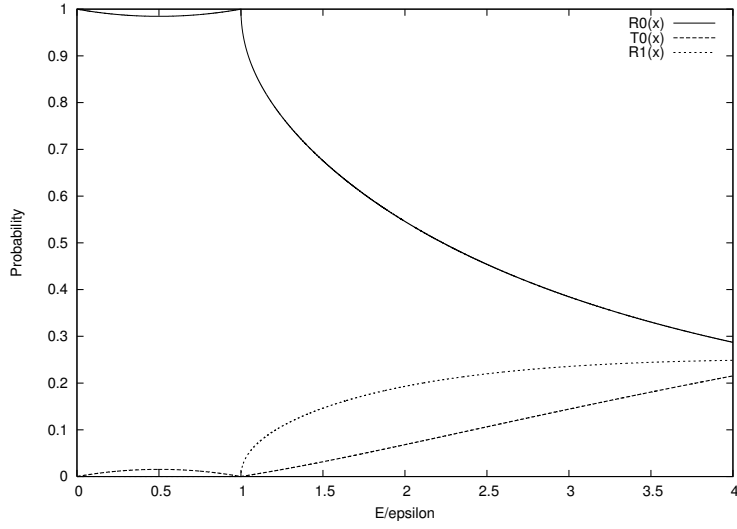
Figure 13.3: The scattering probabilities as a function of the the energy E/ϵ with $m\lambda^2/(2\hbar^2\epsilon) = 4$.

Chapter 14

Periodic potentials

14.1   Introduction

Often we want the eigenstates of a periodic system. For example, a first approximation for the potential seen by valence electrons in a solid is given by the periodic array of nuclei with attached core electrons. After solving this problem, we can add the effects of nuclear motion (i.e. phonons), electron-electron interactions, and changes caused by impurities or dislocations. Similarly, the periodicity of the system can be used when solving for the phonon normal modes of a crystalline solid.

14.2   Translation operator

We will work in just one dimension, but the method easily generalizes to two or three dimensions. A typical Hamiltonian for one particle would be

$$\frac{p^2}{2m} + v(x) \tag{14.1}$$

where the potential is periodic under a translation by a, $v(x + a) = v(x)$.

To make state counting easier, we will initially place our system in a large periodic box so that $-L/2 < x \le L/2$, with periodic boundary conditions so that we identify $x \pm L$ as the same position as x. We take $L = Na$, and let N go to infinity at the end of the calculation.

The translation operator that translates by a is

$$
\begin{aligned}
T &= \int_{-L/2}^{L/2} dx\, |x + a\rangle\langle x| \\
&= \int_{-a/2}^{a/2} dx \sum_{j=-N/2+1}^{N/2} |x + a(j+1)\rangle\langle x + aj|, \tag{14.2}
\end{aligned}
$$

where in the second line the integral over all space has been broken up into a sum of integrals over intervals of length a, and the sum and integral interchanged. Notice that this shows

that the T operator is block diagonal, i.e. it only couples states with the same x value in $|x+aj\rangle$. We can form the matrix elements of these states, writing the nth eigenstates of the translation of states with a given $x+aj$ value

$$|x\ n\rangle = \sum_{j=-N/2+1}^{N/2} \psi_j^{(n)}|x+aj\rangle \tag{14.3}$$

we have

$$\psi_{j+1}^{(n)} = \lambda_n \psi_j^{(n)}\,. \tag{14.4}$$

Enforcing periodic boundary conditions gives the result as we found in the problem sets,

$$\psi_j^{(n)} = e^{i\frac{2\pi n j}{N}} \tag{14.5}$$

with $\lambda_n = e^{i\frac{2\pi n}{N}}$, and $-N/2 < n \le N/2$. We have not yet normalized anything. The eigenstates of T are therefore

$$|x\ n\rangle = \sum_{j=-N/2+1}^{N/2} e^{i\frac{2\pi j n}{N}}|x+ja\rangle \tag{14.6}$$

with eigenvalues $e^{i\frac{2\pi n}{N}}$.

We can write $N = L/a$ so that

$$\frac{2\pi n}{N} = \frac{2\pi n}{L}a \equiv k_n a\,. \tag{14.7}$$

Taking the limit of $N \to \infty$, the spacing of the k_n goes to zero, and the sums over n can be taken to integrals over the k_n value, and we drop the subscript since it no longer has any use. The range of k values is then

$$-\frac{\pi}{a} < k \le \frac{\pi}{a}\,. \tag{14.8}$$

The eigenstates of T can be labeled by the $-a/2 < x \le a/2$ and $-\pi/a < k \le \pi/a$ values

$$|x\ k\rangle = e^{ikx}\sum_j e^{ik(ja)}|x+ja\rangle = \sum_j e^{ik(x+ja)}|x+ja\rangle \tag{14.9}$$

with eigenvalues e^{ika}. The phase of the eigenvectors is our choice. Here we will see that including the additional phase factor e^{ikx} will simplify our later analysis.

The states with the same k but different x values are all degenerate so that we can form an equivalent set of eigenstates of T by taking orthogonal linear combinations of those states. A general linear combination is

$$|\psi_k\rangle = \int_{-a/2}^{a/2} dx'\ u_k(x')\sum_j e^{ik(x'+ja)}|x'+ja\rangle\,. \tag{14.10}$$

Taking the matrix element with $\langle x|$ to give the form of the wave functions, we have

$$
\begin{aligned}
\psi_k(x) &= \langle x|\psi_k\rangle = \int_{-a/2}^{a/2} dx' \, u_k(x') \sum_j e^{ik(x'+ja)} \langle x|x'+ja\rangle \\
&= u_k(x-ja)e^{ikx} \quad \text{for} \; -a/2 + ja < x \le a/2 + ja \,,
\end{aligned} \tag{14.11}
$$

or

$$
\psi_k(x+ja) = u_k(x)e^{ik(x+ja)} \,. \tag{14.12}
$$

Here our x is defined only in the interval $-a/2 < x \le a/2$, so $u_k(x)$ only needs to be defined in that interval. However, if we define $u_k(x)$ outside that interval to be periodic, so that $u_k(x+ja) = u_k(x)$, where $x + ja$ is outside the original interval, we can write

$$
\psi_k(x) = u_k(x)e^{ikx} \tag{14.13}
$$

everywhere. Continuity of $\psi_k(x)$ means that $u_k(x)$ is continuous; $u_k(x)$ is periodic and continuous with period a. Eq. 14.13 is called a Bloch wave. This result is called Bloch's theorem.

If our Hamiltonian commutes with T, that is it has a periodic potential, we can find simultaneous eigenvectors of T and H. Since the eigenvectors of T have the form given by Bloch's theorem, the eigenvectors of H also must have this form.

14.3 Empty Lattice

The potential $v(x) = 0$ is trivially periodic. We know that energy eigenfunctions are $\psi(x) = e^{i\kappa x}$ where $-\infty < \kappa < \infty$. We can see how to rewrite the forms above to give these results. Writing

$$
\kappa = k + \frac{2\pi m}{a} \equiv k + K_m \tag{14.14}
$$

where $m = \text{nint}\frac{\kappa a}{2\pi}$, nint means the nearest integer, and $K_m = 2\pi m/a$ is called a reciprocal lattice vector,

$$
e^{i\kappa x} = e^{ikx}e^{iKx} \tag{14.15}
$$

so that $u_k(x) = e^{iKx}$ which is periodic with period a.

If you Fourier transform a well behaved periodic function like $u_k(x)$, only the wave vectors that correspond to periodic functions with period a will contribute. These are the reciprocal lattice wave vectors. $u_k(x)$ can be written as a Fourier series using the reciprocal lattice vectors.

14.4 Kronig-Penney Model

The Kronig-Penney model originally was a piecewise constant periodic potential. However, the limit is usually taken, as in Baym, where the potential is a periodic array of delta

functions

$$v(x) = v_0 \sum_j \delta(x - ja) \,. \tag{14.16}$$

Following Baym, we define q such that the energy is $E = \hbar^2 q^2 / 2m$, we know that the solutions for the wave functions in the region where the potential is zero are $Ae^{iqx} + Be^{-iqx}$. Therefore the Bloch functions in these regions have the form

$$u_k(x) = Ae^{i(q-k)x} + Be^{-i(q+k)x} \,, \tag{14.17}$$

so this can be the Bloch function in the region $0 < x < a$. In order for this to be periodic and continuous, the values at $x = 0$ and $x = a$ must agree

$$A + B = Ae^{i(q-k)x} + Be^{-i(q+k)x} \tag{14.18}$$

and the slope discontinuity is given by integrating across the delta function

$$0 = -\frac{\hbar^2}{2m} \int_{0-}^{0+} dx \frac{d^2\psi(x)}{dx^2} + v_0\psi(0) = -\frac{\hbar^2}{2m}\left[\left.\frac{d\psi(x)}{dx}\right|_{0+} - \left.\frac{d\psi(x)}{dx}\right|_{0-} \right] + v_0\psi(0) \tag{14.19}$$

or in terms of $u_k(x)$,

$$
\begin{aligned}
\left.\frac{d\psi(x)}{dx}\right|_{0+} &= iqA - iaB \\
\left.\frac{d\psi(x)}{dx}\right|_{0-} &= \left.\frac{de^{ikx}u_k(x)}{dx}\right|_{0-} = iku_k(0) + Ai(q-k)e^{i(q-k)a} - Bi(q+k)e^{-i(q+k)a} \\
&= iq\left[Ae^{i(q-k)a} - Be^{-i(q+k)a} \right]
\end{aligned}
\tag{14.20}
$$

so that

$$iq\left[(A - B - Ae^{i(q-k)a} + Be^{-i(q+k)a} \right] = \frac{2mv_0}{\hbar^2}(A + B) \,. \tag{14.21}$$

Solving we get the condition

$$\cos(ka) - \cos(qa) - \frac{mv_0}{\hbar^2 q}\sin(qa) = 0 \,. \tag{14.22}$$

So given a k value, we can solve for the possible q values which give the possible energies. Since $\cos(qa) - \frac{mv_0}{\hbar^2 q}\sin(qa)$ is not bounded by ± 1, some energy values do not correspond to any possible k value. So there are energy gaps in the expression. A plot as in Baym Fig 4-15 is shown in fig. 14.1 for the case where $mv_0 a/\hbar^2 = 3$. The energies are shown in fig. 14.2. The gnuplot code to produce these plots is

```
set term post eps
set output "kpf1.eps"
set noparametric
```
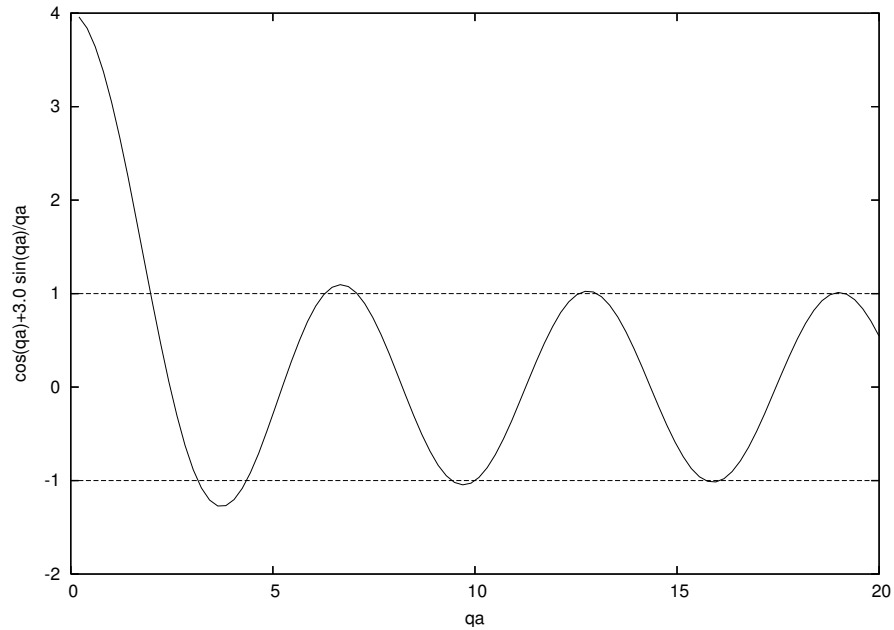
Figure 14.1: A plot of $\cos(qa) + \frac{mv_0a}{\hbar^2}\frac{\sin qa}{qa}$ with $mv_0a/\hbar^2 = 3$ showing the lines ± 1 which bound $cos(ka)$.

```
set nokey
v0=3.0
f(q)=cos(q)+v0*sin(q)/q
set xrange [0:20]
set xlabel "qa"
set ylabel "cos(qa)+3.0 sin(qa)/qa"
plot f(x),1 w l 2,-1 w l 2
set output "kpf2.eps"
set xlabel "ka"
set ylabel "q^2a^2/2"
set xrange [0:pi]
set samples 10000
set trange [0:10]
set parametric
plot acos(f(t)),.5*t**2
```
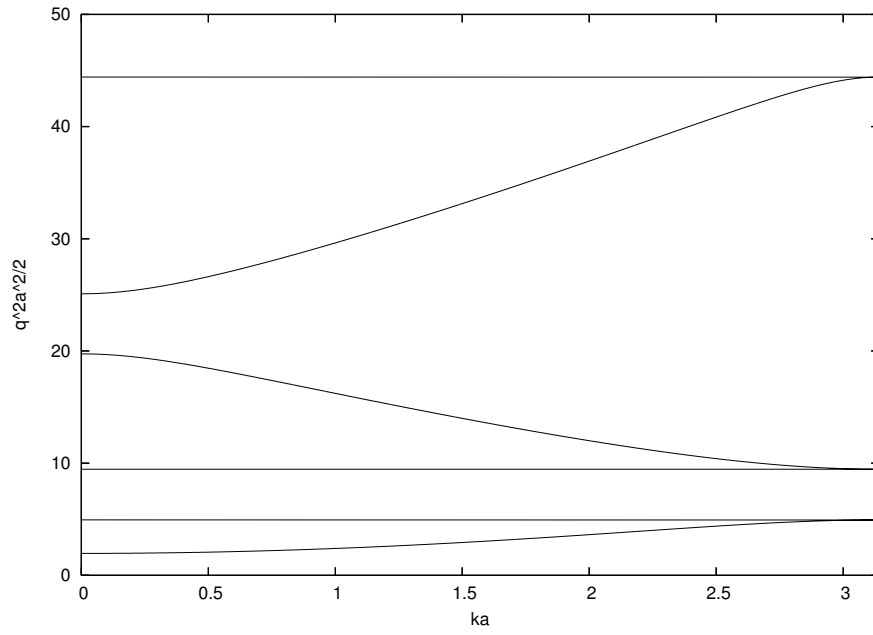
Figure 14.2: A plot of $q^2/2$ versus $ka = \cos^{-1}\left[\cos(qa) + \frac{mv_0 a}{\hbar^2}\frac{\sin qa}{qa}\right]$ with $mv_0 a/\hbar^2 = 3$. The horizontal lines are artefacts of the plotting program.

Chapter 15

Harmonic Oscillator and angular momentum

15.1 Introduction

Any potential that has a strongly attractive minimum can be expanded about the minimum giving an approximate harmonic oscillator potential. In one dimension this is

$$V(x) = V(x_0) + V'(x_0)(x - x_0) + \frac{1}{2}V''(x_0)(x - x_0)^2 + \dots \tag{15.1}$$

and with x_0 a minimum, $V'(x_0) = 0$. Picking the origin at $x = x_0$, and dropping the constant potential gives the usual harmonic oscillator Hamiltonian

$$\frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \,. \tag{15.2}$$

15.2 Natural units

It is convenient to work in natural units where the numerical values of $\hbar$, m and ω are 1. This corresponds to taking m to be the mass unit, ω^{-1} to be the time unit and $\sqrt{\hbar/m\omega}$ to be the length unit.

15.3 Operator solution

The "second quantization" method of solving this starts by noting that if p and x commuted, we could factor the Hamiltonian into factors $x \pm ip$. They don't commute; we can still write

$$\frac{x - ip}{\sqrt{2}}\frac{x + ip}{\sqrt{2}} = \frac{1}{2}\left(p^2 + x^2 + i[x, p]\right) = \frac{p^2}{2} + \frac{x^2}{2} - \frac{1}{2} \tag{15.3}$$

so that

$$H = \frac{p^2}{2} + \frac{x^2}{2} = \frac{x - ip}{\sqrt{2}} \frac{x + ip}{\sqrt{2}} + \frac{1}{2} \,. \tag{15.4}$$

We define the operator

$$a = \frac{x + ip}{\sqrt{2}} \tag{15.5}$$

with a usually taken to be unitless. Taking the adjoint, we have

$$a^\dagger = \frac{x - ip}{\sqrt{2}} \,. \tag{15.6}$$

a and $a^\dagger$ are called lowering and raising or destruction and creation operators. We can calculate the commutation relation $[a, a^\dagger]$ from the commutation relations for x and p

$$[a, a^\dagger] = \left[\frac{x + ip}{\sqrt{2}}, \frac{x - ip}{\sqrt{2}} \right] = -\frac{i}{2}[x, p] + \frac{i}{2}[p, x] = -i[x, p] = 1 \,. \tag{15.7}$$

The Hamiltonian is

$$H = a^\dagger a + \frac{1}{2} \,. \tag{15.8}$$

The eigenvalues of H are just $\frac{1}{2}$ added to the eigenvalues of $a^\dagger a$. $a^\dagger a$ is Hermitian as expected $(a^\dagger a)^\dagger = a^\dagger a$. We call the operator $a^\dagger a \equiv N$ the number operator. We now call the eigenstates of N, $|n\rangle$, so that $N|n\rangle = n|n\rangle$. Since N is Hermitian, the eigenvalues are real. Writing $|\phi\rangle = a|n\rangle$, we get the requirement that

$$n = \langle n|N|n\rangle = \langle n|a^\dagger a|n\rangle = \langle \phi|\phi\rangle \geq 0 \tag{15.9}$$

where n can only equal zero for $a|0\rangle = 0$.

Commuting a and $a^\dagger$ with N, we have

$$
\begin{aligned}
{}[N, a] &= a^\dagger a a - a a^\dagger a = (a^\dagger a - a a^\dagger) a = -a \,. \\
[N, a^\dagger] &= a^\dagger a a^\dagger - a^\dagger a^\dagger a = a^\dagger (a a^\dagger - a^\dagger a) = a^\dagger \,.
\end{aligned} \tag{15.10}
$$

These are both of the form $[N, O] = \alpha O$, where α is a constant; here it is ± 1.

In general if we have an operator N with eigenstates $|n\rangle$, and we have an operator with the commutation relation $[N, O] = \alpha O$, then

$$[N, O]|n\rangle = NO|n\rangle - ON|n\rangle = NO|n\rangle - nO|n\rangle = \alpha O|n\rangle \,. \tag{15.11}$$

Rearranging this is

$$N(O|n\rangle) = (n + \alpha)(O|n\rangle) \,. \tag{15.12}$$

which shows that $O|n\rangle$ is an eigenstate of N with eigenvalue $n + \alpha$, so that $O|n\rangle \propto |n + \alpha\rangle$. It is also possible that O anihilates $|n\rangle$ giving zero. If α is positive, we call O a raising operator and if α is negative we call O a lowering operator.

Here we see that a is a lowering operator and $a^\dagger$ is a raising operator. Taking the eigenstates of N to be normalized $\langle n|n \rangle = 1$, we can calculate the normalization of $a|n\rangle$ and $a^\dagger|n\rangle$,

$$
\begin{aligned}
\langle n|a^\dagger a|n \rangle &= \langle n|N|n \rangle = n \geq 0 \\
\langle n|aa^\dagger|n \rangle &= \langle n|[a, a^\dagger] + N|n \rangle = n + 1
\end{aligned}
\tag{15.13}
$$

so that $a|n\rangle = \sqrt{n}|n-1\rangle$ and $a^\dagger|n\rangle = \sqrt{n+1}|n\rangle$ where we have chosen the phases of the $|n\rangle$ states so that the normalization constant is real. Just as in the power series solution of the differential equation, where only particular eigenvalues lead to normalized wave functions, here we see that if we continue operating with a, n will continue getting smaller, but we get a contradiction if $n < 0$, that the normalization of a state is negative. To get a consistent solution, we must take n to be nonnegative integers, $0, 1, 2, \dots$. We then have $a|0\rangle = 0$ as the smallest possible n state. Since the a and $a^\dagger$ connect all of the integer values of n, we can cosntruct any state from any other by acting on it with raising or lowering operators.

The energy eigenstates are also $|n\rangle$ and the Hamiltonian has eigenvalues of $n + 1/2$. (Recall that the energy unit here is $\hbar\omega$, so converting from natural units gives the familiar $E_n = (n + 1/2)\hbar\omega$.)

We can construct the wave functions by writing

$$
\begin{aligned}
a &= \frac{x + ip}{\sqrt{2}} = \frac{1}{\sqrt{2}} \int dx |x\rangle \left(x + \frac{\partial}{\partial x} \right) \langle x| \\
a^\dagger &= \frac{x - ip}{\sqrt{2}} = \frac{1}{\sqrt{2}} \int dx |x\rangle \left(x - \frac{\partial}{\partial x} \right) \langle x|
\end{aligned}
\tag{15.14}
$$

The ground state is $|0\rangle$, which satisfies $a|0\rangle = 0$. We write

$$
0 = \langle x|a|0\rangle = \langle x|\frac{1}{\sqrt{2}} \int dx' |x'\rangle \left(x' + \frac{\partial}{\partial x'} \right) \langle x'|0\rangle = \frac{1}{\sqrt{2}} \left(x + \frac{\partial}{\partial x} \right) \langle x|0\rangle
\tag{15.15}
$$

This becomes

$$
-x dx = \frac{d\langle x|0\rangle}{\langle x|0\rangle}
\tag{15.16}
$$

and integrating

$$
-\frac{x^2}{2} + \text{Constant} = \ln\left(\langle x|0\rangle \right)
\tag{15.17}
$$

or

$$
\langle x|0\rangle = N e^{-\frac{x^2}{2}} = \frac{1}{\pi^{1/4}} e^{-\frac{x^2}{2}},
\tag{15.18}
$$

where we normalize the state in the last expression. The excited state wave functions can be calculated by operating with $a^\dagger$, which is

$$
\sqrt{n+1}\langle x|n+1\rangle = \frac{1}{\sqrt{2}} \left(x - \frac{d}{dx} \right) \langle x|n\rangle .
\tag{15.19}
$$

139

Since $\langle x|0\rangle$ is a polynomial (in this case a constant) times a gaussian, multiplying by x and taking the derivative means that the excited states continue to be polynomials times gaussians. You can derive the recursion relation for the excited states which you can then relate to the Hermite polynomials. The general form for the wave functions is

$$\langle x|n\rangle = \pi^{-1/4}\frac{1}{\sqrt{2^n n!}}H_n(x)e^{-\frac{x^2}{2}}, \tag{15.20}$$

where the Hermite polynomials $H_n(x)$ satisfy $H_0(x) = 1$, $H_1(x) = 2x$, $H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$, or

$$H_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2}. \tag{15.21}$$

Since 1 dimensional wave functions have units of length to the negative half, and x has units of length, we can convert back to arbitrary units by multiplying x by $\sqrt{\frac{m\omega}{\hbar}}$, and with an overall factor of $(\frac{m\omega}{\hbar})^{1/4}$. For example, the ground state wave function is

$$\psi_0(x) = \langle x|0\rangle = \left(\frac{m\omega}{\hbar}\right)^{1/4}\left(\frac{1}{\pi}\right)^{1/4}\exp\left(-\frac{m\omega}{\hbar}\frac{x^2}{2}\right). \tag{15.22}$$

15.4   Rotations and angular momentum

Since rotations are continuous symmetries, and the limit of zero rotation is the identity operator, rotations must be represented in Hilbert space as a linear unitary transformation. We write the operator that corresponds to a rotation of the coordinate system around $\hat{\boldsymbol{\alpha}}$ by an infinitesimal angle α to be $1 + \frac{i}{\hbar}\boldsymbol{L}\cdot\boldsymbol{\alpha}$, with $\boldsymbol{L}$ Hermitian; $\boldsymbol{L}$ will turn out to be the angular momentum.

The properties of $\boldsymbol{L}$ follow from the known properties of coordinate rotations. Rotations of the coordinates by an angle θ around x, y, and z are given by the matrices

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}$$

$$R_y(\theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}$$

$$R_z(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{15.23}$$

The unit vectors of the transformed system $\hat{\boldsymbol{x}}'$, $\hat{\boldsymbol{y}}'$, $\hat{\boldsymbol{z}}'$ are given in terms of the original system by

$$\begin{pmatrix} \hat{\boldsymbol{x}}' \\ \hat{\boldsymbol{y}}' \\ \hat{\boldsymbol{z}}' \end{pmatrix} = R_j(\theta)\begin{pmatrix} \hat{\boldsymbol{x}} \\ \hat{\boldsymbol{y}} \\ \hat{\boldsymbol{z}} \end{pmatrix} \tag{15.24}$$

140

where j is x, y, or z, and R_j is one of the matrices above.

For infinitesimal rotations $\theta \to 0$ and keeping just the linear terms these become

$$
R_x(\theta) \;\to\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \theta \\ 0 & -\theta & 1 \end{pmatrix}
$$

$$
R_y(\theta) \;\to\; \begin{pmatrix} 1 & 0 & -\theta \\ 0 & 1 & 0 \\ \theta & 0 & 1 \end{pmatrix}
$$

$$
R_z(\theta) \;\to\; \begin{pmatrix} 1 & \theta & 0 \\ -\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{15.25}
$$

We can find the commutation relations for the angular momentum by looking at the infinitesimal rotations given by first rotating around x by ϕ then y by θ then x by $-\phi$ and y by $-\theta$[1]. Multiplying the infinitesimal rotation matrices, we see that such a coordinate transformation is

$$
R_y(-\theta)R_x(-\phi)R_y(\theta)R_x(\phi) \;=\; \begin{pmatrix} 1 & 0 & \theta \\ 0 & 1 & 0 \\ -\theta & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\phi \\ 0 & \phi & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -\theta \\ 0 & 1 & 0 \\ \theta & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \phi \\ 0 & -\phi & 1 \end{pmatrix}
$$

$$
\;=\; \begin{pmatrix} 1 & \phi\theta & 0 \\ -\phi\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = R_z(\phi\theta) . \tag{15.26}
$$

The quantum transformations must have the same property. That is if we rotate by infinitesmals in the same order, we must get the same result. This means that keeping only terms up to ϕ, θ and $\phi\theta$,

$$
\left(1 - \frac{i}{\hbar}\theta L_y\right)\left(1 - \frac{i}{\hbar}\phi L_x\right)\left(1 + \frac{i}{\hbar}\theta L_y\right)\left(1 + \frac{i}{\hbar}\phi L_x\right) = 1 + \frac{\phi\theta}{\hbar^2}\left(L_x L_y - L_y L_x\right) = 1 + \frac{i}{\hbar}\phi\theta L_z \tag{15.27}
$$

or

$$
[L_x, L_y] = i\hbar L_z . \tag{15.28}
$$

Repeating for the other directions, realizing that we can relabel the coordinate axes cyclically, or looking at the more general derivation in chapter 4, we get the set of commutation relations for the angular momentum

$$
[L_\alpha, L_\beta] = i\hbar\epsilon_{\alpha\beta\gamma}L_\gamma . \tag{15.29}
$$

---

[1]It may seem that we should also keep θ^2 and ϕ^2 terms since they are the same numerical order as the $\theta\phi$ terms. In fact, if we do keep them they all cancel. However, another way to see what we are doing here is to equate the derivative $\partial_\phi\partial_\theta$ evaluated at $\phi = \theta = 0$. To get this correct, we only need to keep the constant, ϕ, θ, and $\theta\phi$ terms in the expansion of each piece. Since our separate rotations do not contain $\theta\phi$ terms, we only need to keep the linear terms in their expansions. If you are troubled by this, you can keep all the θ^2 and ϕ^2 terms in both the expansion of the coordinate transformations and in the expansion of the quantum unitary transformations and they will cancel and drop out as noted above.

15.5 Scalar, vector and cartesian tensor operators

The names scalar, vector and tensor describe objects with particular transformation properties under rotations (and often parity). The freshman physics definitions that a scalar has just a magnitude, and a vector has a magnitude and direction are fine for freshman physics, but do not give a clear definition. Is the x-component of an electric field a scalar? (No!).

A scalar is an object that is invariant under rotations. A vector transforms like the coordinate axes, therefore, like a position in coordinate space. It would have 1 index to give its components. An n-rank cartesian tensor has n indices and each index transforms like a vector. A vector is a first rank tensor. A scalar is a zero rank tensor. The moment of inertia tensor is an example of a second rank tensor. It is symmetric under interchange of its indices, so it is also a symmetric second rank tensor.

Objects that change sign under parity but transform like a scalar, vector or tensor under rotations are called pseudoscalars, pseudovectors, and pseudotensors. As we will see, angular momentum is actually an antisymmetric second ranked tensor which transforms like a pseudovector; for example, orbital angular momentum is the cross product $\boldsymbol{r} \times \boldsymbol{p}$, and, unlike a true vector, does not change sign under parity. Here we do not need to worry about this difference since we will only be rotating.

We saw in chapter 4 that under a coordinate transformation, we can transform the state vectors with a unitary transformation $|\psi'\rangle = U|\psi\rangle$. The matrix elements of an operator O in this new frame will therefore be equal to the matrix elements of the operator $U^\dagger O U$ in the original frame. That is we can write

$$O' = U^\dagger O U \tag{15.30}$$

Rotating the coordinates around z by an infinitesimal angle ϕ, a scalar S will be unchanged $S' = S$, while a vector $\boldsymbol{V}$ will have its new components given by

$$
\begin{aligned}
V_x' &= V_x + \phi V_y \\
V_y' &= V_y - \phi V_x \\
V_z' &= V_z \,.
\end{aligned}
\tag{15.31}
$$

The quantum mechanical unitary transformation, keeping linear order in ϕ must agree with these. Therefore

$$S' = \left(1 - \frac{i}{\hbar}\phi L_z\right) S \left(1 + \frac{i}{\hbar}\phi L_z\right) = S - \frac{i}{\hbar}\phi[L_z, S] \tag{15.32}$$

and $[L_z, S] = 0$. Similarly,

$$
\begin{aligned}
V_x' &= \left(1 - \frac{i}{\hbar}\phi L_z\right) V_x \left(1 + \frac{i}{\hbar}\phi L_z\right) = V_x - \frac{i}{\hbar}\phi[L_z, V_x] \\
V_y' &= \left(1 - \frac{i}{\hbar}\phi L_z\right) V_y \left(1 + \frac{i}{\hbar}\phi L_z\right) = V_y - \frac{i}{\hbar}\phi[L_z, V_y] \\
V_z' &= \left(1 - \frac{i}{\hbar}\phi L_z\right) V_z \left(1 + \frac{i}{\hbar}\phi L_z\right) = V_z - \frac{i}{\hbar}\phi[L_z, V_z]
\end{aligned}
\tag{15.33}
$$

tells us that

$$\begin{aligned}
[L_z, V_x] &= i\hbar V_y \\
[L_z, V_y] &= -i\hbar V_x \\
[L_z, V_z] &= 0.
\end{aligned}$$ (15.34)

Again, realizing that there is nothing special about rotations about z, we can make cyclic replacements of the coordinates to get

$$[L_\alpha, S] = 0.$$ (15.35)

All scalars commute with $\boldsymbol{L}$. For the vector case,

$$[L_\alpha, V_\beta] = i\hbar\epsilon_{\alpha,\beta,\gamma}V_\gamma.$$ (15.36)

All operators that transform like vectors must have this commutation relation. We see from Eq. 15.29 that the angular momentum itself transforms like a vector under rotations.

We could continue on and find the commutation relations for other tensors, but we will not do that here.

15.6   Orbital angular momentum

In quantum mechanics we know that we have both orbital and spin angular momentum. Here we have defined the angular momentum to be the generator of rotations. It is reassuring to see that for the orbital case it reverts to the usual $\boldsymbol{r} \times \boldsymbol{p}$. Rotating the coordinates around z by an angle ϕ gives the rotation operator

$$\begin{aligned}
U &= \int dx\,dy\,dz\,|x\cos\phi + y\sin\phi, y\cos\phi - x\sin\phi, z\rangle\langle x, y, z| \\
&= \int dx\,dy\,dz\,|x, y, z\rangle\langle x\cos\phi - y\sin\phi, y\cos\phi + x\sin\phi, z|.
\end{aligned}$$ (15.37)

For an infinitesimal rotation we have

$$\begin{aligned}
U &= \int dx\,dy\,dz\,|x, y, z\rangle\langle x - y\phi, y + x\phi, z| \\
&= \int dx\,dy\,dz\,|x, y, z\rangle\left[1 - y\phi\frac{\partial}{\partial x} + x\phi\frac{\partial}{\partial y}\right]\langle x, y, z| \\
&= 1 + \frac{i}{\hbar}\phi\left[xp_y - yp_x\right].
\end{aligned}$$ (15.38)

Since this must be $1 + \frac{i}{\hbar}\phi L_z$, $L_z = xp_y - yp_x$ for the orbital case, and $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ for a general orbital rotation.

If you are more comfortable using wave functions, the equivalent is to rotate the wave function by an infinitesimal angle ϕ and Taylor expand:

$$
\begin{aligned}
\psi(x - \phi y, y + \phi x, z) &= \psi(x, y, z) + -\phi y \frac{\partial}{\partial x} \psi(x, y, z) + \phi x \frac{\partial}{\partial y} \psi(x, y, z) + ... \\
&= \left[1 + \frac{i}{\hbar} \phi \left(x \frac{\hbar}{i} \frac{\partial}{\partial y} - y \frac{\hbar}{i} \frac{\partial}{\partial x} \right) \right] \psi(x, y, z) \quad (15.39)
\end{aligned}
$$

At this point, since you know the commutation relations for r and p, you can verify by direct calculation that r, p, $L = r \times p$ all satisfy the commutation relations for a vector Eq. 15.36, as we derived generally.

15.7   Angular momentum representations

We can work out the form of the angular momentum operators from their commutation relations. That is we want to work out the matrix elements of L_x, L_y, and L_z. The usual way is to begin by selecting the eigenbasis of one of the operators. The traditional choice is to take the basis where L_z is diagonal.

We have already shown that every scalar commutes with L. Since L transforms under rotations like a vector, we can form the scalar $L^2 = L_x^2 + L_y^2 + L_z^2$, which must commute with L. Since $[L_z, L^2] = 0$, we can simultaneously diagonalize L^2 and L_z. We therefore take our basis to be the eigenvectors of L^2 and L_z.

In the end the result will be the usual $|\ell m\rangle$ states where the eigenvalue of L^2 is $\ell(\ell+1)\hbar^2$, and the eigenvalue of L_z is $m\hbar$. However, to not bias ourselves we will write them as $|\lambda, \lambda_z\rangle$ where λ is the eigenvalue of L^2 and λ_z the eigenvalue of L_z

$$
\begin{aligned}
L^2 |\lambda, \lambda_z\rangle &= \lambda |\lambda, \lambda_z\rangle \\
L_z |\lambda, \lambda_z\rangle &= \lambda_z |\lambda, \lambda_z\rangle . \quad (15.40)
\end{aligned}
$$

Looking at the commutation relations with L_z, we know

$$
\begin{aligned}
[L_z, L_x] &= i\hbar L_y \\
[L_z, L_y] &= -i\hbar L_x \quad (15.41)
\end{aligned}
$$

These are not raising or lowering operator commutation relations, but it should be clear from the form, that if we take a linear combination of L_x and L_y we should be able to get this in the raising/lowering form. Let's look at $L_x + \beta L_y$,

$$
[L_z, L_x + \beta L_y] = i\hbar(L_y - \beta L_x) \stackrel{?}{=} \alpha(L_x + \beta L_y) . \quad (15.42)
$$

Equating coefficients, we have

$$
\begin{aligned}
\alpha &= -i\hbar\beta \\
\alpha\beta &= i\hbar \quad (15.43)
\end{aligned}
$$

with solution $\beta = \pm i$, $\alpha = \pm \hbar$. We therefore define the operators

$$
\begin{aligned}
L_+ &= L_x + iL_y \\
L_- &= L_x - iL_y
\end{aligned}
\tag{15.44}
$$

with the raising and lowering commutation relations

$$
\begin{aligned}
[L_z, L_+] &= \hbar L_+ \\
[L_z, L_-] &= -\hbar L_- \,.
\end{aligned}
\tag{15.45}
$$

L_+ operating on an eigenstate of L_z adds $\hbar$ to the eigenvalue, while L_- subtracts $\hbar$ from the eigenvalue. Just as in the harmonic oscillator, we have the conditions

$$
\langle \lambda, \lambda_z | L_\pm L_\mp | \lambda, \lambda_z \rangle \geq 0
\tag{15.46}
$$

and we can write[2]

$$
L_\pm L_\mp = (L_x \pm iL_y)(L_x \mp iL_y) = L_x^2 + L_y^2 \mp i[L_x, L_y] = L^2 - L_z^2 \pm \hbar L_z \,.
\tag{15.47}
$$

Our condition becomes

$$
\langle \lambda, \lambda_z | L_\pm L_\mp | \lambda, \lambda_z \rangle = \lambda - \lambda_z^2 \pm \hbar \lambda_z \geq 0
\tag{15.48}
$$

Notice that each time we operate with the lowering operator we will make λ_z smaller by $\hbar$. If it goes to a large negative value, the expression becomes negative. Similarly when we operate with the raising operator, λ_z increases by $\hbar$, and if it becomes large and positive, the expression again becomes negative. As in the harmonic oscillator, the only way to have consistency is to have the expression go to zero for some maximum λ_z and some minimum λ_z. These states must be connected by the raising and lowering operators. Solving $\lambda - \lambda_z^2 \pm \hbar \lambda_z = 0$ we get

$$
\begin{aligned}
\lambda_{z\ \mathrm{max}} &= \hbar \left[-\frac{1}{2} \pm \sqrt{\frac{1}{4} + \lambda} \right] \\
\lambda_{z\ \mathrm{min}} &= \hbar \left[\frac{1}{2} \pm \sqrt{\frac{1}{4} + \lambda} \right]
\end{aligned}
\tag{15.49}
$$

The only sign choice that makes $\lambda_{z\ \mathrm{max}} \geq \lambda_{z\ \mathrm{min}}$ is

$$
\begin{aligned}
\lambda_{z\ \mathrm{max}} &= \hbar \left[-\frac{1}{2} + \sqrt{\frac{1}{4} + \lambda} \right] \\
\lambda_{z\ \mathrm{min}} &= \hbar \left[\frac{1}{2} - \sqrt{\frac{1}{4} + \lambda} \right]
\end{aligned}
\tag{15.50}
$$

---

[2]If we had not realized that L^2 commutes with L_z, we would still have tried to define raising and lowering operators. It is at this point where we have to evaluate $L_x^2 + L_y^2$ that we would need to realize that this is $L^2 - L_z^2$ and that since L^2 commutes with L_z we can simultaneously diagonalize them and evaluate L^2 on the states.

This shows that

$$\lambda_{z\ \text{min}} = -\lambda_{z\ \text{max}}. \tag{15.51}$$

Further since the states must be connected by the raising and lowering operators, the difference must be an integer multiple of $\hbar$. Therefore $\lambda_{z\ \text{max}} = \hbar\ell$ where 2ℓ is a nonnegative integer. Substituting back, we see that $\lambda = \ell(\ell+1)\hbar^2$. We now label our eigenstates by their ℓ, m values as $|\ell m\rangle$

We have $L_\pm|\ell m\rangle \propto |\ell, m\pm 1\rangle$. The proportionality constant is given by the normalization we calculated above

$$\langle \ell m|L_\pm L_\mp|\ell m\rangle = \hbar^2\left[\ell(\ell+1) - m(m\mp 1)\right] \tag{15.52}$$

We choose the usual phase convention that the factor is real and positive so that

$$L_\pm|\ell m\rangle = \sqrt{\ell(\ell+1) - m(m\pm 1)}|\ell, m\pm 1\rangle. \tag{15.53}$$

The allowed ℓ values are $0, \frac{1}{2}, 1, \frac{3}{2}, 2, ...$, and since the m values run from $-\ell$ to ℓ with integer spacing there are $2\ell + 1$ possible m values. Half integer angular momentum drops out as one of the possibilities for a consistent rotation operator.

We can write the matrix elements of the operators L_z, $L_\pm$ as

$$\begin{aligned}
\langle \ell m'|L_z|\ell m\rangle &= \delta_{m',m}m\hbar \\
\langle \ell m'|L_\pm|\ell m\rangle &= \delta_{m',m\pm 1}\hbar\sqrt{\ell(\ell+1) - m(m\pm 1)}
\end{aligned} \tag{15.54}$$

and write out the $(2\ell+1) \times (2\ell+1)$ matrices. The L_x and L_y matrices are just $(L_+ + L_-)/2$ and $i(L_- - L_+)/2$.

15.8 Spherical harmonics and orbital angular momentum

Our usual position space basis is normalized so that

$$\begin{aligned}
1 &= \int d^3r|\boldsymbol{r}\rangle\langle\boldsymbol{r}| \\
\langle\boldsymbol{r}'|\boldsymbol{r}\rangle &= \delta^3(\boldsymbol{r} - \boldsymbol{r}')
\end{aligned} \tag{15.55}$$

For cartesian coordinates this is

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz|xyz\rangle\langle xyz| \tag{15.56}$$

and

$$\langle x'y'z'|xyz\rangle = \delta(x - x')\delta(y - y')\delta(z - z') \tag{15.57}$$

Converting to spherical coordinates these become

$$
1 \;=\; \int_0^\infty r^2 dr \int_{-1}^1 d\cos\theta \int_0^{2\pi} d\phi |r\theta\phi\rangle\langle r\theta\phi| = \int_0^\infty r^2 dr \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi |r\theta\phi\rangle\langle r\theta\phi|
$$

$$\tag{15.58}$$

and

$$
\langle r'\theta'\phi'|r\theta\phi\rangle = \frac{\delta(r-r')\delta(\cos\theta - \cos\theta')\delta(\phi-\phi')}{r^2} = \frac{\delta(r-r')\delta(\theta-\theta')\delta(\phi-\phi')}{r^2 \sin\theta}.
$$

$$\tag{15.59}$$

The angular momentum operator in the r-representation is

$$
\begin{aligned}
\int d^3 r |\boldsymbol{r}\rangle \left(-i\hbar \boldsymbol{r} \times \boldsymbol{\nabla} \right) \langle \boldsymbol{r}| &= -i\hbar \int d^3 r |\boldsymbol{r}\rangle \boldsymbol{r} \times \left[\hat{\boldsymbol{r}} \frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\boldsymbol{\phi}} \frac{1}{r\sin\theta} \frac{\partial}{\partial \phi} \right] \langle \boldsymbol{r}| \\
&= -i\hbar \int d^3 r |\boldsymbol{r}\rangle \left[\hat{\boldsymbol{\phi}} \frac{\partial}{\partial \theta} - \hat{\boldsymbol{\theta}} \frac{1}{\sin\theta} \frac{\partial}{\partial \phi} \right] \langle \boldsymbol{r}| \tag{15.60}
\end{aligned}
$$

and substituting

$$
\begin{aligned}
\hat{\boldsymbol{\phi}} &= \cos\phi \hat{\boldsymbol{y}} - \sin\phi \hat{\boldsymbol{x}} \\
\hat{\boldsymbol{\theta}} &= \cos\theta \cos\phi \hat{\boldsymbol{x}} + \cos\theta \sin\phi \hat{\boldsymbol{y}} - \sin\theta \hat{\boldsymbol{z}} \tag{15.61}
\end{aligned}
$$

gives

$$
\begin{aligned}
L_x &= i\hbar \int d^3 r |\boldsymbol{r}\rangle \left[\sin\phi \frac{\partial}{\partial \theta} + \cot\theta \cos\phi \frac{\partial}{\partial \phi} \right] \langle \boldsymbol{r}| \\
L_y &= -i\hbar \int d^3 r |\boldsymbol{r}\rangle \left[\cos\phi \frac{\partial}{\partial \theta} - \cot\theta \sin\phi \frac{\partial}{\partial \phi} \right] \langle \boldsymbol{r}| \\
L_z &= -i\hbar \int d^3 r |\boldsymbol{r}\rangle \frac{\partial}{\partial \phi} \langle \boldsymbol{r}| \tag{15.62}
\end{aligned}
$$

which agress with Baym Eqs. 6-26, 6-31, and 6-32. We can construct the raising and lowering operators as

$$
L_x \pm iL_y = \hbar \int d^3 r |\boldsymbol{r}\rangle e^{\pm i\phi} \left[\pm \frac{\partial}{\partial \theta} + i\cot\theta \frac{\partial}{\partial \phi} \right] \langle \boldsymbol{r}| \tag{15.63}
$$

Since rotating a function of r, θ, ϕ about the origin does not change r, we should have expected that the angular momentum operators would not contain r. The eigenfunctions of L_z and L^2 will be functions of θ and ϕ times any function of r. That is $|r\ell m\rangle$ will be an eigenstate of L_z and L^2 for any value of r. All of these eigenstates with the same ℓ and m are therefore degenerate eigenstates of L^2 and L_z.

147

The eigenfunctions, as we will see below, will turn out to be

$$\langle r\theta\phi|r'\ell m\rangle = \frac{\delta(r-r')}{r^2}Y_{\ell m}(\theta,\phi)\,, \tag{15.64}$$

where $Y_{\ell m}(\theta,\phi)$ are the spherical harmonics. Since the different r' values are degenerate, the general form for the eigenstates of L^2 and L_z are

$$|n\ell m\rangle = \int_0^\infty r'^2 dr' R_{n\ell m}(r')|r'\ell m\rangle \tag{15.65}$$

so that the general eigenfunctions are

$$\psi_{nlm}(r,\theta,\phi) = \langle r\theta\phi|n\ell m\rangle = R_{n\ell m}(r)Y_{\ell m}(\theta,\phi)\,. \tag{15.66}$$

Since the r part separates and plays no role in the angular momentum algebra, and just clutters the expressions, we can drop it in the calculation below, and only add it at the end when we wish to use these states to calculate eigenstates of rotationally invariant systems.

The eigenfunctions of L_z are given by

$$\langle\theta\phi|L_z|\ell m\rangle = \hbar m\langle\theta\phi|\ell m\rangle \tag{15.67}$$

which becomes

$$\frac{\partial}{\partial\phi}\langle\theta\phi|\ell m\rangle = im\langle\theta\phi|\ell m\rangle\,. \tag{15.68}$$

Integrating

$$im\int_0^\phi d\phi = \int \frac{d\langle\theta\phi|\ell m\rangle}{\langle\theta\phi|\ell m\rangle} = \ln\langle\theta\phi|\ell m\rangle - \ln\langle\theta 0|\ell m\rangle \tag{15.69}$$

or

$$\langle\theta\phi|\ell m\rangle = \langle\theta 0|\ell m\rangle e^{im\phi}\,. \tag{15.70}$$

Notice that for the wave function to be single valued, m must be an integer. Notice also that the angular momentum can be defined around any origin, so the value of a nonsingle valued function would have to depend on the path a particle takes around the origin. It is not possible to do this with an arbitrary (and possibly changing) origin. Therefore to keep rotational and translational invariance we require *orbital* angular momenta to have integer ℓ. Notice that it is possible to have a multiple valued function around a particle position, so half integer intrinsic angular momentum (i.e. spin) for a particle does not lead to contradictions.

We can construct the θ dependence exactly as we constructed the ground state of the harmonic oscillator. The raising operator on the state $|\ell\ell\rangle$ gives zero. Therefore

$$\left(\frac{\partial}{\partial\theta} - \ell\cot\theta\right)\langle\theta 0|\ell\ell\rangle = 0 \tag{15.71}$$

Integrating, we get

$$\ln\langle\theta 0|\ell\ell\rangle = \ell\int d\theta\cot\theta = \ell\ln\sin\theta + \text{Constant} \tag{15.72}$$

148

or

$$\langle \theta 0|\ell\ell\rangle \propto \sin^\ell \theta \tag{15.73}$$

We get

$$\langle \theta\phi|\ell\ell\rangle = N e^{i\ell\phi}\sin^\ell \theta \tag{15.74}$$

The normalization integral is

$$1 = |N|^2 \int_0^{2\pi} d\phi \int_{-1}^1 d\cos\theta \sin^{2\ell}\theta \tag{15.75}$$

Normalizing (see the appendix of this chapter) with a phase factor of $(-1)^\ell$ we get

$$\langle \theta\phi|\ell\ell\rangle = Y_{\ell\ell}(\theta,\phi) \tag{15.76}$$

the spherical harmonics. To get the other m values we simply operate with L_-, and divide by $\sqrt{\ell(\ell+1)-m(m-1)}$ factor. This produces all of the spherical harmonics

$$\langle \theta\phi|\ell m\rangle = Y_{\ell m}(\theta,\phi). \tag{15.77}$$

15.9   Orbital L^2 representations

We can write, for the orbital angular momentum,

$$L^2 \;=\; (\boldsymbol{r}_{\mathrm{op}} \times \boldsymbol{p}_{\mathrm{op}}) \cdot (\boldsymbol{r}_{\mathrm{op}} \times \boldsymbol{p}_{\mathrm{op}}). \tag{15.78}$$

Using the summation convention, this becomes

$$
\begin{aligned}
L^2 \;=\;& x_{\alpha\,\mathrm{op}}p_{\beta\,\mathrm{op}}x_{\mu\,\mathrm{op}}p_{\nu\,\mathrm{op}}\epsilon_{\alpha\beta\gamma}\epsilon_{\mu\nu\gamma} \\
=\;& x_{\alpha\,\mathrm{op}}p_{\beta\,\mathrm{op}}x_{\alpha\,\mathrm{op}}p_{\beta\,\mathrm{op}} - x_{\alpha\,\mathrm{op}}p_{\beta\,\mathrm{op}}x_{\beta\,\mathrm{op}}p_{\alpha\,\mathrm{op}} \\
=\;& x_{\alpha\,\mathrm{op}}x_{\alpha\,\mathrm{op}}p_{\beta\,\mathrm{op}}p_{\beta\,\mathrm{op}} - x_{\alpha\,\mathrm{op}}\left[x_{\alpha\,\mathrm{op}},p_{\beta\,\mathrm{op}}\right]p_{\beta\,\mathrm{op}} \\
& -x_{\beta\,\mathrm{op}}p_{\beta\,\mathrm{op}}x_{\alpha\,\mathrm{op}}p_{\alpha\,\mathrm{op}} + x_{\beta\,\mathrm{op}}\left[p_{\beta\,\mathrm{op}},x_{\alpha\,\mathrm{op}}\right]p_{\alpha\,\mathrm{op}} + x_{\alpha\,\mathrm{op}}\left[x_{\beta\,\mathrm{op}},p_{\beta\,\mathrm{op}}\right]p_{\alpha\mathrm{op}} \\
=\;& r_{\mathrm{op}}^2 p_{\mathrm{op}}^2 - i\hbar\boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}} - \boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}}\boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}} - i\hbar\boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}} + 3i\hbar\boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}} \\
=\;& r_{\mathrm{op}}^2 p_{\mathrm{op}}^2 + i\hbar\boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}} - \boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}}\boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}}.
\end{aligned} \tag{15.79}
$$

Rearranging we can write this as

$$p_{\mathrm{op}}^2 \;=\; \;=\; r_{\mathrm{op}}^{-2}\left(\boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}}\boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}} - i\hbar\boldsymbol{r}_{\mathrm{op}}\cdot\boldsymbol{p}_{\mathrm{op}} + L^2\right) \tag{15.80}$$

which shows that L^2/r^2 gives the angular part of the kinetic energy just as in classical mechanics.

An alternative way to get to the same point is to use the results of Eq. 15.62 to write L^2 and verify that these agree with the angular parts of $-\hbar^2 r^2\nabla^2$ in spherical coordinates.

In the $r\ell m$ representation, our result becomes

$$
\begin{aligned}
p_{\text{op}}^2 &= \hbar^2 \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \int_0^{\infty} dr r^2 |r\ell m\rangle \left(-\frac{1}{r}\frac{\partial}{\partial r} r \frac{\partial}{\partial r} - \frac{1}{r}\frac{\partial}{\partial r} + \frac{\ell(\ell+1)}{r^2} \right) \langle r\ell m| \\
&= \hbar^2 \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \int_0^{\infty} dr r^2 |r\ell m\rangle \left(-\frac{1}{r}\frac{\partial^2}{\partial r^2} r + \frac{\ell(\ell+1)}{r^2} \right) \langle r\ell m| \qquad (15.81)
\end{aligned}
$$

15.10 Alternative techniques

Since the spherical harmonics are the eigenfunctions of L_z and L^2, the more usual method is to find the eigenfunctions in a more brute force method. Since L_z and L^2 commute, you can first find the eigenfunctions of L_z. As above, the these are $e^{im\phi}$ multiplying any function of θ. Substituting this into the differential equation for L^2 is equivalent to writing L^2 in this basis. L^2 must be block diagonal in this basis with each block corresponding to a different m value. The result for a given m is the corresponding associated Legendre equation. The normalizable solutions then give the spherical harmonics up to an arbitrarily chosen phase.

If you want to be able to use the raising and lowering operators with Eq. 15.53 *you must make sure that the phase convention you use is consistent with that equation.* For example, our usual phase convention makes

$$
\langle \theta\phi|\ell 0\rangle = Y_{\ell 0}(\theta, \phi) = \sqrt{\frac{2\ell+1}{4\pi}} P_\ell(\cos\theta) \qquad (15.82)
$$

where $P_\ell(x)$ is a Legendre polynomial with phase convention $P_\ell(1) = 1$. From Eq. 15.63 we see that the differential expression for L_- is the negative complex conjugate of the differential expression for L_+, so that

$$
\langle \theta\phi|(L_+)^m|\ell 0\rangle = (\langle \theta\phi|(-L_-)^m|\ell 0\rangle)^* \qquad (15.83)
$$

which shows that by choosing the factors in Eq. 15.53 to be real and positive, we must choose

$$
Y_{\ell m}(\theta, \phi) = (-1)^m Y_{\ell,-m}^*(\theta, \phi). \qquad (15.84)
$$

Often in the mathematics literature the choice is $Y_{\ell m}(\theta, \phi) = Y_{\ell,-m}^*(\theta, \phi)$ which means that the sign of one of the sqare root factors in Eq. 15.53 would need to be chosen negative.

15.11 Harmonic Polynomials

You probably learned in electromagnetism that $r^\ell Y_{\ell m}(\theta, \phi)$ are solutions to Laplace's equation that are well behaved at the origin. These are called harmonic polynomials because

they are polynomials that solve Laplace's equation. We can check this by looking at our analysis above. There we found

$$
\begin{aligned}
r^\ell Y_{\ell\ell}(\theta,\phi) &= N r^\ell e^{i\ell\phi} \sin^\ell\theta = N(re^{i\phi}\sin\theta)^\ell \\
&= N(r\cos\phi\sin\theta + ir\sin\phi\sin\theta)^\ell = N(x+iy)^\ell\,.
\end{aligned}
\tag{15.85}
$$

This shows that $r^\ell Y_{\ell\ell}(\theta,\phi)$ is a polynomial in x, y, and z. We can write the raising and lowering operators as

$$
L_\pm = \pm i\hbar\left[z(\partial_x \pm i\partial_y) - (x\pm iy)\partial_z\right]
\tag{15.86}
$$

which shows that operating on a polynomial $r^\ell Y_{\ell m}$, gives another polynomial. Realizing that the angular momentum does not change r, the r^ℓ will not affect the results.

Often using the harmonic polynomials makes the calculation easier to do and to think about. The spherical harmonics multiplied by r^ℓ for $\ell = 0, 1, 2$ are

$$
\begin{aligned}
Y_{00} &= \frac{1}{\sqrt{4\pi}} \\
rY_{10} &= \sqrt{\frac{3}{4\pi}}\,r\cos\theta = \sqrt{\frac{3}{4\pi}}\,z \\
rY_{1,\pm 1} &= \mp\sqrt{\frac{3}{8\pi}}\,re^{\pm i\phi}\sin\theta = \mp\sqrt{\frac{3}{8\pi}}\,(x\pm iy) \\
r^2 Y_{20} &= \sqrt{\frac{5}{4\pi}}\,r^2\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right) = \sqrt{\frac{5}{4\pi}}\frac{1}{2}\left(2z^2 - x^2 - y^2\right) \\
r^2 Y_{2,\pm 1} &= \mp\sqrt{\frac{15}{8\pi}}\,r^2\sin\theta\cos\theta e^{\pm i\phi} = \mp\sqrt{\frac{15}{8\pi}}\left[z(x\pm iy)\right] \\
r^2 Y_{2,\pm 2} &= \sqrt{\frac{15}{32\pi}}\,r^2\sin^2\theta e^{\pm i2\phi} = \sqrt{\frac{15}{32\pi}}\,(x\pm iy)^2\,.
\end{aligned}
\tag{15.87}
$$

Often in electronic structure and especially in molecular structure where the electronic angular momentum is not a good quantum number (the whole system of electrons and nuclei is rotationally invariant, but the electron Hamiltonian is not separately rotationally invariant), the normalized real and imaginary parts of the spherical harmonics are used instead. For $\ell = 0$ we have the same constant for the s state. The $\ell = 1$ states are called

the p_x, p_y, and p_z states while the $\ell = 2$ states are d_z, d_{xy} d_{xz}, d_{yz}, and $d_{x^2-y^2}$ with

$$
\begin{aligned}
rp_x &= \sqrt{\frac{3}{4\pi}}x \\
rp_y &= \sqrt{\frac{3}{4\pi}}y \\
rp_z &= \sqrt{\frac{3}{4\pi}}z \\
r^2 d_z &= \sqrt{\frac{5}{4\pi}}\frac{1}{2}\left(2z^2 - x^2 - y^2\right) \\
r^2 d_{x^2-y^2} &= \sqrt{\frac{15}{16\pi}}(x^2 - y^2) \\
r^2 d_{xy} &= \sqrt{\frac{15}{4\pi}}xy \\
r^2 d_{xz} &= \sqrt{\frac{15}{4\pi}}xz \\
r^2 d_{yz} &= \sqrt{\frac{15}{4\pi}}yz
\end{aligned}
\tag{15.88}
$$

Notice that the $m = 0$ spherical harmonics, which are real, become the z terms, while the others are just $\sqrt{2}$ multiplied by the real or imaginary parts of the $Y_{\ell m}$ for positive m. The phase is usually chosen arbitrarily, so be careful.

Knowing the structure of the $Y_{\ell m}$ and their real and imaginary parts, especially for $\ell \leq 2$, can make calculating matrix elements and understanding when matrix elements are zero much easier.

15.12 Spin 1/2

Let's look at the angular momentum for the case where $\ell = 1/2$. m has only two possible values $\pm 1/2$. Writing out the 2×2 matrices from Eq. 15.54 and calling them S_z, $S_\pm$, we

have

$$S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$S_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$S_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

$$S_x = \frac{1}{2}(S_+ + S_-) = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$S_y = \frac{i}{2}(S_- - S_+) = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{15.89}$$

Writing

$$\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\sigma} = \frac{\hbar}{2}(\hat{\boldsymbol{x}}\sigma_x + \hat{\boldsymbol{y}}\sigma_y + \hat{\boldsymbol{z}}\sigma_z) \tag{15.90}$$

we get the Pauli spin matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{15.91}$$

15.13   Appendix

Spherical harmonics normalization

The normalization integral for the spherical harmonics that we need is

$$1 = |N|^2 \int_0^{2\pi} d\phi \int_{-1}^{1} d\cos\theta \sin^{2\ell}\theta. \tag{15.92}$$

One way to do this is to convert it into gaussian integrals by multiplying and dividing by $\int_0^\infty dr\, r^{2(\ell+1)} e^{-\alpha r^2}$. This makes the numerator an integral over three-dimensions, which can then be written in cylindrical coordinates

$$1 = |N|^2 \frac{\int d^3 r\,(r\sin\theta)^{2\ell} e^{-\alpha r^2}}{\int_0^\infty dr\, r^{2(\ell+1)} e^{-\alpha r^2}} = 2\pi|N|^2 \frac{\int_0^\infty d\rho\, \rho^{2\ell+1} e^{-\alpha\rho^2} \int_{-\infty}^{\infty} dz\, e^{-\alpha z^2}}{\int_0^\infty dr\, r^{2(\ell+1)} e^{-\alpha r^2}}$$

$$= 4\pi|N|^2 \frac{\int_0^\infty d\rho\, \rho^{2\ell+1} e^{-\alpha\rho^2} \int_0^\infty dz\, e^{-\alpha z^2}}{\int_0^\infty dr\, r^{2(\ell+1)} e^{-\alpha r^2}} \tag{15.93}$$

We can write the needed integrals as

$$\int_0^\infty d\rho\, \rho^{2\ell+1} e^{-\alpha\rho^2} = \frac{1}{2}\int_0^\infty d\rho^2 (\rho^2)^\ell e^{-\alpha\rho^2} = \frac{1}{2}\left(-\frac{\partial}{\partial\alpha}\right)^\ell \int_0^\infty du\, e^{-\alpha u}$$

$$= \frac{1}{2}\left(-\frac{\partial}{\partial\alpha}\right)^\ell \alpha^{-1} = \frac{\ell!}{2\alpha^{\ell+1}} \tag{15.94}$$

153

$$\int_0^\infty dr \, r^{2(\ell+1)} e^{-\alpha r^2} = \left(-\frac{\partial}{\partial \alpha}\right)^{\ell+1} \int_0^\infty dr \, e^{-\alpha r^2}$$

$$= \left(-\frac{\partial}{\partial \alpha}\right)^{\ell+1} \sqrt{2\pi} \alpha^{-1/2} = \sqrt{2\pi} \alpha^{-\ell-3/2} \frac{(2\ell+1)!!}{2^{\ell+1}} \qquad (15.95)$$

$$\int_0^\infty dz \, e^{-\alpha z^2} = \sqrt{2\pi} \alpha^{-1/2} \,. \qquad (15.96)$$

Combining we have

$$1 = 4\pi |N|^2 \frac{\ell! 2^\ell}{(2\ell+1)!!} = 4\pi |N|^2 \frac{(\ell!)^2 2^{2\ell}}{(2\ell)!(2\ell+1)} \qquad (15.97)$$

and the normalization is

$$|N| = \frac{1}{2^\ell \ell!} \sqrt{\frac{2\ell+1}{4\pi} (2\ell)!} \qquad (15.98)$$

Power series solution of the harmonic oscillator

Since the harmonic oscillator potential grows like x^2 for large $|x|$, we expect the eigenfunctions to decay like $\exp(-\alpha x^2)$ for large $|x|$. There is some hope for an analytic power series solution of the form $\sum_j a_j x^j e^{-\alpha x^2}$ since the potential is also a power. Taking natural units with $\hbar = m = \omega = 1$, the time-independent Schrödinger equation becomes

$$\left[\frac{d^2}{dx^2} - x^2 + 2E\right] \psi_n(x) = 0 \qquad (15.99)$$

plugging in the power series form gives

$$\sum_j a_j \left[j(j-1)x^{j-2} + (4\alpha^2 - 1)x^{j+2} + 2(E - j - \alpha)x^j\right] = 0 \,. \qquad (15.100)$$

We can eliminate the x^{j+2} term by choosing $\alpha = \frac{1}{2}$. Equating the coefficients of the powers of x gives the relation

$$a_{j+2}(j+2)(j+1) + (2E - 2j - 1)a_j = 0 \,. \qquad (15.101)$$

Let's see under what conditions we can obtain a solution that matches the boundary conditions. First, only nonnegative integer j values will give a wave function that has continuous derivatives at the origin, as required by the differential equation, and a normalizable solution. We see that since the a_j term is related to the a_{j+2} term, the solutions divide naturally into even and odd parity solutions as expected since parity commutes with the Hamiltonian.

154

If the series does not truncate, for large j we will have $a_{j+2} \sim \frac{2}{j}a_j$ and the power series for large j will go like

$$\psi(x) \sim \sum_m \frac{x^{2m}}{m!}e^{-\frac{x^2}{2}} \sim e^{\frac{x^2}{2}} \tag{15.102}$$

which diverges and is not normalizable. Therefore to have a normalizable result the power series must truncate. This will happen only if $2E - 2j - 1 = 0$, which gives the usual energy eigenvalues $E = \left(j + \frac{1}{2}\right)$ and the corresponding eigenfunctions.

15.14 Problems

Problem 15.1. In this problem you are to use the eigenstates of the harmonic oscillator Hamiltonian

$$H_2 = \frac{p^2}{2m} + \frac{1}{2}m\omega x^2. \tag{15.103}$$

to find the eigenvalues and eigenvectors of the Hamiltonian

$$H_4 = \frac{p^2}{2m} + \frac{1}{4}\alpha x^4. \tag{15.104}$$

a. Using second quantized notation calculate the matrix elements of $\langle m|x^4|n\rangle$ where $|n\rangle$ are the harmonic oscillator eigenstates. Demonstrate that the representation of the H_4 operator in the harmonic oscillator basis is block diagonal in the even and odd harmonic oscillator states.

b. Use the results of part a and Problem 6.6 to calculate the matrix elements of H_4, and evaluate $\langle 0|H_4|0\rangle$. Find the value of the harmonic oscillator ω that minimizes this energy expectation value. Use the value of ω which gives a minimum energy for the rest of the problem. Calculate the value of $\langle 0|H_4|0\rangle$ with this value of ω.

c. Define the eigenstates of H_4 to be $|\Psi_m\rangle$, so that $H_4|\Psi_m\rangle = E_m|\Psi_m\rangle$. Truncate the matrix, formed from the matrix elements in part b, to the first N harmonic oscillator eigenstates. Use whatever numerical tools you like to set up and diagonalize your matrix representation of H_4. Use a numerical value of N large enough so that you get the first 10 eigenvalues of H_4 correct to about 5 decimal places. Be sure to write your results for arbitrary values of α and m. About how large must N be to achieve this accuracy?

d. A system obeys the Schrödinger equation, $H(t)|\psi(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi(t)\rangle$, with the Hamiltonian

$$H(t) = \begin{cases} H_2 & |t| > T/2 \\ H_4 & |t| < T/2 \end{cases}, \tag{15.105}$$

with ω given by the value found in part b. For $t \to -\infty$ the system is in the ground state of H_2. Show that the probability of finding the system in the ground state of H_2 for $t \to \infty$ is

$$P(T) = \left| \sum_M |\langle 0|\Psi_M\rangle|^2 e^{-\frac{i}{\hbar}E_M T} \right|^2 \tag{15.106}$$

where E_M are the eigenvalues of H_4.

e. Use your numerical results to explain why the smallest probability for the particle to still be in the H_2 ground state at large t satisfies

$$P_s \geq \left(2|\langle 0|\Psi_0\rangle|^2 - 1 \right)^2 \tag{15.107}$$

and give the numerical value for the bound on P_s.

Problem 15.2.   a. Write the property of the destruction operator $a|0\rangle = 0$ in the x basis, solve the resulting first order differential equation, and normalize to get $\langle x|0\rangle$. Choose the usual phase convention that the ground state wave function is real and positive. That is verify Baym Eq. 4-151.

b. Show that

$$\langle x|x|n\rangle = x\langle x|n\rangle = \left(\frac{\hbar}{2m\omega}\right)^{1/2} \langle x|(a + a^+)|n\rangle \tag{15.108}$$

and use this to obtain the recursion relation

$$\langle x|n\rangle = \left(\frac{m\omega}{\hbar}\right)^{1/2} \sqrt{\frac{2}{n}} x\langle x|n - 1\rangle - \sqrt{1 - \frac{1}{n}}\langle x|n - 2\rangle \tag{15.109}$$

c. Use your numerical results from problem 15.1 and the results of parts a and b to construct and then plot the second excited state wave function of H_4, $\Psi_2(x) = \langle x|\Psi_2\rangle$. Plot $\Psi_2(x) \left(\frac{\hbar^2}{m\alpha}\right)^{1/12}$ versus $x \left(\frac{m\alpha}{\hbar^2}\right)^{1/6}$. As a check, you can compare your results with the results using the applet on the class web page.

Problem 15.3.   a. Prove the Jacobi identity for any three operators A, B and C,

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \tag{15.110}$$

b. A set of N operators satisfies commutation relations

$$[O_\alpha, O_\beta] = i \sum_\mu C_{\alpha\beta}^\mu O_\mu \tag{15.111}$$

with $C_{\alpha\beta}^\mu$ numbers (they are called the structure constants). Show that the $N \times N$ matrices M_α defined to be

$$(M_\alpha)_{jk} = -iC_{k\alpha}^j, \tag{15.112}$$

with j the row index and k the column index, are an $N \times N$ representation of the operators O_α. You can assume that each O_α is nontrivial, i.e. it has some nonzero matrix elements in some representation. The matrices M_α are called the adjoint representation of the algebra.

c. Write out the three matrices M_1, M_2 and M_3 for the adjoint representation of the angular momentum algebra

$$[L_\alpha, L_\beta] = i\hbar\epsilon_{\alpha\beta\gamma}L_\gamma. \tag{15.113}$$

What value of ℓ do they represent? Find a unitary transformation that brings them into the standard form derived in class.

Problem 15.4.   a. Calculate the commutation relation $[L_\alpha, T_{\beta\gamma}]$ if $T_{\beta\gamma}$ is a second-rank tensor operator, i.e. that each of its indices transforms under rotations like a vector operator.

b. An antisymmetric second-rank tensor operator satisifies $T_{\beta\gamma} = -T_{\gamma\beta}$. Using the commutation relation derived in part a, show that that $P_\eta = \epsilon_{\beta\gamma\eta}T_{\beta\gamma}$ transforms under rotations like a vector operator.

Chapter 16

Quantum normal modes

16.1 Hamiltonian

Often the Hamiltonian for a many-particle system can be written as the sum of the kinetic energies and a potential that depends on the positions of the particles. Usually the potential has an absolute minimum; the attractive coulomb potential is an example where the potential instead goes to $-\infty$. If we Taylor series expand the potential around its minimum, keeping just the second derivative terms, the potential will have the form

$$V = V_{\min} + \frac{1}{2} \sum_{jk} K_{jk} x_j x_k + ... \tag{16.1}$$

The spring constant matrix K_{jk} is the second derivative of the potential with respect to the displacements x_j and x_k evaluated at the minimum. Notice that K_{jk} is symmetric $K_{jk} = K_{kj}$. Truncating the potential, i.e. keeping just the quadratic terms, will be a good approximation if the energies are such that the particles have little or no probability of being in regions where the neglected terms are important. If they are small, the effects of these higher order terms can be included later using perturbation theory. Since $V_{\min}$ is an overall constant, it does not affect the results and the potential energy is usually taken to be zero at the minimum of V.

Our Hamiltonian is then

$$H = \sum_j \frac{p_j^2}{2m_j} + \frac{1}{2} \sum_{jk} x_j K_{jk} x_k \,, \tag{16.2}$$

with the usual canonical commutation relations

$$
\begin{aligned}
{[x_k, p_j]} &= i\hbar \delta_{kj} \\
{[x_k, x_j]} &= [p_k, p_j] = 0 \,.
\end{aligned}
\tag{16.3}
$$

Note that the x_j can indicate one of the $3N$ components of the displacement of N particles in 3-dimensions, so while our Hamiltonian looks like that of a set of particles in 1-dimension, it can also represent 3-dimensional systems. To make this clear, we take the sums to go from 1 to N_d where N_d is the number of degrees of freedom.

16.2 Normal mode transformation

As in the classical problem, we need to bring both the kinetic and potential energies into the form of a set of uncoupled harmonic oscillators. The kinetic energy is most simply dealt with by taking mass-scaled coordinates. That is we first take new momenta $p_j/\sqrt{m_j}$. To satisfy the canonical commutation relations, the displacements conjugate to $p_j/\sqrt{m_j}$ must be $\sqrt{m_j}x_j$. The Hamiltonian is then

$$H = \sum_j \frac{1}{2}\left(\frac{p_j}{\sqrt{m_j}}\right)^2 + \frac{1}{2}\sum_{jk}\left(\sqrt{m_j}x_j\right)A_{jk}\left(\sqrt{m_k}x_k\right) \tag{16.4}$$

with

$$A_{jk} = \frac{K_{jk}}{\sqrt{m_j m_k}}. \tag{16.5}$$

At this point since the momenta are now just squared and summed, any orthogonal transformation of the momenta will also give the sum of the squares. We therefore look for the orthogonal transformation of the displacements that diagonalizes the potential energy so that it is also a sum of squares.

To diagonalize A we calculate the eigenvalues and normalized eigenvectors

$$\sum_k A_{jk}\psi_k^{(n)} = \omega_n^2\psi_j^{(n)} \tag{16.6}$$

and write the transformation

$$X_n = \sum_j \psi_j^{(n)}\sqrt{m_j}x_j. \tag{16.7}$$

The conjugate momentum then must be

$$P_n = \sum_j \psi_j^{(n)}\frac{p_j}{\sqrt{m_j}}. \tag{16.8}$$

We can verify that these satisfy the canonical commutation relations by direct substitution. Since all the x_j commute with each other and all the p_j commute with each other, we immediately get that the X_n commute with each other and the P_n commute with each other. The $[X_n, P_m]$ commutator is

$$\begin{aligned}
[X_n, P_m] &= \sum_{jk}\psi_j^{(n)}\sqrt{m_j}\frac{\psi_k^{(m)}}{\sqrt{m_k}}[x_j, p_k] \\
&= \sum_{jk}\psi_j^{(n)}\sqrt{m_j}\frac{\psi_k^{(m)}}{\sqrt{m_k}}i\hbar\delta_{jk} \\
&= i\hbar\sum_j \psi_j^{(n)}\psi_j^{(m)} = i\hbar\delta_{nm}
\end{aligned} \tag{16.9}$$

160

where in the last line we use the orthogonality of the eigenvectors of a symmetric matrix.

Multiplying X_n and P_n by $\psi_j^{(n)}$ and summing over n using the orthogonality of the $\psi_j^{(n)}$, we get the result

$$\sqrt{m_j}x_j = \sum_n \psi_j^{(n)} X_n$$
$$\frac{p_j}{\sqrt{m_j}} = \sum_n \psi_j^{(n)} P_n . \tag{16.10}$$

Substituting these into our original Hamiltonian, we get the expected result

$$H = \sum_j \left(\frac{1}{2}P_n^2 + \frac{1}{2}\omega_n^2 X_n^2 \right) . \tag{16.11}$$

Since we expanded around the minimum of the potential, the potential energy must be positive for any set of displacements. We were therefore justified in writing the eigenvalues as ω_n^2 since they must all be positive.

16.3 Wave Functions

Since the Hamiltonian has now separated, we can calculate the usual eigenstate wavefunctions as a product of the wave functions for each separated piece. These wave functions are labeled by the state n_j of each of the oscillators, so we have

$$E_{n_1,n_2,...,n_{N_d}} = \sum_{j=1}^{N_d} \hbar\omega_j \left(n_j + \frac{1}{2} \right)$$
$$\Psi_{n_1,n_2,...,n_{N_d}}(X_1, ..., X_{N_d}) = \prod_{j=1}^{N_d} \omega_j^{1/4} \phi_{n_j}(X_j\sqrt{\omega_j}) \tag{16.12}$$

where $\phi_n(x)$ is the harmonic oscillator wave function with $\omega = m = 1$. We can transform this back to the original variables by using Eq. 16.7. In matrix elements and for normalization, we would integrate over $dX_1...dX_{N_d}$, so that changing variables we must calculate the Jacobian

$$dX_1...dX_{N_d} = \left| \frac{\partial(X_1, X_2, ..., X_{N_d})}{\partial(x_1, x_2, ..., x_{N_d})} \right| dx_1 dx_2...dx_{N_d} . \tag{16.13}$$

This is most easily calculated by using the chain rule

$$\frac{\partial(X_1, X_2, ..., X_{N_d})}{\partial(x_1, x_2, ..., x_{N_d})} = \frac{\partial(X_1, X_2, ..., X_{N_d})}{\partial(\sqrt{m_1}x_1, \sqrt{m_2}x_2, ..., \sqrt{m_{N_d}}x_{N_d})} \frac{\partial(\sqrt{m_1}x_1, \sqrt{m_2}x_2, ..., \sqrt{m_{N_d}}x_{N_d})}{\partial(x_1, x_2, ..., x_{N_d})}$$
$$\tag{16.14}$$

The elements of the first Jacobian are just the $\psi_j^{(n)}$, which is the orthogonal transformation so that that determinant is unity. The second term is the determinant of a diagonal matrix with $\sqrt{m_j}$ on the diagonals, so the Jacobian is

$$\frac{\partial(X_1, X_2, ..., X_{N_d})}{\partial(x_1, x_2, ..., x_{N_d})} = \prod_{j=1}^{N_d} \sqrt{m_j}. \tag{16.15}$$

Our normalized wave functions in terms of the original coordinates are therefore

$$\Phi_{n_1,n_2,...,n_{N_d}}(x_1, x_2, ..., x_{N_d}) = \prod_{j=1}^{N_d} (m_j\omega_j)^{1/4} \phi_{n_j} \left(\sum_k \psi_k^{(j)} x_k \sqrt{m_k \omega_j} \right). \tag{16.16}$$

16.4   Second Quantization

For many calculations, second quantized notation is much easier to use. For the normal-mode problem we can define the destruction operators for each mode exactly as for the single oscillator

$$a_n = \sqrt{\frac{\omega_n}{2\hbar}} X_n + i \frac{P_n}{\sqrt{2\hbar\omega_n}} \tag{16.17}$$

The Hamiltonian becomes

$$H = \sum_n \hbar\omega_n \left(a_n^+ a_n + \frac{1}{2} \right) \tag{16.18}$$

Taking sums and differences of a_n and a_n^+, we have

$$X_n = \sqrt{\frac{\hbar}{2\omega_n}}(a_n^+ + a_n)$$
$$P_n = i\sqrt{\hbar\omega_n}(a_n^+ - a_n) \tag{16.19}$$

Substituting Eqs. 16.7 and 16.8, multiplying by $\psi_k^{(n)}$ and summing over n, gives

$$x_k = \sum_n \psi_k^{(n)} \sqrt{\frac{\hbar}{2\omega_n m_k}}(a_n^+ + a_n)$$
$$p_k = i \sum_n \psi_k^{(n)} \sqrt{\frac{\hbar\omega_n m_k}{2}}(a_n^+ - a_n). \tag{16.20}$$

The eigenstates and energies are written as

$$E_{n_1,n_2,...,n_{N_d}} = \sum_{j=1}^{N_d} \hbar\omega_j \left(n_j + \frac{1}{2} \right)$$
$$|n_1, n_2, ..., n_{N_d}\rangle = \prod_{j=1}^{N_d} (a_j^+)^{n_j} |0, 0, ..., 0\rangle \equiv \prod_{j=1}^{N_d} \frac{(a_j^+)^{n_j}}{\sqrt{n_j!}} |0\rangle \tag{16.21}$$

where $|0\rangle$ is defined to be the ground state.

16.5   Example Calculation

As we will learn later, the electric dipole transition rate is proportional to the transition electric-dipole matrix element squared. The dipole operator is

$$\boldsymbol{d} = \sum_j q_j \boldsymbol{r}_j \tag{16.22}$$

so that the electric-dipole operator matrix elements are linear combinations of the the matrix elements x_j. For example, if ω_1 is the lowest frequency, the first excited state has the normal mode oscillator 1 in its first excited state and all the others in their ground state. The matrix element of x_j between this state and the ground state is then

$$\langle 0 | x_j a_1^+ | 0 \rangle = \langle 0 | \sum_n \psi_j^{(n)} \sqrt{\frac{\hbar}{2\omega_n m_j}} (a_n^+ + a_n) a_1^+ | 0 \rangle = \sum_n \psi_j^{(n)} \sqrt{\frac{\hbar}{2\omega_n m_j}} \delta_{n1} = \psi_j^{(1)} \sqrt{\frac{\hbar}{2\omega_1 m_j}} \, . \tag{16.23}$$

16.6 Problems

Problem 16.1. The Hamiltonian for a set of particles is given by

$$H = \sum_{j=1}^{N} \frac{p_j^2}{2m_j} + \frac{1}{2} \sum_{j,k=1}^{N} x_j K_{jk} x_k \,. \tag{16.24}$$

The symmetric matrix $A_{jk} = K_{jk}/\sqrt{m_j m_k}$ has normalized eigenvectors $\psi_j^{(n)}$ and eigenvalues,

$$\sum_{k=1}^{N} A_{jk} \psi_k^{(n)} = \omega_n^2 \psi_j^{(n)} \,. \tag{16.25}$$

The system is in its ground state. Calculate the probability density $P_j(x)$ defined such that the probability of finding displacement x_j between x and $x + dx$ is $P_j(x)dx$.

Chapter 17

Central potentials

The Hamiltonian

$$H = \frac{p^2}{2m} + v(r) \tag{17.1}$$

is rotationally invariant; therefore, it commutes with the angular momentum

$$
\begin{aligned}
[H, \boldsymbol{L}] &= 0 \\
[H, L^2] &= 0
\end{aligned} \tag{17.2}
$$

so that we can find simultaneous eigenvectors of H, L_z and L^2. The eigenvectors $|r\ell m\rangle$ of L^2 and L_z are all degenerate, so the most general form for the eigenvectors of H that are also eigenvectors of L^2 and L_z are arbitrary linear combinations of the degenerate eigenvectors

$$|\psi_{n\ell m}\rangle = \int_0^\infty r^2 dr R_{n\ell m}(r) |r\ell m\rangle . \tag{17.3}$$

The eigenfunctions then have the form

$$
\begin{aligned}
\psi_{n\ell m}(r, \theta, \phi) &= \langle r, \theta, \phi | \psi_{n\ell m} \rangle = \int_0^\infty dr' r'^2 R_{n\ell m}(r') \langle r\theta\phi | r'\ell m \rangle \\
&= \int_0^\infty dr' r'^2 R_{n\ell m}(r') \frac{\delta(r - r')}{r^2} Y_{\ell m}(\theta, \phi) = R_{n\ell m}(r) Y_{\ell m}(\theta, \phi)
\end{aligned} \tag{17.4}
$$

Using our results for the angular momentum eigenstates, we can write the Hamiltonian as

$$H = \sum_{\ell=0}^\infty \sum_{m=-\ell}^\ell \int_0^\infty dr |r\ell m\rangle \left(-\frac{\hbar^2}{2m} \left[\frac{1}{r} \frac{\partial^2}{\partial r^2} r - \frac{\ell(\ell+1)}{r^2} \right] + v(r) \right) \langle r\ell m| . \tag{17.5}$$

Operating on Eq. 17.3, we get the differential equation for the radial function

$$\left(-\frac{\hbar^2}{2m} \left[\frac{1}{r} \frac{\partial^2}{\partial r^2} r - \frac{\ell(\ell+1)}{r^2} \right] + v(r) \right) R_{n\ell m}(r) = E_{n\ell m} R_{n\ell m}(r) . \tag{17.6}$$

Since this does not depend on m, we usually drop the m index since the radial functions can be taken to be the same for every m value corresponding to the same ℓ. Multiplying through by r, we see that $rR_{n\ell}(r)$ satisfies a 1-dimensional time-independent Schrödinger equation with the effective radial potential

$$\frac{\hbar^2 \ell(\ell+1)}{2mr^2} + v(r),\tag{17.7}$$

analogously to the classical central force problem.

The boundary conditions for $R_{n\ell}(r)$ must be chosen so that the wave function is normalizable, and some special attention has to be paid to the origin where spherical coordinates have a coordinate singularity. Except for the coulomb potential, the origin is not a particularly special place, and the nearly all the potentials we deal with will be well behaved there. For these potentials, we can expand the potential in a Taylor series around the origin, and combine the constant part with the energy eigenvalue. For a small enough region around the origin, our solutions will be

$$R_{n\ell}(r) \to Ar^\ell + Br^{-\ell-1}\tag{17.8}$$

For $\ell > 0$, the $r^{-\ell-1}$ is not normalizable at the origin, so that $B = 0$. For $\ell = 0$, both terms are normalizable, but the B term will give a wave function proportional to r^{-1}. We know from electromagnetism that this is the potential of a point charge, and Poisson's equation tells us, $-\nabla^2 r^{-1} = 4\pi\delta^3(\boldsymbol{r})$. Therefore a r^{-1} wave function would require a (rather strange) delta function term in the potential. Since our potential has no delta function, we again reject the B term. Since R goes to either zero or a constant at the origin, rR goes to zero there, and the solution to the effective 1-dimensional radial equation has the boundary condition that rR is zero at the origin.

Potentials that are strongly repulsive at the origin, will necessarily give wave functions that are smaller there. They will also have rR going to zero at the origin.

The attractive coulomb potential is an interesting special case. For $\ell > 0$, the effective potential at the origin is dominated by the L^2/r^2 part, so the boundary condition is the same and the radial function will go like r^ℓ. For $\ell = 0$, we require

$$-\frac{\hbar^2}{2m}(rR)'' - \frac{Ze^2}{r}rR = ErR\tag{17.9}$$

Trying $rR = r^\alpha$, we have

$$-\frac{\hbar^2}{2m}\alpha(\alpha-1)r^{\alpha-2} - Ze^2 r^{\alpha-1} = Er^\alpha\tag{17.10}$$

and the dominant term at the origin will be the first. Just as for the constant potential, it requires $\alpha = 0, 1$, we can again reject the $\alpha = 0$ term, since there is no cancelling delta function anywhere.

17.1   The spherical cavity

Probably the simplest system is a particle in a spherical cavity, so that

$$v(r) = \left\{ \begin{array}{ll} 0 & r < a \\ \infty & r > a \end{array} \right. . \tag{17.11}$$

The radial equation is the spherical Bessel equation, and the solutions that are well behaved at the origin are $j_\ell(kr)$, where $E = \hbar^2 k^2/2m$. The infinite potential at $r = a$ requires the wave function to be zero there, so $j_\ell(k_{n\ell}a) = 0$. Writing the zeroes of the spherical Bessel functions as $\gamma_{n\ell}$, we have $k_{n\ell} = \gamma_{n\ell}/a$, with

$$\begin{aligned} \psi_{n\ell m}(r) &= N_{n\ell} j_\ell(\gamma_{n\ell} r/a) Y_{\ell m}(\theta, \phi) \\ E_n &= \frac{\hbar^2 \gamma_{n\ell}^2}{2ma^2} . \end{aligned} \tag{17.12}$$

The normalization $N_{n\ell}$ can be evaluated using the Lommel integral method.

The energy levels are shown in Table 17.1. I have chosen to label n starting from 1, so $n = 1$ is the first zero of the corresponding Bessel function. The square well shows that in the normal case, the only degeneracy is that given by the L_z quantum number. Each of the energy levels is $2\ell + 1$ degenerate. There is no degeneracy between energy levels with different total angular momentum.

17.2   The harmonic oscillator

The isotropic harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 r^2 \tag{17.13}$$

can be separated into three independent harmonic oscillators in cartesian coordinates. The eigenfunctions are just the products of the eigenfunctions for these three oscillators, and the energies are the sum of the oscillator energies. The possible energies are then

$$E_{n_x n_y n_z} = \hbar\omega \left(n_x + n_y + n_z + \frac{3}{2} \right) \tag{17.14}$$

or nonnegative integers plus three halves times $\hbar\omega$. The degeneracy of the energy level $\hbar\omega(k + 3/2)$ can be calculated by realizing that n_x can run from 0 to k, n_y can run from 0 to $k - n_x$, with n_z given by $k - n_x - n_y$. The degeneracy d_k is

$$d_k = \sum_{n_x=0}^{k} (k - n_x + 1) = \frac{(k+2)(k+1)}{2} . \tag{17.15}$$

The degeneracies $d_0 = 1$, $d_1 = 3$, $d_2 = 6$, $d_3 = 10$ are not equal to $2\ell + 1$, so we know that there are more degneracies here than expected for a general central potential.

n	ℓ	$\gamma_{n\ell}$	$\frac{2ma^2}{\hbar^2}E_n$
1	0	3.1416	9.8696
1	1	4.4934	20.191
1	2	5.7635	33.218
2	0	6.2832	39.479
1	3	6.9879	48.831
2	1	7.7253	59.680
1	4	8.1826	66.955
2	2	9.0950	82.719
1	5	9.3558	87.531
3	0	9.4248	88.827
2	3	10.4171	108.52
1	6	10.5128	110.52
3	1	10.9041	118.89
1	7	11.6570	135.89
2	4	11.7049	137.00
3	2	12.3229	151.85
4	0	12.5664	157.91

Table 17.1: The zeroes of the spherical Bessel functions and the corresponing energies for a cavity of radius a shown to 4 places after the decimal point.

It is convenient to use natural units as we did for the one-dimensional oscillator. The radial equation then becomes

$$-R'' - 2r^{-1}R' + \ell(\ell+1)r^{-2}R + r^2 R - 2ER = 0 \tag{17.16}$$

We can use our knowledge of the cartesian solution which must be polynomials times $e^{-r^2/2}$, to write $R(r) = P(r)e^{-r^2/2}$. Substituting, we have

$$-P'' + 2rP' - 2r^{-1}P' + \ell(\ell+1)r^{-2}P - (2E-3)P = 0 \,. \tag{17.17}$$

Substituting a power series $P = \sum_j a_j r^j$, where $j = \ell, \ell+1, \dots$ from our general analysis above of well behaved potentials, and equating coefficients

$$-a_{j+2}(j+2)(j+1) + a_j 2j - a_{j+2}(j+2)2 + \ell(\ell+1)a_{j+2} - (2E-3)a_j = 0 \tag{17.18}$$

or

$$a_{j+2} = a_j \frac{2j - 2E + 3}{\ell(\ell+1) - (j+2)(j+3)} \tag{17.19}$$

If the series does not truncate, for large j the factor goes like $(j/2)^{-1}$, so the series terms will go like $r^2/(j/2)!$, and for large r, the sum will go like e^{r^2}. Multiplying by the factor $e^{-r^2/2}$ to obtain R, it is not normalizable. Therefore the series must truncate. This requires

$2j - 2E + 3 = 0$ for some value of j. The possible energies are $E = (k + 3/2)$, with k a nonnegative integer. Notice that the series can only truncate for either even or odd terms, therefore the radial function must contain just even or just odd terms in r. The denominator goes to zero if $\ell = j + 1$. Therefore $\ell \leq k$. Putting these restrictions together, we find that $k = 0, 1, 2, ...$, $E = (k + 3/2)\hbar\omega$, and for each k value, we have $k = \ell$, $k = \ell - 2$, $k = \ell - 4$ etc. as long as $\ell \geq 0$. Let's check out the degeneracies. The $k = 0$ level must be $\ell = 0$ and there is just 1 state. The $k = 1$ level must be $\ell = 1$ and the degeneracy is $2\ell + 1$ or 3. The $k = 2$ level must be $\ell = 2$ or $\ell = 0$. The degeneracy is therefore 5 for the $\ell = 2$ states and 1 for the $\ell = 0$ state giving a total degeneracy of 6. The k=3 level has $\ell = 3$ and $\ell = 1$ for a total degeneracy of $7 + 3 = 10$. These agree with our counting from the cartesian levels. The polynomials we constructed are associated Laguerre polynomials.

17.3 Hydrogen atom

The hydrogen atom problem begins as the problem of the quantum states of the two particle system of an electron and a proton. Here we will assume that we have already changed to relative and center of mass coordinates. The relative coordinate Hamiltonian is

$$H = \frac{p^2}{2m} - \frac{e^2}{r} \tag{17.20}$$

where m is the reduced mass and e is the magnitude of the electric charge ($e > 0$). This is a central potential.

It is convenient to work in the natural units of this problem. These are $m = \hbar = e = 1$. If instead of the reduced mass, we use the electron mass, these are called atomic units. For most of our work, the difference between the reduced and electron mass can be neglected, so I will call these atomic units. One extra advantage of atomic units is that you can use them whether you start with SI or Gaussian (or any other) units. We are using Gaussian units here, so the potential is $-e^2/r$, but if you use SI units, where the potential is $-(\mu_0 c^2/4\pi)e^2/r = -e^2/4\pi\epsilon_0 r$, atomic units means that you have chosen units where m, $\hbar$ and $e^2\mu_0 c^2/4\pi$ are all 1.

In atomic units, the unit of mass is the electron mass, the unit of length is the Bohr radius $a_0 = \hbar^2/me^2$ and the unit of time is $\hbar^3/me^4$. The conversion factors are shown in Table 17.2.

The radial equation in atomic units is

$$-\frac{1}{2r}\frac{\partial^2}{\partial r^2}rR_{n\ell}(r) + \frac{\ell(\ell+1)}{2r^2}R_{n\ell}(r) - \frac{1}{r}R_{n\ell}(r) - E_{n\ell}R_{n\ell}(r) = 0 \tag{17.21}$$

To begin with we are most interested in the bound states which have a simple analytic form. For negative energies, and large r the r^{-2} and r^{-1} will go to zero, and the solutions will become $e^{\pm\alpha r}$, with $\alpha = \sqrt{-2E}$. Only the negative sign gives a normalizable wave function, so the bound state radial functions will go like $e^{-\sqrt{-2E}r}$ at large r.

quantity	Name	Symbol	Gaussian value	SI value
a.u. mass	electron mass	m or m_e	9.1094×10^{-28} g	9.1094×10^{-31} kg
a.u. length	Bohr radius	a_0	$\frac{\hbar^2}{me^2} = 0.5292 \times 10^{-8}$cm	$\frac{4\pi}{\mu_0 c^2}\frac{\hbar^2}{me^2} = 0.5292 \times 10^{-10}$m
a.u. time			$\frac{\hbar^3}{me^4} = 2.4189 \times 10^{-17}$ s	$\left(\frac{4\pi}{\mu_0 c^2}\right)^2 \frac{\hbar^3}{me^4} = 2.4189 \times 10^{-17}$s
a.u. energy	Hartree	E_h	$\frac{e^4 m}{\hbar^2} = 4.3597 \times 10^{-11}$ergs	$\left(\frac{\mu_0 c^2}{4\pi}\right)^2 \frac{e^4 m}{\hbar^2} = 4.3597 \times 10^{-18}$J
			$= 27.2114$ eV	$= 27.2114$ eV

Table 17.2: Atomic units and conversion factors.

Noting that every nonderivative term multiplies by a power of r, we try a power series

$$R_{n\ell} = \sum_{j=\ell}^{\infty} a_j r^j e^{-\sqrt{-2E}r} \tag{17.22}$$

where our general arguments above restrict the form at the origin. Substituting into the differential equation, and dividing the result by $e^{-\sqrt{-2E}r}$, which can never be exactly zero,

$$\sum_{j=\ell}^{\infty} \left\{ -\frac{1}{2r} \left[j(j+1)r^{j-1} - 2\sqrt{-2E}(j+1)r^j - 2Er^{j+1} \right] + \frac{\ell(\ell+1)}{2}r^{j-2} - r^{j-1} - Er^j \right\} a_n = 0 \tag{17.23}$$

Setting the coefficient of each power of r to zero gives the recursion relation

$$a_{j+2} \left[\frac{j(j+1) - \ell(\ell+1)}{2} \right] - [\alpha(j+1) - 1]a_{j+1} . \tag{17.24}$$

Taking j large, we find that if the series does not truncate, the terms go like $a_j = \frac{2\sqrt{-2E}}{j}a_{j-1}$, or $a_j \propto (2\sqrt{-2E})^j/j!$. Since the power series for large r is dominated by these terms, and $\sum_j \frac{(2\sqrt{-2E}r)^j}{j!}e^{-\sqrt{-2E}r} = e^{\sqrt{-2E}r}$, the function diverges at large r and is not normalizable. Therefore, the series must truncate for some j value. This means that $\sqrt{-2E} = 1/(j+1)$ for some value of j. Traditionally, the n value is taken to be this $j+1$ value, and since j begins at 0, the smallest n value is 1. From the $j(j+1) - \ell(\ell+1)$ factor, $0 \le \ell < n$ are all possible. The energy only depends on n and

$$E_{n\ell} = E_n = -\frac{1}{2n^2} . \tag{17.25}$$

The recursion relation can be written as

$$a_j = a_{j-1} \frac{\sqrt{-2E_n}(\ell+j) - 1}{(\ell+j)(\ell+j+1) - \ell(\ell+1)} , \tag{17.26}$$

which gives polynomials. These can be written in terms of the the the associated Laguerre

170

polynomials, which have the Rodriguez formula definitions

$$
\begin{aligned}
L_\lambda^\mu(x) &= \frac{d^\mu}{dx^\mu} L_\lambda(x) \\
L_\lambda(x) &= e^x \frac{d^\lambda}{dx^\lambda} e^{-x} x^\lambda .
\end{aligned}
\tag{17.27}
$$

Normalizing, the radial functions become

$$
R_{n\ell}(r) = -\sqrt{\frac{(n-\ell-1)!}{[(n+\ell)!]^3}} \frac{2}{n^2} \left(\frac{2r}{n}\right)^\ell L_{n+\ell}^{2\ell+1}\left(\frac{2r}{n}\right) e^{-\frac{r}{n}} .
\tag{17.28}
$$

The radial functions for the first few states are

$$
\begin{aligned}
R_{10}(r) &= 2e^{-r} \\
R_{20}(r) &= 2^{-3/2}(2-r)e^{-r/2} \\
R_{21}(r) &= 2^{-3/2}\frac{r}{\sqrt{3}}e^{-r/2}
\end{aligned}
\tag{17.29}
$$

The electron in a hydrogen atom can also be unbound, with a positive energy. The radial differential equation is the same, but E is positive. This makes $\sqrt{-2E}$ imaginary, and gives oscillatory solutions as we should expect. The analysis can be done nearly identically, however, and we are lead to the same recursion relations. This time however, there is no divergence since both signs for $\exp(\pm\sqrt{-2E}r)$ give solutions, and the power series does not need to truncate. This should make sense, since an unbound electron should be able to have any positive energy. The power series solution converges, albeit very slowly for large r, so it can be used to define the solutions. Alternatively, for numerical work, the radial differential equation can be solved directly. Writing $E = \frac{k^2}{2}$, the power series is a confluent hypergeometric function, with the result

$$
\begin{aligned}
R_{k\ell}(r) &= \frac{2\sqrt{k}}{\sqrt{(1-e^{-2\pi k^{-1}})}} \prod_{j=1}^\ell \sqrt{j^2+k^{-2}} \frac{(2kr)^{-\ell-1}}{(2\ell+1)!} e^{ikr} F(ik^{-1}+\ell+1, 2\ell+2, 2ikr) \\
&= \frac{\sqrt{k}}{\pi\sqrt{1-e^{-2\pi k^{-1}}}} \prod_{j=1}^\ell \sqrt{j^2+k^{-2}}(2kr)^{-\ell-1} \oint dz e^{-2ikrz}(z+\tfrac{1}{2})^{-ik^{-1}-\ell-1}(z-\tfrac{1}{2})^{ik^{-1}-\ell-1},
\end{aligned}
\tag{17.30}
$$

and

$$
F(a,b,x) = 1 + \frac{a}{b\cdot 1!}x + \frac{a(a+1)}{b(b+1)\cdot 2!}x^2 + \dots .
\tag{17.31}
$$

In the integral representation in the second expression, the two branch points at $z \pm \frac{1}{2}$ are joined by a branch cut and the integration is closed around this cut in a counterclockwise direction.

171

The states are normalized here so that

$$
\begin{aligned}
\langle n\ell m|n'\ell'm'\rangle &= \delta_{nn'}\delta_{\ell\ell'}\delta_{mm'} \\
\langle k\ell m|k'\ell'm'\rangle &= \delta(k-k')\delta_{\ell\ell'}\delta_{mm'} \\
\langle r\theta\phi|n\ell m\rangle &= R_{n\ell}(r)Y_{\ell m}(\theta,\phi) \\
\langle r\theta\phi|k\ell m\rangle &= R_{k\ell}(r)Y_{\ell m}(\theta,\phi)
\end{aligned}
\tag{17.32}
$$

and the completeness relation is

$$
\sum_{n\ell m}|n\ell m\rangle\langle n\ell m| + \sum_{\ell m}\int_0^\infty dk|k\ell m\rangle\langle k\ell m|\,.
\tag{17.33}
$$

17.4 Degeneracies

The $2\ell+1$ degeneracy of all central force problems comes from the conservation of angular momentum. L^2, L_z and H all commute so we simultaneously diagonalized them to get the $|n\ell m\rangle$ states with

$$
\begin{aligned}
H|n\ell m\rangle &= E_{n\ell m}|n\ell m\rangle \\
L_z|n\ell m\rangle &= m\hbar|n\ell m\rangle \\
L^2|n\ell m\rangle &= \ell(\ell+1)\hbar^2|n\ell m\rangle\,.
\end{aligned}
\tag{17.34}
$$

However, H commutes with L_x and L_y as well, and therefore $L_\pm$. Applying $L_\pm$ to the Hamiltonian equation, we get

$$
L_\pm H|n\ell m\rangle = H L_\pm|n\ell m\rangle = E_{n\ell m}L_\pm|n\ell m\rangle
\tag{17.35}
$$

which shows immediately that $L_\pm|n\ell m\rangle \propto |n\ell m \pm 1\rangle$ is an eigenstate of H with eigenvalue $E_{n\ell m}$. The energy therefore must be independent of m, and the states are $2\ell+1$ fold degenerate.

The extra degeneracies of the isotropic harmonic oscillator and the hydrogen atom can be understood from other symmetries of the system. In the harmonic oscillator we can operate with a raising operator on one component and the lowering operator on another. These operators will commute with H. Operating on the H eigenstate equation gives the other degenerate harmonic oscillator states.

The huge degeneracy of the hydrogen atom can be traced to the extra conserved quantity, the Runge-Lenz vector as described in Baym. Since the corresponding Runge-Lenz operator commutes with the Hamiltonian, you can again apply it to the H eigenstate equations to generate the other degenerate states. Pauli used this extra symmetry, as described in Baym, to solve for the bound state energies of the hydrogen atom using Heisenberg's matrix mechanics before Schrödinger had written down his differential equation. Pauli's demonstration that Heisenberg's method gave the correct hydrogen spectrum (Pauli also included

static electric and magnetic fields to get the Stark effect etc. too) was a major reason why people believed Heisenberg's quantum mechanics was likely correct.

Be sure to compare the energy levels of the hydrogen atom to the energy levels of the spherical cavity. The cavity levels are more representative of potentials without any special symmetry.

17.5   Problems

Problem 17.1.   a. A *classical* particle is described in the lab frame by the Lagrangian $\mathcal{L}_0(\boldsymbol{r}, \boldsymbol{v}) = \frac{1}{2}mv^2 - v(r)$. Show that the Hamiltonian in a frame rotating with angular velocity $\boldsymbol{\Omega}$ has the form

$$H_R(\boldsymbol{r}, \boldsymbol{p}) = H_0(\boldsymbol{r}, \boldsymbol{p}) - \boldsymbol{L} \cdot \boldsymbol{\Omega} \tag{17.36}$$

where in Eq. 17.36 $\boldsymbol{r}$, $\boldsymbol{p}$, and $\boldsymbol{L}$ are the rotating frame position and momentum and angular momentum.

   b. In the same system described quantum mechanically, the state in the rotating frame would be given by rotating the lab frame state by an angle $|\boldsymbol{\Omega}|t$ around $\hat{\boldsymbol{\Omega}}$, or $|\psi_R(t)\rangle = e^{\frac{i}{\hbar}\boldsymbol{L} \cdot \boldsymbol{\Omega} t}|\psi_0(t)\rangle$. If the lab frame Hamiltonian is H, which commutes with $\boldsymbol{L}$, show that the rotating frame propagator is $\exp\left(-\frac{i}{\hbar}[H - \boldsymbol{L} \cdot \boldsymbol{\Omega}]t\right)$, so that the the rotating frame Hamiltonian has the same form as the classical case.

   You can assume in both cases that the frames coincide at $t = 0$.

Problem 17.2. A particle is trapped in a cavity of radius a, so that its Hamiltonian is

$$H = \frac{p^2}{2m} + v(r) \tag{17.37}$$

with

$$v(r) = \begin{cases} 0 & r < a \\ \infty & r \geq a \end{cases}. \tag{17.38}$$

The particle is in thermal equilibrium with the walls, which are very cold so that the particle remains in its ground state in the frame of the walls.

   a. Show that if the cavity is rotated about the origin with angular velocity $\boldsymbol{\Omega}$, that the particle, when viewed in the lab frame, will remain in an angular momentum zero state until Ω reaches a critical angular speed $\alpha\hbar/ma^2$, and find the numerical value of α.

   b. Write a few sentences that *qualitatively* compare your result with the vortex formation results in Bose-Einstein condensates described in Physics Today, Vol. 53, No. 8, 19 (2000). The online version is available from the past issues link at http://www.physicstoday.org/

   Note, it requires a full many-body treatment to make quantitative predictions. You will not be able to predict any numbers from our simplified problem.

   Note: If you are wondering how the angular momentum gets transferred, we require some imperfections in the cavity wall, away from perfect rotational invariance. This is implicit in my statement that the particle comes into equilibrium with the rotating walls. These imperfections, when small, do not change the equilibrium result, they only change the transition rates and therefore the time to reach equilibrium.

Chapter 18

Relative and center of mass coordinates

18.1 Introduction

Changing to relative and center of mass coordinates for a pair of particles should be familiar from classical mechanics. Perhaps the most straightforward way of getting the quantum result is to transform the classical Hamiltonian and then quantize it. However, since quantum mechanics is the more fundamental theory, it is instructive to make the transformation quantum mechanically.

18.2 Two particle Hamitonian with a relative coordinate potential

A two particle Hamiltonian

$$\frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + v(\boldsymbol{r}_1 - \boldsymbol{r}_2) \tag{18.1}$$

is translationally invariant if we translate the entire coordinate system, or translate both particles together. The generator of the translation of $\boldsymbol{r}_1$ is $\boldsymbol{p}_1$, while the generator of translations of $\boldsymbol{r}_2$ is $\boldsymbol{p}_2$. The translation operator for translation of both particles is the product of the translation operators for each separately. Using infinitesimal translations by $\boldsymbol{a}$, we have

$$T(\boldsymbol{a}) = \left(1 + \frac{i}{\hbar}\boldsymbol{p}_1 \cdot \boldsymbol{a}\right)\left(1 + \frac{i}{\hbar}\boldsymbol{p}_2 \cdot \boldsymbol{a}\right) = \left(1 + \frac{i}{\hbar}\left[\boldsymbol{p}_1 + \boldsymbol{p}_2\right] \cdot \boldsymbol{a}\right) \tag{18.2}$$

so that the generator of translations of both particles is $\boldsymbol{p}_1 + \boldsymbol{p}_2$. This is the total momentum $\boldsymbol{P} = \boldsymbol{p}_1 + \boldsymbol{p}_2$ which will also turn out to be the center of mass momentum. We define the relative coordinate as $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$, to simplify the potential. We can check that all of the

components of $\boldsymbol{P}$ commute with all of the components of $\boldsymbol{r}$. Therefore $[H, \boldsymbol{P}] = 0$, and we can simultaneously find eigenstates of H and $\boldsymbol{P}$.

The conjugate position to $\boldsymbol{P}$ and the conjugate momentum to $\boldsymbol{r}$ must satisfy the canonical commutation relations. Up to unimportant transformations (like translations and rotations) of the overall coordinate system, we can take $X = \alpha x_1 + \beta x_2$, and $p_x = \gamma p_{1x} - \delta p_{2x}$, with α, β, γ, δ constants, and similarly for the y and z components. Substituting into the canonical commutation relations

$$
\begin{aligned}
[P_x, X] = -i\hbar &\Rightarrow \alpha + \beta = 1 \\
[p_x, x] = -i\hbar &\Rightarrow \gamma + \delta = 1 \\
[p_x, X] = 0 &\Rightarrow \gamma\alpha - \delta\beta = 0
\end{aligned}
\tag{18.3}
$$

solving in terms of the constant α, we get the possibilities

$$
\begin{aligned}
X &= \alpha x_1 + (1-\alpha)x_2 \\
p_x &= (1-\alpha)p_{1x} - \alpha p_{2x} \\
p_{1x} &= \alpha P_x + p_x \\
p_{2x} &= (1-\alpha)P_x - p_x .
\end{aligned}
\tag{18.4}
$$

Any choice of α would give us a valid transformation into some relative and "center of mass" coordinates. To find the useful one, we plug our expressions into the Hamiltonian

$$
\begin{aligned}
\frac{p_{1x}^2}{2m_1} + \frac{p_{2x}^2}{2m_2} &= \frac{(\alpha P_x + p_x)^2}{2m_1} + \frac{[(1-\alpha)P_x - p_x]^2}{2m_2} + v(\boldsymbol{r}) \\
&= \frac{\alpha^2 P_x^2}{2m_1} + \frac{\alpha P_x p_x}{m_1} + \frac{p_x^2}{2m_1} + \frac{(1-\alpha)^2 P_x^2}{2m_2} - \frac{(1-\alpha)P_x p_x}{m_2} + \frac{p_x^2}{2m_2} .
\end{aligned}
\tag{18.5}
$$

If we pick α such that the cross terms are zero, the Hamiltonian will separate. We solve $\alpha/m_1 - (1-\alpha)/m2 = 0$ to get

$$
\alpha = \frac{m_1}{m_1 + m_2}
\tag{18.6}
$$

This gives the usual center of mass coordinates. Other choices of α correspond to other systems with different m_1 and m_2. We have

$$
\begin{aligned}
\boldsymbol{p}_1 &= \frac{m_1}{m_1 + m_2}\boldsymbol{P} + \boldsymbol{p} \\
\boldsymbol{p}_2 &= \frac{m_2}{m_1 + m_2}\boldsymbol{P} - \boldsymbol{p} \\
\boldsymbol{r}_1 &= \boldsymbol{R} + \frac{m_2}{m_1 + m_2}\boldsymbol{r} \\
\boldsymbol{r}_2 &= \boldsymbol{R} - \frac{m_1}{m_1 + m_2}\boldsymbol{r}
\end{aligned}
\tag{18.7}
$$

and

$$
\begin{aligned}
\boldsymbol{P} &= \boldsymbol{p}_1 + \boldsymbol{p}_2 \\
\boldsymbol{p} &= \frac{m_2}{m_1 + m_2}\boldsymbol{p}_1 - \frac{m_1}{m_1 + m_2}\boldsymbol{p}_2 \\
\boldsymbol{R} &= \frac{m_1}{m_1 + m_2}\boldsymbol{r}_1 + \frac{m_2}{m_1 + m_2}\boldsymbol{r}_2 \\
\boldsymbol{r} &= \boldsymbol{r}_1 - \boldsymbol{r}_2 .
\end{aligned}
\tag{18.8}
$$

The Hamiltonian is

$$
H = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + v(\boldsymbol{r}_1 - \boldsymbol{r}_2) = \frac{P^2}{2M} + \frac{p^2}{2m} + v(\boldsymbol{r})
\tag{18.9}
$$

where

$$
\begin{aligned}
M &= m_1 + m_2 \\
\frac{1}{m} &= \frac{1}{m_1} + \frac{1}{m_2}
\end{aligned}
\tag{18.10}
$$

18.3   Example problem

An atom consists of a tritium nucleus (a proton and two neutrons) and an electron with the electron in its 1s ground state with the atom at rest. The nucleus beta decays to becomes ^{3}He with two protons and one neutron emitting an electron and an antineutrino. Assume that this process can be modeled as the nucleus instantaneously changing from charge e to charge 2e and instantly transferring a momentum $m_N\boldsymbol{v}$ to the remaining helium nucleus of mass m_N

What is the probability that the electron is in the 1s ground state around the moving helium nucleus? You can ignore the size of the nucleus and assume the mass of both nuclei are so large that the reduced mass is well approximated by the electron mass.

18.4   Example solution

The hydrogenic 1s state is

$$
\psi_{1s}(r) = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{3/2} e^{-Zr/a_0}
\tag{18.11}
$$

where Ze is the nuclear charge, and a_0 the Bohr radius for the reduced mass.

The initial wave function before the decay is $\psi_{1s}^H(\boldsymbol{r}_e - \boldsymbol{r}_N)$ where the superscript H is for hydrogen, and $\boldsymbol{r}_e$ and $\boldsymbol{r}_N$ are the electron and nuclear coordinates. Just after the decay, the nucleus has gained momentum $m_N\boldsymbol{v}$, but there is no time for the electron-nuclear part of the wave function to change. We therefore simply operate on the state with the momentum translation operator $e^{\frac{i}{\hbar}m_N\boldsymbol{v}\cdot\boldsymbol{r}_{N\mathrm{op}}}$; the total wave function becomes

$$
\psi_{1s}^H(|\boldsymbol{r}_e - \boldsymbol{r}_N|)e^{\frac{i}{\hbar}m_N\boldsymbol{v}\cdot\boldsymbol{r}_N} = \psi_{1s}^H(r)e^{\frac{i}{\hbar}m_N\boldsymbol{v}\cdot\boldsymbol{R}}e^{-\frac{i}{\hbar}\frac{m_e m_N}{m_e + m_N}\boldsymbol{v}\cdot\boldsymbol{r}}
\tag{18.12}
$$

where in the second expression I have written the wave function in terms of the relative and center of mass coordinates. The desired final state wave function has the center of mass with the same momentum as above and the relative coordinate in its ground state

$$e^{\frac{i}{\hbar}m_N \boldsymbol{v}\cdot\boldsymbol{R}}\psi_{1s}^{He}(r)\,. \tag{18.13}$$

The overlap squared gives the probability. Since the center of mass states are the same, we have, writing $\hbar q = m_e v$,

$$
\begin{aligned}
P &= \left| \int d^3r\, e^{-\frac{i}{\hbar}\frac{m_e m_N}{m_e+m_N}\boldsymbol{v}\cdot\boldsymbol{r}}\psi_{1s}^{H}(r)\psi_{1s}^{He}(r) \right|^2 = \left| \frac{\sqrt{8}}{\pi} \int d^3r\, e^{-\frac{i}{\hbar}m_e\boldsymbol{v}\cdot\boldsymbol{r}}e^{-3r/a_0} \right|^2 \\
&= \left| \frac{4\sqrt{8}}{q} \int_0^\infty dr\, r \sin(qr) e^{-3r/a_0} \right|^2 = \left| \frac{24\sqrt{8}}{(9+q^2 a_0^2)^2} \right|^2 = \frac{8^3 \cdot 9}{(9+q^2 a_0^2)^4} = \frac{8^3 \cdot 9}{(9+\frac{m_e^2 v^2 a_0^2}{\hbar^2})^4}
\end{aligned}
\tag{18.14}
$$

Chapter 19

Time independent perturbation theory

19.1 Introduction

Often we can break the Hamiltonian of a system into two pieces,

$$H = H_0 + V \tag{19.1}$$

where we know how to solve for the eigenstates and eigenvalues of H_0. H_0 is called the unperturbed Hamiltonian, and V is the perturbation, which must be small in comparison. We then expand in powers of V which gives us perturbation theory and, at the same time, tells us what V small means.

Traditionally, there are two routes to deriving perturbation theory. These are called Brillouin-Wigner and Rayleigh-Schrödinger perturbation theory. In the end they give the same result. Rayleigh-Schrödinger is formulated like classical perturbation theory. The Brillouin-Wigner formalism is faster and more simply related to both time-dependent perturbation theory and scattering theory, so we will begin there.

19.2 Brillouin-Wigner perturbation theory

Our task is to calculate the eigenstates and energies of H in terms of the eigenstates and energies of H_0. We define

$$
\begin{aligned}
H_0|n\rangle &= \epsilon_n|n\rangle \\
H|N\rangle &= E_n|N\rangle
\end{aligned}
\tag{19.2}
$$

where we use lower case letters for the H_0 states, and the corresponding upper case letter for the H state. If we slowly turn on V, we expect that the H_0 states will change into the corresponding H state, so that $|n\rangle \to |N\rangle$, and $\epsilon_n \to E_n$[1].

---

[1]Notice that we use the lower case index on E_n. If this bothers you, then use some other notation for the H eigenstates like $|n^{(H)}\rangle$ instead of $|N\rangle$.

We now expand $|N\rangle$ in the eigenstates of H_0. We expect $|N\rangle$ to be approximately $|n\rangle$, so we write

$$|N\rangle = |n\rangle + \sum_{m\neq n} |m\rangle\langle m|N\rangle . \qquad (19.3)$$

$|N\rangle$ is *not* normalized to 1. You, of course, could normalize it to 1, but taking the normalization $\langle n|N\rangle = 1$ simplifies the expressions. We write the time-independent Schrödinger equation as

$$(H_0 + V)|N\rangle = E_n|N\rangle . \qquad (19.4)$$

Operating on the left with $\langle m|$,

$$\epsilon_m\langle m|N\rangle + \langle m|V|N\rangle = E_n\langle m|N\rangle \qquad (19.5)$$

or

$$(E_n - \epsilon_m)\langle m|N\rangle = \langle m|V|N\rangle \qquad (19.6)$$

Writing this result for $m = n$, and substituting it into 19.3 for $m \neq n$, gives the main equations of Brillouin-Wigner perturbation theory

$$\begin{aligned} E_n &= \epsilon_n + \langle n|V|N\rangle \\ |N\rangle &= |n\rangle + \sum_{m\neq n} \frac{|m\rangle\langle m|V|N\rangle}{E_n - \epsilon_m} . \end{aligned} \qquad (19.7)$$

We can now iterate these equations. That is, we substitute the right hand side of the $|N\rangle$ equation for the last $|N\rangle$ of that equation,

$$|N\rangle = |n\rangle + \sum_{m\neq n} \frac{|m\rangle\langle m|V|n\rangle}{E_n - \epsilon_m} + \sum_{m\neq n, j\neq n} \frac{|m\rangle\langle m|V|j\rangle\langle j|V|N\rangle}{(E_n - \epsilon_m)(E_n - \epsilon_j)} \qquad (19.8)$$

Plugging that result into the energy equation

$$E_n = \epsilon_n + \langle n|V|n\rangle + \sum_{m\neq n} \frac{\langle n|V|m\rangle\langle m|V|n\rangle}{E_n - \epsilon_m} + \sum_{m\neq n, j\neq n} \frac{\langle n|V|m\rangle\langle m|V|j\rangle\langle j|V|N\rangle}{(E_n - \epsilon_m)(E_n - \epsilon_j)} \qquad (19.9)$$

We see the big advantage of Brillouin-Wigner perturbation theory. The equations are very easy to iterate to high order. The disadvantage is that the perturbed energies E_n appear, which we don't know. However, we can substitute the energy equation to evaluate them.

Before going on, we see from these equations that to be able to drop the higher order terms, we need the matrix elements of $\langle m|V|j\rangle$ between the unperturbed states to be much smaller than $E_n - \epsilon_m$.

Since the energies E_n are also a series in powers of V, we can expand our expressions in powers of V. This will give us the Rayleigh Schrödinger result. Let's first calculate $|N\rangle$ to first order in V. The first term that contains an E_n is already first order in V. Therefore, we can substitutine $E_n = \epsilon_n$ in that term since the correction to E_n will give a second order in V term. The perturbed state to first order in V is then

$$|N\rangle = |n\rangle + \sum_{m \neq n} \frac{|m\rangle\langle m|V|n\rangle}{\epsilon_n - \epsilon_m} + ... \tag{19.10}$$

Substituting this into the energy equation gives a second order result for the energy

$$E_n = \epsilon_n + \langle n|V|n\rangle + \sum_{m \neq n} \frac{\langle n|V|m\rangle\langle m|V|n\rangle}{\epsilon_n - \epsilon_m} + ... \tag{19.11}$$

The vast majority of perturbation calculations stop at this level.

We can continue to higher order. Keeping second order in V in the state expansion, we see that in Eq. 19.8 we can replace $|N\rangle$ with $|n\rangle$ and E_n with ϵ_n in the last term. However in the next to last term the numerator has one power of V, so we need to keep the linear term in V in the denominator. We get

$$
\begin{aligned}
|N\rangle &= |n\rangle + \sum_{m \neq n} \frac{|m\rangle\langle m|V|n\rangle}{\epsilon_n - \epsilon_m + \langle n|V|n\rangle} + \sum_{m \neq n, j \neq n} \frac{|m\rangle\langle m|V|j\rangle\langle j|V|n\rangle}{(\epsilon_n - \epsilon_m)(\epsilon_n - \epsilon_j)} + ... \\
&= |n\rangle + \sum_{m \neq n} \frac{|m\rangle\langle m|V|n\rangle}{(\epsilon_n - \epsilon_m)\left(1 + \frac{\langle n|V|n\rangle}{\epsilon_n - \epsilon_m}\right)} + \sum_{m \neq n, j \neq n} \frac{|m\rangle\langle m|V|j\rangle\langle j|V|n\rangle}{(\epsilon_n - \epsilon_m)(\epsilon_n - \epsilon_j)} + ... \\
&= |n\rangle + \sum_{m \neq n} \frac{|m\rangle\langle m|V|n\rangle}{\epsilon_n - \epsilon_m} + \sum_{m \neq n, j \neq n} \frac{|m\rangle\langle m|V|j\rangle\langle j|V|n\rangle}{(\epsilon_n - \epsilon_m)(\epsilon_n - \epsilon_j)} - \sum_{m \neq n} \frac{|m\rangle\langle m|V|n\rangle\langle n|V|n\rangle}{(\epsilon_n - \epsilon_m)^2} + ...
\end{aligned}
\tag{19.12}
$$

The corresponding third order energy is

$$
\begin{aligned}
E_n &= \epsilon_n + \langle n|V|n\rangle + \sum_{m \neq n} \frac{\langle n|V|m\rangle\langle m|V|n\rangle}{\epsilon_n - \epsilon_m} + \sum_{m \neq n, j \neq n} \frac{\langle n|V|m\rangle\langle m|V|j\rangle\langle j|V|n\rangle}{(\epsilon_n - \epsilon_m)(\epsilon_n - \epsilon_j)} \\
&\quad - \sum_{m \neq n} \frac{\langle n|V|m\rangle\langle m|V|n\rangle\langle n|V|n\rangle}{(\epsilon_n - \epsilon_m)^2} + ...
\end{aligned}
\tag{19.13}
$$

19.3 Using perturbation theory

One of the main uses of perturbation theory is to give a quantitative reason to neglect terms that are at an unphysical energy scale. For example, in solving the hydrogen atom, we took

the proton to be a point particle. We know that the proton is a composite of three quarks held together by gluons. Its radius is of the order of Femtometers. The lowest lying nucleon excitation is the delta resonance which is about 200 MeV higher in energy than the proton. For more complicated nuclei, the nuclear radius is still some Femtometers and the nuclear excitations are measured in MeV. The electron charge density at the origin for hydrogen is roughly $\rho \sim -ea_0^{-3}$ so the matrix elements of the perturbation from the finite size of the proton are of order $V \sim e^2/R(R/a_0)^3$ where R is the proton radius. Since e^2/a_0 is 27.2 eV, and $R/a_0 \sim 10^{-15}$, we see that the matrix element over the energy denominator for atomic states is $\sim 10^{-5}$ and for nuclear excitations it is six or more orders of magnitude smaller.

Perturbation theory makes quantitative our usual arguments about separation of energy scales. Often, like in our example, you do not have to make accurate calculations; orders of magnitude are good enough to see which terms can be ignored.

19.4   Ground state Energy

The first order energy is the result we would get if we calculated the expectation value of H using the ground state, $|0\rangle$, of H_0. From the Rayleigh-Ritz variational principle, we know that this is an upper bound to the ground state energy. If V is small enough that the second order energy was accurate, the second order term would necessarily have to be negative. For the ground state, we see that the numerator of the second order term

$$\langle 0|V|n\rangle\langle n|V|0\rangle = |\langle 0|V|n\rangle|^2 \geq 0. \tag{19.14}$$

and the denominator $\epsilon_0 - \epsilon_n$ must be negative since ϵ_0 is the lowest energy eigenvalue of H_0. The second order correction to the ground-state energy is always negative.

19.5   Two state system

Obviously, we can solve a two state system exactly by diagonalizing the 2×2 Hamiltonian. It is instructive to look at the perturbation expansion. The H_0 matrix is

$$\begin{pmatrix} \langle 1|H_0|1\rangle & \langle 1|H_0|2\rangle \\ \langle 2|H_0|1\rangle & \langle 2|H_0|2\rangle \end{pmatrix} = \begin{pmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_2 \end{pmatrix} \tag{19.15}$$

and the perturbation will be a general Hermitian matrix

$$\begin{pmatrix} \langle 1|V|1\rangle & \langle 1|V|2\rangle \\ \langle 2|V|1\rangle & \langle 2|V|2\rangle \end{pmatrix}. \tag{19.16}$$

Let's calculate the energy E_1. The state sums are over $m \neq 1$, but there are only 2 states, so this just means $m = 2$ here. The result is

$$E_1 = \epsilon_1 + \langle 1|V|1\rangle + \frac{\langle 1|V|2\rangle\langle 2|V|1\rangle}{E_1 - \epsilon_2}\left[1 + \frac{\langle 2|V|2\rangle}{E_1 - \epsilon_2} + \left(\frac{\langle 2|V|2\rangle}{E_1 - \epsilon_2}\right)^2 + \left(\frac{\langle 2|V|2\rangle}{E_1 - \epsilon_2}\right)^3 + ...\right].$$

$$\tag{19.17}$$

The series in the brackets is a geometric series $1 + x + x^2 ... = 1/(1 - x)$,

$$
E_1 \;=\; \epsilon_1 + \langle 1|V|1\rangle + \frac{\langle 1|V|2\rangle\langle 2|V|1\rangle}{E_1 - \epsilon_2} \frac{1}{1 - \frac{\langle 2|V|2\rangle}{E_1 - \epsilon_2}} = \epsilon_1 + \langle 1|V|1\rangle + \frac{\langle 1|V|2\rangle\langle 2|V|1\rangle}{E_1 - \epsilon_2 - \langle 2|V|2\rangle}
$$

$$(19.18)$$

Multiplying through by $E_1 - \epsilon_2 - \langle 2|V|2\rangle$ gives the equation

$$
\left(\epsilon_1 + \langle 1|V|1\rangle - E_1\right)\left(\epsilon_2 + \langle 2|V|2\rangle - E_1\right) - \langle 1|V|2\rangle\langle 2|V|1\rangle = 0\,. \tag{19.19}
$$

This is just the equation for the determinant of $H - \lambda I$, where $\lambda = E_1$, which is the exact result. However, to get the exact result, even for the 2 state system, we needed to calculate perturbation theory to all orders.

If we keep just the second order energy, we have

$$
E_1 = \epsilon_1 + \langle 1|V|1\rangle + \frac{|\langle 1|V|2\rangle|^2}{\epsilon_1 - \epsilon_2}\,. \tag{19.20}
$$

This is the result we would get by calculating the determinant and keeping terms with diagonal and just the pair of off-diagonal elements. The generalization to larger systems is similar. The second-order energy is the term in the determinant from taking all off diagonal elements that couple to the desired state.

19.6 Static electric polarizability of hydrogen

The static dielectric constant of a gas can be calculated from the average static polarizability, γ, of the molecules using the Clausius-Mossotti equation (or the equivalent Lorentz-Lorenz equation).

For low densities where the induced dipoles of the other molecules in the gas do not screen the applied field $\boldsymbol{\mathcal{E}}$, each molecule gets an induced dipole $\gamma\boldsymbol{\mathcal{E}}$. The dipole moment per unit volume is the polarizability $\boldsymbol{P}$, so that $\epsilon = 1 + 4\pi n\gamma$, where n is the gas density.

It is difficult to make a gas of atomic hydrogen since the atoms will combine to become H_2 molecules. Spin polarized atoms do not form molecules, so in principle, a gas of spin polarized hydrogen atoms can be made if spin flips and recombination from interaction with the walls can be limited. In any case, since we know the unperturbed states of hydrogen, it is a natural system for us to try perturbation theory.

A constant in time and space electric field $\boldsymbol{\mathcal{E}}$ can be described by the electric potential $\Phi = -\boldsymbol{r}\cdot\boldsymbol{\mathcal{E}}$. The hydrogen atom Hamiltonian becomes

$$
\frac{p_p^2}{2m_p} + \frac{p_e^2}{2m_e} - \frac{e^2}{|\boldsymbol{r}_e - \boldsymbol{r}_p|} + e(\boldsymbol{r}_e - \boldsymbol{r}_p)\cdot\boldsymbol{\mathcal{E}}\,. \tag{19.21}
$$

Converting to relative and center of mass coordinates, the Hamiltonian becomes

$$
\frac{P^2}{2M} + \frac{p}{2m} - \frac{e^2}{r} + e\boldsymbol{r}\cdot\boldsymbol{\mathcal{E}}\,. \tag{19.22}
$$

Not surprisingly, the perturbation acts only on the relative Hamiltonian.

We want to apply the electric field, and calculate the induced dipole moment. The dipole operator for a group of particles is the sum of their charges times positions, $\sum_i q_i \boldsymbol{r}_i$, so for the hydrogen atom

$$\boldsymbol{d} = -e(\boldsymbol{r}_e - \boldsymbol{r}_p) = -e\boldsymbol{r} \,. \tag{19.23}$$

Since the center of mass state is unchanged, it factors out of the calculation, and we can deal with just the relative coordinate.

The expectation value of the induced dipole in the ground state is

$$\frac{\langle \Psi_0(\boldsymbol{\mathcal{E}}) | \boldsymbol{d} | \Psi_0(\boldsymbol{\mathcal{E}}) \rangle}{\langle \Psi_0(\boldsymbol{\mathcal{E}}) | \Psi_0(\boldsymbol{\mathcal{E}}) \rangle} \tag{19.24}$$

where $|\Psi_0(\boldsymbol{\mathcal{E}})\rangle$ is the perturbed ground state in the presence of the electric field. We are interested here in the small field limit when the induced dipole moment is proportional to the field. This tells us that we only need to keep linear order terms in the field to get the linear term in the dipole moment. The perturbation expansion is

$$|\Psi_0(\boldsymbol{\mathcal{E}})\rangle = |0\rangle + \sum_{m \neq 0} \frac{|m\rangle \langle m | e\boldsymbol{r} \cdot \boldsymbol{\mathcal{E}} | 0 \rangle}{\epsilon_0 - \epsilon_m} + \ldots \tag{19.25}$$

The normalization is $1 + O(\mathcal{E}^2)$, so keeping just linear order in the field, this is 1. The ground state expectation value of $\boldsymbol{r}$ is zero, so the result for the induced dipole

$$\langle \boldsymbol{d} \rangle = \sum_{m \neq 0} \frac{\langle 0 | e\boldsymbol{r} \cdot \boldsymbol{\mathcal{E}} | m \rangle \langle m | e\boldsymbol{r} | 0 \rangle + \langle 0 | e\boldsymbol{r} | m \rangle \langle m | e\boldsymbol{r} \cdot \boldsymbol{\mathcal{E}} | 0 \rangle}{\epsilon_m - \epsilon_0} \tag{19.26}$$

We can pick our coordinate system so that the electric field points along z. The zero field ground state wave function is rotationally invariant, so in the r basis, the wave function will be a radial function. Multiplying by $e\boldsymbol{r} \cdot \boldsymbol{\mathcal{E}} = \mathcal{E}z = \mathcal{E}r\cos\theta$, we see that the resulting state has a $\cos\theta$ or z angular dependence. Since this is an $\ell = 1$, $m = 0$, or p_z state, only those states will give non zero matrix element. Since the intermediate state must be a p_z state, only the z component of the dipole operator will give a non zero matrix element. The resulting induced dipole is therefore along the field, and given by

$$\langle \boldsymbol{d} \rangle = 2e^2 \mathcal{E} \hat{\boldsymbol{z}} \sum_m \frac{|\langle 0 | z | mp_z \rangle|^2}{\epsilon_m - \epsilon_0} \,. \tag{19.27}$$

The polarizability is then

$$\gamma = 2e^2 \sum_m \frac{|\langle 0 | z | mp_z \rangle|^2}{\epsilon_m - \epsilon_0} \,. \tag{19.28}$$

The state sum must include both bound and unbound states.

Before evaluating this expression, we will rederive it from the energy expression.

19.7 Hellman-Feynman theorem and responses from energy derivatives

In classical mechanics and statistical mechanics, we often calculate quantities by relating them to derivatives of the energy. Since these other fields are a consequence of quantum mechanics, we might expect similar methods to work here too.

We begin with a Hamiltonian that is a function of a parameter $H(\lambda)$. For the polarizability, the parameter will be the magnitude of the electric field. The eigenstates of $H(\lambda)$ are $|\Psi_n(\lambda)\rangle$, with energy $E_n(\lambda)$. We can write

$$E_n(\lambda) = \frac{\langle \Psi_n(\lambda)|H(\lambda)|\Psi_n(\lambda)\rangle}{\langle \Psi_n(\lambda)|\Psi_n(\lambda)\rangle} \tag{19.29}$$

Taking the derivative with respect to λ, and denoting it by a prime

$$
\begin{aligned}
E_n'(\lambda) &= \frac{\langle \Psi_n(\lambda)|H'(\lambda)|\Psi_n(\lambda)\rangle}{\langle \Psi_n(\lambda)|\Psi_n(\lambda)\rangle} + \frac{\langle \Psi_n'(\lambda)|H(\lambda)|\Psi_n(\lambda)\rangle}{\langle \Psi_n(\lambda)|\Psi_n(\lambda)\rangle} + \frac{\langle \Psi_n(\lambda)|H(\lambda)|\Psi_n'(\lambda)\rangle}{\langle \Psi_n(\lambda)|\Psi_n(\lambda)\rangle} \\
&\quad - \frac{\langle \Psi_n(\lambda)|H(\lambda)|\Psi_n(\lambda)\rangle}{\langle \Psi_n(\lambda)|\Psi_n(\lambda)\rangle} \left[\frac{\langle \Psi_n'(\lambda)|\Psi_n(\lambda)\rangle}{\langle \Psi_n(\lambda)|\Psi_n(\lambda)\rangle} + \frac{\langle \Psi_n(\lambda)|\Psi_n'(\lambda)\rangle}{\langle \Psi_n(\lambda)|\Psi_n(\lambda)\rangle} \right] \\
&= \frac{\langle \Psi_n(\lambda)|H'(\lambda)|\Psi_n(\lambda)\rangle}{\langle \Psi_n(\lambda)|\Psi_n(\lambda)\rangle} \tag{19.30}
\end{aligned}
$$

where by operating with the Hamiltonian either to the left or right on the state without a derivative to give $E_n(\lambda)$ times that state, all of the derivative with respect to the wave function terms cancel. This is the Hellman-Feynman theorem. The result should not be a surprise. We already knew that the first order perturbation of the energy is given by the expectation value of the perturbation in the unperturbed state. If we have $H_0 = H(\lambda)$, and $V = H(\lambda + \Delta\lambda) - H(\lambda)$, for $\Delta\lambda \to 0$, $V = H'(\lambda)\Delta\lambda + ...$, and perturbation theory tells us that for normalized $|\Psi_n(\lambda)\rangle$

$$E_n(\lambda + \Delta\lambda) = E_n(\lambda) + \langle \Psi_n(\lambda)|H'(\lambda)|\Psi_n(\lambda)\rangle \Delta\lambda + ... \tag{19.31}$$

rearranging,

$$E_n'(\lambda) = \lim_{\Delta\lambda \to 0} \frac{E_n(\lambda + \Delta\lambda) - E_n(\lambda)}{\Delta\lambda} = \langle \Psi_n(\lambda)|H'(\lambda)|\Psi_n(\lambda)\rangle . \tag{19.32}$$

The Hamiltonian for the hydrogen atom in an electric field along z is $H(\mathcal{E}) = H(0) + ez\mathcal{E}$ and $H'(\mathcal{E}) = ez = -d_z$ where d_z is the z component of the dipole operator. The ground-state polarizability is then

$$
\begin{aligned}
\gamma &= \frac{\partial \langle d_z \rangle}{\partial \mathcal{E}} = -\frac{\partial^2 E_0(\mathcal{E})}{\partial \mathcal{E}^2}\bigg|_{\mathcal{E}=0} = -\frac{\partial^2}{\partial \mathcal{E}^2} \sum_m \frac{|\langle m|e\mathcal{E}z|0\rangle|^2}{\epsilon_0 - \epsilon_m} \\
&= 2e^2 \sum_m \frac{|\langle 0|z|mp_z\rangle|^2}{\epsilon_m - \epsilon_0}, \tag{19.33}
\end{aligned}
$$

as before.

19.8   Bounds on the polarizability

Each of the terms in the perturbation sum for the ground-state polarizability is positive. Therefore any partial sum will give a lower bound on the polarizability. Polarizability has units of volume; the dipole moment has units of charge times distance, and the electric field has units of charge divided by distance squared, so their ratio has units of distance cubed. We can work in atomic units and multiply our answer by a_0^3 at the end.

The first term in the sum comes from the matrix element

$$\langle 2p_z|z|1s\rangle = \int d^3r\, \psi_{1s}(\boldsymbol{r})z\psi_{2p_z}(\boldsymbol{r}) \tag{19.34}$$

where in atomic units

$$\psi_{1s}(\boldsymbol{r}) = \frac{1}{\sqrt{\pi}}e^{-r}$$

$$\psi_{2p_z}(\boldsymbol{r}) = \frac{1}{\sqrt{32\pi}}ze^{-r/2}. \tag{19.35}$$

The angular integral over z^2 gives $4\pi r^2/3$, so that

$$\langle 2p_z|z|1s\rangle = \frac{1}{3\sqrt{2}}\int_0^\infty r^4 e^{-3r/2} = \frac{1}{3\sqrt{2}}\left(\frac{2}{3}\right)^5 \underbrace{\int_0^\infty x^4 e^{-x}}_{4!} = \frac{2^8}{\sqrt{2}3^5}. \tag{19.36}$$

with $\epsilon_{2p} = -0.125$, and $\epsilon_{1s} = -0.5$, we have

$$2e^2\frac{|\langle 0|z|2p_z\rangle|^2}{\epsilon_{2p}-\epsilon_{1s}} = 2.9596...a_0^3 \tag{19.37}$$

We can obtain an upper bound by using the completeness relation. Since the z operator only couples the p_z states, we can change the sum to a sum over all states without changing the result. If the energy denominator were constant, we could then use the completeness relation to perform the sum. We rewrite the polarizability as

$$\begin{aligned}
\gamma &= \frac{|\langle 2p_z|z|1s\rangle|^2}{\epsilon_{2p}-\epsilon_{1s}} + \sum_{m\neq 2p}\frac{|\langle m|z|1s\rangle|^2}{\epsilon_m-\epsilon_{1s}} \\
&\leq \frac{|\langle 2p_z|z|1s\rangle|^2}{\epsilon_{2p}-\epsilon_{1s}} + \sum_{m\neq 2p}\frac{|\langle m|z|1s\rangle|^2}{\epsilon_{3p}-\epsilon_{1s}},
\end{aligned} \tag{19.38}$$

where by replacing the energy in the denominator with the $3p$ energy, we overestimate the contributions from the $4p$, $5p$, etc. states. Including the $2p$ contribution in the sum and subtracting it off,

$$\begin{aligned}
\gamma &\leq \frac{|\langle 2p_z|z|1s\rangle|^2}{\epsilon_{2p}-\epsilon_{1s}} + \sum_m\frac{|\langle m|z|1s\rangle|^2}{\epsilon_{3p}-\epsilon_{1s}} - \frac{|\langle 2p_z|z|1s\rangle|^2}{\epsilon_{3p}-\epsilon_{1s}} \\
\gamma &\leq \frac{|\langle 2p_z|z|1s\rangle|^2}{\epsilon_{2p}-\epsilon_{1s}} - \frac{|\langle 2p_z|z|1s\rangle|^2}{\epsilon_{3p}-\epsilon_{1s}} + \sum_m\frac{\langle 1s|z^2|1s\rangle}{\epsilon_{3p}-\epsilon_{1s}}.
\end{aligned} \tag{19.39}$$

The ground state expectation of z^2 in atomic units is

$$\langle 1s|z^2|1s\rangle = \frac{4}{3}\int_0^\infty dr r^2 e^{-2r} = \frac{1}{6}\underbrace{\int_0^\infty dx x^3 e^{-x}}_{3!} = 1 \tag{19.40}$$

and our bounds are

$$2.9596...a_0^3 \le \gamma \le 4.962...a_0^3. \tag{19.41}$$

What happens when we add more states? If you add to these sums *all* of the bound states of hydrogen, the bounds improve, but are still not very good. They become

$$3.663...a_0^3 \le \gamma \le 4.797...a_0^3. \tag{19.42}$$

Something is odd. Why do we need to include the continuum states to describe the ground-state polarizability of the hydrogen atom? The answer is clear if you imagine what the solution for a hydrogen atom in a weak electric field should look like. You would expect the electron cloud around the nucleus to shift slightly. But look at the change in the wave function when we include the $2p_z$ term in the expansion. It gives a contribution that is proportional to $e^{-r/2a_0}$. That is, the wave function contribution extends twice as far from the nucleus as the original $1s$ state. If we try to correct this with the $3p_z$ state, it extends three times as far. Each correction term gives a contribution to the polarizability, but adds these long range pieces that the next term in the expansion must try to remove. This continues even into the continuum states. A different expansion is needed for rapid convergence.

19.9 Exact calculation of the hydrogen polarizability

The hydrogen atom Hamiltonian can be separated in parabolic coordinates, and the properties of the ground state in a static electric field solved analytically. An alternative is to try to write a perturbation expansion of the Schrödinger equation without using the unperturbed basis and uses our physical understanding that the field does not grossly change the electronic density. From our perturbation expansion, we know that the first order ground state, in addition to the zeroth order ground state contains linear combination of states with angular momentum $\ell = 1$ and for a field along z, $m = 0$. This tells us that the wave function must look like (in atomic units)

$$\psi(\boldsymbol{r}) = \frac{1}{\sqrt{\pi}}e^{-r}\left[1 + \mathcal{E}\cos\theta f(r) + ...\right] \tag{19.43}$$

where we factor out the exponential from the ground state. We also know from our perturbation expansion that the energy change is second order in the field. Keeping just linear terms in the Schrödinger equation

$$\left[-\frac{\nabla^2}{2} - \frac{1}{r} + \mathcal{E}r\cos\theta\right]\psi(\boldsymbol{r}) = -\frac{1}{2}\psi(\boldsymbol{r}) \tag{19.44}$$

and equating the linear term in $\mathcal{E}$ on both sides, we have

$$-\frac{1}{2}f''(r) + f'(r)(1 - r^{-1}) + f(r)r^{-2} + r = 0 \,. \tag{19.45}$$

We can try a power series. Since I know the result, I will just try $f = ar + br^2$. Substituting, we find a solution $b = -1/2$, $a = 2b$. Therefore, the exact ground-state wave function for the hydrogen atom in an electric field to linear order in the field is

$$\psi(\boldsymbol{r}) = e^{-r}\left(1 - \mathcal{E}r\left[1 + \frac{r}{2}\right]\cos\theta\right) + O(\mathcal{E}^2) \tag{19.46}$$

The dipole moment expection value gives $\gamma = 4.5a_0^3$.

Notice that the form of the wave function shows why our expansion in the usual p states was slowly convergent.

19.10   van der Waals interaction example

Neutral molecules in their ground states which have no permanent dipole moment attract each other through a van der Waals interaction. As we will see, this interaction goes like R^{-6} where R is the distance between the molecules.

Before calculating the interaction, lets make a hand waving argument. We begin by looking at two hydrogen atoms separated by a distance R much larger than the size of the atoms, i.e. $R \gg a_0$. If the atoms are in their ground states initially, they would have spherically symmetric electronics charge distributions, and the interaction energy would be zero. Now what happens if one of the atoms has a fluctuation (e.g. its electron spends a little more time on one side of its proton than the other)? The electronic density fluctuation would give the atom a dipole moment, $\boldsymbol{d}_1$. This produces a dipole field that is proportional to d_1/R^3 at the other atom. We have already seen that when we apply an electric field to an atom it gets a dipole moment $\boldsymbol{d}_2 = \gamma\boldsymbol{E}$, and from electrostatics, the energy of this dipole in an electric field is

$$W = -\boldsymbol{d}_2 \cdot \boldsymbol{E} = -\gamma E^2 \propto -\frac{1}{R^6} \,. \tag{19.47}$$

Now let's see how this comes out of the mathematics. We want to place two hydrogen atoms a distance R apart. For convenience, lets take the coordinate system so that the hydrogen atoms lie along the z axis. Since the hydrogen atoms are separated by much more than their size, we can identify each electron with its proton. That is, we do not have to worry about the fact that the two electrons are indistinguishable. Here we can distinguish them by saying that electron 1 forms a hydrogen atom with proton 1, and electron 2 forms a hydrogen atom with proton 2. When we say the hydrogen atoms are a distance R apart, we really mean their center's of mass are a distance R apart. Here we will make a further approximation, that since the proton's mass is much greater than the electron mass, we can take the proton mass to infinity. The protons will then be fixed in space a distance R apart.

It is convenient to choose the origin of the coordinate systems for each electron to have its origin around the corresponding proton. We write the Hamiltonian as

$$
\begin{aligned}
H &= H_0 + V \\
H_0 &= \frac{p_1^2}{2m} - \frac{e^2}{r_1} + \frac{p_2^2}{2m} - \frac{e^2}{r_2} \\
V &= \frac{e^2}{R} - \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{R}|} - \frac{e^2}{|\boldsymbol{r}_2 + \boldsymbol{R}|} + \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2 - \boldsymbol{R}|}
\end{aligned}
\tag{19.48}
$$

where the four terms of V are just the coulomb interactions between proton 1 and proton 2; electron 1 and proton 2; electron 2 and proton 1; and electron 1 with electron 2. H_0 is the sum of two noninteracting hydrogen atoms, so the eigenfunctions will be

$$
\begin{aligned}
&|n_1\ell_1 m_1; n_2\ell_2 m_2\rangle \\
\Psi_{n_1\ell_1 m_1; n_2\ell_2 m_2}(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \langle \boldsymbol{r}_1 \boldsymbol{r}_2 | n_1\ell_1 m_1; n_2\ell_2 m_2\rangle = \langle \boldsymbol{r}_1 | n_1\ell_1 m_1\rangle\langle \boldsymbol{r}_2 | n_2\ell_2 m_2\rangle \\
&= R_{n_1\ell_1}(r_1)Y_{\ell_1 m_1}(\theta_1, \phi_1)R_{n_2\ell_2}(r_2)Y_{\ell_2 m_2}(\theta_2, \phi_2)
\end{aligned}
\tag{19.49}
$$

with energies $\epsilon_{n_1} + \epsilon_{n_2}$ the sum of the two separate hydrogen atom eigenstate energies.

We now can apply perturbation theory to find the change in energy of any of the states of H_0 when V is included. Our hand waving calculation indicated that since R is large, we probably want to make an expansion of V in powers of R^{-1}. This will make the matrix elements easier to calculate and it will illuminate the physics, showing why we can still think of the van der Waals interaction as a dipole-dipole interaction.

You can expand V using any of the methods that you learned in classical electromagnetism. However, since we just want to get the lowest order term or terms, we can just use cartesian coordinates and Taylor series expand. We can write

$$
V = \Phi(0,0) - \Phi(\boldsymbol{r}_1, 0) - \Phi(0, \boldsymbol{r}_2) + \Phi(\boldsymbol{r}_1, \boldsymbol{r}_2)
\tag{19.50}
$$

with

$$
\begin{aligned}
\Phi(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2 - R\hat{\boldsymbol{z}}|} = \frac{1}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2 - R)^2}} \\
&= R^{-1}\left(1 + R^{-2}\left[(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - 2R(z_1 - z_2)\right]\right)^{-1/2}
\end{aligned}
\tag{19.51}
$$

We use the Taylor series

$$
(1 + x)^{-1/2} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 + \dots
\tag{19.52}
$$

to write

$$
\begin{aligned}
\Phi(\boldsymbol{r}_1, \boldsymbol{r}_2) &= R^{-1} + (z_1 - z_2)R^{-2} \\
&\quad + R^{-3}\left(-\frac{1}{2}\left[(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2\right] + \frac{3}{8}[2(z_1 - z_2)]^2\right) + \dots \\
&= R^{-1} + (z_1 - z_2)R^{-2} + \frac{1}{2}R^{-3}[2(z_1 - z_2)^2 - (x_1 - x_2)^2 - (y_1 - y_2)^2] + \dots
\end{aligned}
\tag{19.53}
$$

Substituting into Eq. 19.50, the first three terms cancel contributions that do not contain both $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ components. Therefore only the cross terms survive,

$$V = -\frac{e^2}{R^3} \left(2z_1z_2 - x_1x_2 - y_1y_2\right) + \dots . \tag{19.54}$$

Recall that the dipole moment operator for a single hydrogen atom was $\boldsymbol{d} = -e\boldsymbol{r}$, and here $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are the relative coordinates for our two hydrogen atoms. Since we chose our coordinates so that $\hat{\boldsymbol{R}}$ is along z, the lowest order surviving term in V has the form

$$-\frac{e^2}{R^3} \left(2z_1z_2 - x_1x_2 - y_1y_2\right) = -\frac{1}{R^3} \left(3\boldsymbol{d}_1 \cdot \hat{\boldsymbol{R}} \boldsymbol{d}_2 \cdot \hat{\boldsymbol{R}} - \boldsymbol{d}_1 \cdot \boldsymbol{d}_2\right) \tag{19.55}$$

which is the standard dipole-dipole interaction term with the dipole moments replaced with dipole operators.

We want to calculate the change of the ground-state energy with V. We can see that our dipole-dipole term does not contribute to a change of energy at lowest order. The second order term will give us a R^{-6} term as we expected from our hand-waving argument. To be complete, we should check that the higher terms in the expansion of V also do not contribute at first order. This is easy, since the charge distribution from the $1s$ electronic state is spherically symmetric and we assumed the charge distributions do not overlap. Therefore, the full V will have zero expectation for the unperturbed ground state, and we do not have to consider the first order term at all.[2]

Let's look at the matrix elements that will contribute at second order. Our initial state is $|1s1s\rangle$, and we will multiply this by $2z_1z_2$, x_1x_2 or y_1y_2. Since both atomic wave functions are multiplied by x, y, or z, they will be in $\ell = 1$ states. You can of course use the good ℓ, m states as we wrote above, but this will be substantially uglier; instead it is simpler to use the p_x, p_y, and p_z combinations. Multiplying by z_1z_2 will require the intermediate state to be a p_zp_z combination, the only part of the potential that has a matrix element between these states and the ground state is the z_1z_2 piece. Similarly for the x_1x_2 and y_1y_2 parts. Furthermore, since we are using the p_x, p_y, p_z states, the matrix elements are equal

$$\langle np_z|z|1s\rangle = \langle np_y|y|1s\rangle = \langle np_x|x|1s\rangle . \tag{19.56}$$

Therefore we can write

$$\begin{aligned}
\Delta E &= \sum_{I \neq 1s^2} \frac{\langle 1s1s|V|I\rangle \langle I|V|1s1s\rangle}{2\epsilon_{1s} - \epsilon_I} \\
&= -\frac{6e^2}{R^6} \sum_{n,m} \frac{|\langle np_z mp_z|z_1z_2|1s1s\rangle|^2}{\epsilon_{np} + \epsilon_{mp} - 2\epsilon_{1s}} \\
&= -\frac{6e^2}{R^6} \sum_{n,m} \frac{|\langle np_z|z|1s\rangle|^2 |\langle mp_z|z|1s\rangle|^2}{\epsilon_{np} + \epsilon_{mp} - 2\epsilon_{1s}}
\end{aligned} \tag{19.57}$$

---

[2]The first order term calculated for arbitrary R (assuming that the electrons have opposite spins so that we do not have to antisymmetrize their spatial wave functions) is straightforward to calculate. The result in given in appendix B.

A variational calculation gives a strength of 6.499 atomic units[3]

$$V(R) = -\frac{\hbar^{10}}{m^5 e^8} \frac{6.499}{R^6} \tag{19.58}$$

I confined an electron to a radius of 10 a.u. around the proton, and using lowest 10 radial solutions for the $\ell = 1$ states, I get a numerical value of 6.497 a.u.

19.11 Degenerate perturbation theory

Degeneracies, particularly with systems with symmmetries, are fairly common. Standard perturbation theory breaks down when applied to degenerate states. You can see that from Eq. 19.6 or Eq. 19.10 where the energy denominators can become small or diverge if the state $|m\rangle$ has an energy equal to or close to the energy of $|n\rangle$.

To use perturbation theory for these degenerate states, we must modify the procedure to eliminate the terms with small or zero energy denominators.

It is helpful to make a thought experiment (or thought calculation in this case). Imagine that we simply wrote out the whole Hamiltonian H in the basis of the eigenstates of H_0 and told our computer to diagonalize that matrix. To simplify the notation we write the set of degenerate states whose perturbation we wish to calculate as $\{n_i\}$, so summing over those states will be written as $\sum_{j=\{n_i\}}$, and summing over all other states will be written as $\sum_{j\neq\{n_i\}}$. The eigenstates of H corresponding to the degenerate states of H_0 would then be

$$|N_i\rangle = \sum_{j=\{n_i\}} |j\rangle\langle j|N_i\rangle + \sum_{m\neq\{n_i\}} |m\rangle\langle m|N_i\rangle. \tag{19.59}$$

Notice that if could somehow manage to find the coefficients $\langle j|N_i\rangle$ for the degenerate states, we could equally well use that combination and any other orthogonal combinations for the remaining degenerate eigenstates of H_0. Let's define those combinations as

$$|\psi_i^{(0)}\rangle = \sum_{j=\{n_i\}} |j\rangle\langle j|N_i\rangle \tag{19.60}$$

and the orthogonal combinations as $|\psi_j^{(0)}\rangle$ with $j \neq i$.

Our equation becomes

$$|N_i\rangle = |\psi_i^{(0)}\rangle + \sum_{m\neq\{n_i\}} |m\rangle\langle m|N_i\rangle. \tag{19.61}$$

Operating with $\langle m|H$ on the left, where $m \neq \{n_i\}$, just as before, we get the equation

$$\langle m|N_i\rangle = \frac{\langle m|V|N_i\rangle}{E_i - \epsilon_m}. \tag{19.62}$$

---

[3]L. Pauling and J.Y. Beach, *The van der Waals interaction of hydrogen atoms*, Phys. Rev. **47**, 686 (1935)

We get a well behaved perturbation expansion

$$|N_i\rangle \;=\; |\psi_i^{(0)}\rangle + \sum_{m \neq \{n_i\}} \frac{|m\rangle\langle m|V|N_i\rangle}{E_i - \epsilon_m} \,. \tag{19.63}$$

We just need to figure out how to calculate the $|\psi_i^{(0)}\rangle$ in perturbation theory.

Operating with $\langle\psi_j^{(0)}|H$, on Eq. 19.61 we get

$$E_i \delta_{ij} = \langle\psi_j^{(0)}|H_0 + V|N_i\rangle \,. \tag{19.64}$$

Notice that this says that the $|\psi_i^{(0)}\rangle$ give a matrix $M_{ji} = \langle\psi_j^{(0)}|H_0 + V|N_i\rangle$ that must have zero off diagonal elements. That is, it must be diagonal.

Substituting our expression, Eq. 19.63, for $|N_i\rangle$ and iterating, we get

$$
\begin{aligned}
M_{ji} &= \langle\psi_j^{(0)}|H_0 + V|N_i\rangle \\
&= \langle\psi_j^{(0)}| \left[H_0 + V + \sum_{m \neq \{n_i\}} \frac{V|m\rangle\langle m|V}{E_i - \epsilon_m} + \sum_{m,k \neq \{n_i\}} \frac{V|m\rangle\langle m|V|k\rangle\langle k|V}{(E_i - \epsilon_m)(E_i - \epsilon_k)} + \ldots \right] |\psi_i^{(0)}\rangle \\
&\equiv \langle\psi_j^{(0)}|M|\psi_i^{(0)}\rangle
\end{aligned}
\tag{19.65}
$$

where the last line defines the operator M.

To find the $|\psi_i^{(0)}\rangle$, we therefore need to find the eigenstates of the operator M.

In the usual case, we keep only the lowest order term that couples the degenerate states. If that is the first or second order term, we can substitute ϵ_i for E_i in the expansion. Since this is vast majority of cases, let's analyze it first, and then come back to the more general case where $E_i \neq \epsilon_i$ in M.

We write the matrix $\langle j|M|m\rangle$ where j and m are states from the original H_0 degnerate set. We diagonalize this matrix, and its eigenvectors are the coefficients $\langle j|\psi_i^{(0)}\rangle$. The eigenvalue corresponding to $|\psi_i^{(0)}\rangle$ is $\langle\psi_i^{(0)}|H_0 + V|N_i\rangle$, so the eigenvalues give the perturbed energy for that degenerate state. Since the unperturbed energies are all the same, we obtain all of the perturbed energy levels and the states $|\psi_i^{(0)}\rangle$ from this diagonalization. We can then use these $|\psi_i^{(0)}\rangle$ in Eq. 19.63 to get the eigenstate expansion.

For the far less common case, where higher order terms are kept, since E_i appears in the operator M, we must either substitute or iterate the perturbed energy to the correct order for E_i. One way would be to begin with E_i set to the unperturbed energy. We then calculate the matrix elements of M with this approximate value of E_i, and diagonalize M. We then pick the state we are interested in, and substitute its current eigenvalue for E_i, recalculate the matrix elements of M, and rediagonalize. We repeat this until we have either convergence or have the correct result to the order in V we require. Notice that we must repeat this whole diagonalization and iteration calculation, for each of the different E_i that appear, to get all of the degenerate states.

To truncate degenerate perturbation theory at a particular order for the energy, we simply truncate the expression for M to that order. So for first order degenerate perturbation theory,

$M = H_0 + V$. As mentioned above, typically, degenerate perturbation theory is truncated at the lowest nonvanishing order. So if the first order term is nonzero, we usually stop there. If it is zero, then higher order degenerate perturbation theory must be used. This corresponds to taking more terms in the expression for M.

19.12   n=2 Stark effect in hydrogen

Baym looks at the Stark effect (i.e. the perturbation caused by a constant electric field) for hydrogen atoms in the $n = 2$ states. He calculates using the $m = \pm 1$, $m = 0$ $\ell = 1$ states, and the $2s$ state. We can equivalently use the $2p_x$, $2p_y$, and $2p_z$ states. Taking the coordinates so that the electric field is along z, the Hamiltonian is given by Eq. 19.22. Since H_0 is a constant in this basis, for the lowest order contribution, we can diagonalize $ez\mathcal{E}$ in the basis of these four states. Since the perturbation contains a z, we must have an additional z factor in the matrix elements in order to get a nonzero result. Therefore, the only two states coupled by this perturbation are $|2p_z\rangle$ and $|2s\rangle$. The perturbation matrix is

$$e\mathcal{E} \begin{pmatrix} \langle 2s|z|2s\rangle & \langle 2s|z|2p_z\rangle \\ \langle 2p_z|z|2s\rangle & \langle 2p_z|z|2p_z\rangle \end{pmatrix} = e\mathcal{E} \begin{pmatrix} 0 & \langle 2s|z|2p_z\rangle \\ \langle 2p_z|z|2s\rangle & 0 \end{pmatrix}. \tag{19.66}$$

The nonzero matrix element using our standard phase convention is[4]

$$\langle 2s|z|2p_z\rangle = -3a_0 \tag{19.68}$$

The perturbed eigenstates are $2^{-1/2}[|2s\rangle \pm |2p_z\rangle]$ with energies $-\frac{e^2}{a_0}\frac{1}{8} \mp 3a_0\mathcal{E}$. The $|2p_x\rangle$ and $|2p_y\rangle$ remain unperturbed with energy $-\frac{e^2}{a_0}\frac{1}{8}$.

19.13   Nearly degenerate perturbation theory

We may find that we have a set of states with energies that are nearly equal, but distinct, so that low order perturbation theory either does not converge or converges slowly. Perturbation theory tells us what "nearly degenerate" means. It means that the matrix elements of the potential V are not small compared to the energy difference.

Having dealt with degenerate states, the generalization to nearly degenerate states is straightforward. In fact all of the equations go through, and the nearly degenerate case is treated identically to the way we treated the degenerate case.

---

[4]In atomic units

$$\begin{aligned} \psi_{2s}(\boldsymbol{r}) &= \frac{1}{\sqrt{32\pi}}(2 - r)e^{-r/2}, \ \psi_{2p_z}(\boldsymbol{r}) = \frac{1}{\sqrt{32\pi}}ze^{-r/2} \\ \langle 2p_z|z|2s\rangle &= \frac{1}{32\pi}\int d^3r z^2(2 - r)e^{-r} = \frac{1}{24}\underbrace{\int_0^\infty r^4(2 - r)e^{-r}}_{(2\cdot 4! - 5!) = -72} = -3. \end{aligned} \tag{19.67}$$

The matrix element has units of length, so converting back to physical units gives $-3a_0$.

19.14 Integrating out degrees of freedom

Sometimes you will hear people talk about "integrating out" degrees of freedom, and thereby replacing the original Hamiltonian with a different, effective, Hamiltonian. Nearly degenerate perturbation theory is one way to do this. Summing over the higher energy states gives the effective Hamiltonian M that acts just on the set of low energy states that are the nearly degenerate terms. If high order is required, we can run into the $E_i \neq \epsilon_i$ complication noted above; this effect is sometimes known as "energy dependent interactions." Since these energy dependent effective Hamiltonians are more difficult to use, various approximations have been developed to obtain an energy independent Hamiltonian. One example goes by the name Lee-Suzuki method[5], but is beyond the scope of this lecture.

19.15 Tight binding example

The tight-binding approximation comes about in condensed matter by thinking of assembling a solid by bringing atoms in from a large distance apart.

Here we can imagine assembling hydrogen atoms evenly spaced along a line as a prelude to building a solid. We are first interested in the noninteracting electronic states; we can imagine including the electron-electron interaction later as a perturbation. Since the states of a single hydrogen atom form a complete set, we could use the states on any one atom as the basis for a one electron state, however a linear combination of the states centered on one atom will *not* converge quickly to the states on another atom. Instead, if we imagine the protons as being far apart, the ground state of a single electron would be very close to the electron being in the ground state around one of the protons. Since each of the protons would give a ground state with the same energy, for large separations, the ground state would have a degeneracy given by the number of protons. One problem is that when we bring the protons closer together, these $1s$ states around each proton overlap, and they do not form an orthogonal set. We can use Gram-Schmidt or other orthogonalization methods to form an orthogonal set of (nearly) degenerate orbitals.

One choice which is like the Wannier method, is to assume the overlaps are small and only the nearest neighbor overlap is important. We can obtain a set of approximately orthogonal orbitals by writing our original orbitals as

$$\langle r|I\rangle = \phi_I(\boldsymbol{r}) = \phi(\boldsymbol{r} - \boldsymbol{R}_I) \tag{19.69}$$

where $\boldsymbol{r}$ is the electron coordinate and R_I are the nuclear positions. For simplicity I will assume that ϕ is real. The nearest neighbor overlap will be

$$\langle I|I\pm 1\rangle = \int d^3r \langle I|\boldsymbol{r}\rangle\langle \boldsymbol{r}|I\pm 1\rangle = \int d^3r \phi^*(\boldsymbol{r} - R_i)\phi(\boldsymbol{r} - R_{I\pm 1}) \equiv \alpha, \tag{19.70}$$

---

[5]For a discussion see, e.g. B.K. Jennings, *Projection operator formalisms and the nuclear shell mode*, Europhys. Lett. **72**, 211 (2005).

where we assume α is small. The approximately orthogonal combinations are

$$|i\rangle = |I\rangle - \frac{\alpha}{2}(|I+1\rangle + |I-1\rangle) . \tag{19.71}$$

This will give us an approximate Hamiltonian of the form

$$\langle i|H_0|j\rangle = -t\left[\delta_{i,j+1} + \delta_{i,j-1}\right] + C\delta_{ij} . \tag{19.72}$$

This is a simple tight-binding Hamilton. It is of the same form as the Hamiltonian Baym writes in chapter 3 when motivating the Schrödinger equation. Since C is a constant added to H, we can drop it without any physical consequences. It just shifts all the energies by C.

We have already solved this Hamiltonian. It is translationally invariant by relabeling the indices. Taking a large system with periodic boundary conditions, we get

$$\langle j|\psi_n\rangle = \frac{1}{\sqrt{N}}e^{i\frac{2\pi jn}{N}} \tag{19.73}$$

and multiplying by H_0, the energies are

$$\epsilon_n = -2t\cos\left(\frac{2\pi n}{N}\right) . \tag{19.74}$$

It is sometimes convenient to take the spacing of the atoms to be a, the size of the system is $L = Na$. Writing $k = \frac{2\pi n}{L}$, and $x_j = ja$, in the limit of $N \to \infty$, we have

$$\langle x_j|k\rangle = \sqrt{\frac{a}{L}}e^{ikx_j}$$
$$\epsilon_n = -2t\cos(ka) \tag{19.75}$$

with $-\pi/a < k \le \pi/a$.

Now let's look at the diatomic case. We will approximate the change in the tight-binding Hamiltonian by an additional term

$$V = v_i\delta_{ij}$$
$$v_i = \begin{cases} 0 & i \text{ even} \\ v & i \text{ odd} \end{cases} . \tag{19.76}$$

To simplify the mathematics we will take N to be even, so the resulting Hamiltonian is translationally invariant by two sites.

This corresponds to a chain of atoms of alternating types. The energy of the two atoms differs by v, and the hopping amplitude from one kind of atom to the other is $-t$.

We first look at the weak coupling limit where v can be taken to be a perturbation. We begin by calculating the matrix elements of the perturbation between the eigenstates of H_0.

For $2N$ even,

$$
\begin{aligned}
\langle \psi_n | V | \psi_m \rangle &= \frac{v}{N} \sum_{j \text{ odd}} e^{i \frac{2\pi(m-n)j}{N}} \\
&= x^{-N/2+1} + x^{-N/2+3} + ... + x^{-1} + x + x^3 + x^5 + ... + x^{N/2-1} \quad \text{with } x = e^{i \frac{2\pi(m-n)}{N}} \\
&= \frac{x^{-N/2+1} - x^{N/2+1}}{1 - x^2} = \frac{x^{-N/2} - x^{N/2}}{x^{-1} - x} = \frac{v}{N} \frac{\sin(\pi(m-n))}{\sin\left(\frac{2\pi(m-n)}{N}\right)} .
\end{aligned} \tag{19.77}
$$

This numerator is zero for all values of m and n. The denominator is zero for $m - n$ an integer multiple $N/2$. Therefore taking n and m in the range $-N/2 < n, m \le N/2$, the only nonzero matrix elements are for $n - m = 0, N/2, -N/2$. It is simpler to go back to the original sum for these cases. The exponential is 1 if $m = n$, and -1 if the differ by $N/2$, so we have

$$
\langle \psi_n | V | \psi_m \rangle = \begin{cases} \frac{v}{2} & m = n \\ -\frac{v}{2} & m = n \pm \frac{N}{2} \\ 0 & \text{otherwise} \end{cases} \tag{19.78}
$$

The states $|\psi_n\rangle$ and $|\psi_{-n}\rangle$ of H_0 are degenerate. We see that the perturbation couples the degenerate states at $n = \pm N/2$, so we must use degenerate perturbation for those states, and for the nearby states. We could use normal perturbation theory for states away from these values of n. The nearly degenerate perturbation theory result requires us to diagonalize the Hamiltonian

$$
\begin{pmatrix} \epsilon_n + \langle \psi_n | V | \psi_n \rangle & \langle \psi_n | V | \psi_{n-N/2} \rangle \\ \langle \psi_{n-N/2} | V | \psi_n \rangle & \epsilon_{n-N/2} + \langle \psi_{n-N/2} | V | \psi_{n-N/2} \rangle \end{pmatrix} = \begin{pmatrix} \epsilon_n + \frac{v}{2} & -\frac{v}{2} \\ -\frac{v}{2} & \epsilon_{n-N/2} + \frac{v}{2} \end{pmatrix} \tag{19.79}
$$

We can simplify this a bit by realizing that $\epsilon_{n-N/2} = -\epsilon_n$.

Diagonalizing this 2×2 matrix gives the eigenvalues

$$
E_\pm(n) = \frac{v}{2} \pm \sqrt{\epsilon_n^2 + \frac{v^2}{4}} \tag{19.80}
$$

For the degenerate states, $n = \pm N/2$, the energies are $E_+(N/2) = v$, $E_-(N/2) = 0$, and we see that an energy gap opens up just like we saw from the Kronig-Penney model. Since the perturbation is in fact block diagonal with these 2×2 blocks, this nearly degenerate perturbation calculation is exact for this case. Substituting our expression for ϵ, and changing to using k, these expressions simplify to

$$
E_\pm(k) = \frac{v}{2} \pm \sqrt{4t^2 \cos^2(ka) + \frac{v^2}{4}} \tag{19.81}
$$

Sometimes the part of the Hamiltonian that we thought would be small turns out to be even larger than H_0. When this happens, we can sometimes still do perturbation theory by

first diagonalizing V and then using the original H_0 as a perturbation. This is called the strong coupling limit. For our case, V is diagonal in the original $|j\rangle$ basis, with eigenvalue v for j odd, and eigenvalue 0 for j even. Notice the huge degeneracy in this case. Looking at the form of H_0, we see that it couples even states to odd states. Therefore it does not couple the degenerate states at first order. To break the degeneracy, we need to go to second order degenerate perturbation theory. The matrix we need to diagonalize is, for j and k even,

$$\langle j| \left[H_0 + \sum_{m \text{ odd}} \frac{H_0|m\rangle\langle m|H_0}{-v} \right] |k\rangle = \begin{cases} -\frac{t^2}{v} & j = k \pm 2 \\ \\ -\frac{2t^2}{v} & j = k \end{cases}$$

Repeating for j and k odd gives

$$\langle j| \left[H_0 + \sum_{m \text{ even}} \frac{H_0|m\rangle\langle m|H_0}{-v} \right] |k\rangle = \begin{cases} +\frac{t^2}{v} & j = k \pm 2 \\ \\ v + \frac{2t^2}{v} & j = k \end{cases}$$

These are of the form of the original H_0 except only the just even or just odd terms are included, and the effective hopping amplitude between the sites is now $\pm t^2/v$. The eigenvalues for the even and odd terms are

$$
\begin{aligned}
E_e(k) &= -\frac{2t^2}{v}[1 + \cos(2ka)] \\
E_o(k) &= \frac{2t^2}{v}[1 + \cos(2ka)] + v\,.
\end{aligned}
\tag{19.82}
$$

We can check that this strong coupling result is correct by comparing to the strong coupling limit of the exact expression Eq. 19.80. Expanding we get

$$
\begin{aligned}
E_\pm(k) &= \frac{v}{2} \pm \left[4t^2 \cos^2(ka) + \frac{v^2}{4} \right]^{1/2} \\
&= \frac{v}{2} \pm \frac{v}{2} \left[1 + \frac{8t^2}{v^2}[1 + \cos(2ka)] \right]^{1/2} \\
&= \frac{v}{2} \pm \frac{v}{2} \pm \frac{2t^2}{v}[1 + \cos(2ka)] + \dots\,,
\end{aligned}
\tag{19.83}
$$

which agrees with our strong coupling calculation.

Since the full Hamiltonian is translationally invariant by two sites, we could have used that symmetry without perturbation theory to get a complete solution. The translation operator by two sites couples only the even numbered states together and the odd numbered states together. Therefore it is block diagonal in these states. The form the translation operator in the blocks is the same as translation by one operator for $N/2$ states however the state labels change by 2 when we translate. The normalized even site eigenvectors are

$$
\langle j|\psi^{e(n)}\rangle = \begin{cases} \sqrt{\frac{2}{N}} e^{i\frac{2\pi nj}{N}} & j \text{ even} \\ 0 & j \text{ odd} \end{cases}
\tag{19.84}
$$

197

and the odd site eigenvectors are

$$\langle j|\psi^{o(n)}\rangle = \begin{cases} \sqrt{\frac{2}{N}}e^{i\frac{2\pi nj}{N}} & j \text{ odd} \\ 0 & j \text{ even} \end{cases} \tag{19.85}$$

The states with the same n values have the same tranlation operator eigenvalues. Therefore the Hamiltonian can couple them. The part H_0 only couples the even site states to the odd site states, while V is nonzero only for the odd states. We therefore get 2×2 blocks

$$\begin{pmatrix} 0 & \langle \psi^{e(n)}|H_0|\psi^{o(n)}\rangle \\ \langle \psi^{o(n)}|H_0|\psi^{e(n)}\rangle & v \end{pmatrix} = \begin{pmatrix} 0 & \epsilon_n \\ \epsilon_n & v \end{pmatrix} \tag{19.86}$$

with eigenvalues $E_\pm = \frac{v}{2} \pm \sqrt{\epsilon_n^2 + \frac{v^2}{4}}$ exactly as before.

19.16 Born-Oppenheimer approximation

We have already tacitly used the Born-Oppenheimer approximation when we calculated the van der Waals interaction. The idea is to split the system into fast and slow (or high and low energy) degrees of freedom.

For example, let's look at a hydrogen atom moving in an electric field which is constant in time, but changes with position. Hamiltonian can be written as

$$\begin{aligned} H &= \frac{p_p^2}{2m_p} + \frac{p_e^2}{2m_e} - \frac{e^2}{|\boldsymbol{r}_e - \boldsymbol{r}_p|} + e\Phi(\boldsymbol{r}_p) - e\Phi(\boldsymbol{r}_e) \\ &= \frac{P^2}{2M} + \frac{p^2}{2m} - \frac{e^2}{r} + e\Phi\left(\boldsymbol{R} - \frac{m_e}{m_p + m_e}\boldsymbol{r}\right) - e\Phi\left(\boldsymbol{R} + \frac{m_p}{m_p + m_e}\boldsymbol{r}\right). \end{aligned} \tag{19.87}$$

Laboratory fields will be very slowly varying over the size of the hydrogen atom, so we can expand the potentials in a Taylor series in $\boldsymbol{r}$,

$$\begin{aligned} e\Phi\left(\boldsymbol{R} - \frac{m_e}{m_p + m_e}\boldsymbol{r}\right) - e\Phi\left(\boldsymbol{R} + \frac{m_p}{m_p + m_e}\boldsymbol{r}\right) &= e\left[\Phi(\boldsymbol{R}) - \frac{m_e}{m_p + m_e}\boldsymbol{r} \cdot \boldsymbol{\nabla}_R\Phi(\boldsymbol{R}) + ...\right] \\ &\quad -e\left[\Phi(\boldsymbol{R}) + \frac{m_p}{m_p + m_e}\boldsymbol{r} \cdot \boldsymbol{\nabla}_R\Phi(\boldsymbol{R}) + ...\right] \\ &= -e\boldsymbol{r} \cdot \boldsymbol{\nabla}_R\Phi(R) = e\boldsymbol{r} \cdot \boldsymbol{\mathcal{E}}(R), \tag{19.88} \end{aligned}$$

which is the usual interaction of a dipole in an electric field $-\boldsymbol{d} \cdot \boldsymbol{\mathcal{E}}(R)$, with $\boldsymbol{d} = -e\boldsymbol{r}$.

Our Hamiltonian is now

$$H = \frac{P^2}{2M} + \frac{p^2}{2m} - \frac{e^2}{r} + e\boldsymbol{r} \cdot \boldsymbol{\mathcal{E}}(R). \tag{19.89}$$

Unlike the case where the field was constant, here the nonconstant field couples the relative and center of mass parts of the Hamiltonian so that they cannot be separated. Physically, we expect the electric field to polarize the atom, giving it a dipole moment. This dipole moment will interact with the field which will give a force on the entire atom, changing its center of mass trajectory. The atom will be attracted to regions with a high field. With a slowly changing field, it seems that a good approximation should be to calculate the dipole moment from the relative coordinate given the field at the current center of mass position, and then use that dipole moment to calculate the center of mass Hamiltonian. This is the Born-Oppenheimer approximation.

Let's see how to make this precise. If we calculate the eigenstates of the relative Hamiltonian with the center of mass position fixed, we are using center of mass position eigenstates. The only part of our Hamiltonian that does not commute with $\boldsymbol{R}$ is $P^2/2M$. Therefore, the first part of the Born-Oppenheimer approximation is to write

$$
\begin{aligned}
H &= H_0 + V \\
H_0 &= \frac{p^2}{2m} - \frac{e^2}{r} + e\boldsymbol{r} \cdot \boldsymbol{\mathcal{E}}(R) \\
V &= \frac{P^2}{2M} \,.
\end{aligned}
\tag{19.90}
$$

The eigenstates of H_0 will be an outer product of the position eigenstates of the center of mass and the relative eigenstates at that center of mass position,

$$
H_0|n(\boldsymbol{R}), \boldsymbol{R}\rangle = E_n(\boldsymbol{R})|n(\boldsymbol{R}), \boldsymbol{R}\rangle \,,
\tag{19.91}
$$

where I have written $n(\boldsymbol{R})$ to stress that the relative coordinate eigenstates change as $\boldsymbol{R}$ changes.

Usually we are interested in the case where the atom is in its electronic ground state. The excited states are separated by approximately the energy difference between the $n = 1$ and $n = 2$ levels of hydrogen which is $\Delta E = 10.2$ eV. Recalling that room temperature is 300 K which corresponds to $1/40$ eV, this ΔE corresponds to about 10^5 K. Therefore the usual hydrogen atom kinetic energies will be small compared to this as will the energy differences between the $E_n(\boldsymbol{R})$ at different positions. We must use nearly degenerate perturbation theory to calculate the states. The lowest order contributions will be given by diagonalizing $H_0 + V$ in the basis of the nearly degenerate states. That is, we should diagonalize the operator whose matrix elements are

$$
\langle n(\boldsymbol{R}')\boldsymbol{R}'|H|n(\boldsymbol{R})\boldsymbol{R}\rangle
\tag{19.92}
$$

which corresponds to finding the eigenstates of the Hamiltonian

$$
\frac{P^2}{2M} + \langle n(\boldsymbol{R})|\boldsymbol{P}|n(\boldsymbol{R})\rangle \cdot \frac{\boldsymbol{P}}{M} + \frac{1}{2M}\langle n(\boldsymbol{R})|P^2|n(\boldsymbol{R})\rangle + E_n(\boldsymbol{R}) \,.
\tag{19.93}
$$

For bound light particles, the expectation value of the momentum in an energy eigenstate is zero.[6] The expectation value of P^2 when written in terms of the derivatives corresponds to moving the heavy particles. This therefore corresponds to moving all of the light particles in the opposite direction, and this will be of order $\frac{m}{M}$ multiplied by the light particle kinetic energy, and if $m \ll M$ it can be dropped. If it is kept, it corresponds to an additional effective potential which gives the lowest order reduced mass corrections to the energy $E_n(R)$.

Dropping these two terms, the Born-Oppenheimer Hamiltonian is then taken to be

$$H_{BO} = \frac{P^2}{2M} + E_n(\boldsymbol{R}).$$ (19.94)

If the motion of the center of mass is in the regime where classical mechanics will work, we can simply take the classical limit of the corresponding Born-Oppenheimer Schrödinger equation, and $E_n(\boldsymbol{R})$ becomes the classical potential for the center of mass motion.

For our example, for normal laboratory fields, we can calculate $E_0(\boldsymbol{R})$ using our perturbation expressions which gave

$$E_0(\boldsymbol{R}) = \epsilon_{1s} - \frac{\gamma}{2}\mathcal{E}^2(\boldsymbol{R})$$ (19.95)

where γ is the ground state polarizability $4.5a_0^3$, and the effective Hamiltonian for the center of mass for the electronic ground-state becomes

$$H_{BO} = \frac{P^2}{2M} + \epsilon_{1s} - \frac{\gamma}{2}\mathcal{E}^2(\boldsymbol{R}).$$ (19.96)

The ϵ_{1s} term is a constant potential with no physical significance and can be dropped. Notice that our result automatically gives the factor of half for the energy of an induced dipole, without any additional assumptions. You can compare this with the discussion in Jackson[2] in his section 4.7 around his equation 4.94.

Going beyond Born-Oppenheimer simply requires us to go to higher order in the perturbation theory. This can bring in terms where the excited electronic states contribute. This leads to phenomena like collisional excitations, etc.

The generalization to other degrees of freedom is straightforward. For atomic and molecular interactions, we again take the nuclear kinetic energy term to be the perturbation. Solving the electronic structure at fixed nuclear positions then gives the effective potential between the nuclei which can then be used for either quantum or classical calculations for the atomic and molecular positions. We have "integrated out" the electronic degrees of freedom.

When researchers solve Newton's equations to try to understand protein folding, the potentials are, in principle, calculated or fit to the expected electronic structure energy, and the Born-Oppenheimer approximation is used for the nuclei. For many regimes, the classical limit is appropriate and the Born-Oppenheimer Hamiltonian can be interpreted classically and solved by integrating Newton's equations.

Similar techniques are used to obtain potentials between heavy quarks by integrating out the gluon contributions or between protons and neutrons by integrating out the meson fields.

---

[6]You can show this by taking the expectation value of the commutator $\frac{im}{\hbar}[H_{\text{light}}, x_{\text{light}}] = p_{x\ \text{light}}$. As long as the expectation value of x_{light} is finite, we can expand the commutator and (dropping the light subsscript) write $\langle n|[H, x]|n\rangle = (E_n - E_n)\langle n|x|n\rangle = 0$.

19.17 Appendix A. – Rayleigh Schrödinger perturbation theory

Rayleigh Schrödinger perturbation theory is derived exactly the same way that perturbation theory in classical mechanics is derived. You write the perturbation with a counting factor λ,

$$H = H_0 + \lambda V . \tag{19.97}$$

You now write the state $|N\rangle$ and the energy E_n as a power series in λ,

$$\begin{aligned}
|N\rangle &= |n\rangle + \lambda|N^{(1)}\rangle + \lambda^2|N^{(2)}\rangle + ... \\
E_n &= \epsilon_n + E_n^{(1)}\lambda + E_n^{(2)}\lambda^2 + ...
\end{aligned} \tag{19.98}$$

where as before we choose the normalization so that all of the $|N^{(i)}\rangle$ are orthogonal to $|n\rangle$. Substitute into the time-independent Schrödinger equation,

$$\begin{aligned}
(H_0 + \lambda V)\left(|n\rangle + \lambda|N^{(1)}\rangle + \lambda^2|N^{(2)}\rangle + ...\right) = \\
\left(\epsilon_n + E_n^{(1)}\lambda + E_n^{(2)}\lambda^2 + ...\right)\left(|n\rangle + \lambda|N^{(1)}\rangle + \lambda^2|N^{(2)}\rangle + ...\right) .
\end{aligned} \tag{19.99}$$

Equating the coefficients of the powers of λ gives

$$\begin{aligned}
H_0|n\rangle &= \epsilon_n|n\rangle \\
H_0|N^{(1)}\rangle + V|n\rangle &= \epsilon_n|N^{(1)}\rangle + E_n^{(1)}|n\rangle \\
H_0|N^{(2)}\rangle + V|N^{(1)}\rangle &= \epsilon_n|N^{(2)}\rangle + E_n^{(1)}|N^{(1)}\rangle + E_n^{(2)}|n\rangle
\end{aligned} \tag{19.100}$$

Operating with $\langle n|$ from the left, these become

$$\begin{aligned}
\langle n|V|n\rangle &= E_n^{(1)} \\
\langle n|V|N^{(1)}\rangle &= E_n^{(2)}
\end{aligned} \tag{19.101}$$

which is exactly our same result as with Brillouin-Wigner perturbation theory. Operating with $m \neq n$ and rearranging we get

$$\begin{aligned}
(\epsilon_n - \epsilon_m)\langle m|N^{(1)}\rangle &= \langle m|V|n\rangle \\
(\epsilon_n - \epsilon_m)\langle m|N^{(2)}\rangle &= \langle m|V|N^{(1)}\rangle - E_n^{(1)}\langle m|N^{(1)}\rangle
\end{aligned} \tag{19.102}$$

or

$$\begin{aligned}
|N^{(1)}\rangle &= \sum_{m \neq n} \frac{|m\rangle\langle m|V|n\rangle}{\epsilon_n - \epsilon_m} \\
|N^{(2)}\rangle &= \sum_{m \neq n} \frac{|m\rangle\langle m|V|N^{(1)}\rangle - E_n^{(1)}|m\rangle\langle m|N^{(1)}\rangle}{\epsilon_n - \epsilon_m} \\
&= \sum_{m \neq n} \frac{|m\rangle\langle m|V|j\rangle\langle j|V|n\rangle}{(\epsilon_n - \epsilon_m)(\epsilon_n - \epsilon_j)} - \frac{|m\rangle\langle m|V|n\rangle\langle n|V|n\rangle}{(\epsilon_n - \epsilon_m)^2}
\end{aligned} \tag{19.103}$$

which is, of course, the same result we obtained by expanding the Brillouin-Wigner result in powers of V.

19.18   Appendix B. – First order term in our two-hydrogen interaction.

The first order term is

$$\Delta E^{(1)} = \langle 1s1s|V|1s1s \rangle \tag{19.104}$$

with V given by Eq. 19.48. Working in atomic units,

$$\langle \boldsymbol{r}|1s \rangle = \psi_{1s}(\boldsymbol{r}) = \frac{1}{\sqrt{\pi}} e^{-r} . \tag{19.105}$$

The expectation value of V is then

$$
\begin{aligned}
\langle 1s1s|V|1s1s \rangle &= \int d^3 r_1 d^3 r_2 \langle 1s1s|\boldsymbol{r}_1 \boldsymbol{r}_2 \rangle \langle \boldsymbol{r}_1 \boldsymbol{r}_2|V|1s1s \rangle \\
&= \frac{1}{\pi^2} \int d^3 r_1 d^3 r_2 e^{-2r_1 - 2r_2} \left[\frac{1}{R} - \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{R}|} - \frac{1}{|\boldsymbol{r}_2 + \boldsymbol{R}|} + \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2 - \boldsymbol{R}|} \right]
\end{aligned}
\tag{19.106}
$$

You can either directly do these integrals directly, or you can make it a little easier if you use what you learned in electrostatics. Let's do it the easier way. Notice that the integrals we must do are exactly the interaction of the charge density of atom 1 with that of atom 2. That is, our integrals above are

$$
\begin{aligned}
\langle 1s1s|V|1s1s \rangle &= \int d^3 r_1 \int d^3 r_2 \frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2 - \boldsymbol{R}|} \\
\rho(\boldsymbol{r}) &= \delta^3(\boldsymbol{r}) - \frac{1}{\pi} e^{-2r}
\end{aligned}
\tag{19.107}
$$

In electrostatics, we would probably first calculate the potential of one charge distribution and use that to calculate the interaction energy

$$
\begin{aligned}
\langle 1s1s|V|1s1s \rangle &= \int d^3 r_1 \rho(\boldsymbol{r}_1)\Phi(\boldsymbol{r}_1 - \boldsymbol{R}) \\
\Phi(\boldsymbol{r}) &= \int d^3 r' \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}
\end{aligned}
\tag{19.108}
$$

A quick way to calculate the potential is to use Gauss's law, since the charge density is spherically symmetric around the proton position,

$$
\begin{aligned}
Q(r) &= 4\pi \int_0^r dr' r'^2 \rho(r') = e^{-2r} \left[1 + 2r + 2r^2 \right] \\
E(r) &= r^{-2} Q(r) = e^{-2r} \left[r^{-2} + 2r^{-1} + 2 \right] \\
\Phi(r) &= \int_r^\infty dr' E(r') = e^{-2r} \left[\frac{1}{r} + 1 \right] .
\end{aligned}
\tag{19.109}
$$

We can check this by calculating, $-\nabla^2 \Phi = 4\pi\rho$. Plugging in the potential, the delta function can be integrated immediately

$$\langle 1s1s|V|1s1s\rangle \;=\; \int d^3r_1 \rho(\boldsymbol{r}_1)\Phi(\boldsymbol{r}_1 - \boldsymbol{R}) = \Phi(R) - \frac{1}{\pi}\int d^3r_1 e^{-2r_1}\Phi(|\boldsymbol{r}_1 - \boldsymbol{R}|)\,.$$

(19.110)

Taking the z axis along R for the r_1 integration coordinates, and calling the variable $|\boldsymbol{r}_1 - \boldsymbol{R}| = r'$, we can write

$$r_1^2 - 2Rr_1\cos\theta_1 + R^2 = r'^2 \tag{19.111}$$

so that

$$-2Rr_1 d\cos\theta_1 = 2r'dr'\,. \tag{19.112}$$

Changing integration variables from $\cos\theta_1$ to r' and doing the ϕ_1 integration, we have

$$
\begin{aligned}
\langle 1s1s|V|1s1s\rangle \;=\;& \Phi(R) - \frac{2}{R}\int_0^\infty dr_1 r_1 e^{-2r_1}\int_{|r_1-R|}^{r_1+R} dr' r' \Phi(r')\\
=\;& \Phi(R) + \frac{2}{R}\int_0^\infty dr_1 r_1 e^{-2r_1}\left(e^{-2(r_1+R)}\left[\frac{3}{4} + \frac{r_1+R}{2}\right] - e^{-2(|r_1-R|)}\left[\frac{3}{4} + \frac{|r_1-R|}{2}\right]\right)\\
=\;& \Phi(R) + \frac{2e^{-2R}}{R}\int_0^\infty dr_1 r_1 e^{-4r_1}\left[\frac{3}{4} + \frac{r_1+R}{2}\right]\\
& - \frac{2e^{-2R}}{R}\int_0^R dr_1 r_1\left[\frac{3}{4} + \frac{R-r_1}{2}\right] - \frac{2e^{2R}}{R}\int_R^\infty dr_1 r_1 e^{-4r}\left[\frac{3}{4} + \frac{r_1-R}{2}\right]\\
=\;& e^{-2R}\left[\frac{1}{R} + \frac{5}{8} - \frac{3R}{4} - \frac{R^2}{6}\right]\,.
\end{aligned}
$$

(19.113)

As expected this decays exponentially, and can be ignored at large distances compared to the R^{-6} van der Waals interaction.

An alternative method of doing the integration is to notice that the integrations are convolutions. Writing a tilde to indicate the Fourier transform, we have

$$
\begin{aligned}
\langle 1s1s|V|1s1s\rangle \;=\;& \int d^3r_1 \int d^3r_2 \frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2 - \boldsymbol{R}|}\\
=\;& \int d^3r_1 \int d^3r_2 \left[\int \frac{d^3k_1}{(2\pi)^3}\tilde\rho(k_1)e^{i\boldsymbol{k}_1\cdot\boldsymbol{r}_1}\right]\left[\int \frac{d^3k_2}{(2\pi)^3}\tilde\rho(k_2)e^{i\boldsymbol{k}_2\cdot\boldsymbol{r}_2}\right]\\
& \cdot\left[\int \frac{d^3k_3}{(2\pi)^3}\tilde v(k_3)e^{i\boldsymbol{k}_3\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2-\boldsymbol{R})}\right]\\
=\;& \int \frac{d^3k_1}{(2\pi)^3}\tilde\rho(k_1)\int \frac{d^3k_2}{(2\pi)^3}\tilde\rho(k_2)\int \frac{d^3k_3}{(2\pi)^3}\tilde v(k_3)(2\pi)^3\delta(\boldsymbol{k}_1+\boldsymbol{k}_3)(2\pi)^3\delta(\boldsymbol{k}_2-\boldsymbol{k}_3)e^{i\boldsymbol{k}_3\cdot\boldsymbol{R}}\\
=\;& \int \frac{d^3k_1}{(2\pi)^3}\tilde\rho(k_1)\tilde\rho(-\boldsymbol{k}_1)\tilde v(-k_1)e^{-i\boldsymbol{k}_1\cdot\boldsymbol{R}}\,.
\end{aligned}
$$

(19.114)

The transforms we need can be derived from the Yukawa potential $e^{-\alpha r}/r$ transform,

$$\int d^3r \frac{e^{-\alpha r}}{r} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} = \frac{4\pi}{k} \int_0^\infty dr \sin(kr) e^{-\alpha r} = -\frac{2\pi i}{k} \int_0^\infty dr \left[e^{(ik-\alpha)r} - e^{(-ik-\alpha)r} \right] = \frac{4\pi}{k^2 + \alpha^2}$$

$$(19.115)$$

The Fourier transform of the charge density and coulomb potentials are

$$\begin{aligned}
\tilde{\rho}(\boldsymbol{k}) &= \int d^3r \rho(\boldsymbol{r}) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} = 1 - \frac{1}{\pi} \int d^3r e^{-2r} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \\
&= 1 + \frac{1}{\pi}\frac{\partial}{\partial\alpha} \int d^3r \frac{e^{-\alpha r}}{r} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \bigg|_{\alpha=2} = 1 + \frac{\partial}{\partial\alpha}\frac{4}{k^2+\alpha^2}\bigg|_{\alpha=2} = \frac{k^2(k^2+8)}{(k^2+4)^2} \\
\tilde{v}(\boldsymbol{k}) &= \int d^3r \frac{1}{r} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} = \frac{4\pi}{k^2}
\end{aligned}$$

$$(19.116)$$

where the Coulomb interaction is the Yukawa with $\alpha = 0$. The matrix element is then

$$\begin{aligned}
\langle 1s1s|V|1s1s\rangle &= \int \frac{d^3k}{(2\pi)^3} \tilde{\rho}(\boldsymbol{k})\tilde{\rho}(-\boldsymbol{k})\frac{4\pi}{k^2} e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} = \int \frac{d^3k}{(2\pi)^3} \frac{4\pi k^2(k^2+8)^2}{(k^2+4)^4} e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \\
&= \frac{2}{\pi} \int_0^\infty dk k^2 \frac{\sin(kR)}{kR}\frac{k^2(k^2+8)^2}{(k^2+4)^4} = \frac{1}{\pi R}\mathrm{Im}\int_{-\infty}^\infty dk e^{ikr}\frac{k^3(k^2+8)^2}{(k^2+4)^4} .
\end{aligned}$$

$$(19.117)$$

This last integral can be done by closing the contour with a large semicircle in the upper half plane. The e^{ikr} factor ensures that this semicircle gives no contribution. The contour encloses a fourth order pole at $k = 2i$ with residue,

$$\text{residue} = \frac{1}{3!}\frac{\partial^3}{\partial k^3}\frac{1}{\pi R}\frac{k^3(k^2+8)^2 e^{ikr}}{(k+2i)^4}\bigg|_{k=2i} = \frac{e^{-2R}}{\pi}\left[\frac{1}{2R} + \frac{5}{16} - \frac{3R}{8} - \frac{R^2}{12} \right]$$

$$(19.118)$$

and the integral is then $\mathrm{Im}[2\pi i \text{ residue}]$ or

$$\langle 1s1s|V|1s1s\rangle = e^{-2R}\left[\frac{1}{R} + \frac{5}{8} - \frac{3R}{4} - \frac{R^2}{6} \right]$$

$$(19.119)$$

in agreement with our direct integration above.

The exponential makes this term go rapidly to zero with increasing separation. Beyond about 2 Angstrom separation, the van der Waals interaction dominates the charge density overlap contribution.

19.19 Problems

Problem 19.1. The rotational and vibrational spectrum of the H_2 molecule in its electronic ground state is reasonably well described by assuming that the hydrogen atoms interact via a Morse potential. The Hamiltonion is

$$H = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) + v(|\boldsymbol{r}_1 - \boldsymbol{r}_2|) \tag{19.120}$$

with

$$v(r) = V_0\{\exp[-2(r - r_0)/a] - 2\exp[-(r - r_0)/a]\} \tag{19.121}$$

and the mass m is the mass of a hydrogen atom. The parameters of the Morse potential are fit to experiment (using essentially the method we will develop below) and these values are

$$
\begin{aligned}
V_0 &= 4.72 \text{ eV} \\
r_0 &= 0.74 \text{ Angstrom} \\
a &= 0.50 \text{ Angstrom}.
\end{aligned} \tag{19.122}
$$

The wave function separates into center of mass and relative terms. The relative wave function further separates into an angular and radial part. Write the radial Schrödinger equation for this system.

Problem 19.2. Taylor series expand the Morse potential in powers of $r - r_{\min}$ keeping terms through $(r - r_{\min})^4$ where $r_{\min}$ is the position of the minimum of the Morse potential. Similarly expand the $\ell(\ell + 1)/r^2$ term around this same $r_{\min}$ value keeping terms through $(r - r_{\min})^2$. Identify the terms that give a 1 dimensional harmonic oscillator in the radial coordinate when $\ell = 0$. Since the potential is quite repulsive at r=0, you can extend allowed r values to negative infinity without changing the low lying states. This harmonic oscillator Hamiltonian is your unperturbed Hamiltonian. The remaining terms, including the terms with $\ell \neq 0$, are your perturbation.

Problem 19.3. Traditionally the energy eigenvalues are expanded as a power series in $(n + 1/2)$ and $\ell(\ell + 1)$ with the lowest order terms

$$E(n, \ell) = E_0 + E_v\left(n + \frac{1}{2}\right) + E_vX\left(n + \frac{1}{2}\right)^2 + E_{rot}\ell(\ell + 1) + \alpha\ell(\ell + 1)\left(n + \frac{1}{2}\right) \tag{19.123}$$

where n is the vibrational quantum number and ℓ is the angular momentum quantum number. These terms are respectively a constant, the harmonic vibrational energy, an anharmonic term, the rotational term and the ro-vibrational coupling. The form of the terms and their interpretation is derived using perturbation theory. You are to calculate values in eV for E_0, E_v, E_{rot}, and α, and a dimensionless value for X.

Begin by calculating the lowest order nonzero corrections to the energy from the $(r - r_{min})^3$ and $(r - r_{min})^4$ terms from the potential expansion.

Next calculate the lowest order nonzero terms from the expansion of $\ell(\ell + 1)/r^2$ that produce terms of the form given in Eq. 19.123. You should notice that the $r - r_{\min}$ term

gives no contribution. The lowest order term where this contributes is the second order energy from this centrifugal $r - r_{min}$ term and the $(r - r_{min})^3$ term from the potential expansion. Include this last term too.

Note that you should write the energy as a polynomial in $\ell(\ell+1)$ and $(n+1/2)$, dropping all terms that are not included in 19.123. For example terms like $(n+1/2)^3$ should be dropped; they would be included in a higher order expansion.

Problem 19.4. Compare your results with the experimental values below taken from the values tabulated in Gerhard Herzberg, "Molecular Spectra and Molecular Structure," (Van Nostrand, New York, 1991).

$$
\begin{aligned}
D_0 &= 4.476 \text{ eV} \\
E_v &= 0.545 \text{ eV} \\
\alpha &= -3.71 \times 10^{-4} \text{ eV} \\
X &= -2.68 \times 10^{-2} \\
E_{rot} &= 7.54 \times 10^{-3} \text{ eV}
\end{aligned}
\tag{19.124}
$$

where D_0 is the energy required to just dissociate the ground state of the molecule (i.e. to separate the atoms giving them zero energy).

Problem 19.5. A rigid rotor is described by the unperturbed Hamiltonian

$$
H_0 = -\frac{\hbar^2}{2I} \frac{\partial^2}{\partial \phi^2}
\tag{19.125}
$$

with a perturbation

$$
V = \frac{\hbar^2}{2I} \lambda \cos \phi .
\tag{19.126}
$$

(a) Solve for the energies and degeneracies of the unperturbed Hamiltonian.

(b) Use sine and cosine basis functions which are even or odd under $\phi \to -\phi$. Show that the perturbation only couples even states to even states and odd states to odd states. Calculate the lowest nonzero perturbation to the energy of the three lowest energy states (two of them are degenerate when unperturbed).

(c) Use $e^{im\phi}$ basis states. Again calculate the lowest nonzero perturbation to the three lowest energy states. You will have to use second order degenerate perturbation theory to calculate the energy. Verify your result agrees with part (b).

Problem 19.6. A simplified model of an electron trapped at an interface (for example between GaAs and AlGaAs) could be to have the electron of mass m and charge $-e$ trapped in the x direction in an infinite square well potential

$$
V(x) = \begin{cases} 0 & |x| < \frac{L}{2} \\ \infty & |x| > \frac{L}{2} \end{cases}
\tag{19.127}
$$

and able to move freely in the y and z directions. The electron is in its ground state.

a. Show that standard perturbation theory gives a polarizability α for an electric field applied along x as

$$\alpha = \frac{me^2 L^4}{\hbar^2} \frac{2^{10}}{\pi^6} \sum_{n=1}^{\infty} \frac{n^2}{(4n^2 - 1)^5} . \tag{19.128}$$

b. Evaluate this sum to a precision of 5 decimal places. Use the applet on the class web page to estimate the quadratic term in the change in energy when an electric field is applied to the infinite square well, and verify that it is in agreement with your result above.

Problem 19.7. Find the polarizability of problem 19.6 above by using the method we used in class for the hydrogen atom. Write the first order wave function as $\psi(x) = [1 + \mathcal{E}F(x)]\psi_0(x)$, solve the resulting differential equation for $F(x)$, and calculate the induced dipole to first order to get the exact polarizability analytically. Verify that it agrees with your numerical result of problem 19.6.

Problem 19.8. To test the Born-Oppenheimer approximation, we take the case of two heavy particles of mass M which each interact via a spring with a light particle of mass m. For simplicity, all three particles are restricted to move in one dimension. The Hamiltonian is

$$\frac{P_1^2}{2M} + \frac{P_2^2}{2M} + \frac{p^2}{2m} + \frac{1}{2}k(x - X_1)^2 + \frac{1}{2}k(x - X_2)^2 \tag{19.129}$$

where P_1, X_1, and P_2 X_2 are the canonical coordinates for the two heavy particles respectively, and p, x are the canonical coordinates for the light particle.

a. Calculate the energy eigenvalues of this Hamiltonian.

b. Use the Born-Oppenheimer approximation assuming $m \ll M$ to calculate the approximate energy eigenvalues of this Hamiltonian. Show that the Born-Oppenheimer approximation agrees with the exact calculation up to terms of order $\frac{m}{M}$ in the energy.

Problem 19.9. You may have noticed that the usual Born-Oppenheimer approximation does not include the correct total mass in the center of mass term. For example, if we have a model atom

$$H = \frac{P^2}{2M} + \frac{p^2}{2m} + v(\boldsymbol{r} - \boldsymbol{R}) \tag{19.130}$$

with $m \ll M$, the Born-Oppenheimer approximation would solve for the eigenvalues E_n and vectors $\psi_n(\boldsymbol{r} - \boldsymbol{R})$ of

$$H_{\text{light}} = \frac{p^2}{2m} + v(\boldsymbol{r} - \boldsymbol{R}) \tag{19.131}$$

and then use

$$H_{BO} = \frac{P^2}{2M} + E_n \tag{19.132}$$

as the heavy particle Hamiltonian. However, separating into relative and center of mass motion, with the light particle in state n would give an effective center of mass Hamiltonian

$$H_{CM} = \frac{P_{cm}^2}{2(M+m)} + E_n' \tag{19.133}$$

where E_n' would be the relative energies with the reduced mass relative Hamiltonian. Let's investigate how, in perturbation theory, the additional mass is generated.

a. Show that for $m \ll M$, $\frac{P^2}{2(M+m)} = \frac{P^2}{2M}\left[1 - \frac{m}{M} + ...\right]$.

b. We showed in the text that none of the first-order nearly degenerate perturbation terms produces an additional P^2 term. Therefore we must go to second order. Show that the second order corrections gives just the term needed to give the correction of part a. *Hint: The important matrix elements to include correspond to terms like*

$$\frac{1}{M}\langle m(R)|\boldsymbol{P}|n(R)\rangle \cdot \boldsymbol{P}. \tag{19.134}$$

Problem 19.10. Calculate the matrix elements of z between the ground state and all the p_z bound state levels of the hydrogen atom. Show that the result is

$$\langle np_z|z|1s\rangle = \sqrt{\frac{n^2-1}{3n}} \frac{16n^4}{(n^2-1)^3} \left(\frac{n-1}{n+1}\right)^n a_0. \tag{19.135}$$

Numerically sum the contribution to the completeness relation to get an idea of the relative contribution of the continuum states, i.e. calculate

$$C = 1 - \frac{\sum_{n=2}^{\infty} |\langle np_z|z|1s\rangle|^2}{\langle 1s|z^2|1s\rangle} \tag{19.136}$$

Numerically sum the expression for the polarizability and obtain upper and lower bounds to the polarizability when the continuum states are ignored.

Problem 19.11.   a. Use the integral form of the continuum hydrogenic orbitals, Eq. 17.30, to calculate the matrix elements

$$\langle kp_z|z|1s\rangle \equiv \langle k, \ell = 1, m = 0|z|n = 1, \ell = 0, m = 0\rangle. \tag{19.137}$$

b. Integrate to calculate the contribution of the continuum states to the ground-state polarizability. My integral is

$$a_0^3 \frac{2^{10}}{3} \int_0^\infty dk \frac{k}{(1+k^2)^6} \frac{e^{-4\tan^{-1}(k)/k}}{1 - e^{-2\pi/k}} \tag{19.138}$$

c. Perform the integral numerically and show that when added to the bound-state contribution your result for the ground-state polarizability is $4.5a_0^3$.

Problem 19.12. For three hydrogen atoms in their ground state, the main part of the long range interaction between them is the sum of the two-body interactions. However, there is an additional three-body interaction known as the Axilrod-Teller triple-dipole interaction that comes from the third order perturbation energy of the three dipole-dipole perturbations, B.M. Axilrod and E. Teller, J. Chem. Phys. **11**, 299 (1943).

Show that this interaction has the form

$$V_3 = \frac{3C}{r_{12}^3 r_{13}^3 r_{23}^3} \left[3\cos\theta_1 \cos\theta_2 \cos\theta_3 + 1\right] \tag{19.139}$$

where r_{ij} is the distance between the protons i and j. The lines joining the protons make a triangle whose angle at proton i is θ_i. Write an expression for C in terms of a sum over a function of single hydrogen atom matrix elements and single hydrogen atom energies. Do not include any terms that are zero.

Problem 19.13. The 3s, 3p, 3d (i.e. the $n = 3$, $\ell = 0, 1, 2$) levels of hydrogen are degenerate. A uniform electric field $\boldsymbol{\mathcal{E}}$ is applied. Calculate the shift of the energy levels and their degeneracies to linear order in $\mathcal{E}$.

Problem 19.14. An external static electric field is applied to a hydrogen atom.

a. If the field is slowly varying over the hydrogen atom, show that the perturbation $V = e\Phi(\boldsymbol{r}_p) - e\Phi(\boldsymbol{r}_e)$ can be approximated by

$$V = -\sum_\alpha d_\alpha \mathcal{E}_\alpha(\boldsymbol{R}) - \frac{1}{6}\frac{m_p - m_e}{m_p + m_e}\sum_{\alpha\beta}\frac{\partial\mathcal{E}_\alpha(\boldsymbol{R})}{\partial R_\beta}Q_{\alpha\beta} + ... \tag{19.140}$$

where m_e is the electron mass, m_p is the proton mass, $\boldsymbol{\mathcal{E}} = -\boldsymbol{\nabla}\Phi$ is the electric field,

$$
\begin{aligned}
d_\alpha &= -er_\alpha \\
Q_{\alpha\beta} &= -e(3r_\alpha r_\beta - r^2\delta_{\alpha\beta}),
\end{aligned}
\tag{19.141}
$$

$\boldsymbol{r}$ and $\boldsymbol{R}$ are the relative and center of mass coordinates and the subscripts indicate the cartesian indices.

b. A field like that of part a is applied to a hydrogen atom initially in its ground state. You can assume that the atom is well enough localized so that you can consider its center of mass to be fixed. You can also ignore the difference between the reduced and electron mass i.e. you can now assume the proton has infinite mass. The magnitude of the electric field and its gradient are weak so that we only need to keep linear terms in the field and its gradient. Show using perturbation theory that the induced quadrupole can be written as

$$\langle Q_{\alpha\beta}\rangle = \sum_{\mu\nu} C_{\alpha\beta\mu\nu}\frac{\partial\mathcal{E}_\mu}{\partial R_\nu} \tag{19.142}$$

and show that

$$C_{\alpha\beta\mu\nu} = \gamma_Q \left[\delta_{\alpha\mu}\delta_{\beta\nu} + \delta_{\alpha\nu}\delta_{\beta\mu} - \frac{2}{3}\delta_{\alpha\beta}\delta_{\mu\nu} \right] . \tag{19.143}$$

Write an expression for the quadrupole polarizability, γ_Q; Verify that your expression is equivalent to

$$\gamma_Q = 3e^2 \sum_m \frac{|\langle 1s|xy|md_{xy}\rangle|^2}{\epsilon_m - \epsilon_{1s}} \tag{19.144}$$

where $|md_{xy}\rangle$ state is a bound or continuum state with angular part given by the d_{xy} function of the angular momentum notes.

c. Using just the lowest energy matrix element that contributes, calculate a lower bound for the quadrupole polarizability. Use the completeness relation to calculate an upper bound. Give your final results as numerical values multiplied by a_0^5.

Problem 19.15. Using the results of problem 19.14, and the differential equation perturbation method, calculate the exact value of the quadrupole polarizability, γ_Q for hydrogen.

Problem 19.16. We have learned that a change of gauge changes the Hamiltonian and wave function, but does not change the physically measurable results. Since energy differences are measurable, we should expect that the change in energy on applying a magnetic field will be gauge independent.

a. Prove the identities

$$
\begin{aligned}
xp_y &= \frac{L_z}{2} + \frac{im}{2\hbar}[H, xy] \\
xy &= = \frac{1}{2i\hbar}[L_z, x^2] \\
x^2 - y^2 &= \frac{1}{2\hbar^2}[L_z, [L_z, x^2]]
\end{aligned}
\tag{19.145}
$$

where

$$H = \frac{p^2}{2m} + v(r) \tag{19.146}$$

b. A particle of charge q, moves in a rotationally invariant potential, with energy eigenstates that have good L_z. Using the gauge suggested in Baym Chapter 11, problem 11, $\boldsymbol{A} = -\boldsymbol{r} \times \boldsymbol{\mathcal{H}}/2$, show that the change in energy of the eigenstate $|n\rangle$, when a constant magnetic field $\mathcal{H}\hat{\boldsymbol{z}}$ is turned on, that is correct to second order in $\mathcal{H}$, is given by the expectation value

$$\Delta E = \langle n| \left(-\frac{q\mathcal{H}}{2mc}L_z + \frac{q^2\mathcal{H}^2}{8mc^2}[x^2 + y^2] \right) |n\rangle \tag{19.147}$$

c. Write the expression for this energy change in the gauge $\boldsymbol{A} = \mathcal{H}x\hat{\boldsymbol{y}}$. Use the results of part a to show that this expression is identical to the expression in part b.

d. Convince yourself that you can show, in perturbation theory, that an arbitrary time independent change of gauge does not change this energy shift.

Problem 19.17. Calculate the ground-state energy of the Hamiltonian

$$
\begin{aligned}
H &= H_0 + V \\
H_0 &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \\
V &= \lambda \left[\frac{1}{6}\hbar^{-1}m^2\omega^3 x^4 - \frac{1}{2}m\omega^2 x^2 \right]
\end{aligned}
\tag{19.148}
$$

using perturbation theory. Give the first order, second order, and third order in λ corrections to the energy. Compare the result for $\lambda = 1$ to the energy we calculated for the pure x^4 Hamiltonian.

Problem 19.18. To better understand how higher-order degenerate perturbation theory works, let's look at a three-state system with unperturbed Hamiltonian matrix elements

$$
\hbar\omega \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\tag{19.149}
$$

and a perturbation with matrix elements

$$
\hbar\omega\lambda \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 0 \end{pmatrix}.
\tag{19.150}
$$

a. Calculate the eigenvalues of $H = H_0 + V$ correct to order λ^3 by writing the exact determinant expression, and solving it by iteration or expansion. Use these energies along with a standard method to find the normalized eigenvectors to obtain the eigenvectors correct to order λ^2.

b. Repeat the calculation of part a. by using nondegenerate perturbation theory for the state with unperturbed eigenvalue $\hbar\omega$, and degenerate perturbation for the two states with unperturbed eigenvalue zero. Calculate to the same order as in part a, and show that the results are identical.

Problem 19.19. The 2001 Nobel prize in physics was awarded to Wolfgang Ketterle, Eric Cornell, and Carl Wieman for observing Bose-Einstein condensation of cold atoms in a harmonic trap made by a laser. Let's see how we could begin to look at particles in a harmonic trap using perturbation theory. The Hamiltonian for two noninteracting particles in a harmonic trap is

$$
H_0 = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \frac{1}{2}m\omega^2 r_1^2 + \frac{1}{2}m\omega^2 r_2^2.
\tag{19.151}
$$

Assume here that the particles are not identical, so symmetry or antisymmetry under interchange can be ignored (our ground state will be symmetric under interchange, so it will describe bosons, but some of our excited states would be excluded for bosons). In addition, the particles interact with a potential which we model as a gaussian

$$V = \lambda \hbar \omega e^{-\alpha \frac{m\omega}{\hbar} |r_1 - r_2|^2} \; . \tag{19.152}$$

a. What are the eigenvalues of H_0 and their degeneracies?

b. Calculate the first and second order in λ corrections to the unperturbed ground-state energy. Simplify your results as much as you can. My result is[7]

$$\Delta E_0^{(1)} = \frac{\lambda \hbar \omega}{(2\alpha + 1)^{3/2}}$$

$$\Delta E_0^{(2)} = -\lambda^2 \hbar \omega \sum_{k=1}^{\infty} \frac{(2k+1)! \alpha^{2k}}{2k(k!)^2 (2\alpha + 1)^{2k+3}}$$

$$= -\hbar \omega \frac{\lambda^2}{(2\alpha + 1)^3} \left[\frac{1 + 2\alpha - \sqrt{1 + 4\alpha}}{\sqrt{1 + 4\alpha}} + \frac{1}{2} \ln \frac{(1 + 2\alpha - \sqrt{1 + 4\alpha})(2\alpha + 1)^2}{(1 + 2\alpha + \sqrt{1 + 4\alpha})\alpha^2} \right]$$

$$\tag{19.154}$$

c. Calculate numerically, the total ground-state energy to second order in λ for $\lambda = 0.1$, 0.5, and $\alpha = 0.1$, 1.0, 10.0. Compare your results with the exact energies (written to 4 places after the decimal point)

$$E_0(\lambda = 0.1, \alpha = 0.1) = 3.0759 \ \hbar \omega, \quad E_0(\lambda = 0.5, \alpha = 0.1) = 3.3772 \ \hbar \omega,$$
$$E_0(\lambda = 0.1, \alpha = 1.0) = 3.0191 \ \hbar \omega, \quad E_0(\lambda = 0.5, \alpha = 1.0) = 3.0919 \ \hbar \omega,$$
$$E_0(\lambda = 0.1, \alpha = 10.0) = 3.0010 \ \hbar \omega, \quad E_0(\lambda = 0.5, \alpha = 10.0) = 3.0051 \ \hbar \omega,$$

$$\tag{19.155}$$

---

[7]I found the integral

$$\int_0^\infty dx e^{-x^2} H_{2m}(xy) = \frac{\sqrt{\pi}}{2} \frac{(2m)!}{m!} (y^2 - 1)^m \tag{19.153}$$

helpful, where $H_n(x)$ is a Hermite polynomial.

Chapter 20

Variational method

20.1   Rayleigh-Ritz variational principle

From its name, it should be clear that this principle was first used for other problems. It was developed to approximately calculate the vibrational frequencies of things like drum heads. We can calculate the expectation value of $H - E_0$ where E_0 is the ground-state energy, and insert a complete set of energy eigenstates. We find that its expectation value

$$\langle\Psi|(H - E_0)|\Psi\rangle = \sum_n \langle\Psi|(H - E_0)|n\rangle\langle n|\Psi\rangle = \sum_n |\langle n|\Psi\rangle|^2(E_n - E_0) \geq 0\,, \qquad (20.1)$$

is positive or zero since both $E_n - E_0$ and $|\langle n|\Psi\rangle|^2$ must be greater than or equal to zero. Rearranging we then have

$$\langle\Psi|(H - E_0)|\Psi\rangle = \langle\Psi|H|\Psi\rangle \geq E_0\langle\Psi|\Psi\rangle \qquad (20.2)$$

or

$$E_0 \leq \frac{\langle\Psi_T|H|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} \qquad (20.3)$$

where $|\Psi_T\rangle$ is any trial function.

20.2   Using the variational principle

A typical way to use the variational principle is to guess a form for the trial function using your knowledge of the important physics, with some unknown parameters. The energy is calculated and the parameters varied to obtain the lowest energy expectation value. The resulting form of the wave function and energy then can be used to approximate other properties.

20.3   Expected errors

If we write the trial function as

$$|\Psi_T\rangle = |\Psi_0\rangle + \lambda|\delta\Psi\rangle \tag{20.4}$$

with $\langle\Psi_0|\delta\Psi\rangle = 0$, we see that the calculated variational energy is second order in the error, $E_0 + O(\lambda^2)$, while other expectation values are linear in the error. We therefore can expect fairly accurate energies, but less accuracy for other properties.

20.4   Simple examples

Let's see how this method works for an example where we know the correct answer. We know that the ground-state energy of the 1-d harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 = \int dx |x\rangle \left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2 \right] \langle x| \tag{20.5}$$

is $\frac{1}{2}\hbar\omega$ and the ground-state wave function is a gaussian. To simplify our expressions, we can use natural units where the numerical values are $\hbar = m = \omega = 1$. An extraordinarily crude wave function would be the triangular function

$$\psi_T(x) = \langle x|\Psi_T\rangle = \begin{cases} a - |x| & -a < x < a \\ 0 & |x| > a \end{cases} \tag{20.6}$$

The variational energy is

$$
\begin{aligned}
E_T &= \frac{\langle\Psi_T|H|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} \\
&= \frac{\int dx \langle\Psi_T|x\rangle \left[-\frac{1}{2}\frac{d^2}{dx^2} + \frac{1}{2}x^2 \right] \langle x|\Psi_T\rangle}{\int dx \langle\Psi_T|x\rangle\langle x\Psi_T\rangle} \\
&= \frac{1}{2}\frac{\int_{-a}^{a} dx (a - |x|) \left[2\delta(x) + x^2(a - |x|) \right]}{\int_{-a}^{a} dx (a - |x|)^2} \\
&= \frac{1}{2}\frac{2a + \frac{a^5}{15}}{\frac{2a^3}{3}} = \frac{3}{2a^2} + \frac{a^2}{20}
\end{aligned} \tag{20.7}
$$

Setting the derivative with respect to a to zero, gives $-3/a^3 + a/10 = 0$, $a = (30)^{1/4}$ or

$$E_T = \sqrt{\frac{3}{10}} \simeq 0.5477 \tag{20.8}$$

compared to the correct answer of 0.5. Even with this really poor choice for a trial function, we found an energy off by just 10 percent.

Let's try a system where we don't have an analytic answer. After the hydrogen atom, the next simplest atom is the He atom with two electrons. Taking the nuclear mass to be large compared to the electrons, we can ignore the nuclear motion and fix the nucleus at the origin. The Hamiltonian is

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \frac{2e^2}{|\boldsymbol{r}_1|} - \frac{2e^2}{|\boldsymbol{r}_2|} + \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}. \tag{20.9}$$

It will be convenient to use atomic units where $\hbar = e = m = 1$. The Hamiltonian is then

$$H = \int d^3r_1 d^3r_1 |\boldsymbol{r}_1\boldsymbol{r}_2\rangle \left[-\frac{\nabla_1^2 + \nabla_2^2}{2} - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \right] \langle \boldsymbol{r}_1\boldsymbol{r}_2|. \tag{20.10}$$

We now need to come up with a reasonable trial function. If the electron-electron interaction is ignored, the Hamiltonian separates into two hydrogenic (with Z=2) Hamiltonians, one for each particle. The ground states of these Hamiltonians have the form $\exp(-Zr)$, so the unnormalized ground state without the electron-electron interaction will be $\exp[-2(r_1+r_2)]$. We could try this as a variational function, but let's first think about the effect of the coulomb repulsion of the electrons. If one of the electrons is far from the nucleus and the other is closer, we would expect the far electron to see an effective charge less than 2. In addition, we would expect that the wave function should become small if the two electrons are close together since this is a strongly repulsive region. This second effect gives trial functions that are harder to integrate analytically, therefore, as a first try, we might just try adjusting the effective charge in the trial function as a variational parameter. That is, a very simple trial function would be

$$\langle \boldsymbol{r}_1\boldsymbol{r}_2|\Psi_T\rangle = e^{-\alpha(r_1+r_2)} \tag{20.11}$$

with α a variational parameter.

We can write the Laplacians as

$$\nabla_1^2 = \frac{1}{r_1} \frac{d^2}{dr_1^2} r_1 - \frac{L_1^2}{\hbar^2 r_1^2} \tag{20.12}$$

and since the trial function is invariant under rotations of r_1, the angular part gives zero. ∇_2^2 behaves identically. We then have

$$\nabla_1^2 e^{-\alpha r_1} = -\frac{2\alpha}{r} + \alpha^2 \tag{20.13}$$

and our variational integral becomes

$$\langle H \rangle = -\alpha^2 + (2\alpha - 4)\frac{\int d^3r\, r^{-1} e^{-2\alpha r}}{\int d^3r\, e^{-2\alpha r}} + \tag{20.14}$$

Complete calculation – save Hyleraas for problems?

20.5 Excited states

The Rayleigh-Ritz variational method can also be applied to some excited states. The most common method is when the Hamiltonian has a symmetry so that it commutes with an easily diagonalized operator. For example, the helium atom electronic Hamiltonian commutes with rotations of both electrons. Therefore it commutes with the total electronic angular momentum. It is relatively easy to write down trial functions that are eigenstates of the total electronic angular momentum. The one we wrote above is an eigenstate with angular momentum zero. It is also straightforward to write states that are eigenstates of the total electronic angular momentum with other L^2 eigenvalues. The Rayleigh-Ritz procedure would then lead to the an upper bound for the lowest energy state with this L^2 eigenvalue.

In principle, if an accurate ground state is found, we could orthogonalize our trial function to this accurate ground state and then by optimizing this constrained trial function we would get a bound on the first excited state. In practice this tends to be numerically unstable. Generally, it is better to simply form a parameterized basis and diagonalize in this basis, varying the parameters to optimize the basis.

20.6 Feynman spectrum of liquid helium

Richard Feynman used physical arguments to guess an approximate form of the excited state wave function for liquid helium.[1] Inelastic neutron scattering measurements from low temperature liquid helium show an energy and momentum loss of the neutrons with a strong peak which suggests that the neutrons are exciting a single particle like excitation of the helium. A plot of the peak positions as an excitation energy as a function of momentum is shown in fig. 20.1. The low momentum excitations have a linear dispersion, i.e. they are phonons, while at higher momentum, the curve turns over and has a minimum. For historical reasons the excitations around this minimum are called rotons. Lev Landau predicted such an excitation that would have a gap and quadratic momentum dependence and thought they would have rotational motion – in fact the flow pattern around such an excitation does look like a smoke ring.

At low temperatures, the helium atoms can be viewed as elementary particles. That is, we have used the Born-Oppenheimer approximation, with the electronic structure integrated out to give an effective potential between the helium atoms. The Hamiltonian for liquid helium at low temperatures is approximately

$$\sum_i \frac{p_i^2}{2m} + \sum_{i<j} v(|\boldsymbol{r}_i - \boldsymbol{r}_j|) \tag{20.15}$$

where $v(r)$ is the two-body potential between the atoms. More accurate calculations add a three-body potential as well. The two-body potential has a repulsive core from the Pauli exclusion principle when the electrons of the different atoms overlap, and has a van der

---

[1]R.P. Feynman, *Atomic theory of the two-fluid model of liquid helium*, Phys. Rev. **94**, 262 (1954).
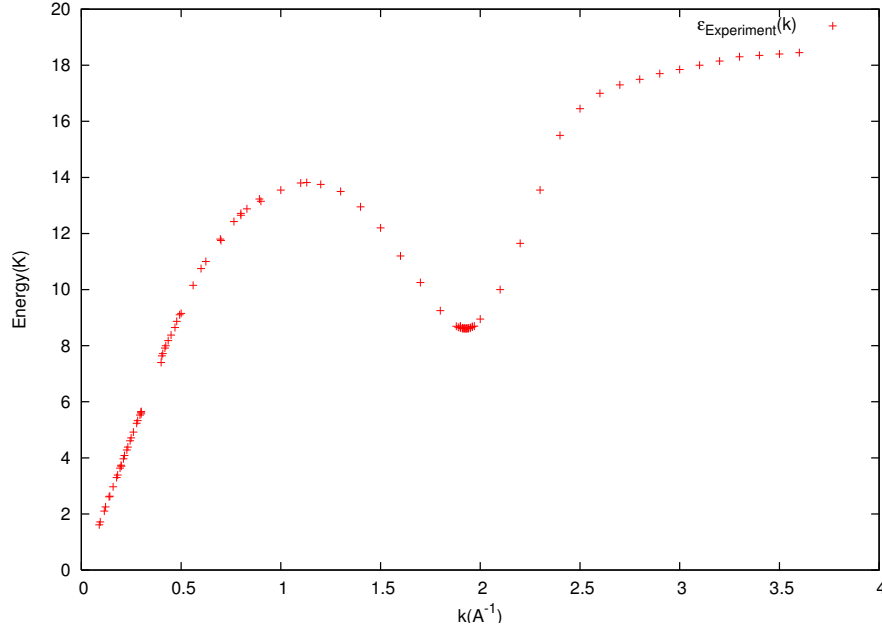
Figure 20.1: The experimental single particle excitation spectrum of liquid ^{4}He at low temperature extracted from neutron scattering energy loss as a function of momentum transfer. The data is from R.J. Donnelly, J.A. Donnelly, and R.N. Hills,*Specific heat and dispersion curve for helium II*, J. Low Temp. Phys., **44**, 471 (1981).

Waals attraction at large distances. It can be calculated from the electronic structure, or fit to dilute helium gas data.

Feynman simply stated that there was a ground-state wave function, $\Psi_0(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N)$ which we will write as $\Psi_0(R)$ for compactness. The ground state energy is E_0. Naturally occuring helium is practically all ^{4}He which has 2 electrons and a nucleus with 2 protons and 2 neutrons – it is a boson. The wave function must be symmetric under interchange of particles. We also expect the ground-state to be an eigenstate of total momentum with eigenvalue zero. The Hamiltonian is time-reversal invariant (i.e. the differential equation is real), so the nondegenerate ground state wave function is real. During neutron scattering, the neutrons transfer momentum and energy to the helium atoms. We can therefore imagine writing the wave function where we transfer momentum $\hbar\boldsymbol{q}$ to particle 1 as the ground-state with particle one's momentum increased by $\hbar\boldsymbol{q}$. This would be

$$\Phi(R) = e^{i\boldsymbol{q}\cdot\boldsymbol{r}_1}\Psi_0(R)\,. \tag{20.16}$$

However, since the particles are bosons, the wave function must be symmetric under interchange. Feynman therefore symmetrized this choice and wrote

$$\Phi_q(R) = \sum_j e^{i\boldsymbol{q}\cdot\boldsymbol{r}_j}\Psi_0(R) \equiv F(R)\Psi_0(R)\,. \tag{20.17}$$

Since this function has momentum $\hbar\boldsymbol{q}$, it is orthogonal to the ground state, and we can

217

use it as a trial function for momentum $\hbar \boldsymbol{q}$. Of course, we might expect the lowest energy such state to be the flow state where the momentum is shared by all the particles, but this form should be nearly orthogonal to the flow state, and much closer to the true single particle like excitation seen in neutron scattering.

The hope, in the days before computer simulations were easily feasible, was to write expressions in terms of ground-state expectation values which might be measured.

The straightforward way of calculating the kinetic energy expectation is to write out the direct expression and integrate by parts. The terms where the Hamiltonian operates on the ground state give the ground state energy. In addition there are other terms which include various derivatives. These can be integrated by parts to move the derivatives to the operator F. A somewhat faster way to this point is to write the ground-state expectation value of the double commutator

$$
\begin{aligned}
\frac{1}{2} & \frac{\langle \Psi_0 | [F^*(R), [H, F(R)]] | \Psi_0 \rangle}{\langle \Psi_0 | F^*(R) F(R) | \Psi_0 \rangle} \\
&= \frac{1}{2} \frac{\langle \Psi_0 | F^*(R) H F(R) - F^*(R) F(R) H - H F(R) F^*(R) + F(R) H F^*(R) | \Psi_0 \rangle}{\langle \Psi_0 | F^*(R) F(R) | \Psi_0 \rangle} \\
&= \langle H \rangle - E_0
\end{aligned}
\tag{20.18}
$$

where

$$
\langle H \rangle = \frac{\langle \Psi_0 | F^*(R) H F(R) | \Psi_0 \rangle}{\langle \Psi_0 | F^*(R) F(R) | \Psi_0 \rangle}
\tag{20.19}
$$

and we can combine the terms since both the ground-state wave function and the expectation value of H are real. The double-commutator is readily calculated to be

$$
\frac{1}{2} [F^*(R), [H, F(R)]] = -\hbar^2 \sum_j \boldsymbol{\nabla}_j F^*(R) \cdot \boldsymbol{\nabla}_j F(R) .
\tag{20.20}
$$

For the Feynman choice, this becomes

$$
\langle H \rangle - E_0 = \frac{\hbar^2 k^2}{2 m I(k)}
\tag{20.21}
$$

where the integral $I(k)$ is

$$
I(k) = \frac{1}{N} \sum_{nm} \frac{\int dR |\Psi(R)|^2 e^{i \boldsymbol{k} \cdot (\boldsymbol{r}_n - \boldsymbol{r}_m)}}{\int dR |\Psi(R)|^2} = 1 + \frac{1}{N} \sum_{n \neq m} \frac{\int dR |\Psi(R)|^2 e^{i \boldsymbol{k} \cdot (\boldsymbol{r}_n - \boldsymbol{r}_m)}}{\int dR |\Psi(R)|^2}
\tag{20.22}
$$

Except at $k = 0$, $I(k) = S(k)$ the static structure factor, and of course the Feynman wave function only makes sense for an excitation with $k \neq 0$. The static structure factor can be measured by x-ray or neutron scattering. Before looking at those results, let's try to estimate what it should look like. The two-body distribution function for an isotropic liquid like ^{4}He is only dependent on the magnitude of $\boldsymbol{r}$,

$$
g(r) = \frac{L^3}{N^2} \left\langle \sum_{i \neq j} \delta^3(\boldsymbol{r} - \boldsymbol{r}_i + \boldsymbol{r}_j) \right\rangle
\tag{20.23}
$$

where L^3 is the volume and N the number of particles. It is proportional to the probability of finding two particles separated by $\boldsymbol{r}$. The normalization is chosen so that $g(r) \to 1$ For $r \to \infty$. Subtracting this long range 1 to give a well behaved function, multiplying by $\frac{N}{L^3}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$, and integrating over $\boldsymbol{r}$ gives

$$
\begin{aligned}
\frac{N}{L^3}\int d^3r[g(r)-1]e^{i\boldsymbol{k}\cdot\boldsymbol{r}} &= \frac{1}{N}\sum_{n\neq m}\left\langle e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_n-\boldsymbol{r}_m)}\right\rangle - N\delta^3_{\boldsymbol{k},0} \\
&= I(k)-1-N\delta^3_{\boldsymbol{k},0}\equiv S(k)-1 \quad\quad (20.24)
\end{aligned}
$$

which defines the static structure factor $S(k)$. Since the Feynman excitation energy requires $k\neq 0$, we can write the Feynman energy as

$$
\epsilon_{\text{Feynman}}(k) = \frac{\hbar^2 k^2}{2mS(k)} . \quad\quad (20.25)
$$

Since the helium atoms have a hard core, we expect that $g(r)$ will go rapidly to zero as the cores come in contact. The attractive van der Waals interaction will favor having pairs of particles in their attractive well. We therefore expect $g(r)$ to be zero at the origin, and remain zero for 2 Angstroms or so. It then will have a peak which will be greater than 1 (since 1 corresponds to the average density) around where the potential is most attractive. Since we have more particles at this distance, we would expect that we would not find particles within the hard core diameter of these and therefore a minimum should follow, and $g(r)$ will oscillate. However, since this is a liquid, there will not be strong long range order and these oscillations will die out. Fourier transforming we would expect $S(k)$ to have a peak corresponding to this oscillation and the experimental result does look like that. The static structure factor, as we will see when we study scattering in detail, can be measured from neutron or x-ray scattering. The results of an x-ray scattering experiment are shown in fig. 20.2.

Plugging this into the Feynman form gives the excitation spectrum shown in fig. 20.3 The Feynman form overestimates the excitation energy as you might expect for a variational calculation. However, it gives all of the main physical features. At low momentum, the excitations are phonons, with the correct velocity of sound. The momentum of the roton minimum is correct. The main defect is that the roton energy is overestimated as is the curvature at the minimum. Feynman and Cohen[2] later identified the missing ingredient. When a helium atom moves through the liquid, the other helium atoms respond by moving back around it – exactly analogous to classical backflow of a liquid. By including these kinds of correlations they were able to get substantial agreement with experiment.

20.7 Lower bounds

It is sometimes possible to find lower bounds to the energy, and in this way, the correct energy can be bracketed. Lower bounds tend to be much less useful since they are usually

---

[2]R.P. Feynman and M. Cohen, *Energy spectrum of the excitations in liquid helium*, **102**, 1189 (1956).
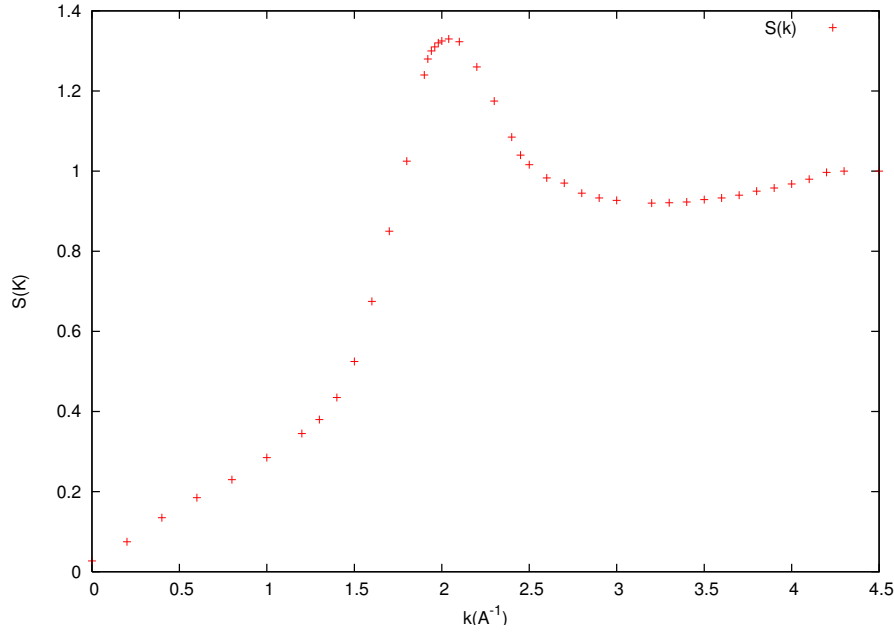
Figure 20.2: The static structure factor $S(k)$ measured with x-ray scattering at $T = 0.79$K. The data is from F.K. Achter and L. Meyer, *X-ray scattering from helium*, Phys. Rev. **188**, 291 (1969).
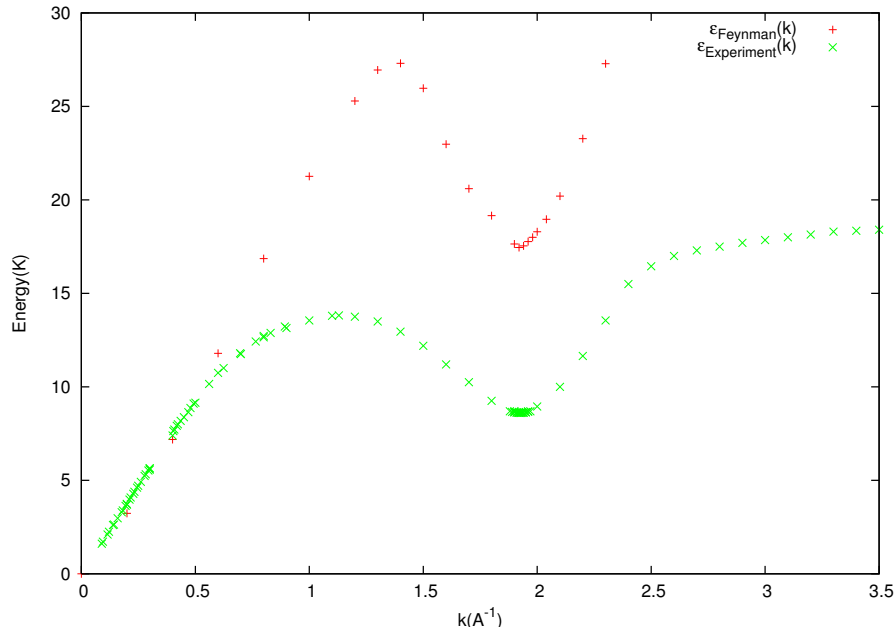
Figure 20.3: A comparison of the Feynman form to the experimental results for the single-particle-excitation spectrum of liquid ^{4}He. The Feynman form has the correct qualitative behavior and gives the roton minimum at the correct wave number. It overestimates the excitation energy because of its variational character.

inaccurate. The Temple lower bound uses the expression

$$\langle \Psi_T | (H - E_0)(H - E_1) | \Psi_T \rangle \geq 0 \tag{20.26}$$

where E_1 is the first excited state (or the first excited state contained in $|\Psi_T\rangle$). It is convenient to write

$$\langle O \rangle \equiv \frac{\langle \Psi_T | O | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \tag{20.27}$$

so that we obtain

$$\langle H^2 \rangle - \langle H \rangle^2 + (E_1 - \langle H \rangle)(E_0 - \langle H \rangle) \geq 0 \,. \tag{20.28}$$

If we can find a trial function that is good enough to satisfy $E_1 > \langle H \rangle$, then we can divide through by the positive quantity $E_1 - \langle H \rangle$ and obtain the Temple bound

$$E_0 \geq \langle H \rangle - \frac{\langle H^2 \rangle - \langle H \rangle^2}{E_1 - \langle H \rangle} \quad \text{if } E_1 > \langle H \rangle \tag{20.29}$$

The Stephenson lower bound starts by noting that

$$\langle \Psi_T | (H - C)^2 | \Psi_T \rangle \geq \langle \Psi_T | (E_0 - C)^2 | \Psi_T \rangle \quad \text{if } C < \frac{E_1 + E_0}{2} \,. \tag{20.30}$$

You can check this by expanding $|\Psi_T\rangle$ in energy eigenstates. The two sides are equal for the ground-state component. The first excited state component will satisfy the inequality if $C < \frac{E_1 + E_0}{2}$, and with this constraint all of the higher eigenstate components satisfy the inequality. Therefore, for the case where $\langle H \rangle < \frac{E_1 + E_0}{2}$, we can replace C with $\langle H \rangle$, to give

$$\langle (H - \langle H \rangle)^2 \rangle = \langle H^2 \rangle - \langle H \rangle^2 \geq (E_0 - \langle H \rangle)^2 \quad \text{if } \langle H \rangle < \frac{E_1 + E_0}{2} \,. \tag{20.31}$$

Taking the square root and rearranging, this becomes the Stephenson bound

$$E_0 \geq \langle H \rangle - \sqrt{\langle H^2 \rangle - \langle H \rangle^2} \quad \text{if } \langle H \rangle < \frac{E_1 + E_0}{2} \,. \tag{20.32}$$

Both of the lower bounds require some knowledge of the first excited state energy or the energy of the first excited state contained in the trial function. The Temple bound is usually better if E_1 can be accurately estimated. A bound on E_1 is in any case needed to employ Eq. 20.29. The Stephenson bound is somewhat easier to use since no estimate of the first excited state energy is needed. However, you must be confident that the expectation value $\langle H \rangle$ is less than the average of the ground and the first excited state energies.

20.8 Bracketing example

The ground state of the three-dimensional infinite square well of radius a has energy $E_0 = \hbar^2 \pi^2 / 2ma^2$ and angular momentum $\ell = 0$. The first $\ell = 0$ excited state has energy $E_1 = 4E_0$.

Working in natural units where $\hbar = m = a = 1$, a simple variational trial function with the proper behavior at the origin and goes to zero at $r = a$ (i.e. $r = 1$), with no free parameters would be the quadratic form $1 - r^2$. We can keep this simple polynomial form and add one variational parameter b by taking

$$\psi_T(r) = (1 - r^2)(1 + br^2) \tag{20.33}$$

The expectation value of H and H^2, using hermiticity of H to calculate H^2, and performing the trivial angular integrals, then the polynomial radial integrals

$$
\begin{aligned}
\langle H \rangle &= -\frac{1}{2} \frac{\int_0^1 dr\, r \psi_T(r) \frac{d^2}{dr^2} r \psi_T(r)}{\int_0^1 dr (r\psi_T(r))^2} = \frac{11}{4} \frac{23b^2 + 54b + 63}{5b^2 + 22b + 33} \\
\langle H^2 \rangle &= \frac{1}{4} \frac{\int_0^1 dr \left(\frac{d^2}{dr^2} r \psi_T(r) \right)^2}{\int_0^1 dr (r\psi_T(r))^2} = \frac{495}{8} \frac{37b^2 + 42b + 21}{5b^2 + 22b + 33}.
\end{aligned}
\tag{20.34}
$$

Notice if we used just the quadratic form for $\psi_T(r)$ with $b = 0$, the upper bound on the energy would be 5.25 compared to the exact value of $\frac{\pi^2}{2} \cong 4.9348$ or about 6 pecent high. Taking the derivative with respect to b, setting it to zero gives a quadratic equation with solutions $b = \frac{-111 \pm 36\sqrt{5}}{59}$. Plugging into $\langle H \rangle$ shows that taking positive sign gives a minimum, with values

$$
\begin{aligned}
\langle H \rangle &= \frac{165}{2} \frac{69 - 16\sqrt{5}}{450 + 47\sqrt{5}} \cong 4.93769 \\
\langle H^2 \rangle - \langle H \rangle &= \frac{825}{8} \frac{324750 - 144773\sqrt{5}}{(450 + 47\sqrt{5})^2} \cong 0.34396
\end{aligned}
\tag{20.35}
$$

The Stephenson lower bound is then 4.3512, and using the exact value for E_1, the Temple bound is 4.9145. The Rayleigh-Ritz upper bound is 0.06 percent high, while the Stephenson lower bound is 13 percent low, and the Temple bound even using the exact E_1 is 0.4 percent low. In practice, for a problem with an unknown solution, we would expect our bound on E_1 to be somewhat worse and we see that the lower bound is an order of magnitude worse than the upper bound. This sort of behavior is typical. Lower bounds tend to be useless unless the quality of the trial function is very high so that the energy variance is low.

20.9 Upper bounds for excited states

MacDonald[3] showed that upper bounds for excited states also can be obtained. In the simplest version of this method, a set of N orthogonal trial functions $|\psi_n^T\rangle$ are used, with

---

[3]J.K.L. MacDonald, *Succesive approximations by the Rayleigh-Ritz Variation method* Phys. Rev. **43**, 830 (1933).

$\langle \Psi_n^T | \Psi_m^T \rangle = \delta_{mn}$. MacDonald's result says that the if the eigenvalues are ordered from smallest to largest as 1 to N, the mth eigenvalue of the matrix

$$h_{ij} = \langle \Psi_i^T | H | \Psi_j^T \rangle \tag{20.36}$$

is an upper bound to the mth eigenvalue of H.

The proof is by induction. The ground-state Rayleigh-Ritz upper bound that we have already shown is the $N = 1$ case. If we assume that MacDonald's theorem is true for $N - 1$, then we take for the N case the set of $|\Psi_n^T\rangle$ the eigenvectors found for the $N - 1$ case, plus a new orthogonal eigenvector $|\Psi_N^T\rangle$. Writing the eigenvalues for the $N - 1$ case as $\lambda_n^{(N-1)}$, the matrix h_{ij} has the form

$$\begin{pmatrix} \lambda_1^{(N-1)} & 0 & 0 & 0 & \dots & h_{1N} \\ 0 & \lambda_2^{(N-1)} & 0 & 0 & \dots & h_{2N} \\ 0 & 0 & \lambda_3^{(N-1)} & 0 & \dots & h_{3N} \\ \vdots & \vdots & \ddots & \ddots & \dots & \vdots \\ h_{1N}^* & h_{2N}^* & h_{3N}^* & h_{4N}^* & \dots & h_{NN} \end{pmatrix}. \tag{20.37}$$

The eigenvalue equation is $0 = \det|h - \lambda I|$ gives the equation

$$\det(\lambda) = (h_N N - \lambda) \prod_{i=1}^{N-1} (\lambda_i^{(N-1)} - \lambda) - \sum_{j=2}^{N-1} |h_{jN}|^2 \prod_{i=1, i\neq j}^{N-1} (\lambda_i^{(N-1)} - \lambda) = 0 \tag{20.38}$$

Since the left hand side is an Nth order polynomial, it has N roots giving the N eigenvalues. Therefore, assuming no degeneracies for the moment, if we evaluate the polynomial $\det(\lambda)$ at the $\lambda_i^{(N-1)}$ values, the first term is zero, and the second term's sign is $(-1)^i$. Evaluating at $\lambda \leftarrow -\infty$ the first term dominates and the sign is positive. The first term dominates again when $\lambda \to \infty$ and the sign is $(-1)^N$. Therefore the sign changes between $-\infty$ and λ_1^{N-1}, between each of the λ_i^{N-1} and λ_{i+1}^{N-1}, and between λ_{N-1}^{N-1} and ∞. This is exactly the number of sign changes needed to get N roots, so each of the new eigenvalues is less than or equal to each of the corresponding previous eigenvalues, except for the last which is greater than or equal to the last. A quick review of the argument in case of degeneracies will show that it still goes through. Since taking the limit of $N \to \infty$ will converge to the exact result, MacDonald's theorem is proved. The nth eigenvalue of a matrix constructed as the matrix elements of the Hamiltonian of a set of orthogonal basis functions will be an upper bound to the exact nth eigenvalue.

20.10 More complicated examples

Helium liquid?

20.11 Problems

Problem 20.1. An unnormalized variational trial function for the ground state of a two electron He like atom or ion can be written as a linear combination of the two functions $\Psi_T = \psi_0 + b\psi_1$,

$$\psi_0(\boldsymbol{r}_1, \boldsymbol{r}_2) = e^{-\alpha(r_1+r_2)/a_0}$$
$$\psi_1(\boldsymbol{r}_1, \boldsymbol{r}_2) = \alpha \frac{r_{12}}{a_0} e^{-\alpha(r_1+r_2)/a_0} \tag{20.39}$$

where $r_1 = |\boldsymbol{r}_1|$, $r_2 = |\boldsymbol{r}_2|$, $r_{12} = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$, and $a_0 = \hbar^2/me^2$ is the usual Bohr radius. The first term has the form of two hydrogenic ground-state orbitals, while the second includes a correlation that favors keeping the two electrons apart. The Hamiltonian for a two-electron atom or ion is

$$H = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{r_{12}} . \tag{20.40}$$

Calculate the expectation value of the Hamiltonian,

$$\langle H \rangle = \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \tag{20.41}$$

as a function of the variational parameters α and b.[4] Your result should be

$$\langle H \rangle = \frac{e^2}{a_0} \left\{ \frac{\alpha^2(16 + 50b + 64b^2) - \alpha[32Z - 10 + (120Z - 32)b + (144Z - 35)b^2]}{16 + 70b + 96b^2} \right\} . \tag{20.42}$$

---

[4]There are many ways to perform the needed integrals. For the most general ones, one way is to first show that

$$I = \int d^3r_1 \int d^3r_2 f(r_1, r_2, r_{12}) = 8\pi^2 \int_0^\infty dr_1 r_1 \int_0^\infty dr_2 r_2 \int_{|r_1-r_2|}^{r_1+r_2} dr_{12} r_{12} f(r_1, r_2, r_{12}) .$$

You can then show that with a further change of variables to $u = r_1 + r_2$, $v = r_1 - r_2$ along with the requirement that f is symmetric under interchange of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, i.e. even in v, this reduces to

$$I = 2\pi^2 \int_0^\infty du \int_0^u dv(u^2 - v^2) \int_v^u dr_{12} r_{12} f((u+v)/2, (u-v)/2, r_{12}) .$$

All the integrals needed are a sum of terms with

$$f = \frac{u^n v^m r_{12}^p}{u^2 - v^2} e^{-2\alpha u} .$$

These integrals become

$$I(n, m, p) = 2\pi^2 \int_0^\infty du\, e^{-2\alpha u} u^n \int_0^u dv\, v^m \int_v^u dr_{12} r_{12}^{p+1} = 2\pi^2 \frac{\Gamma(m + n + p + 4)}{(2\alpha)^{p+m+n+4}(m+1)(p+m+3)} .$$

Find the $\alpha_{\min}(b)$ value that gives a minimum energy and substitute this into your expressions for the potential and kinetic energy expectation values. What are the relationships between the expectation value of the energy, potential energy and kinetic energy for this value of α? (This is a special case of the virial theorem.) Verify that for $b = 0$ and $Z = 2$ the result is equivalent to Baym's result Eq. 20-17. Plot your expression for the energy for $Z = 2$ as a function of b and locate the minimum energy from your plot. Compare this to the experimental value Baym Eq. 20-13.

Show that the H$^-$ ion is bound by plotting the energy as a function of b for $Z = 1$, and give an estimate of the electron affinity. Note that the ion is not bound with the uncorrelated wave function, i.e. when $b = 0$. Compare your result with the experimental electron affinity of the H$^-$ ion of 0.75419 eV[5].

[5]K.R. Lykke, K.K. Murray and W.C. Lineberger, *Threshold Photodetachment of H$^-$*, Phys. Rev. A **43**, 6104 (1991).

Chapter 21

Time dependent perturbation theory

21.1 Pictures or Representations

We have been working in what is called the Schrödinger representation (or picture). The time dependence of our quantum state is given by the Schrödinger equation

$$H_S|\psi(t)\rangle_S = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi(t)\rangle_S \,, \tag{21.1}$$

and operators like $x_{\text{op }S}$, $p_{\text{op }S}$, etc. have no time dependence. Here I write a subscript S to remind us that this is in the Schrödinger representation. Other representations or pictures can sometimes simplify our calculations. We will find that perturbation theory can be written in a simplified form in what is known as the interaction representation (or picture).

To begin, we will assume that all of our operators in the Schrödinger representation, and in particular the Hamiltonian, are independent of time. In that case, the formal solution to the Schrödinger equation is

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}H_S t}|\psi(0)\rangle_S \,. \tag{21.2}$$

All physical quantities can be derived from matrix elements of the operators. We can write the time dependence of a matrix element of the operator O_S as

$$_S\langle\phi(t)|O_S|\psi(t)\rangle_S \;\; = \;\; {}_S\langle\phi(0)|e^{\frac{i}{\hbar}H_S t}O_S e^{-\frac{i}{\hbar}H_S t}|\psi(0)\rangle_S \equiv {}_H\langle\phi|O_H(t)|\psi\rangle_H \tag{21.3}$$

where

$$\begin{aligned}
|\psi\rangle_H &= |\psi(0)\rangle_S \\
O_H(t) &= e^{\frac{i}{\hbar}H_S t}O_S e^{-\frac{i}{\hbar}H_S t} = e^{\frac{i}{\hbar}H_S t}O_H(0)e^{-\frac{i}{\hbar}H_S t} \,. \tag{21.4}
\end{aligned}$$

Here the subscript H indicates the Heisenberg representation or picture. Instead of propagating the states forward in time, we include the propagator terms in the operators. The operators are then time dependent and the state is independent of time. The matrix elements are unchanged, so all physical results are the same.

One way to think about this is that we are performing a unitary transformation of both the states and the operators. The unitary transformation removes the time dependence from the states and transfers it to the operators.

Heisenberg based his quantum mechanics on the time dependent matrix elements, while Schrödinger developed his wave equation which propagated the wave function in time. This difference led to the standard naming convention when dealing with states and operators.

One place where the Heisenberg representation is superior is when dealing with relativistic problems. By putting all of the space-time dependence together in the operators it is easier to obtain a manifestly covariant formalism.

Let's look at the Heisenberg Hamiltonian,

$$H_H(t) = e^{\frac{i}{\hbar}H_S t} H_S e^{-\frac{i}{\hbar}H_S t} = H_S \tag{21.5}$$

so that if the Schrödinger Hamiltonian is independent of time, so is the Heisenberg Hamiltonian, $H_H(t) = H_H(0) = H_S$. We will therefore just write H for all of these.

Since the state is independent of time in the Heisenberg representation, we no longer have a Schrödinger equation for the time dependence of the Heisenberg states. Instead, we can write a differential equation for the operators,

$$
\begin{aligned}
\frac{dO_H(t)}{dt} &= \frac{d}{dt}\left(e^{\frac{i}{\hbar}Ht}O_S e^{-\frac{i}{\hbar}Ht}\right) = \frac{i}{\hbar}H\left(e^{\frac{i}{\hbar}Ht}O_S e^{-\frac{i}{\hbar}Ht}\right) - \frac{i}{\hbar}\left(e^{\frac{i}{\hbar}Ht}O_S e^{-\frac{i}{\hbar}Ht}\right)H \\
&= \frac{i}{\hbar}[H, O_H(t)].
\end{aligned}
\tag{21.6}
$$

These commutators are no harder to calculate than in the Schrödinger represenation. For example

$$
\begin{aligned}
[A_H(t), B_H(t)] &= A_H(t)B_H(t) - B_H(t)A_H(t) \\
&= e^{\frac{i}{\hbar}Ht}A_S e^{-\frac{i}{\hbar}Ht}e^{\frac{i}{\hbar}Ht}B_S e^{-\frac{i}{\hbar}Ht} - e^{\frac{i}{\hbar}Ht}B_S e^{-\frac{i}{\hbar}Ht}e^{\frac{i}{\hbar}Ht}A_S e^{-\frac{i}{\hbar}Ht} \\
&= e^{\frac{i}{\hbar}Ht}[A_S, B_S]e^{-\frac{i}{\hbar}Ht}
\end{aligned}
\tag{21.7}
$$

where the last expression is the Heisenberg representation of Schrödinger operator commutator.

When you solved the Schrödinger equation, you had to give the initial state of the system. Similarly, here you must specify what the operators do at $t = 0$, and the Heisenberg equation of motion Eq. 21.6 gives the time dependence.

Let's compare the calculations we would do in the two representations:
Schrödinger

- Specify the initial state of the system, usually by giving the coefficients in some basis.

- Specify the operators, usually by giving their matrix elements in the basis. The matrix elements of the usual time independent operators will not change.

- Propagate the state in time using the solution to the Schrödinger equation, Eq. 21.1

- Form expectation values, etc., of our operators and states to calculate physical properties.

Heisenberg

- Specify the state of the system, usually by giving the coefficients in some basis. This will not change.

- Specify the initial operators at $t = 0$, usually by giving their matrix elements in the basis.

- Propagate the operators in time using the solution to the Heisenberg Equations of motion, Eq. 21.6.

- Form expectation values, etc., of our operators and states to calculate physical properties.

21.2 The harmonic oscillator example

As an example, let's look at the time dependence of some quantities for the simple harmonic oscillator. The Hamiltonian, x and p in the Schrödinger representation are

$$
\begin{aligned}
H &= \hbar\omega(a_S^+ a_S + 1/2) \\
x_S &= \sqrt{\frac{\hbar}{2m\omega}}(a_S^+ + a_S) \\
p_S &= \sqrt{\frac{m\hbar\omega}{2}}i(a_S^+ - a_S).
\end{aligned}
\tag{21.8}
$$

These are then the Heisenberg operators at $t = 0$,

$$
\begin{aligned}
H &= \hbar\omega(a_H(0)^+ a_H(0) + 1/2) \\
a_H(0) &= a_S \\
a_H^+(0) &= a_S^+ \\
x_H(0) &= \sqrt{\frac{\hbar}{2m\omega}}(a_H^+(0) + a_H(0)) \\
p_H(0) &= \sqrt{\frac{m\hbar\omega}{2}}i(a_H^+(0) - a_H(0)).
\end{aligned}
\tag{21.9}
$$

Usually we would need to choose a basis and solve the Heisenberg equations of motion for the operator matrix elements. The harmonic oscillator is particularly simple, and we can use the Heisenberg equation of motion to solve for the time dependence directly

$$
\begin{aligned}
\frac{da_H(t)}{dt} &= \frac{i}{\hbar}[H, a_H(t)] = i\omega\left(a_H^+(t)a_H(t)a_H(t) - a_H(t)a_H^+(t)a_H(t)\right) \\
&= i\omega[a_H^+(t), a_H(t)]a_H(t) = -i\omega a_H(t)
\end{aligned}
\tag{21.10}
$$

with solution

$$a_H(t) = e^{-i\omega t} a_H(0). \tag{21.11}$$

Either repeating for $a_H^+(t)$ or taking the hermitian conjugate,

$$a_H^+(t) = e^{i\omega t} a_H^+(0). \tag{21.12}$$

The position and momentum operators are then

$$
\begin{aligned}
x_H(t) &= \sqrt{\frac{\hbar}{2m\omega}}(a_H^+(t) + a_H(t)) = \sqrt{\frac{\hbar}{2m\omega}}(e^{i\omega t}a_H^+(0) + e^{-i\omega t}a_H(0)) \\
&= \sqrt{\frac{\hbar}{2m\omega}}\left[\cos\omega t \left(a_H^+(0) + a_H(0)\right) + i\sin\omega t \left(a_H^+(0) - a_H(0)\right)\right] \\
&= x_H(0)\cos\omega t + \frac{p_H(0)}{m\omega}\sin\omega t \\
p_H(t) &= \sqrt{\frac{m\hbar\omega}{2}}i(a_H^+(t) - a_H(t)) \\
&= p_H(0)\cos\omega t - m\omega x_H(0)\sin\omega t \tag{21.13}
\end{aligned}
$$

As practice, take an initial state of $|\psi(0)\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$, where $|0\rangle$ is the ground state and $|1\rangle$ the first excited state. Calculate the expectation value of x and p as a function of time using both the Schrödinger represention, where you would first calculate $|\psi(t)\rangle$ and then form the matrix elements, and the Heisenberg representation, where you calculate $x_H(t)$ and $p_H(t)$ as above, then form the expectation value. You then need to calculate the matrix elements of $x_H(0)$ and $p_H(0)$ to obtain the time dependence. Both methods will give the same result.

21.3 Operators with explicit time dependence

In our previous examples, our operators in the Schrödinger representation were independent of time. We say that they have no explicit time dependence. The Heisenberg operators depended on time. They had implicit time dependence through the equation of motion. Since these terms are sometimes confusing, let's review how this works in classical physics.

Let's imagine that we have a classical particle that moves in a classical electromagnetic field which, at position $\boldsymbol{r}$ and time t, is $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$. If we sit at a fixed position $\boldsymbol{r}$ and the fields change with t, we say these fields have explicit time dependence. The force on a particle of charge q whose trajectory is $\boldsymbol{r}_p(t)$, is

$$\boldsymbol{F}_p(t) = q\left[\boldsymbol{E}(\boldsymbol{r}_p(t),t) + \frac{1}{c}\frac{d\boldsymbol{r}_p(t)}{dt} \times \boldsymbol{B}(\boldsymbol{r}_p(t),t)\right]. \tag{21.14}$$

The fields at the particle's position can change with time through two mechanisms. First, the field can change with time as before even if the particle does not move, this is the explicit

time dependence. But now the value of the field used to calculate the force can change because the particle is at a different position. So even if the field were static, the force can have a time dependence through the equations of motion. This is implicit time dependence. It is implicit here because one of the arguments, in this case the position, $\boldsymbol{r}_p(t)$, changes with time.

In the Schrödinger representation, all of the time dependence from the equations of motion are isolated in the propagation of the state and the operators can have only explicit time dependence since they do not depend on the equations of motion.

The Heisenberg representation corresponds to the classical evaluation at the particle position. Operators have implicit time dependence through the propagator factors. The operators can also have explicit time dependence, where they change with time in the Schrödinger representation. In that case, the operators will have both explicit and implicit time dependence in the Heisenberg representation.

Let's now take a Hamiltonian with explicit time dependence. For example, we could imagine having a Hydrogen atom placed between the plates of a capacitor, and we have a control that changes the voltage across the capacitor so that the electric field applied to the hydrogen atom changes with time in a known way. We then have a time dependent Hamiltonian $H_S(t)$. The Schrödinger equation becomes

$$H_S(t)|\psi(t)\rangle_S = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi(t)\rangle_S. \tag{21.15}$$

In principle we could use Euler's method to solve this equation. Euler's method says that for small enough Δt

$$|\psi(t+\Delta t)\rangle_S = |\psi(t)\rangle + \Delta t|\psi'(t)\rangle_S = \left(1 - \frac{i}{\hbar}H_S(t)\Delta t\right)|\psi(t)\rangle_S \tag{21.16}$$

so that

$$
\begin{aligned}
|\psi(t)\rangle_S &= U_S(t,t_0)|\psi(t_0)\rangle_S \\
U_S(t,t_0) &= \lim_{\Delta t\to 0}\left[1 - \frac{i}{\hbar}H_S(t-\Delta t)\Delta t\right]\left[1 - \frac{i}{\hbar}H_S(t-2\Delta t)\Delta t\right]\dots \\
&\quad \dots\left[1 - \frac{i}{\hbar}H_S(t_0+\Delta t)\Delta t\right]\left[1 - \frac{i}{\hbar}H_S(t_0)\Delta t\right]
\end{aligned}
\tag{21.17}
$$

Alternatively we can use the exponential form

$$
\begin{aligned}
U_S(t,t_0) &= \lim_{\Delta t\to 0}\exp\left[-\frac{i}{\hbar}H_S(t-\Delta t)\Delta t\right]\exp\left[-\frac{i}{\hbar}H_S(t-2\Delta t)\Delta t\right]\dots \\
&\quad \dots\exp\left[-\frac{i}{\hbar}H_S(t_0+\Delta t)\Delta t\right]\exp\left[-\frac{i}{\hbar}H_S(t_0)\Delta t\right],
\end{aligned}
\tag{21.18}
$$

which has nicer properties. Expanding each exponential to first order in Δt shows that the two forms are equivalent. Both define the propagator U_S. The propagator is still unitary. This is easiest to see from the exponential form where each term is unitary. Forming the hermitian conjugate, the order of the terms is reversed and i becomes $-i$. Multiplying the propagator by its hermitian conjugate, each exponential cancels in turn leaving the identity. Since the hermitian conjugate (or adjoint) is the inverse, U_S is unitary. Notice that if H is independent of time that $U_S(t, t_0) = e^{-\frac{i}{\hbar}[t-t_0]H}$ giving our previous results.

The time derivative of the propagator is readily evaluated from Eq. 21.18

$$
\begin{aligned}
\frac{\partial}{\partial t} U_S(t, t_0) &= \lim_{\delta t \to 0} \frac{U_S(t + \delta t, t_0) - U_S(t, t_0)}{\delta t} = \lim_{\delta t \to 0} \frac{e^{-\frac{i}{\hbar} H_S(t) \delta t} U_S(t, t_0) - U_S(t, t_0)}{\delta t} \\
&= \lim_{\delta t \to 0} \frac{[1 - \frac{i}{\hbar} H_S(t) \delta t] U_S(t, t_0) - U_S(t, t_0)}{\delta t} = -\frac{i}{\hbar} H_S(t) U_S(t, t_0) \, . \quad (21.19)
\end{aligned}
$$

Alternatively you can write the Schrödinger equation as

$$
H_s(t) U_S(t, t_0) |\psi(t_0)\rangle = -\frac{\hbar}{i} \frac{\partial}{\partial t} U_S(t, t_0) |\psi(t_0)\rangle \quad (21.20)
$$

and since this must be true independent of $|\psi(t_0)\rangle$, it immediately gives the result of Eq. 21.19.

Taking the hermitian conjugate of both sides, we have

$$
\frac{\partial}{\partial t} [U_S(t, t_0)]^+ = \frac{i}{\hbar} [U_S(t, t_0)]^+ H_S(t) \, . \quad (21.21)
$$

Evaluating a time dependent matrix element allows us to convert from the Schrödinger to the Heisenberg represenations,

$$
\begin{aligned}
{}_S\langle \phi(t) | O_S(t) | \psi(t) \rangle_S &= {}_S\langle \phi(0) | [U_S(t, 0)]^+ O_S(t) U_S(t, 0) | \psi(0) \rangle_S \equiv {}_H\langle \phi | O_H(t) | \psi \rangle_H \\
|\psi\rangle_H &= |\psi(0)\rangle \\
O_H(t) &= [U_S(t, 0)]^+ O_S(t) U_S(t, 0) \, . \quad (21.22)
\end{aligned}
$$

We can now take the time derivative of $O_H(t)$. We get

$$
\begin{aligned}
\frac{dO_H(t)}{dt} &= \frac{i}{\hbar} [U_S(t, 0)]^+ H_S(t) O_S(t) U_S(t, 0) - \frac{i}{\hbar} [U_S(t, 0)]^+ O_S(t) H_S(t) U_S(t, 0) \\
&\quad + [U_S(t, 0)]^+ \frac{dO_S(t)}{dt} U_S(t, 0) \\
&= \frac{i}{\hbar} [H_H(t), O_H(t)] + [U_S(t, 0)]^+ \frac{dO_S(t)}{dt} U_S(t, 0) \\
&\equiv \frac{i}{\hbar} [H_H(t), O_H(t)] + \frac{\partial O_H(t)}{\partial t} \quad (21.23)
\end{aligned}
$$

where in the last line we have *defined* what we mean by the partial derivative of a Heisenberg operator as a function of time. Just as in the classical case, this corresponds to the explicit time dependence. If the operator has no time dependence in the Schrödinger represenation, its partial time derivative is zero in the Heisenberg representation.

21.4 Ehrenfest's theorem

Ehrenfest's theorem relates the time derivative of the expectation value to the commutator with the Hamiltonian. In the Heisenberg representation it is simply the expectation value of the operator equation of motion Eq. 21.22,

$$_H\langle\psi(t)|\frac{dO_S(t)}{dt}|\psi(t)\rangle_H =_H \langle\psi|\frac{i}{\hbar}[H_H(t), O_H(t)]|\psi\rangle_H +_H \langle\psi|\frac{\partial O_H(t)}{\partial t}|\psi\rangle_H . \quad (21.24)$$

Since the expectation values have the same form in the Schrödinger representation

$$_S\langle\psi(t)|\frac{dO_S(t)}{dt}|\psi(t)\rangle_S =_S \langle\psi(t)|\frac{i}{\hbar}[H_S(t), O_S(t)]|\psi(t)\rangle_S +_S \langle\psi(t)|\frac{\partial O_S(t)}{\partial t}|\psi(t)\rangle_S . \quad (21.25)$$

You can, of course, derive this directly in the Schrödinger representation by simply taking the time derivative of $_S\langle\psi(t)|O_S(t)|\psi(t)\rangle_S$. The derivatives of the states gives the commutator, and the derivative with respect to the explicit time dependence of $O_S(t)$ gives the last term.

In particular, for a Hamiltonian of the form $\frac{p^2}{2m} + V(x)$ and generalizations to more particles in more dimensions, the expection values of p and x satisfy the classical equations of motion where the force is the expectation value of the negative gradient of the potential.

21.5 Classical limits and Poisson brackets

In classical mechanics, we can have quantities like $F(x, p, t)$ that we evaluate along a particle's trajectory $x_p(t), p_p(t)$. To write the equation of motion, we write the total time derivative

$$\frac{dF(x_p(t), p_p(t), t)}{dt} = \left.\frac{\partial F(x, p, t)}{\partial x}\right|_{x=x_p(t), p=p_p(t)} \frac{dx_p(t)}{dt}$$
$$+ \left.\frac{\partial F(x, p, t)}{\partial p}\right|_{x=x_p(t), p=p_p(t)} \frac{dp_p(t)}{dt} + \left.\frac{\partial F(x, p, t)}{\partial t}\right|_{x=x_p(t), p=p_p(t)} \quad (21.26)$$

Hamilton's equations give us

$$\frac{dx_p(t)}{dt} = \left.\frac{\partial H(x, p, t)}{\partial p}\right|_{x=x_p(t), p=p_p(t)}$$
$$\frac{dp_p(t)}{dt} = -\left.\frac{\partial H(x, p, t)}{\partial x}\right|_{x=x_p(t), p=p_p(t)} \quad (21.27)$$

so that we have

$$\frac{dF(x_p(t), p_p(t), t)}{dt} = \left[\frac{\partial F(x, p, t)}{\partial x}\frac{\partial H(x, p, t)}{\partial p} - \frac{\partial F(x, p, t)}{\partial p}\frac{\partial H(x, p, t)}{\partial x} + \frac{\partial F(x, p, t)}{\partial t}\right]_{x=x_p(t), p=p_p(t)} . \quad (21.28)$$

The Poisson bracket is defined as the Jacobian,

$$\{A, B\} = \frac{\partial(A, B)}{\partial(x, p)} = \frac{\partial A}{\partial x}\frac{\partial B}{\partial p} - \frac{\partial A}{\partial p}\frac{\partial B}{\partial q} \tag{21.29}$$

and our classical time derivative becomes

$$\frac{dF(x_p(t), p_p(t), t)}{dt} = \left[\{F(x, p, t), H(x, p, t)\} + \frac{\partial F(x, p, t)}{\partial t}\right]_{x=x_p(t), p=p_p(t)}. \tag{21.30}$$

It looks suspiciously like the classical limit of our commutator is the Poisson bracket. This is, in fact, true. We can see this by looking at two operators $A(x_{op}, p_{op})$ and $B(x_{op}, p_{op})$. First we must define a function of two noncommutating operators a little more carefully. We need to make sure that the operator order is specified. One way is for any operator function, you rewrite it by commuting x_{op} and p_{op} until the x_{op} are to the left of all the p_{op}. Equivalently we require that the mixed matrix elements are

$$\langle x|A(x_{op}, p_{op})|p\rangle = A(x, p). \tag{21.31}$$

This imposes no real restriction since we can take any operator combination and perform the necessary commutations.

In the Heisenberg formalism, the classical limit occurs when the matrix elements are calculated in a state that can be described classically. That is a state where the uncertainty bounds on Δx and Δp have no physical consequences. For example, a basketball with a 0.6 kilogram mass being thrown at 10 m/s, has a momentum of 6 kg m/s. $\hbar$ is 1.054×10^{-34} kg m^2/s, so measuring the basketball's position within a nanometer gives a quantum uncertainty of roughly $\Delta p \sim \hbar/\Delta x \sim 10^{-25}$ kg m/s, which is 25 orders of magnitude smaller than the momentum. The quantum effects are unmeasurable. In this case we can approximate the operators by using completeness relations to write

$$
\begin{aligned}
A(x_{op}, p_{op}) &= \int dx \int \frac{dp}{(2\pi\hbar)}|x\rangle\langle x|A(x_{op}, p_{op})|p\rangle\langle p| \\
&= \int dx \int \frac{dp}{(2\pi\hbar)}|x\rangle\langle x|A(x, p)|p\rangle\langle p|
\end{aligned} \tag{21.32}
$$

The function $A(x, p)$, when evaluated in a state, that represents a system that is well approximated by classical physics, and is peaked around $x = x_0$ and $p = p_0$ can be expanded as

$$A(x, p) = A(x_0, p_0) + \frac{\partial A(x_0, p_0)}{\partial x_0}(x - x_0) + \frac{\partial A(x_0, p_0)}{\partial p_0}(p - p_0) + \dots \tag{21.33}$$

substituting into the expression above, we can write

$$
\begin{aligned}
&A(x_{op}, p_{op}) \\
&= \int dx \int \frac{dp}{(2\pi\hbar)}|x\rangle\langle x|\left(A(x_0, p_0) + \frac{\partial A(x_0, p_0)}{\partial x_0}(x - x_0) + \frac{\partial A(x_0, p_0)}{\partial p_0}(p - p_0) + \dots\right)|p\rangle\langle p| \\
&= \int dx \int \frac{dp}{(2\pi\hbar)}|x\rangle\langle x|\left(A(x_0, p_0) + \frac{\partial A(x_0, p_0)}{\partial x_0}(x_{op} - x_0) + \frac{\partial A(x_0, p_0)}{\partial p_0}(p_{op} - p_0) + \dots\right)|p\rangle\langle p|
\end{aligned} \tag{21.34}
$$

where we can replace the x and p in the expansion by x_{op} and p_{op}. With this replacement, we again have completeness relations so that we have derived the approximation

$$A(x_{\mathrm{op}}, p_{\mathrm{op}}) \;=\; A(x_0, p_0) + \frac{\partial A(x_0, p_0)}{\partial x_0}(x_{\mathrm{op}} - x_0) + \frac{\partial A(x_0, p_0)}{\partial p_0}(p_{\mathrm{op}} - p_0) + \ldots$$

$$(21.35)$$

The commutator of two such operator expansions is then

$$[A(x_{\mathrm{op}}, p_{\mathrm{op}}), B(x_{\mathrm{op}}, p_{\mathrm{op}})] \;=\; \left(\frac{\partial A(x_0, p_0)}{\partial x_0}\frac{\partial B(x_0, p_0)}{\partial p_0} - \frac{\partial B(x_0, p_0)}{\partial x_0}\frac{\partial A(x_0, p_0)}{\partial p_0}\right)[x_{\mathrm{op}}, p_{\mathrm{op}}] + \ldots$$

$$=\; i\hbar\, \{A(x, p), B(x, p)\}|_{x=x_0, p=p_0} + \ldots \qquad (21.36)$$

where, when we form the matrix element of the commutator, the corresponding Poisson bracket is evaluated at the x and p values where the classical limit wave packet is peaked. Substituting this into the Heisenberg equation of motion for the operators gives the classical Poisson bracket expression with the $\hbar$ factors cancelling. The higher order terms that we dropped correspond to higher powers of p and x in the commutators, which give more powers of $\hbar$. If kept, these would give quantum corrections to the classical motion.

21.6   The interaction representation

As we will see, the mixed representation where both the operators and states change with time simplifies the notation for time dependent perturbation theory. In the usual case, we write the Schrödinger representation Hamiltonian as

$$H_S(t) = H_0 + V_S(t)\,. \qquad (21.37)$$

As in time independent perturbation theory, we assume that we can solve H_0. We will assume that H_0 has no explicit time dependence, but $V_S(t)$ does change with time. The Schrödinger propagator given by Eq. 21.18. The matrix elements are as before

$$\begin{aligned}
{}_S\langle\phi(t)|O_S(t)|\psi(t)\rangle_S &= {}_S\langle\phi(0)|[U_S(t,0)]^+ O_S(t) U_S(t,0)|\psi(0)\rangle_S \equiv {}_I\langle\phi(t)|O_I(t)|\psi(t)\rangle_I \\
|\psi(t)\rangle_I &= e^{\frac{i}{\hbar}H_0 t}U_S(t,0)|\psi(0)\rangle \\
O_I(t) &= e^{\frac{i}{\hbar}H_0 t}O_S(t)e^{-\frac{i}{\hbar}H_0 t}
\end{aligned} \qquad (21.38)$$

We see that the interaction representation operators are the Heisenberg representation operators for the Hamiltonian H_0. The interaction representation state looks complicated, but taking its time derivative gives a simple result,

$$\begin{aligned}
-\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi(t)\rangle_I &= -\frac{\hbar}{i}\frac{\partial}{\partial t}e^{\frac{i}{\hbar}H_0 t}U_S(t,0)|\psi(0)\rangle_I \\
&= e^{\frac{i}{\hbar}H_0 t}(-H_0)U_S(t,0)|\psi(0)\rangle_I + e^{\frac{i}{\hbar}H_0 t}H_S(t)U_S(t,0)|\psi(0)\rangle_I \\
&= e^{\frac{i}{\hbar}H_0 t}V_S(t)U_S(t,0)|\psi(0)\rangle_I \\
&= e^{\frac{i}{\hbar}H_0 t}V_S(t)e^{-\frac{i}{\hbar}H_0 t}e^{\frac{i}{\hbar}H_0 t}U_S(t,0)|\psi(0)\rangle_I \\
&= V_I(t)|\psi(t)\rangle_I\,.
\end{aligned} \qquad (21.39)$$

We see that the interaction representation state satisfies a Schrödinger equation with the Hamiltonian replaced by $V(t)$ in the interaction representation.

21.7   Time dependent perturbation theory

In time dependent perturbation theory, we wish to calculate how a system changes with time under the influence of a time dependent perturbation.

For example, we might have a system in an eigenstate of H_0 with $V_S(t)$ zero initially. $V_S(t)$ is nonzero for some time and then for large t is zero again. What is the probability of finding our system in an excited state of H_0.

Another example is to calculate the dielectric constant of a material as a function of the frequency ω. We would apply an electric field oscillating at frequency ω and calculate the induced dipole which would, in the linear response regime, also oscillate at frequency ω.

If the perturbation $V_S(t)$ is small in some sense, we can expand in powers of $V_S(t)$. This will give us time-dependent perturbation theory. By looking at the higher order terms, we can determine exactly what $V_S(t)$ small means.

21.7.1   Schrödinger representation

In the Schrödinger represenation, the propagator is given by Eq. 21.18. Each of the exponential terms has the form

$$e^{-\frac{i}{\hbar}[H_0+V_S(t)]\Delta t} \tag{21.40}$$

Since H_0 and $V_S(t)$ do not commute, we cannot split the exponential. That is

$$e^{A\Delta t+B\Delta t} \neq e^{A\Delta t}e^{B\Delta t}. \tag{21.41}$$

This is readily seen by making a power series expansion of the exponentials (this always converges) and noting that on the right all of the A operators are to the left of all the B operators, while on the left they are intermixed. The Baker-Hausdorff formula gives the correct result in terms of an expansion in nested commutators

$$e^{A\Delta t}e^{B\Delta t} = e^{(A+B)\Delta t+\frac{\Delta t^2}{2}[A,B]+\frac{\Delta t^3}{12}([A,[A,B]]+[[A,B],B])+\cdots} \tag{21.42}$$

Notice that if we can ignore Δt^2 terms in comparison to Δt, then we can approximate this by the first term. Since we are taking the limit $\Delta t \to 0$ in our propagator expression, we can use this lowest order breakup. Therefore in our expansion we can write

$$e^{-\frac{i}{\hbar}[H_0+V_S(t)]\Delta t} = e^{-\frac{i}{\hbar}H_0\Delta t}\left[1-\frac{i}{\hbar}V_S(t)\Delta t\right] + O(\Delta t^2). \tag{21.43}$$

The Schrödinger propagator becomes

$$\begin{aligned}
U_S(t,t_0) &= \lim_{\Delta t\to 0} e^{-\frac{i}{\hbar}H_0\Delta t}\left[1-\frac{i}{\hbar}V_S(t-\Delta t)\Delta t\right]e^{-\frac{i}{\hbar}H_0\Delta t}\left[1-\frac{i}{\hbar}V_S(t-2\Delta t)\Delta t\right] \\
&\quad \cdots e^{-\frac{i}{\hbar}H_0\Delta t}\left[1-\frac{i}{\hbar}V_S(t_0+\Delta t)\Delta t\right]e^{-\frac{i}{\hbar}H_0\Delta t}\left[1-\frac{i}{\hbar}V_S(t_0)\Delta t\right].
\end{aligned} \tag{21.44}$$

It is helpful to define $\Delta t = (t - t_0)/M$, and $t_j = t_0 + j\Delta t$, so that

$$
\begin{aligned}
U_S(t, t_0) &= \lim_{M \to \infty} e^{-\frac{i}{\hbar} H_0 \Delta t} \left[1 - \frac{i}{\hbar} V_S(t_{M-1}) \Delta t \right] e^{-\frac{i}{\hbar} H_0 \Delta t} \left[1 - \frac{i}{\hbar} V_S(t_{M-2}) \Delta t \right] \\
&\quad \dots e^{-\frac{i}{\hbar} H_0 \Delta t} \left[1 - \frac{i}{\hbar} V_S(t_1) \Delta t \right] e^{-\frac{i}{\hbar} H_0 \Delta t} \left[1 - \frac{i}{\hbar} V_S(t_0) \Delta t \right] .
\end{aligned}
\tag{21.45}
$$

We now expand this and collect the terms containing zero, one, two, etc. factors of V_S,

$$
\begin{aligned}
U_S(t, t_0) &= \lim_{M \to \infty} \left(e^{-\frac{i}{\hbar} H_0 (t - t_0)} + \left(-\frac{i}{\hbar} \right) \sum_{j=0}^{M-1} \Delta t \, e^{-\frac{i}{\hbar} H_0 (t - t_j)} V_S(t_j) e^{-\frac{i}{\hbar} H_0 (t_j - t_0)} \right. \\
&\quad \left. + \left(-\frac{i}{\hbar} \right)^2 \sum_{j=1}^{M-1} \Delta t \sum_{k=0}^{j-1} \Delta t \, e^{-\frac{i}{\hbar} H_0 (t - t_j)} V_S(t_j) e^{-\frac{i}{\hbar} H_0 (t_j - t_k)} V_S(t_k) e^{-\frac{i}{\hbar} H_0 (t_k - t_0)} + \dots \right) \\
&= e^{-\frac{i}{\hbar} H_0 (t - t_0)} + \left(-\frac{i}{\hbar} \right) \int_{t_0}^{t} dt' \, e^{-\frac{i}{\hbar} H_0 (t - t')} V_S(t') e^{-\frac{i}{\hbar} H_0 (t' - t_0)} \\
&\quad + \left(-\frac{i}{\hbar} \right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \, e^{-\frac{i}{\hbar} H_0 (t - t')} V_S(t') e^{-\frac{i}{\hbar} H_0 (t' - t'')} V_S(t'') e^{-\frac{i}{\hbar} H_0 (t'' - t_0)} + \dots
\end{aligned}
\tag{21.46}
$$

21.7.2 Interaction representation

We can obtain our results directly in the interaction representation. Integrating the differential equation,

$$
\begin{aligned}
|\psi(t)\rangle_I &= \psi(0)\rangle_I + \int_0^t dt' \frac{d}{dt'} |\psi(t')\rangle_I \\
&= |\psi(0)\rangle_I + \left(-\frac{i}{\hbar} \right) \int_0^t dt' V_I(t') |\psi(t')\rangle_I .
\end{aligned}
\tag{21.47}
$$

We now iterate this equation

$$
\begin{aligned}
|\psi(t)\rangle_I &= |\psi(0)\rangle_I + \left(-\frac{i}{\hbar} \right) \int_0^t dt' V_I(t') |\psi(0)\rangle_I + \left(-\frac{i}{\hbar} \right)^2 \int_0^t dt' \int_0^{t'} dt'' V_I(t') V_I(t'') |\psi(t'')\rangle_I \\
|\psi(t)\rangle_I &= |\psi(0)\rangle_I + \left(-\frac{i}{\hbar} \right) \int_0^t dt' V_I(t') |\psi(0)\rangle_I + \left(-\frac{i}{\hbar} \right)^2 \int_0^t dt' \int_0^{t'} dt'' V_I(t') V_I(t'') |\psi(0)\rangle_I \\
&\quad + \left(-\frac{i}{\hbar} \right)^3 \int_0^t dt' \int_0^{t'} dt'' \int_0^{t''} dt''' V_I(t') V_I(t'') V_I(t''') |\psi(0)\rangle_I + \dots
\end{aligned}
\tag{21.48}
$$

The last line shows that the interaction representation propagator is

$$
\begin{aligned}
U_I(t, t_0) &= 1 + \left(-\frac{i}{\hbar} \right) \int_{t_0}^{t} dt' V_I(t') + \left(-\frac{i}{\hbar} \right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' V_I(t') V_I(t'') \\
&\quad + \left(-\frac{i}{\hbar} \right)^3 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \int_{t_0}^{t''} dt''' V_I(t') V_I(t'') V_I(t''') + \dots
\end{aligned}
\tag{21.49}
$$

21.7.3 Time ordered notation

For the problems we will mostly look at, we will use the perturbation expansions of the propagators as derived above. For more complicated problems when many terms are analyzed, or when manifest Lorentz invariance is desired, the time ordering operator T is introduced. T operating on a product of operators, that are functions of time, puts them in the order so that the earliest time is on the right, followed by the next earliest, continuing until the operator with the latest time is on the left. That is

$$
\begin{aligned}
TA(t) &= A(t) \\
TA(t)B(t') &= \theta(t - t')A(t)B(t') + \theta(t' - t)B(t')A(t)
\end{aligned}
\tag{21.50}
$$

and so on for more operators.

We can write Eq. 21.49 as

$$
U_I(t, t_0) = 1 + \sum_{n=1}^{\infty} \left(-\frac{i}{\hbar}\right)^n \frac{1}{n!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 ... \int_{t_0}^{t} dt_n T[V_I(t_1)V_I(t_2)...V_I(t_n)],
\tag{21.51}
$$

where we have extended the integrations to all times. This gives $n!$ orderings of the factors, so we divide by $n!$. Sometimes this is written

$$
U_I(t, t_0) = T \exp\left(-\frac{i}{\hbar} \int_{t_0}^{t} dt' V_I(t')\right)
\tag{21.52}
$$

which is only defined by making a power series expansion of the exponential and applying T to each term to recover Eq. 21.51.

Analysis of this series in many-body and relativistic systems will give a diagrammatic expansion that leads to Feynman diagrams.

Often for relativistic theories, the $V_I(t)$ can be written as an integral over positions $V_I(t) = \int d^3r v(\boldsymbol{r}, t)$ where $v(\boldsymbol{r}, t)$ is a Lorentz scalar. Combining the d^3r with the dt of the expansion gives a manifestly invariant expansion except for the time ordering operator. However, for time like intervals, the time ordering is invariant under Lorentz transformation. If two events are within their light cones, they can communicate by sending a signal from one to the other. All observers must agree on who is the sender and who is the receiver, so the time ordering of the sending event and the receiving event is invariant. The expression will not necessarily be Lorentz invariant for space-like intervals. In that case two different frames will give two different orderings. For the theory to be Lorentz invariant, these two ordering should give the same answer. If the $v(\boldsymbol{r}, t)$ commute with each other for space-like intervals the theory will be Lorentz invariant. As Steven Weinberg says[1]

> It is this condition that makes the combination of Lorentz invariance and quantum mechanics so restrictive.

---

[1]S. Weinbger, *The quantum theory of fields*, (Cambridge University Press, Cambridge, 1995).

21.8 The frequency dependent polarizability of hydrogen

Using time independent perturbation theory, we calculated the polarizability of the hydrogen atom by applying a static electric field and calculating the expectation value for the induced dipole moment. We now have the tools to extend this calculation to the frequency dependent polarizability.

Our H_0 is the hydrogen atom Hamiltonian

$$\frac{P^2}{2M} + \frac{p^2}{2m} - \frac{e^2}{r} \tag{21.53}$$

and, as before, we assume that the field gradients times the size of the atom is much smaller than the field itself (so we are restricted to wave lengths much longer than the atomic size). We can write the perturbation as the negative of the dipole operator times the electric field. At some distant time in the past, we turn on a laser with a frequency ω, so the electric field is

$$\boldsymbol{\mathcal{E}}(t) = \mathcal{E}_0 \cos \omega t. \tag{21.54}$$

Before proceeding we need to understand a little about the transient behavior of a driven system. Imagine that we had a classical ideal harmonic oscillator with frequency ω_0, at rest then at $t = 0$ suddenly turned on a sinusoidal force. We could solve the classical equations of motion for $t > 0$ by finding a particular solution with $\cos(\omega t + \phi)$ time dependence. We would add to this a solution of the homogeneous equation, i.e. $A\cos(\omega_0 t + \phi_0)$ so that we matched the boundary conditions at $t = 0$. Since the ideal oscillator has no damping, the solution will oscillate at both the driving frequency ω, and the natural frequency ω_0 forever. In a physical harmonic oscillator, small amounts of friction will damp the solutions to the homogeneous equations. So after many damping times, only the solution that oscillates at the driving frequency will remain.

An exactly analogous process happens in our quantum system. If we suddenly turn on the electric field, we will get a solution that contains both oscillations at the driving frequency ω, and oscillations that correspond to admixtures of states excited by the sudden turn on. A physical atom is coupled to the rest of the universe; for example, we have neglected the process where an atom in its excited state emits a photon and decays, because we have omitted the coupling between the atom and the photon electromagnetic field. Just as in the harmonic oscillator case, these will cause the excited states to be damped out as they decay, and the physical response at long times in the linear response regime will be just at the driving frequency.

Rather than include these damping processes, it is often easier to modify the way we turn on the interaction so that these terms are not excited. In both the classical driven harmonic oscillator and the quantum system, this can be done by turning on the interaction very slowly, that is, adiabatically. We therefore often multiply the interaction by the factor $\exp(\eta t)$, for $t < 0$, where the limit $\eta \to 0^+$ is taken. This turns the interaction on slowly, and around $t = 0$ the interaction has its full value. We then measure the response around

$t = 0$. You can also define the factor to be 1 for $t > 0$ in which case you can calculate for any time $t > 0$. For scattering problems we often use the factor $\exp(-\eta|t|)$, which turns off the interaction between the projectile and target for both large positive and large negative times. This is, of course, unphysical but since the projectile wave packet is far from the target at large positive and negative times, we still get the physical result and can use the methods of time dependent perturbation theory to derive the results.

For our polarizability calculation, we then take $t_0 \to -\infty$ in our propagator, and we take the interaction to be

$$V_S(t) = e z \mathcal{E}_0 \cos(\omega t) e^{\eta t} = \frac{e z \mathcal{E}_0}{2} \left[e^{i(\omega - i\eta)t} + e^{-i(\omega + i\eta)t} \right] \tag{21.55}$$

for a field along z and $t < 0$.

Taking the hydrogen atom to be in its ground state in the distant past, the interaction representation state, keeping up to linear terms in the perturbation, is for $t \leq 0$,

$$|\psi(t)\rangle_I = |1s\rangle - \frac{i}{\hbar} \int_{-\infty}^{t} dt' e^{\frac{i}{\hbar}H_0 t'} \frac{e z \mathcal{E}_0}{2} \left[e^{i(\omega - i\eta)t'} + e^{-i(\omega + i\eta)t'} \right] e^{-\frac{i}{\hbar}H_0 t'} |1s\rangle. \tag{21.56}$$

To evaluate the H_0 operators, we insert a complete set of H_0 eigenstates so that

$$
\begin{aligned}
|\psi(t)\rangle_I &= |1s\rangle - \frac{i}{\hbar} \sum_m \int_{-\infty}^{t} dt' |m\rangle\langle m| e^{\frac{i}{\hbar}\epsilon_m t'} \frac{e z \mathcal{E}_0}{2} \left[e^{i(\omega - i\eta)t'} + e^{-i(\omega + i\eta)t'} \right] e^{-\frac{i}{\hbar}\epsilon_{1s} t'} |1s\rangle \\
&= |1s\rangle - \frac{e \mathcal{E}_0}{2} \sum_m |m\rangle\langle m|z|1s\rangle \left[\frac{i}{\hbar} \int_{-\infty}^{t} dt' e^{\frac{i}{\hbar}(\epsilon_m - \epsilon_{1s} + \hbar\omega - i\eta)t'} + \frac{i}{\hbar} \int_{-\infty}^{t} dt' e^{\frac{i}{\hbar}(\epsilon_m - \epsilon_{1s} - \hbar\omega - i\eta)t'} \right] \\
&= |1s\rangle - \frac{e \mathcal{E}_0}{2} \sum_m |m\rangle\langle m|z|1s\rangle \left[\frac{e^{\frac{i}{\hbar}(\epsilon_m - \epsilon_{1s} + \hbar\omega)t}}{\epsilon_m - \epsilon_{1s} + \hbar\omega - i\eta} + \frac{e^{\frac{i}{\hbar}(\epsilon_m - \epsilon_{1s} - \hbar\omega)t}}{\epsilon_m - \epsilon_{1s} - \hbar\omega - i\eta} \right] \tag{21.57}
\end{aligned}
$$

where we can drop the η factors in the exponents. Since η is an infinitesimal, I drop any positive factors that multiply η, so the η in the last expression is $\hbar$ times the original η.

The only matrix elements that survive are when $|m\rangle$ is a p_z state. Just as in the static case, the p_z states will give nonzero matrix elements only for the z component of the dipole operator. The z component of the dipole operator is

$$d_{zI}(t) = e^{\frac{i}{\hbar}H_0 t}(-ez)e^{-\frac{i}{\hbar}H_0 t} \tag{21.58}$$

and its expectation value keeping terms linear in $\mathcal{E}_0$,

$$
\begin{aligned}
\langle d_z \rangle(t) &= \frac{{}_I\langle \psi(t)|d_{zI}(t)|\psi(t)\rangle_I}{{}_I\langle \psi(t)|\psi(t)\rangle_I} \\
&= \text{Re } \mathcal{E}_0 e^{-i\omega t} \left(e^2 \sum_m |\langle mp_z|z|1s\rangle|^2 \left[\frac{1}{\epsilon_m - \epsilon_{1s} + \hbar\omega + i\eta} + \frac{1}{\epsilon_m - \epsilon_{1s} - \hbar\omega - i\eta} \right] \right). \tag{21.59}
\end{aligned}
$$

This shows that our complex polarizability is the factor in parentheses.

If we take the limit $\omega \to 0$, our result goes over to the static polarizability we derived with time independent perturbation theory. This shows that we can use time dependent perturbation theory to obtain all of the time independent perturbation results. All we need to do is multiply V by the adiabatic factor which slowly turns it on, and calculate the change in an eigenstate of H_0 when the interaction is turned on fully.

The $i\eta$ factors in the denominator do nothing except when $\hbar\omega$ matches one of the transition energies $\epsilon_m - \epsilon_{1s}$. It turns out that if we calculate to higher order, for ω near such a transition, the $i\eta$ factors turn into into $i\Gamma/2\hbar$ where Γ is the transition rate from the mp_z to the $1s$ state.

Sometimes it is convenient to combine the factors. Writing $\epsilon_m - \epsilon_{1s} = \hbar\omega_{m1}$,

$$
\begin{aligned}
\langle d_z \rangle (t) &= \operatorname{Re} \mathcal{E}_0 e^{-i\omega t} \frac{2e^2}{\hbar} \sum_m \frac{|\langle mp_z|z|1s\rangle|^2 \omega_{m1}}{\omega_{m1}^2 - (\omega + i\eta)^2} \\
&= \operatorname{Re} \mathcal{E}_0 e^{-i\omega t} \frac{2e^2}{\hbar} \sum_m \frac{|\langle mp_z|z|1s\rangle|^2 \omega_{m1}}{\omega_{m1}^2 - \omega^2 - i\eta\omega}
\end{aligned}
\tag{21.60}
$$

It is informative to compare these results to those of the classical driven harmonic oscillator. The equation of motion of an oscillator with friction coefficient $m\zeta$, driven by a force $F\cos\omega t$ is

$$
m\ddot{z} = -m\omega_0^2 z - m\zeta\dot{z} + F\cos\omega t \,.
\tag{21.61}
$$

This is a linear equation so we can solve

$$
m\ddot{z} = -m\omega_0^2 z - m\zeta\dot{z} + F e^{-i\omega t} \,.
\tag{21.62}
$$

and take the real part at the end. The friction causes the homogeneous solutions that oscillate at the natural frequency to be damped. For large times after the force is turned on, the solution to the complex equation will be $z_0 e^{-i\omega t}$. Substituting we find

$$
z_0 = \frac{F}{m} \frac{1}{\omega_0^2 - \omega^2 - i\omega\zeta}
\tag{21.63}
$$

and

$$
z(t) = \operatorname{Re} \frac{F}{m} e^{-i\omega t} \frac{1}{\omega_0^2 - \omega^2 - i\omega\zeta} \,.
\tag{21.64}
$$

This has the same form as one of the terms of our quantum result. The η in the quantum result does take the place of the friction in the classical equation. You should have seen this classical form before. If not, review the simple model for $\epsilon(\omega)$ in Jackson's Classical Electrodynamics, section 7.5A.

If we multiply and divide the quantum result by the mass m, the coefficent multiplying $\frac{e^2}{m(\omega_{j1}^2 - \omega^2 - i\eta\omega)}$ is called the oscillator strength f_{j1}. More generally for an arbitrary initial eigenstate of H_0, $|n\rangle$, the oscillator strengths are

$$
f_{jn} = \frac{2m}{\hbar^2} |\langle j|z|n\rangle|^2 (E_j - E_n) \,.
\tag{21.65}
$$

The Thomas-Rieche-Kuhn sum rule or f sum rule was originally found by applying correspondence principles and old quantum theory. In this case it says

$$\sum_j f_{jn} = 1 \,.$$

<div align="right">(21.66)</div>

We can verify this result

$$
\begin{aligned}
\sum_j f_{jn} &= \frac{2m}{\hbar^2} \sum_j \langle n|z|j\rangle \langle j|z|n\rangle (E_j - E_n) \\
&= \frac{m}{\hbar^2} \sum_j \left(\langle n|[z,H]|j\rangle \langle j|z|n\rangle + \langle n|z|j\rangle \langle j|[H,z]|n\rangle \right) \\
&= \frac{m}{\hbar^2} \left(\langle n|[z,H]z|n\rangle + \langle n|z[H,z]|n\rangle \right) \\
&= \frac{m}{\hbar^2} \langle n|[[z,H],z]|n\rangle \,.
\end{aligned}
$$

<div align="right">(21.67)</div>

The commutator is

$$[[z,H],z] = -\frac{\hbar}{im}[p_z, z] = \frac{\hbar^2}{m}$$

<div align="right">(21.68)</div>

proving the f sum rule.

This sum rule played a major role in the development of quantum mechanics. Heisenberg rederived it in his first paper on quantum mechanics, and used it as the quantum condition. Born and Jordan showed soon after that this was equivalent to taking the commutator $[x, p] = i\hbar$ which is the canonical quantization condition.

21.9 Transition probabilities

In some experiments, after placing our system in an eigenstate of H_0, we turn on a time dependent perturbation. This modifies the system. The perturbation is turned off and we make a measurement. These types of experiments are often called pump probe experiments. The initial perturbation is the pump stage that changes the system and the probe stage is a measurement.

One possibility is to measure the probability of the system being in a particular eigenstate of H_0 after the perturbation. Let's calculate the probability of starting in the state $|n\rangle$ and ending in the state $|m\rangle$ to lowest order in the perturbation. If the perturbation is turned on at time t_0 and turned off at time t, the probability amplitude at lowest order is

$$
\begin{aligned}
\langle m|\psi(t)\rangle &= \langle m|1 - \frac{i}{\hbar} \int_{t_0}^t dt' V_I(t')|n\rangle \\
&= \delta_{mn} - \frac{i}{\hbar} \int_{t_0}^t dt' \langle m|V_I(t')|n\rangle \\
&= \delta_{mn} - \frac{i}{\hbar} \int_{t_0}^t dt' e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m)t'} \langle m|V_S(t')|n\rangle \,.
\end{aligned}
$$

<div align="right">(21.69)</div>

The lowest order transition probability for $m \neq n$ is

$$
\begin{aligned}
P_{nm} &= \left| -\frac{i}{\hbar} \int_{t_0}^{t} dt' \langle m | V_I(t') | n \rangle \right|^2 \\
&= \left| -\frac{i}{\hbar} \int_{t_0}^{t} dt' e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m)t'} \langle m | V_S(t') | n \rangle \right|^2 .
\end{aligned}
\tag{21.70}
$$

To calculate to higher order you simply take more terms in the expansion.

21.10 Hydrogen atom moving through a parallel plate capacitor

Again taking our hydrogen atom example, we can imagine sending the hydrogen atom through a parallel plate capacitor. The hydrogen atom starts far from the capacitor, it reaches the capacitor and enters between its plates at time $-T/2$, and exits from the capacitor at time $T/2$. In the rest frame of the hydrogen atom, it sees a changing electric field which goes rapidly from zero to the internal capacitor field and then drops again to zero. We model this as a hydrogen atom which sees a uniform time dependent electric field

$$
\mathcal{E}(t) = \begin{cases} -\hat{z}\mathcal{E}_0 & |t| < \frac{T}{2} \\ 0 & |t| > \frac{T}{2} \end{cases} .
\tag{21.71}
$$

We have ignored the smooth change from the fringing field of the capacitor and modeled it as an abrupt change, and we ignore the magnetic field that would be seen in the rest frame of the hydrogen atom.

The relative Hamiltonian is then

$$
\begin{aligned}
H_0 &= \frac{p^2}{2m} - \frac{e^2}{r} \\
V_S(t) &= \begin{cases} ez\mathcal{E}_0 & |t| < \frac{T}{2} \\ 0 & |t| > \frac{T}{2} \end{cases} .
\end{aligned}
\tag{21.72}
$$

This is a special case of a perturbation where the time dependence factors out, so that it is in the form of a time independent operator multiplying a time dependent function

$$
V_S(t) = f(t) v_S .
\tag{21.73}
$$

Since $f(t)$ is zero outside the time when the perturbation is turned on, we can extend the time integration from $-\infty$ to ∞. The probability of making a transition from state $|n\rangle$ to $|m\rangle$ becomes

$$
\begin{aligned}
P_{nm} &= \left| -\frac{i}{\hbar} \int_{-\infty}^{\infty} dt' e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m)t'} f(t) \langle m | v_S | n \rangle \right|^2 \\
&= \frac{1}{\hbar^2} \left| \tilde{f}\left(\frac{\epsilon_m - \epsilon_n}{\hbar} \right) \right|^2 |\langle m | v_S | n \rangle|^2 ,
\end{aligned}
\tag{21.74}
$$

where $\tilde{f}(\omega)$ is the Fourier transform of $f(t)$

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} dt e^{i\omega t} f(t) \,. \tag{21.75}$$

Our general result shows that if the time dependence of the $f(t)$ is slow enough that it has only has large nonzero Fourier transform for ω values much smaller than $\Delta E/\hbar$, the transition probability is very low. This is the adiabatic limit that we used for the dynamic polarizability.

For our hydrogen atom problem we can perform the integral. At lowest order, only those states coupled by the z operator can occur. We know that these will be p_z states. The transition rate to a $2p_z$ state is then

$$
\begin{aligned}
P &= \left| -\frac{i}{\hbar} \int_{-T/2}^{T/2} dt' e^{-\frac{i}{\hbar}(\epsilon_{1s}-\epsilon_{2p})t'} \langle 2p_z | e\mathcal{E}_0 z | n \rangle \right|^2 \\
&= 4e^2\mathcal{E}_0^2 \frac{\sin^2\left(\frac{(\epsilon_{1s}-\epsilon_{2p})T}{2\hbar}\right)}{(\epsilon_{1s}-\epsilon_{2p})^2} |\langle 2p_z | z | 1s \rangle|^2 \\
&= 2e^2\mathcal{E}_0^2 \frac{|\langle 2p_z | z | 1s \rangle|^2}{(\epsilon_{1s}-\epsilon_{2p})^2} \left[1 - \cos\left(\frac{(\epsilon_{1s}-\epsilon_{2p})T}{\hbar}\right) \right]
\end{aligned}
\tag{21.76}
$$

For small T, the probability is

$$P = \frac{e^2\mathcal{E}_0^2}{\hbar^2} \frac{|\langle 2p_z | z | 1s \rangle|^2}{(\epsilon_{1s}-\epsilon_{2p})^2} T^2 + ... \tag{21.77}$$

For large T, the result oscillates at the frequency corresponding to the energy difference between the states. This result is correct to order $\mathcal{E}_0^2$, so if we ignore higher order in the field transitions, keeping just this order in perturbation is adequate. The physical probability cannot be larger than 1. If the calculated probability becomes larger than one, we obviously must keep higher order terms. Very roughly, for small T, the integrals will be order $\langle m|V|n\rangle T$, so in general this quantity for small T should be small for our expansion to be good. For longer times, the frequency corresponding to the energy difference sets the time scale, and we require $|\langle m|V|n\rangle|/(\epsilon_m - \epsilon_n)$ to be small just as in time independent perturbation theory.

21.11   Transition rates and Fermi's golden rule

In the previous section, we assumed that the states were not degenerate. That is that $\epsilon_n \neq \epsilon_m$. For a system with a continuum of states, we need to modify our calculations a little. A typical case, which we will cover once we have quantized the electromagnetic field, is a hydrogen atom in an excited state along with the electromagnetic field in its ground state (i.e. no photons). The hydrogen atom can decay to its ground state emitting a photon. The energy of the hydrogen ground state plus the photon energy is equal to the hydrogen excited

state energy, so the system of hydrogen atom and electromagnetic field has transitioned from one degenerate state to another. Experimentally, we find that there is a transition rate Γ, that is, a transition probability per unit time $\Gamma = dP/dt$.[2] Our time dependent formalism will allow us to calculate this transition rate.

We wish to calculate the transition rate between two eigenstates of H_0 under the influence of a perturbation v. One way to formulate the problem is calculate the time derivative of the transition probability. To simplify the analysis, we can turn on the interaction adiabatically $V_S(t) = ve^{\eta t}$ for $t < 0$. For infinitesimal η, and times near $t = 0$, we will have the full physical V and the correct transition rate. Taking the derivative of the transition probability will then give us the desired transition rate,

$$
\begin{aligned}
\Gamma_{n\to m} &= \frac{d}{dt}\left|-\frac{i}{\hbar}\int_{-\infty}^{t} dt' e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m + i\eta)t'}\langle m|v|n\rangle\right|^2\Bigg|_{t=0} \\
&= |\langle m|v|n\rangle|^2 \frac{d}{dt}\left(\frac{e^{\frac{i}{\hbar}(\epsilon_n - \epsilon_m - i\eta)t}}{\epsilon_n - \epsilon_m - i\eta}\frac{e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m + i\eta)t}}{\epsilon_n - \epsilon_m + i\eta}\right)\Bigg|_{t=0} \\
&= |\langle m|v|n\rangle|^2 \frac{d}{dt}\frac{e^{\frac{2\eta}{\hbar}t}}{(\epsilon_n - \epsilon_m)^2 + \eta^2}\Bigg|_{t=0} \\
&= |\langle m|v|n\rangle|^2 \frac{2}{\hbar}\frac{\eta}{(\epsilon_n - \epsilon_m)^2 + \eta^2}.
\end{aligned}
\tag{21.78}
$$

We need to evaluate the function of η above when $\eta \to 0$,

$$
\frac{\eta}{x^2 + \eta^2}
\tag{21.79}
$$

If $x \neq 0$, taking η to zero gives 0. However, if $x = 0$, we get $1/\eta$ which goes to infinity. This is the behavior of a constant times a delta function. We integrate over x to find the constant. We make the substitution $x = \eta \tan\theta$,

$$
\int_{-\infty}^{\infty} dx \frac{\eta}{x^2 + \eta^2} = \int_{-\pi/2}^{\pi/2} d\theta\, \eta \sec^2\theta \frac{\eta}{\eta^2\tan^2\theta + \eta^2} = \int_{-\pi/2}^{\pi/2} d\theta = \pi
\tag{21.80}
$$

or

$$
\Gamma_{n\to m} = \frac{2\pi}{\hbar}|\langle m|v|n\rangle|^2\delta(\epsilon_n - \epsilon_m)
\tag{21.81}
$$

which is called Fermi's golden rule.

---

[2]There is a subtlety here. We have seen that at short times, the transition probability grows quadratically with time, so that if an eigenstate of H_0 is really prepared at time 0, the transition rate will not be exponential for very small times. It turns out that for very long times an isolated system also does not have an exponential decay rate because of the possibility of the system making transitions back into the original state. For nearly all experimental situations, only the intermediate time period where the rate is constant is relevant.

The appearance of the delta function in Fermi's golden rule indicates that talking about a transition rate only makes sense if we integrate over either the states n or m. This means that there must be a continuum of states around our initial state for the system to be able to make a transition from our initial state into some other state. In our hydrogen decay problem, the photon energies form a continuum, so we must integrate over the final photon states to get the physical transition rate. The integral over the quantum number m times the delta function is called the density of states

$$\int dm\delta(\epsilon_n - \epsilon_m) \equiv \rho(\epsilon_n) \tag{21.82}$$

so sometimes Fermi's golden rule is written with a density of states factor rather than the delta function.

21.12 Scattering in the Born Approximation

Let's apply Fermi's golden rule to a scattering problem. Far from the target, our projectile is in a wave packet which, since it does not feel the target interaction, can be described by a free particle Hamiltonian. We can write the full Hamiltonian as

$$H = H_0 + V \tag{21.83}$$

where H_0 is the Hamiltonian for the target and projectile with no interaction between them. V is the target-projectile interaction.

In the physical scattering experiment we take the z axis as the direction of the incoming projectile so that its momentum is strongly peaked around $\boldsymbol{k}_0 = \hat{\boldsymbol{z}}k_0$. Our detectors detect particles scattered into a solid angle $d\Omega$ centered on the spherical angles θ and ϕ. With a beam of particles incident on our target, we find that rate of particles into our detector is proportional to the incident flux of particles, where the incident flux is the incident number per unit area per unit time. Experiments therefore report the ratio

$$\frac{d\sigma(\theta, \phi)}{d\Omega} = \frac{\text{Number of particles per unit time into the detector}}{\text{Number of particles per unit area per unit time incident}} = \frac{\text{Transition rate}}{\text{Incident flux}}. \tag{21.84}$$

This has units of area and is called the differential cross section. This area of the incident beam has the same number of particles per unit time through it that is detected in our detector at θ, ϕ subtending a solid angle $d\Omega$.

For definiteness let's look at the case of a projectile of mass m scattering from a potential $v(\boldsymbol{r})$. We have

$$H_0 = \frac{p^2}{2m} \tag{21.85}$$

with eigenstates $|\boldsymbol{k}\rangle$. We have derived our perturbation theory using states normalized to 1, so we take

$$\langle \boldsymbol{r}|\boldsymbol{k}\rangle = L^{-3/2}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{21.86}$$

where we place our system in a large periodic box of side L in order to obtain these normalized states. Our physical results must not depend on L for $L \to \infty$. The detector at θ, ϕ will detect outgoing particles with momentum

$$\boldsymbol{k}' = \sin\theta\cos\phi\hat{\boldsymbol{x}} + \sin\theta\sin\phi\hat{\boldsymbol{y}} + \cos\theta\hat{\boldsymbol{z}} \tag{21.87}$$

with θ and ϕ in the solid angle $d\Omega$. The transition rate from Fermi's golden rule is

$$\Gamma_{\boldsymbol{k}_0 \to \boldsymbol{k}'} = \frac{2\pi}{\hbar}|\langle \boldsymbol{k}'|v(r_{\text{op}})|\boldsymbol{k}_0\rangle|^2 \delta\left(\frac{\hbar^2 k_0^2}{2m} - \frac{\hbar^2 k'^2}{2m}\right) . \tag{21.88}$$

We need to sum over all the states that can go into our detector. The $\boldsymbol{k}' = \frac{2\pi}{L}(n_x\hat{\boldsymbol{x}} + n_y\hat{\boldsymbol{y}} + n_z\hat{\boldsymbol{z}})$. When $L \to \infty$ the state sum goes to an integral

$$\sum_{n_x,n_y,n_z} = \int dn_x dn_y dn_z = \frac{L^3}{(2\pi)^3}\int d^3k' = \frac{L^3}{(2\pi)^3}\int k'^2 dk' d\Omega \tag{21.89}$$

The incident current density $\boldsymbol{j}$ in the z direction is the number of particles per unit area per unit time, so it gives the incident flux. The incident current density in the z direction is (recall Baym chapter 3 problem 10)

$$\frac{1}{2m}\left[L^{-3/2}e^{-i\boldsymbol{k}_0\cdot\boldsymbol{r}}\frac{\hbar}{i}\frac{\partial}{\partial z}L^{-3/2}e^{i\boldsymbol{k}_0\cdot\boldsymbol{r}} - L^{-3/2}e^{i\boldsymbol{k}_0\cdot\boldsymbol{r}}\frac{\hbar}{i}\frac{\partial}{\partial z}L^{-3/2}e^{-i\boldsymbol{k}_0\cdot\boldsymbol{r}}\right] = \frac{\hbar k_0}{mL^3} . \tag{21.90}$$

Our differential cross section is

$$\frac{d\sigma(\theta,\phi)}{d\Omega} = \frac{mL^3}{\hbar k_0}\frac{2\pi}{\hbar}\frac{L^3}{(2\pi)^3}\int_0^\infty dk' k'^2 |\langle \boldsymbol{k}'|v(\boldsymbol{r}_{\text{op}})|\boldsymbol{k}_0\rangle|^2 \delta\left(\frac{\hbar^2 k_0^2}{2m} - \frac{\hbar^2 k'^2}{2m}\right) . \tag{21.91}$$

We can transform the delta function using

$$\delta(f(x)) = \frac{\delta(x - x_0)}{\left|\frac{df(x_0)}{dx_0}\right|} \tag{21.92}$$

so that

$$\delta\left(\frac{\hbar^2 k_0^2}{2m} - \frac{\hbar^2 k'^2}{2m}\right) = \frac{m}{\hbar^2 k_0}\delta(k' - k_0) \tag{21.93}$$

and

$$\frac{d\sigma(\theta,\phi)}{d\Omega} = \frac{m^2}{4\hbar^4\pi^2}L^6|\langle \boldsymbol{k}'|v(\boldsymbol{r}_{\text{op}})|\boldsymbol{k}_0\rangle|^2 . \tag{21.94}$$

The matrix element is

$$\begin{aligned}
\langle \boldsymbol{k}'|v(\boldsymbol{r}_{\text{op}})|\boldsymbol{k}_0\rangle &= \int d^3r \langle \boldsymbol{k}'|\boldsymbol{r}\rangle v(\boldsymbol{r})\langle \boldsymbol{r}|\boldsymbol{k}_0\rangle \\
&= \frac{1}{L^3}\int d^3r e^{-i(\boldsymbol{k}'-\boldsymbol{k}_0)\cdot\boldsymbol{r}}v(\boldsymbol{r}) = L^{-3}\tilde{v}(\boldsymbol{k}' - \boldsymbol{k}_0)
\end{aligned} \tag{21.95}$$

where $\tilde{v}(\boldsymbol{q})$ is the Fourier transform of $v(\boldsymbol{r})$.

Plugging this in to the cross section, all of the L factors cancel as expected, and we get

$$
\begin{aligned}
\frac{d\sigma(\theta, \phi)}{d\Omega} &= \frac{m^2}{4\hbar^4\pi^2}|\tilde{v}(\boldsymbol{k}' - \boldsymbol{k}_0)|^2 \\
&= \frac{m^2}{4\hbar^4\pi^2}|\tilde{v}(k_0[\sin\theta\cos\phi\hat{\boldsymbol{x}} + \sin\theta\sin\phi\hat{\boldsymbol{y}} + (\cos\theta - 1)\hat{\boldsymbol{z}}])|^2 \quad (21.96)
\end{aligned}
$$

Since this is the lowest order calculation, it gives the cross section in the Born approximation. Keeping higher order terms in the perturbation expansion will give higher order results. However, the Born approximation or slight modifications to it, is one of the most popular ways to use scattering to understand the properties of the target.

21.13   Born approximation for electron scattering

In condensed matter and biophysics, electrons and photons are the two most popular probes. We will discuss photon scattering in detail when we quantize electromagnetism. Let's look at electron scattering in the Born approximation. For a model problem, let's look at the scattering of an electron from a hydrogen atom. We will first assume that the projectile electron has an energy much higher than the atomic states of the hydrogen atom. In this case, as we will learn later, we can ignore the identity of the electrons (essentially, the amplitude for the high energy electron to have substantial overlap with the bound atomic state of the hydrogen electron is very small so that the exchange term coming from the identity of the particles can be neglected.) We can write the unperturbed Hamiltonian and the interaction between the target and projectile as

$$
\begin{aligned}
H_0 &= \frac{p_1^2}{2m_e} + \frac{p_2^2}{2m_e} + \frac{p_p^2}{2m_p} + -\frac{e^2}{|\boldsymbol{r}_2 - \boldsymbol{r}_p|} \\
V &= \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} - \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_p|} \quad (21.97)
\end{aligned}
$$

Let's look at scattering from the ground state first. In the lab frame we have the center of mass of the hydrogen atom at rest which means its momentum will be zero, the relative state is the $1s$ ground state, and the incoming projectile has momentum $\hbar\boldsymbol{k}_0$. We can write this state as $|\boldsymbol{k}_0, \boldsymbol{K} = 0, 1s\rangle$

After scattering electron 2 will be in a relative atomic state n, which will be the $1s$ state for the elastic case or some other state in the case of inelastic scattering. The electron and the center of mass momentum can change because of the the scattering. We can write this state as $|\boldsymbol{k}', \boldsymbol{K}', n\rangle$. The incident flux is calculated exactly as before and is $\frac{\hbar k_0}{m_e L^3}$. The cross section for scattering with the electron with outgoing momentum $\boldsymbol{k}'$ into our detector and

not measuring the center of mass of the target is

$$
\frac{d\sigma(\theta,\phi)_{1s\to n}}{d\Omega} = \frac{m_e L^3}{\hbar k_0}\frac{2\pi}{\hbar}\frac{L^3}{(2\pi)^3}\int_0^\infty dk'\frac{L^3}{(2\pi)^3}\int d^3K'|\langle \boldsymbol{k}_0, \boldsymbol{K}=0, 1s|V|\boldsymbol{k}', \boldsymbol{K}', n\rangle|^2 \cdot
$$
$$
\delta\left(\frac{\hbar^2 k_0^2}{2m_e} + \epsilon_{1s} - \frac{\hbar^2 k'^2}{2m_e} - \epsilon_n - \frac{\hbar^2 K'^2}{2(m_e+m_p)}\right) . \tag{21.98}
$$

Before continuing, notice that we have now reduced the calculation to doing integrals. We need to calculate the matrix elements we require, and we need to integrate over the projectile electron's momentum magnitude and the center mass momentum.

The matrix element is

$$
\langle \boldsymbol{k}_0, \boldsymbol{K}=0, 1s|V|\boldsymbol{k}', \boldsymbol{K}', n\rangle = L^{-6}\int d^3r \int d^3R \int d^3r_1 e^{-i(\boldsymbol{k}'-\boldsymbol{k}_0)\cdot\boldsymbol{r}_1}e^{i\boldsymbol{K}'\cdot\boldsymbol{R}}\psi_{1s}^*(\boldsymbol{r})\psi_n(\boldsymbol{r}) \cdot
$$
$$
\left[\frac{e^2}{\left|\boldsymbol{r}_1 - \boldsymbol{R} - \frac{m_p}{m_p+m_e}\boldsymbol{r}\right|} - \frac{e^2}{\left|\boldsymbol{r}_1 - \boldsymbol{R} + \frac{m_e}{m_p+m_e}\boldsymbol{r}\right|}\right] \tag{21.99}
$$

The Fourier transform of the coulomb interaction (see the appendix of the time independent perturbation theory notes or just calculate it) is

$$
\int d^3r \frac{e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}}{r} = \frac{4\pi}{q^2} \tag{21.100}
$$

and writing $\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{k}_0$, we get

$$
\int d^3r \frac{e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}}{|\boldsymbol{r}-\boldsymbol{a}|} = e^{i\boldsymbol{q}\cdot\boldsymbol{a}}\int d^3r \frac{e^{-i\boldsymbol{q}\cdot(\boldsymbol{r}-\boldsymbol{a})}}{|\boldsymbol{r}-\boldsymbol{a}|}
$$
$$
= e^{i\boldsymbol{q}\cdot\boldsymbol{a}}\int d^3r \frac{e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}}{r} = e^{i\boldsymbol{q}\cdot\boldsymbol{a}}\frac{4\pi}{q^2} . \tag{21.101}
$$

The matrix elements become

$$
\langle \boldsymbol{k}_0, \boldsymbol{K}=0, 1s|V|\boldsymbol{k}', \boldsymbol{K}', n\rangle =
$$
$$
L^{-6}\frac{4\pi e^2}{q^2}\int d^3R\, e^{i(\boldsymbol{q}+\boldsymbol{K}')\cdot\boldsymbol{R}}\int d^3r\left[e^{i\frac{m_p}{m_p+m_e}\boldsymbol{q}\cdot\boldsymbol{r}} - e^{-i\frac{m_e}{m_p+m_e}\boldsymbol{q}\cdot\boldsymbol{r}}\right]\psi_{1s}^*(\boldsymbol{r})\psi_n(\boldsymbol{r}) . \tag{21.102}
$$

Before going on, let's deal with the center of mass part. There are two ways to handle the integral. The simplest way is to remember that we are in an extremely large but finite box of side L, with discrete momentum states. Given the $\boldsymbol{q}$ state, we must have $\boldsymbol{K}' = -\boldsymbol{q}$ for the $\boldsymbol{R}$ integral to be nonzero. The $\boldsymbol{R}$ integral over the (with their L^{-3} factor) then gives a Kronecker delta $\delta_{\boldsymbol{q},\boldsymbol{K}'}$, where the Kronecker delta tells us that the integer quantum numbers for the momentum must match. The square of a Kronecker delta, is the Kronecker delta,

so squaring the matrix element doesn't change it. The center of mass wave vector integral should now be rewritten as the original sum. Doing the sum just picks out the $\boldsymbol{K'} = -\boldsymbol{q}$ term. An alternative method is to work in the $L \to \infty$ limit in which case, the integral

$$\int d^3 R e^{i(\boldsymbol{q}+\boldsymbol{K'})\cdot\boldsymbol{R}} = (2\pi)^3 \delta^3(\boldsymbol{q} + \boldsymbol{K'}) . \tag{21.103}$$

Squaring the matrix element will then give a delta function squared. We now need to make sure we know how to interpret a delta function squared. What we have to do is go back to the limit and write

$$|\delta(k_x)|^2 = \frac{1}{2\pi}\int_{-L/2}^{L/2} dx e^{-ik_x x} \frac{1}{2\pi}\int_{-L/2}^{L/2} dx' e^{-ik_x x'} \tag{21.104}$$

The sleazy math method is to do one integral in the limit of $L \to \infty$,

$$\frac{1}{2\pi}\int_{-L/2}^{L/2} dx e^{-ik_x x} = \delta(k_x) \tag{21.105}$$

and then notice that the second integral will be zero unless evaluated at $k_x = 0$, so that it gives a contribution of $L/2\pi$. We get the result

$$|\delta(k_x)|^2 = \frac{L}{2\pi}\delta(k_x) \tag{21.106}$$

More rigorous mathematics requires that we analyze the product carefully. The integrals can be done explicitly to give

$$\int_{-L/2}^{L/2} dx e^{-ik_x x} \int_{-L/2}^{L/2} dx' e^{-ik_x x'} = \frac{4\sin^2(k_x L/2)}{k_x^2} . \tag{21.107}$$

This function becomes strongly peaked around $k_x = 0$ as $L \to \infty$. For $k_x = 0$, it goes to infinity as $L \to \infty$. Therefore it has the properties of a delta function. The coefficient can be calculated by integrating,

$$\frac{1}{2\pi}\int_{-L/2}^{L/2} dx e^{-ik_x x} \frac{1}{2\pi}\int_{-L/2}^{L/2} dx e^{-ik_x x} = \frac{\sin^2(\alpha T/2)}{\pi^2 \alpha^2} \to \frac{L}{2\pi}\delta(k_x) \tag{21.108}$$

exactly as calculated using sleazy math.

The main result is that momentum is conserved. *We did not have to put this into the calculation.* It drops out by itself. Of course if you understand the physics so that you know the momentum will be conserved, you can make sure to arrange to have it drop out early. That is wy we did not start with the $\boldsymbol{r}$ integration in the matrix element.

Our differential cross section is now

$$\frac{d\sigma(\theta,\phi)_{1s\to n}}{d\Omega} = \frac{m_e}{4\pi^2\hbar^4 k_0}\int_0^\infty dk' k'^2 \left(\frac{4\pi e^2}{q^2}\right)^2 \left|\int d^3 r \left[e^{i\frac{m_p}{m_p+m_e}\boldsymbol{q}\cdot\boldsymbol{r}} - e^{-i\frac{m_e}{m_p+m_e}\boldsymbol{q}\cdot\boldsymbol{r}}\right]\psi_{1s}^*(\boldsymbol{r})\psi_n(\boldsymbol{r})\right|^2 .$$
$$\delta\left(\frac{\hbar^2 k_0^2}{2m_e} + \epsilon_{1s} - \frac{\hbar^2 k'^2}{2m_e} - \epsilon_n - \frac{\hbar^2|\boldsymbol{k'}-\boldsymbol{k_0}|^2}{2(m_e+m_p)}\right) . \tag{21.109}$$

Integrating over the k' magnitude removes the delta function. Let's define a transition density as $\rho_{nm} = \psi_n^*(\boldsymbol{r})\psi_m(\boldsymbol{r})$. It's Fourier tranform is

$$\tilde{\rho}_{nm}(\boldsymbol{k}) = \int d^3r\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\psi_n^*(\boldsymbol{r})\psi_m(\boldsymbol{r}) \,. \tag{21.110}$$

Our result is

$$\frac{d\sigma(\theta,\phi)_{1s\to n}}{d\Omega} = \frac{m_e}{4\pi^2\hbar^4 k_0} \frac{k'^2}{\frac{\hbar^2 k'}{m_e} + \frac{\hbar^2}{2(m_e+m_p)}(k' - \hat{\boldsymbol{k}}'\cdot\boldsymbol{k}_0)} \cdot$$
$$\cdot \left(\frac{4\pi e^2}{q^2}\right)^2 \left|\rho_{1s,n}\left(-\frac{m_p}{m_p+m_e}\boldsymbol{q}\right) - \rho_{1s,n}\left(\frac{m_e}{m_p+m_e}\boldsymbol{q}\right)\right|^2 \tag{21.111}$$

where $\boldsymbol{q} = \boldsymbol{k}' - \boldsymbol{k}_0$, and the magnitude of k' has been forced by the delta function to be the value that conserves energy.

Even for the hydrogen atom, the nuclear mass is much larger than the electron mass, so the density term can be very well approximated by $|\rho_{1s,m}(\boldsymbol{q}) - \rho_{1s,m}(0)|^2$. Orthogonality shows that $\rho_{nm}(0) = \delta_{n,m}$. For elastic scattering, the factor inside the magnitude squared, is the Fourier transform of the charge density divided by e. Therefore, we can use scattering to find the magnitude of the Fourier transform of the charge density. We have lost the phase, so to invert to get back the charge density requires the use of a phasing algorithm. One way to do this for an isolated system is to have a bound on its size, so that we know that the charge density is zero outside of that bound. By enforcing the two constraints that the magnitude of the Fourier transform must match that measured by scattering, and the charge density is zero outside the bound, it is possible to solve the phase problem and obtain the the charge density from scattering. There are other ways to phase the data, for example by adding a strong scattering center and using the interference from it to find the phases.

Our Born approximation also gives us inelastic scattering by including the excited states in the matrix element. We do not need to have any new calculational methods. Inelastic scattering allows us to probe the excitations of the target and measure the excitation energies from the energy loss of our projectile.

One of the neat things about the quantum mechanical result is that all of the conservation laws and kinematics are built in. Momentum and energy are automatically conserved in the result. Elastic and inelastic scattering come out naturally without any additional formalism.

Using higher order perturbation theory, we can calculate more accurate cross sections.

21.14 Transitions from harmonic perturbations

We already calculated the effect of a weak harmonic perturbation. Here we only need to extend this to the case where the perturbation can induce transitions. We take the perturbation for $t < 0$ to be

$$V_S(t) = v\cos(\omega t)e^{\eta t} \,. \tag{21.112}$$

The lowest order transition rate at time t is

$$
\begin{aligned}
\Gamma_{n \to m} &= \frac{d}{dt} \left| -\frac{i}{2\hbar} \int_{-\infty}^{t} dt' \left[e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m + \hbar\omega + i\eta)t'} + e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m - \hbar\omega + i\eta)t'} \right] \langle m|v|n \rangle \right|^2 \\
&= \frac{1}{4} |\langle m|v|n \rangle|^2 \frac{d}{dt} \left(\left[\frac{e^{\frac{i}{\hbar}(\epsilon_n - \epsilon_m + \hbar\omega - i\eta)t}}{\epsilon_n - \epsilon_m + \hbar\omega - i\eta} + \frac{e^{\frac{i}{\hbar}(\epsilon_n - \epsilon_m - \hbar\omega - i\eta)t}}{\epsilon_n - \epsilon_m - \hbar\omega - i\eta} \right] \cdot \right. \\
&\qquad \left. \left[\frac{e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m + \hbar\omega + i\eta)t}}{\epsilon_n - \epsilon_m + \hbar\omega + i\eta} + \frac{e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m - \hbar\omega + i\eta)t}}{\epsilon_n - \epsilon_m - \hbar\omega + i\eta} \right] \right) \\
&= \frac{1}{4} |\langle m|v|n \rangle|^2 \frac{d}{dt} \left[\frac{e^{\frac{2\eta}{\hbar}t}}{(\epsilon_n - \epsilon_m + \hbar\omega)^2 + \eta^2} + \frac{e^{\frac{2\eta}{\hbar}t}}{(\epsilon_n - \epsilon_m - \hbar\omega)^2 + \eta^2} \right. \\
&\qquad \left. + \frac{e^{-i2\omega t + \frac{2\eta}{\hbar}t}}{(\epsilon_n - \epsilon_m - \hbar\omega - i\eta)(\epsilon_n - \epsilon_m + \hbar\omega + i\eta)} + \frac{e^{i2\omega t + \frac{2\eta}{\hbar}t}}{(\epsilon_n - \epsilon_m + \hbar\omega + i\eta)(\epsilon_n - \epsilon_m - \hbar\omega - i\eta)} \right]
\end{aligned}
$$

$$(21.113)$$

To measure the terms that oscillate at frequency $2\omega t$, we would have to be able to measure the change in the transition rate at different points during the oscillating fields. I know of no experiments that measure transition rates with this type of accuracy. I also don't know any reason why such a measurement would be interesting. Instead the time averaged transition rate is measured. Averaging over a cycle of the perturbation, we drop these oscillating terms. The remaining two terms give delta functions as before

$$
\Gamma = \frac{\pi}{2\hbar} |\langle m|v|n \rangle|^2 \left[\delta(\epsilon_n - \epsilon_m + \hbar\omega) + \delta(\epsilon_n - \epsilon_m - \hbar\omega) \right] . \tag{21.114}
$$

The oscillating field causes transitions with energy differences $\pm\hbar\omega$. The factor of $\pi/2\hbar$ rather than $2\pi/\hbar$ that we had with the usual Fermi's golden rule comes from the factor of half when writing $\cos(\omega t)$ as two exponentials.

21.15   Higher order transition rates

The simplicity of the perturbation expansion Eq. 21.49 allows us to calculate the higher order terms rather easily. We turn on our perturbation adiabatically $V_S(t) = v e^{\eta t}$ for $t < 0$. Inserting complete sets of eigenstates, $|m\rangle$, of H_0, we have the amplitude for finding a system

252

at time $t < 0$ in state $|m\rangle$ if it is initially in the state $|n\rangle$ as

$$
\begin{aligned}
{}_I\langle m|\psi(t)\rangle_I &= \langle m|\left[1 + \left(-\frac{i}{\hbar}\right)\int_{-\infty}^{t} dt' V_I(t') + \left(-\frac{i}{\hbar}\right)^2 \int_{-\infty}^{t} dt' \int_{t_0}^{t'} dt'' V_I(t')V_I(t'')\right.\\
&\quad \left. + \left(-\frac{i}{\hbar}\right)^3 \int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt'' \int_{-\infty}^{t''} dt''' V_I(t')V_I(t'')V_I(t''') + ...\right]|n\rangle\\
&= \delta_{mn} + \left(-\frac{i}{\hbar}\right)\int_{-\infty}^{t} dt' e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m + i\eta)t}\langle m|v|n\rangle\\
&\quad + \left(-\frac{i}{\hbar}\right)^2 \int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt'' \sum_k \langle m|v|k\rangle e^{-\frac{i}{\hbar}(\epsilon_k - \epsilon_m + i\eta)t'}\langle k|v|n\rangle e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_k + i\eta)t''}\\
&\quad + \left(-\frac{i}{\hbar}\right)^3 \int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt'' \int_{-\infty}^{t''} dt''' \sum_{kj} \langle m|v|k\rangle \cdot\\
&\quad e^{-\frac{i}{\hbar}(\epsilon_k - \epsilon_m + i\eta)t'}\langle k|v|j\rangle e^{-\frac{i}{\hbar}(\epsilon_j - \epsilon_k + i\eta)t''}\langle j|v|n\rangle e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_j + i\eta)t'''} + ...
\end{aligned}
\tag{21.115}
$$

Performing the integrals, the η in the exponent makes the lower limit give zero. The upper limit evaluates the exponential at the same time as the next exponential. The intermediate state energies then cancel in the exponent. We get

$$
\begin{aligned}
{}_I\langle m|\psi(t)\rangle_I &= \delta_{mn} + \left(-\frac{i}{\hbar}\right)\int_{-\infty}^{t} dt' e^{-\frac{i}{\hbar}(\epsilon_n - \epsilon_m + i\eta)t}\left[\langle m|v|n\rangle + \sum_j \frac{\langle m|v|j\rangle\langle j|v|n\rangle}{\epsilon_n - \epsilon_j + i\eta}\right.\\
&\quad \left. + \sum_{jk} \frac{\langle m|v|j\rangle\langle j|v|k\rangle\langle k|v|n\rangle}{(\epsilon_n - \epsilon_j + i\eta)(\epsilon_n - \epsilon_k + i\eta)} + ...\right].
\end{aligned}
\tag{21.116}
$$

This is the Fermi's golden rule result except that $\langle m|v|n\rangle$ has been replaced by the expression (which looks a lot like our time independent perturbation expansion) in brackets.

For a concise notation, we define the T matrix (T matrix stands for transition matrix) as the operator that gives the expression,

$$
T(E) = v + \sum_j \frac{v|j\rangle\langle j|v}{E - \epsilon_j + i\eta} + \sum_{jk} \frac{v|j\rangle\langle j|v|k\rangle\langle k|v}{(E - \epsilon_j + i\eta)(E - \epsilon_k + i\eta)} + ...
\tag{21.117}
$$

and the exact transition rate is

$$
\Gamma_{n\to m} = \frac{2\pi}{\hbar}|\langle n|T(\epsilon_n)|m\rangle|^2 \delta(\epsilon_n - \epsilon_m).
\tag{21.118}
$$

For cases where v does not couple $|n\rangle$ and $|m\rangle$ we must go to second order order or more to find the lowest order nonzero transition rate. We can also rewrite the T matrix series to give the equation

$$
T = v + \sum_j \frac{v|j\rangle\langle j|T}{E - \epsilon_j + i\eta}.
\tag{21.119}
$$

Iterating gives our series expansion. For strong perturbations, this equation can be solved directly as, for example, a matrix equation for the T matrix to find the transition rates nonperturbatively.

21.16 Problems

Problem 21.1. A hydrogen atom in an initial electronic eigenstate $|n\rangle$ has its center of mass at the origin. A charged particle with charge Ze moves in a straight line so that its position is

$$\boldsymbol{R}(t) = vt\hat{\boldsymbol{z}} + b\hat{\boldsymbol{x}}.\qquad(21.120)$$

You should assume that both the charged particle and the hydrogen's proton are so heavy that neither is deflected. We can then view the electromagnetic field from the charged particle as a time-dependent perturbation on the hydrogen atom electronic Hamiltonian H_0.

(a) Show that when the impact parameter, b, is much greater than the size of the hydrogen atom the lowest order (in Ze) interaction representation matrix element of the perturbation

$$\int_{-\infty}^{\infty} dt\langle j|V_I(t)|n\rangle$$

$$= \int_{-\infty}^{\infty} dt\langle j|e^{\frac{i}{\hbar}H_0 t}\left[\frac{e}{2mc}\boldsymbol{p}\cdot\boldsymbol{A}(\boldsymbol{r},t) + \frac{e}{2mc}\boldsymbol{A}(\boldsymbol{r},t)\cdot\boldsymbol{p} - e\Phi(\boldsymbol{r},t)\right]e^{-\frac{i}{\hbar}H_0 t}|n\rangle$$

$$(21.121)$$

reduces to

$$\int_{-\infty}^{\infty} dt\langle j|V_I(t)|n\rangle = e\langle j|\boldsymbol{r}|n\rangle\cdot\tilde{\boldsymbol{E}}\left(\frac{E_j - E_n}{\hbar}\right)\qquad(21.122)$$

where

$$\tilde{\boldsymbol{E}}(\boldsymbol{\omega}) = \int_{-\infty}^{\infty} dt e^{i\omega t}\boldsymbol{E}(0,t)\qquad(21.123)$$

with $\boldsymbol{E}(0,t)$ the charged particle's electric field at the origin (i.e. the proton position). Since we are interested in transitions, the state $|j\rangle$ is not equal to $|n\rangle$. Hint: The commutator $\frac{im}{\hbar}[H_0,\boldsymbol{r}] = \boldsymbol{p}$ is useful.

(b) Calculate the probability that the hydrogen atom makes a transition from its ground state $n = 1, \ell = 0$ to an $n = 2, \ell = 1$ state as a function of v and b. Your result should be

$$P = \frac{2^{11}}{3^8}\frac{Z^2 e^8}{\gamma^2 v^4 \hbar^4}\left[K_1^2\left(\frac{3}{8}\frac{e^4 mb}{\gamma v \hbar^3}\right) + \frac{1}{\gamma^2}K_0^2\left(\frac{3}{8}\frac{e^4 mb}{\gamma v \hbar^3}\right)\right]\qquad(21.124)$$

where K_0 and K_1 are modified Bessel functions and $\gamma = (1 - v^2/c^2)^{-1/2}$.[3]

---

[3]I found Eq. 9.6.25 in Abramowitz and Stegun useful

$$K_\nu(xz) = \frac{\Gamma(\nu + 1/2)(2z)^\nu}{\pi^{1/2}x^\nu}\int_0^\infty \frac{\cos(xt)dt}{(t^2 + z^2)^{\nu+1/2}}.$$

Problem 21.2.   a. In J.D. Jackson, Classical Electrodynamics, Chapter 13, problems 13.2 and 13.3, he describes how to calculate the energy transferred to a classical particle of mass m bound to the origin with a harmonic oscillator potential of angular frequency ω_0 when a charged particle like that of our problem 21.1 passes by. (He uses a harmonic oscillator because it gives a final result like quantum mechanics.) The result is

$$\Delta E = \frac{2Z^2 e^4}{mb^2 v^2} \xi^2 \left[K_1^2(\xi) + \frac{1}{\gamma^2} K_0^2(\xi) \right] , \qquad (21.125)$$

where $\xi = \omega_0 b / \gamma v$, $\gamma = (1 - \frac{v^2}{c^2})^{-1/2}$.

Calculate the expectation value of the energy transferred to a quantum harmonic oscillator (that is the sum over all states of the probability of making a particular transition times the energy transfer of that transition), initially in an energy eigenstate state, and compare with the classical result.

  b. Show that using the same approximations as in problem 21.1, that if $(E_n - E_j)m/\hbar\gamma v \ll 1$ for all matrix elements $\langle j | \boldsymbol{r} | n \rangle$ that are important in the sum, the expectation value of the energy transfer to an arbitrarily bound electron (that is one not necessarily in a harmonic oscillator or hydrogen atom, but any potential) goes to $\Delta E = 2Z^2 e^4/mb^2 v^2$. This is the classical result for $\Delta p^2/2m$ where Δp is the impulse (c.f. Jackson problem 13.1)

Problem 21.3. A hydrogen atom is in its ground state. An electromagnetic field from a laser of angular frequency ω is applied to the atom. Assume the field is well described classically so that the simplest way to include it is to take $\Phi = 0$, and a vector potential

$$\boldsymbol{A} = \hat{\boldsymbol{z}} \frac{c\mathcal{E}}{\omega} \cos\left[\frac{\omega}{c}(x - ct)\right] \qquad (21.126)$$

  a. Verify that this corresponds to a time-harmonic electric field with a peak magnitude $\mathcal{E}$. What is the corresponding magnetic field?

  b. Calculate the ionization rate to a continuum state, assuming that the continuum states can be approximated by free particle states (i.e. plane waves). Do not assume that the proton is infinitely massive.

  c. Now assume that $\frac{\omega a_0}{c} \ll 1$ Repeat your calculation using the correct hydrogenic continuum states. Show that in the appropriate limit your results agree with the free particle final state result of part b. (Hint: use the integral form for the radial functions Eq. 17.30 and perform the r integration first.)

  d. In the long wave length limit (which is often called neglecting retardation or the dipole approximation) of part c., the magnetic field contributions can be neglected. Show that using the result for a constant in space electric field, $\Phi = z\mathcal{E}\cos(\omega t)$, gives the same result as part c. when the correct hydrogenic continuum states are used. Show that using this gauge and approximating the final state by a plane wave gives a matrix

element that does not agree with the correct result in any limit. (The matrix element differs by a factor of 2.)

A uniform in space harmonic in time electric field is applied

$$\boldsymbol{\mathcal{E}} = \hat{\boldsymbol{z}}\mathcal{E}\cos(\omega t)\,, \tag{21.127}$$

with $\hbar\omega$ greater than the electron binding energy[4]. Do not assume that the proton is infinitely massive. binding energy.

a. Calculate the ionization rate to a continuum state, assuming that the continuum states can be approximated by free particle states (i.e. the plane waves).

b. Repeat your calculation using the correct hydrogenic continuum states. Show that, for large ω, your result goes over to that of part a. (Hint: use the integral form for the radial functions Eq. 17.30 and perform the r integration first.)

---

[4]Realize that such a field is not possible in a source free region. The correct solutions to Maxwell's equations without sources are linear combinations of plane waves, and the electric field must be accompanied by a magnetic field. The only way to avoid this would be to add a uniform current density such that it cancels the $\partial_t \boldsymbol{calE}$ term in the magnetic field curl equation. However, the particles that make up this current would interact with the hydrogen atom. Instead, we are really approximating the spatial variation of the plane wave $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ by a constant over the hydrogen atom. This means that the wave length must be much longer than a Bohr radius or $\omega a_0/c \ll 1$.

Chapter 22

Potential Scattering

22.1 Introduction

We have already looked at scattering using Fermi's golden rule. If we include higher order terms, that method becomes completely equivalent to the methods that we will now study. For scattering in the slab geometry (that is 1-dimensional scattering) we began with the experimental conditions, constructed an initial wave packet for the beam, and expanded the packet in eigenstates of the Hamiltonian. We chose these eigenstates to have the boundary condition of no incoming wave from positive infinity since the beam had no such components. We were then able to read off the transmission and reflection amplitudes and calculate the transmission and reflection probabilities for a packet with well defined energy from the corresponding eigenstate of the Hamiltonian.

We now want to repeat this for scattering in three-dimensions, that is, from a localized target.

The simplest nontrivial system is a single particle scattering from a potential which is zero far from the origin,

$$H = \frac{p^2}{2m} + v(\boldsymbol{r}) \tag{22.1}$$

and $v(\boldsymbol{r}) = 0$ for $|\boldsymbol{r}| > R$, where R can be large but is finite. Note this Hamiltonian can also describe the scattering of two particles in the center of mass frame, interacting via $v(\boldsymbol{r})$, with $\boldsymbol{r}$ the relative coordinate, and m the reduced mass of the pair.

We know, in general, how to proceed. Our beam produces a wave packet, and if we remove the target, this packet would be a free particle packet which we know how to propagate in time. We can, in principle, solve for the eigenstates of our Hamiltonian above choosing *any* boundary condition, as long as these are enforced at a distance much farther from the origin than where our experiment takes place. We then expand our wave packet in these eigenstates, multiply each by its $e^{-\frac{i}{\hbar}E_n t}$ factor and look at the result for t large enough that the scattering has taken place. To calculate the cross section we need to calculate the probability of the beam particle ending up at the detector having been scattered through some angle.

Just as in the slab geometry, we will find that we can greatly simplify the calculation by choosing the boundary conditions appropriately.

The biggest difference between this scattering from a localized target and the slab (one-dimensional) scattering is that we cannot divide space into two parts like we did for the slab. Instead we need to think about the eigenstates of H solutions that best correspond to the experiment.

22.2 Energy eigenstates

In order to specify completely the solution of a differential equation we must also specify the boundary conditions so that the integration constants can be determined. To do this generally, it is often simpler to convert the differential equation into an integral equation where the integration constants are included directly. It is convenient to write the energy as

$$E_k = \frac{\hbar^2 k^2}{2m} \tag{22.2}$$

and rearranging a little, the differential equation for the energy eigenfunctions becomes

$$\left[\nabla^2 + k^2\right] \psi_{\boldsymbol{k}}(\boldsymbol{r}) = \frac{2m}{\hbar^2} v(\boldsymbol{r}) \psi_{\boldsymbol{k}}(\boldsymbol{r}) \,. \tag{22.3}$$

Notice this looks like Helmholtz equation with a source term on the right. Of course, here the source depends on $\psi_{\boldsymbol{k}}(\boldsymbol{r})$ so that it is not quite as simple as the case where the source is completely known. However, we can attack the problem the same way. If we view the source as fixed, and had a particular solution to that equation, we could add to it solutions to the homogeneous equation

$$\left[\nabla^2 + k^2\right] \phi_{\boldsymbol{k}}(\boldsymbol{r}) = 0 \tag{22.4}$$

to set the boundary conditions. The solutions to the homogeneous equation are just the free particle solutions, i.e. plane waves $\phi_{\boldsymbol{k}} = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$.

One way to produce a particular solution is to use the Green's function for the Helmholtz equation

$$\left[\nabla^2 + k^2\right] G(\boldsymbol{r}, \boldsymbol{r}') = \delta^3(\boldsymbol{r} - \boldsymbol{r}') \tag{22.5}$$

which we will solve giving explicit forms in the next section. With a known Green's function, we can write the general solution as

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \phi_{\boldsymbol{k}}(\boldsymbol{r}) + \frac{2m}{\hbar^2} \int d^3r' G(\boldsymbol{r}, \boldsymbol{r}') v(\boldsymbol{r}') \psi(\boldsymbol{r}') \,. \tag{22.6}$$

You can verify that this is a solution by operating on it with $\nabla^2 + k^2$. The ϕ term gives zero, and bringing the derivative inside the integral, $\nabla^2 + k^2$ operating on the Green's function gives the delta function which can be integrated immediately to give the original differential equation.

Notice that if $v = 0$, our solution would be $\psi = \phi$. Since we would like to have our solution go over to the corresponding plane wave in this limit, we would like to choose the boundary conditions on the Green's function so that $\phi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$. In general, for regions outside the potential, we expect a general Green's function to include both outgoing and incoming waves. However, if the Green's function includes incoming waves, we would have to modify the ϕ to cancel them if we want the incoming waves to look like those of a plane wave. The simplest choice that will lead to the states of most physical interest will be to require the Green's function to have only outgoing waves. In this case we can write solutions

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \frac{2m}{\hbar^2} \int d^3 r' G(\boldsymbol{r}, \boldsymbol{r}') v(\boldsymbol{r}') \psi_{\boldsymbol{k}}(\boldsymbol{r}') \tag{22.7}$$

where G contributes only outgoing waves.

22.3 Helmholtz Green's function

Since this appears repeatedly in electromagnetism, you no doubt have derived it before[1]. Probably the simplest way to find the Green's function is to change variables to $\boldsymbol{R} = \boldsymbol{r} - \boldsymbol{r}'$, so that the origin of $\boldsymbol{R}$ is at $\boldsymbol{r}'$. The resulting equation only depends on $\boldsymbol{R}$ so we can write

$$\left[\nabla_R^2 + k^2\right] G(\boldsymbol{R}) = \delta^3(\boldsymbol{R}) \tag{22.8}$$

This equation is rotationally invariant. We can therefore write the solution as a linear combination of radial functions times spherical harmonics. The delta function is spherically symmetric, so the radial functions for $\ell > 0$ just satisfy the homogeneous Helmholtz equation and can be dropped since those parts could be added when we add the homogeneous solution to the whole differential equation[2]. We therefore can take $G(\boldsymbol{R}) = G(R)$ to be a radial function. Writing out ∇^2, we get for $R \neq 0$,

$$\frac{1}{R}\frac{d^2}{dR^2} RG(R) + k^2 G(R) = 0 \tag{22.9}$$

or

$$\frac{d^2}{dR^2} RG(R) = -k^2 RG(R) \tag{22.10}$$

which has the general solution

$$RG(R) = Ae^{ikR} + Be^{-ikR} \tag{22.11}$$

---

[1]See for example. J.D. Jackson's book[2], where he derives the Helmholtz Green's function three different ways in section 6.4, section 9.6, and section 12.11.

[2]More rigorously, the radial functions that are well behaved at the origin and which give the solutions of the Helmholtz equation for $\ell > 0$, will be proportional to $j_\ell(kR) = \frac{1}{2}[h^{(1)}(kR) + h^{(2)}(kR)]$ and are a mixture of incoming and outgoing waves. Our final boundary condition restricting ourselves to outgoing waves only, will require that we drop these terms.

or

$$G(R) = A\frac{e^{ikR}}{R} + B\frac{e^{-ikR}}{R}\,. \tag{22.12}$$

We can enforce the delta function condition by using the divergence theorem

$$\int_V d^3R\nabla^2 G(R) = \int_S d\boldsymbol{S}\cdot\nabla G(R) = \int_V d^3R\left[\delta^3(\boldsymbol{R}) - k^2 G(R)\right]\,. \tag{22.13}$$

Taking the volume to be a sphere of radius $a \to 0$, the integral over $k^2 G(R)$ goes to zero. For small R, $G(R \to 0) \to \frac{A+B}{R}$, and the integral gives

$$-4\pi(A + B) = 1\,. \tag{22.14}$$

Alternatively, you can notice that for any $k > 0$, for small R, the Greens function goes to $(A + B)/R$, and if you recall that the coulomb Green's function solves $\nabla^2 r^{-1} = -4\pi\delta^3(\boldsymbol{r})$ you can write Eq. 22.14 immediately.

Any choice of A and B that satisfy Eq. 22.14 can be used.

The Green's function with only outgoing waves will contain only the e^{ikR} term. Our choice in the previous section leads us to choose

$$G(R) = -\frac{e^{ikR}}{4\pi R} \tag{22.15}$$

or

$$G(\boldsymbol{r},\boldsymbol{r}') = -\frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{4\pi|\boldsymbol{r}-\boldsymbol{r}'|} \tag{22.16}$$

22.4   Lippmann-Schwinger equation

The integral equation that we have derived is a version of the Lippmann-Schwinger equation,

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{m}{2\pi\hbar^2}\int d^3r'\frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}v(\boldsymbol{r}')\psi_{\boldsymbol{k}}(\boldsymbol{r}') \tag{22.17}$$

If v is not too large, you could solve this equation by iteration. That is ignoring v, $\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$. You then plug that into the right hand side of Eq. 22.17 to get the next approximation for $\psi_{\boldsymbol{k}}(\boldsymbol{r})$. Continuing this, if it converges, will give a solution. Numerical solutions often change the integration into a sum over a set of points of $\boldsymbol{r}_j$ with weights. You need to use a little care to have an integration formula that can properly integrate the $1/|\boldsymbol{r}-\boldsymbol{r}'|$ singularity[3], but with that in hand, you will get a set of linear equations of the form

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}_j) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}_j} + \sum_n G_{jn}v(\boldsymbol{r}_n)\psi_{\boldsymbol{k}}(\boldsymbol{r}_n) \tag{22.18}$$

---

[3]There is no real singularity as you can see if you shift the origin of $\boldsymbol{r}'$ to $\boldsymbol{r}$. The $r'^2 dr'$ factor in the volume element will cancel the r'^{-1} divergence of the integrand.

which can be solved by standard matrix techniques[4].

The point here is that we can solve this equation by various standard techniques. What we want to show in the next sections is how to extract scattering data from the solution.

22.5 The large r form of the solutions

Since initially the beam's wave packet will be far from the target as will be the detector, we want to examine the behavior of the solutions to the Lippmann-Schwinger equation in the limit of large r. Notice that in the integral, $\boldsymbol{r}'$ only contributes if the potential is not zero. Therefore, r' will be of order the size of the target, while r is much larger. We can expand in powers of r'/r, and drop terms that go to zero in the limit of $r \to \infty$,

$$
\begin{aligned}
|\boldsymbol{r} - \boldsymbol{r}'| &= \sqrt{r^2 + r'^2 - 2\boldsymbol{r} \cdot \boldsymbol{r}'} = r\sqrt{1 + \frac{r'^2}{r^2} - 2r^{-1}\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} \\
&= r\left[1 + \frac{1}{2}\left(\frac{r'^2}{r^2} - 2r^{-1}\hat{\boldsymbol{r}} \cdot \boldsymbol{r}' + ...\right)\right] \\
&= r - \hat{\boldsymbol{r}} \cdot \boldsymbol{r}'.
\end{aligned}
\tag{22.19}
$$

We therefore can write

$$
\lim_{r \to \infty} \frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} = \frac{e^{ikr-ik\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'}}{r - \hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} = \frac{e^{ikr}}{r}e^{-ik\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'},
\tag{22.20}
$$

where in the last line, we can drop the $\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{r}}'$ in the denominator since it is much smaller than r. We cannot drop it in the phase factor, since, even though it is much smaller than r, it will change the phase unless $k\hat{\boldsymbol{r}} \cdot \boldsymbol{r}' \ll 1$. Whenever you expand phase factors, it is not enough for one factor to be much smaller than another; a factor can only be neglected if it is small compared to 1.

At large r, the solution to the Lippman-Schwinger equation has the form

$$
\psi_{\boldsymbol{k}}(\boldsymbol{r} \to \infty) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \frac{e^{ikr}}{r}f_{\boldsymbol{k}}(\hat{\boldsymbol{r}})
\tag{22.21}
$$

with

$$
f_{\boldsymbol{k}}(\hat{\boldsymbol{r}}) = -\frac{m}{2\pi\hbar^2}\int d^3r' e^{-ik\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'}v(\boldsymbol{r}')\psi_{\boldsymbol{k}}(\boldsymbol{r}').
\tag{22.22}
$$

$f_{\boldsymbol{k}}(\hat{\boldsymbol{r}})$ is called the scattering amplitude.

As we had constructed it, our eigenfunctions have plane waves plus the additional outgoing waves from the scattering.

---

[4]For example, you can define $M_{jn} \equiv \delta_{jn} - G_{jn}v(\boldsymbol{r}_n)$ so that the Lippmann-Schwinger equation becomes $\sum_n M_{jn}\psi_{\boldsymbol{k}}(\boldsymbol{r}_n) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}_j}$, with solution $\psi_{\boldsymbol{k}}(\boldsymbol{r}_n) = M_{nj}^{-1}e^{i\boldsymbol{k}\cdot\boldsymbol{r}_j}$. The inverse could be calculated using gauss elimination.

22.6   Wave packet dynamics

We assume, as usual, that a particle in our projectile beam is in a wave packet strongly peaked around some momentum $\hbar \boldsymbol{k}_0$, and broadly peaked around some position far enough away from the target that the packet moves like a free particle wave packet before encountering the target. It will be convenient for our subsequent analysis to take the origin of time and of the coordinate system so that the free packet would be peaked at the origin at time $t = 0$, and take $\boldsymbol{k}_0$ along z. In addition, it will be simplest to assume that the beam extends very far in the x and y directions, so that we can take its value independent of x and y.

So at large negative t, we can write a free particle wave packet

$$\psi_0(\boldsymbol{r}, t) = \int \frac{d^3k}{(2\pi)^3} a_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} e^{-\frac{i}{\hbar}E_k t} . \tag{22.23}$$

The coefficients $a_{\boldsymbol{k}}$ are strongly peaked around $\boldsymbol{k} = \boldsymbol{k}_0$, and could be determined by measuring the projectile beam properties. Expanding the packet as usual (review our slab geometry calculations if you need to) and using our assumptions we have

$$\psi_0(\boldsymbol{r}, t) = N\psi_0(x, y, z - vt, 0) \tag{22.24}$$

where N is an overall phase factor that will not contribute when we square to get the probability, and $v = \frac{dE_{k_0}}{\hbar dk_0}$.

Let's see if our construction of $\psi_{\boldsymbol{k}}(\boldsymbol{r})$ makes sense. We look at the interacting wave packet

$$\psi(\boldsymbol{r}, t) = \int \frac{d^3k}{(2\pi)^3} a_{\boldsymbol{k}} \psi_{\boldsymbol{k}}(\boldsymbol{r}) e^{-\frac{i}{\hbar}E_k t} . \tag{22.25}$$

We want to analyze this for large negative time before the scattering has taken place when the packet should be peaked at large $|\boldsymbol{r}|$, and then at large positive time after the scattering has taken place when again the packet should be peaked at large $|\boldsymbol{r}|$. We can therefore use our large r expansions of $\psi_{\boldsymbol{k}}$. These give

$$\begin{aligned}
\psi(\boldsymbol{r} \to \infty, t) &= N\psi_0(x, y, z - vt, 0) + \frac{f_{\boldsymbol{k}_0}(\hat{\boldsymbol{r}})}{r} \int \frac{d^3k}{(2\pi)^3} a_{\boldsymbol{k}} e^{ikr} e^{-\frac{i}{\hbar}E_k t} \\
&= N\psi_0(x, y, z - vt, 0) + N' \frac{f_{\boldsymbol{k}_0}(\hat{\boldsymbol{r}})}{r} \psi_0(0, 0, r - vt, 0) , \tag{22.26}
\end{aligned}$$

where N and N' are phase factors. Recalling that ψ_0 was constructed to be peaked around all of its arguments 0, we see that at large negative time $r - vt$, both r and $-vt$ are large and positive so that $r - vt$ is always very large and positive and the second term does not contribute. Therefore our wave packet constructed from the Lippmann-Schwinger equation gives the correct initial wave packet, and since it satisfies the time dependent Schrödinger equation, it must give the correct packet at all times. For large positive times, both terms contribute.

22.7 Calculating the cross section

In a scattering experiment, the incident beam has a finite extent in the x-y directions, so that it does not interfere with the scattered beam at the detectors except in the forward direction. If the detectors are constructed so that the incident beam can interfere with the scattered waves in other directions, we get holograms. Holograms can give additional information since the interference can give direct dependence on the scattering phase, but they are also much more difficult experiments to do, so we will not consider them here. For normal scattering, the probability of finding the particle in a detector that subsumes the solid angle $d\Omega$ at the angle $\hat{\boldsymbol{r}}$, is

$$P = d\Omega \int r^2 dr \left| \frac{N' f_{\boldsymbol{k}_0}(\hat{\boldsymbol{r}}) \psi_0(0,0,r-vt,0)}{r^2} \right|^2 \tag{22.27}$$

while the probability per unit area of the incident beam is

$$\frac{P_i}{A} = \int dz \, |N\psi_0(0,0,z-vt,0)|^2 \ . \tag{22.28}$$

The ratio is the differential cross section

$$\frac{d\sigma}{d\Omega} = |f_{\boldsymbol{k}_0}(\hat{\boldsymbol{r}})|^2 \ . \tag{22.29}$$

We have a result much like in the slab geometry. Once we have formulated the scattering problem properly, we can solve the time independent Schrödinger equation with the appropriate boundary conditions. Looking at these solutions far from the target, we can pick off the coefficient (here the scattering amplitude) and immediately calculate the cross section.

22.8 Born approximation

We have already obtained the Born approximation result using Fermi's golden rule. The Born approximation assumes that the potential gives a small perturbation. That is the potential must be much smaller than the incident energy, so it is sometimes called a high-energy approximation.

In the Born approximation we plug the plane wave into the integral and stop at that order. The scattering amplitude becomes

$$\begin{aligned} f_{\boldsymbol{k}}^B(\hat{\boldsymbol{r}}) &= -\frac{m}{2\pi\hbar^2} \int d^3r' e^{-ik\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} v(\boldsymbol{r}') e^{i\boldsymbol{k}\cdot\boldsymbol{r}'} \\ &= -\frac{m}{2\pi\hbar^2} \int d^3r e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}'} v(\boldsymbol{r}') \end{aligned} \tag{22.30}$$

where we can write $\boldsymbol{k}' = \hat{\boldsymbol{r}}k$, which gives the wave vector in the direction of $\hat{\boldsymbol{r}}$. Writing the momentum transfer as $\hbar\boldsymbol{q} = \hbar(\boldsymbol{k}' - \boldsymbol{k})$, the Born approximation scattering amplitude is proportional to the Fourier transform of the potential

$$f_{\boldsymbol{k}}^B(\hat{\boldsymbol{r}}) = -\frac{m}{2\pi\hbar^2} \int d^3r e^{-i\boldsymbol{q}\cdot\boldsymbol{r}'} v(\boldsymbol{r}') = -\frac{m}{2\pi\hbar^2} \tilde{v}(\boldsymbol{q}) \ . \tag{22.31}$$

The Born approximation differential cross section is then

$$
\begin{aligned}
\frac{d\sigma}{d\Omega} &= \frac{m^2}{4\pi^2\hbar^4} |\tilde{v}(\boldsymbol{q})|^2 \\
&= \frac{m^2}{4\pi^2\hbar^4} |\tilde{v}\left(k\left[\sin\theta\cos\phi\hat{\boldsymbol{x}} + \sin\theta\sin\phi\hat{\boldsymbol{y}} + (\cos\theta - 1)\hat{\boldsymbol{z}}\right]\right)|^2 \quad (22.32)
\end{aligned}
$$

which agrees with our result using the lowest order Fermi's golden rule as expected since they both keep terms second order in the potential in the cross section.

22.9 Optical theorem

The total cross section, σ, is the integrated differential cross section

$$
\sigma = \int d\Omega \frac{d\sigma}{d\Omega} . \quad (22.33)
$$

The optical theorem relates the total cross section to the forward scattering amplitude

$$
\sigma = \frac{4\pi}{k} \mathrm{Im} f(0) \quad (22.34)
$$

where $f(0)$ is the scattering amplitude in the forward direction $\theta = 0$. It should not be surprising that something like the optical theorem must exist since experimentally we could measure the total cross section by measuring the scattering into all of the detectors, or we could look just in the forward direction where we would measure the fraction of particles scattered out of the beam. Since the number of particles is conserved, these two measurements must agree. The number scattered out of the beam will be given by the eigenstate evaluated in the forward direction where a little reflection will tell you that only the interference term can contribute in the limit that the solid angle of the detector goes to zero. A detailed calculation then gives the optical theorem.

A more mundane approach just uses mathematics. Writing

$$
\begin{aligned}
f(\Omega) &= -\frac{m}{2\pi\hbar^2} \int d^3r' e^{-ik\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} v(\boldsymbol{r}')\psi_{\boldsymbol{k}}(\boldsymbol{r}') \\
\sigma &= \int d\Omega |f(\Omega)|^2 \quad (22.35)
\end{aligned}
$$

The calculation of the cross section requires the integration

$$
\int d\Omega e^{ik\hat{\boldsymbol{r}}\cdot\boldsymbol{r}''} e^{-ik\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} = 2\pi \int_{-1}^{1} d\cos(\theta) e^{ik\cos(\theta)|\boldsymbol{r}'-\boldsymbol{r}''|} = \frac{4\pi\sin(k|\boldsymbol{r}'-\boldsymbol{r}''|)}{k|\boldsymbol{r}'-\boldsymbol{r}''|} \quad (22.36)
$$

so that the total cross section becomes

$$
\sigma = \frac{m^2}{\pi\hbar^4 k} \int d^3r' \int d^3r'' \frac{\sin(k|\boldsymbol{r}'-\boldsymbol{r}''|)}{|\boldsymbol{r}'-\boldsymbol{r}''|} v(\boldsymbol{r}')\psi_{\boldsymbol{k}}(\boldsymbol{r}')v(\boldsymbol{r}'')\psi_{\boldsymbol{k}}^*(\boldsymbol{r}'') . \quad (22.37)
$$

The forward scattering amplitude is evaluated with $\hat{\boldsymbol{r}} = \hat{\boldsymbol{z}}$. We can then use the Lippman-Schwinger equation to substitute for $\exp(-ikz')$,

$$f(0) = -\frac{m}{2\pi\hbar^2} \int d^3r' \left[\psi_{\boldsymbol{k}}^*(\boldsymbol{r}') + \frac{m}{2\pi\hbar^2} \int d^3r'' \frac{e^{-ik|\boldsymbol{r}'-\boldsymbol{r}''|}}{|\boldsymbol{r}'-\boldsymbol{r}''|} v(\boldsymbol{r}'') \psi_{\boldsymbol{k}}^*(\boldsymbol{r}'') \right] v(\boldsymbol{r}') \psi_{\boldsymbol{k}}(\boldsymbol{r}') \quad (22.38)$$

Taking its complex conjugate does not change the term containing $\int d^3r' \psi \psi_{\boldsymbol{k}}^*(\boldsymbol{r}') v(\boldsymbol{r}') \psi_{\boldsymbol{k}}(\boldsymbol{r}')$, so it is real. Taking the complex conjugate and interchanging the labels of the two dummy integration variables then shows that

$$\mathrm{Im} f(0) = \frac{m^2}{4\pi^2\hbar^4} \int d^3r' \int d^3r'' \frac{\sin(k|\boldsymbol{r}'-\boldsymbol{r}''|)}{|\boldsymbol{r}'-\boldsymbol{r}''|} v(\boldsymbol{r}') \psi_{\boldsymbol{k}}(\boldsymbol{r}') v(\boldsymbol{r}'') \psi_{\boldsymbol{k}}^*(\boldsymbol{r}''). \quad (22.39)$$

Comparing we get the optical theorem

$$\sigma = \frac{4\pi}{k} \mathrm{Im} f(0). \quad (22.40)$$

22.10 Failure of the optical theorem in the Born approximation

In forward scattering, the momentum transfer is zero, so the Born approximation forward scattering amplitude is

$$f(0) = -\frac{m}{2\pi\hbar} \int d^3r \, v(\boldsymbol{r}) \quad (22.41)$$

which is real. Therefore the optical theorem with the Born scattering amplitude gives zero total cross section, which is obviously wrong. The reason should be clear, the Born approximation keeps just linear terms in the potential in the scattering amplitude. The cross section must have at least quadratic terms. Since the Born amplitude has no quadratic terms in the potential, it gives zero for the total cross section if you use the optical theorem and the Born forward scattering amplitude.

22.11 Partial wave expansion and phase shifts

We will now specialize to the case where $v(\boldsymbol{r})$ is a central potential $v(r)$. In this case, we know that the Hamiltonian is rotationally invariant.

Therefore we can write a general solution as

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_{\ell m} a_{\ell m} R_\ell(r) Y_{\ell m}(\theta, \phi). \quad (22.42)$$

We can now plug this into either the differential equation or the Lippmann-Schwinger integral equation. Since solving differential equations is probably more familiar, we can begin there.

Writing $\nabla^2 = \frac{1}{r}\frac{d^2}{dr^2}r - \frac{\ell(\ell+1)}{r^2}$, the radial functions satisfy

$$\left[\frac{d^2}{dr^2} - \frac{\ell(\ell+1)}{r^2} + k^2 - \frac{2m}{\hbar^2}v(r)\right] rR_\ell(r) = 0\,. \tag{22.43}$$

While there can be bound state solutions, and they would need to be included in the completeness relations, they will be localized around the target, and will not contribute to the expansion of the beam wave packet which is initally localized far from the target. Therefore, for the scattering problem we only need to find the positive energy solutions. For a fixed energy, k is fixed, and Eq. 22.43 is an ordinary second order differential equation. The boundary condition that the solution is well behaved at the origin, means that rR will go to zero there, and we can imagine integrating the the differential equation numerically as explained in the next section using a method like Runge-Kutta out until we get past the range R of the potential, where the potential is either zero or small enough to ignore. At this point, our function must satisfy the equation with $v = 0$. This is the spherical Bessel equation with solutions given by linear combinations of $j_\ell(kr)$ and $n_\ell(kr)$. It is useful here to use instead the spherical Hankel functions[5]

$$\begin{aligned} h_\ell^{(1)}(kr) &= j_\ell(kr) + in_\ell(kr) \\ h_\ell^{(2)}(kr) &= j_\ell(kr) - in_\ell(kr)\,. \end{aligned} \tag{22.44}$$

The large argument expansions are

$$\begin{aligned} h_\ell^{(1)}(kr \to \infty) &= (-i)^{\ell+1}\frac{e^{ikr}}{kr} \\ h_\ell^{(2)}(kr \to \infty) &= i^{\ell+1}\frac{e^{-ikr}}{kr}\,. \end{aligned} \tag{22.45}$$

The $h^{(1)}$ Hankel functions therefore represent outgoing waves, while the $h^{(2)}$ Hankel functions represent incoming waves. At large r, our radial functions must be a linear combination which is convenient to write as

$$R_\ell(r > R) = N_\ell(k)\left[h_\ell^{(2)}(kr) + S_\ell(k)h_\ell^{(1)}(kr)\right] \tag{22.46}$$

After integrating our differential equation to obtain $R_\ell(r)$, we can solve for the complex coefficients $N_\ell(k)$ and $S_\ell(k)$.

The key observation to make is that, since, for large r,

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) \to e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \frac{e^{ikr}}{r}f_{\boldsymbol{k}}(\hat{\boldsymbol{r}}) \tag{22.47}$$

the only incoming waves, and therefore the only component of $h_\ell^{(2)}(kr)$ comes from the plane wave.

---

[5]For real x, $h_\ell^{(2)}(x) = h_\ell^{(1)*}(x)$, so we could just write this using the complex conjugate, however, if you ever need to analytically continue to complex k, the standard notation is a little easier to use.

Expanding the plane wave in spherical harmonics gives

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = 4\pi \sum_{\ell m} i^\ell j_\ell(kr) Y_{\ell m}(\hat{\boldsymbol{r}}) Y_{\ell m}^*(\hat{\boldsymbol{k}}) = \sum_\ell i^\ell (2\ell+1) j_\ell(kr) P_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}) \qquad (22.48)$$

and using $j_\ell(x) = \frac{1}{2}\left[h_\ell^{(1)}(kr) + h_\ell^{(2)}(kr) \right]$ the solution at large r is

$$
\begin{aligned}
\psi_{\boldsymbol{k}}(\boldsymbol{r}) \;\rightarrow\; & \sum_\ell i^\ell \frac{2\ell+1}{2} P_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}) \left[h_\ell^{(2)}(kr) + S_\ell(k) h_\ell^{(1)}(kr) \right] \\
=\; & \sum_\ell i^\ell \frac{2\ell+1}{2} P_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}) \left[2 j_\ell(kr) + (S_\ell(k)-1) h_\ell^{(1)}(kr) \right] \\
\rightarrow\; & e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \frac{e^{ikr}}{r} \sum_\ell (-i) \frac{2\ell+1}{2k} (S_\ell(k)-1) P_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}) \qquad (22.49)
\end{aligned}
$$

so that the scattering amplitude is

$$f_{\boldsymbol{k}}(\hat{\boldsymbol{r}}) = \frac{1}{2ik} \sum_\ell (2\ell+1)(S_\ell(k)-1) P_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}) \qquad (22.50)$$

We can calculate the total cross section by integrating $|f|^2$. The orthogonality of the P_ℓ gives

$$\int d\Omega P_\ell(\cos\theta) P_{\ell'}(\cos\theta) = \frac{4\pi}{2\ell+1}\delta_{\ell,\ell'} \qquad (22.51)$$

while $P_\ell(1) = 1$,

$$
\begin{aligned}
\sigma \;=\; & \int d\Omega |f_{\boldsymbol{k}}(\hat{\boldsymbol{r}})|^2 = \sum_\ell |S_\ell(k)-1|^2 \frac{\pi(2\ell+1)}{k^2} = \sum_\ell \left(1 + |S_\ell(k)|^2 - 2\mathrm{Re}S_\ell(k)\right) \frac{\pi(2\ell+1)}{k^2} \\
\sigma \;=\; & \frac{4\pi}{k} \mathrm{Im} f_{\boldsymbol{k}}(1) = \sum_\ell \frac{2\pi(2\ell+1)}{k^2} \left(1 - \mathrm{Re}S_\ell\right) . \qquad (22.52)
\end{aligned}
$$

In order for these to agree, the magnitude of $S_\ell(k)$ must be unity. $S_\ell(k)$ must be a phase factor. We write

$$S_\ell(k) \equiv e^{i2\delta_\ell(k)} \qquad (22.53)$$

where $\delta_\ell(k)$ is called the phase shift, and the factor of 2 is conventional to make the final equations look a little nicer.

The physical reason that $S_\ell(k)$ is a phase is that it is the coefficient of the outgoing wave with angular momentum ℓ when the incoming wave has a coefficient of one. Since angular momentum is conserved, the scattering cannot remove probability from angular momentum ℓ. The magnitude of the outgoing wave is therefore fixed and the scattering can only delay or advance the position of the outgoing wave which corresponds to shifting its phase. We will see later that all scattering processes can be described by phase shifts since the operator

269

relating the incoming to outgoing waves (the S-matrix) must be unitary since it too must conserve probability. The eigenvalues of the S-matrix are the phase shifts. The scattering is completely described by giving the eigenvectors and eigenvalues of the S-matrix. For central potentials, the eigenvectors are the angular momentum states, so the only unknowns are the phase shifts.

We can write

$$\frac{S_\ell(k) - 1}{i} = 2e^{i\delta_\ell(k)}\frac{e^{i\delta_\ell(k)} - e^{-i\delta_\ell(k)}}{2i} = e^{i\delta_\ell(k)}\sin\delta_\ell(k).$$ (22.54)

The scattering amplitude and total cross sections in terms of the phase shifts are

$$f_{\boldsymbol{k}}(\hat{\boldsymbol{r}}) = \frac{1}{k}\sum_\ell (2\ell + 1)e^{i\delta_\ell(k)}\sin\delta_\ell(k)P_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}})$$

$$\sigma = \sum_\ell \sigma_\ell = \frac{4\pi}{k^2}\sum_\ell (2\ell + 1)\sin^2\delta_\ell(k),$$ (22.55)

where we have defined the total cross section for each angular momentum channel σ_ℓ.

In looking at the expression for σ_ℓ, we see that it has a maximum value when $\sin^2\delta_\ell(k) = 1$. This is sometimes called the unitarity limit, since having a larger value would require a complex $\delta_\ell(k)$ which would violate conservation of particle number (and therefore unitarity of the S-matrix). This result can be understood from classical physics. If we take a classical particle with momentum $\hbar k\hat{\boldsymbol{z}}$ the perpendicular distance from the origin that allows $\boldsymbol{r}\times\boldsymbol{p}$ values between $\hbar(\ell - \frac{1}{2})$ and $\hbar(\ell + \frac{1}{2})$ is

$$\frac{2\ell - 1}{2k} < r_\perp < \frac{2\ell + 1}{2k}$$ (22.56)

Therefore the area of the classical beam with angular momentum $\hbar(\ell \pm \frac{1}{2})$ that could scatter is

$$\pi\frac{(2\ell + 1)^2 - (2\ell - 1)^2}{4k^2} = \frac{2\pi\ell}{k^2}$$ (22.57)

which shows the same main terms, although the overall value is smaller by about a factor of 4 due to quantum effects.

22.12   Numerical method

Our radial equation for positive energies is a second-order ordinary differential equation. We can integrate it numerically using a second order method like Numerov's method, or rewrite the second-order differential equation as two coupled first-order equations and apply a first order method like Runge-Kutta or even Euler's method. Since you are probably more familiar with the first-order methods, let's rewrite the equation for $u_\ell(r) \equiv rR_(r)$, with $E = \hbar^2k^2/2m$

$$\frac{d^2u_\ell(r)}{dr^2} = \left[\frac{2m}{\hbar^2}v(r) + \frac{\ell(\ell + 1)}{r^2} - k^2\right]u_\ell(r)$$ (22.58)

as

$$\frac{du'_\ell(r)}{dr} = \left[\frac{2m}{\hbar^2}v(r) + \frac{\ell(\ell+1)}{r^2} - k^2\right]u_\ell(r)$$

$$\frac{du_\ell(r)}{dr} = u'_\ell(r).$$

(22.59)

We now have two coupled first order equations for the functions $u_\ell(r)$ and $u'_\ell(r)$.

We can now apply any first order method. Some times there are numerical subtleties that must be handled if the potentials are very strong, but let's assume that we have a "normal" case. For $\ell = 0$, we can start the integration at the origin. Since $u_0(r) = rR_0(r)$, and $R_0(r)$ is well behaved, $u_0(0) = 0$. Similarly, the derivate $u'_0(0) = R_0(0)$, but this is just a normalization. Multiplying $R_0(0)$ by a constant will simply multiply our whole solution by that constant. Therefore we can take any convenient nonzero value for $u'_0(0)$, e.g. $u'_0(0) = 1$. For $\ell > 0$, the $\frac{\ell(\ell+1)}{r^2}$ term will dominate, and we know that the solution near the origin will be $R_\ell(r \to 0) = Nr^\ell$ with N a normalization. To avoid evaluating exactly at the origin, we start a small distance Δr away and $u_\ell(\Delta r) = N(\Delta r)^{\ell+1}$, $u'_\ell(\Delta r) = N(\ell+1)(\Delta r)^\ell$, with any convenient value for N.

Our numerical integration will then step the solution out from the origin. For example, if we used the crudest Euler's method, and a step size of h, we would just add the derivatives times h to each function to move it to $r + h$,

$$u'_\ell(r+h) = u'_\ell(r) + \left[\frac{2m}{\hbar^2}v(r) + \frac{\ell(\ell+1)}{r^2} - k^2\right]u_\ell(r)h$$

$$u_\ell(r+h) = u_\ell(r) + u'_\ell(r)h.$$

(22.60)

More accurate methods, like Runge-Kutta, have somewhat more complicated expressions for the right hand sides, but the principle is the same.

We continue integrating until the potential is either zero, or small enough to ignore (we can test this by increasing the point R where we assume the potential is zero and see if the results change). We can test the integration accuracy by changing h and repeating the calculation.

At the point $r = R$, we need to fit our solution to

$$u_\ell(r > R) = rN_\ell\left[h_\ell^{(2)}(kr) + e^{2i\delta_\ell(k)}h_\ell^{(1)}(kr)\right]$$

(22.61)

Since our differential equations are completely real, it is convenient to write the spherical Bessel functions back in terms of $j_\ell(kr)$ and $n_\ell(kr)$,

$$u_\ell(r > R) = rN_\ell e^{i\delta_\ell(k)}\left[\cos\delta_\ell(k)j_\ell(kr) - \sin\delta_\ell(k)n_\ell(kr)\right]$$

(22.62)

Evaluating this for $Ru'_\ell(R)/u_\ell(R)$ makes the unimportannt normalization drop out,

$$\frac{Ru'_\ell(R)}{u_\ell(R)} = 1 + kR\frac{\cos\delta_\ell(k)j'_\ell(kr) - \sin\delta_\ell(k)n'_\ell(kr)}{\cos\delta_\ell(k)j_\ell(kr) - \sin\delta_\ell(k)n_\ell(kr)}.$$

(22.63)

We can divide the numerator and denominator by $\cos \delta_\ell(k)$ to get an expression for $\tan \delta_\ell(k)$ and solve,

$$\tan \delta_\ell(k) = \frac{kRj'_\ell(kR) + \left(1 - \frac{Ru'_\ell(R)}{u_\ell(R)}\right) j_\ell(kR)}{kRn'_\ell(kR) + \left(1 - \frac{Ru'_\ell(R)}{u_\ell(R)}\right) n_\ell(kR)}, \tag{22.64}$$

where $j'_\ell(x) = \frac{dj_\ell(x)}{dx}$, $n'_\ell(x) = \frac{dn_\ell(x)}{dx}$.

We need to repeat this for each ℓ value to get the phase shifts at the chosen scattering energy. Formally, the sum over ℓ is infinite. However, as ℓ increases, eventually, the $\ell(\ell+1)/r^2$ term will dominate the $v(r)$ term for the energies of interest[6]. You can test that you have taken a large enough maximum ℓ value by making sure the phase shifts at large ℓ are small, and that adding additional terms does not change your results within the desired numerical accuracy.

Given the phase shifts, we then construct the scattering amplitude, and the differential cross sections by calculating the Legendre polynomials. These can be readily calculated using their recursion relation,

$$\ell P_\ell(x) = (2\ell - 1)x P_{\ell-1}(x) - (\ell - 1)P_{\ell-2}(x) \tag{22.65}$$

starting with $P_0(x) = 1$, and $P_1(x) = x$.

Our numerical algorithm for calculating the scattering amplitude is to first calculate the phase shifts:

1. Set $k^2 = \frac{2mE}{\hbar^2}$, where E is the projectile energy.

2. Choose a step size h that seems physically reasonable, in any case $kh \ll 1$. Start with $\ell = 0$, then increment ℓ. Integrate using your favorite method out from the origin to a point $r = R$ where the potential is zero (or sensibly so), and calculate the ratio $u'_\ell(R)/u_\ell(R)$.

3. Calculate the spherical Bessel functions at kR and solve Eq. 22.64 for the phase shifts $\delta_\ell(k)$.

4. For $\ell > kR$, the phase shifts should drop to zero, and you can truncate at a maximum ℓ which gives terms less than your desired accuracy.

You should check that changing the integration step size does not change your result.

Given the phase shifts, the scattering amplitude, can be calculated from Eq. 22.55. The differential cross section is its magnitude squared. You can check your numerics by adjusting the step size and the maximum ℓ value.

---

[6]In the classical limit a particle with angular momentum more than $\hbar kR$ never sees the potential and does not scatter. The angular momentum of scattered quantum particles can exceed this value, but not by a large factor.

22.13 Analytic example

While numerical calculations for the scattering from central potentials are straightforward, the potentials with analytic solutions are rather limited. Even for those, the resulting expressions are often not very illuminating unless they are evaluated numerically and plotted.

An example with a straightforward analytic solution is the delta-shell potential,

$$H = \frac{p^2}{2m} + \lambda\delta(r-a) \,. \tag{22.66}$$

Note, this is *not* a three dimensional delta function. It is a delta function in the radial coordinate only, and the delta function forms a sort of spherical shell at a distance $r = a$ from the origin. The radial equation becomes

$$\left[\frac{d^2}{dr^2} - \frac{\ell(\ell+1)}{r^2} + k^2 - \frac{2m\lambda}{\hbar^2}\delta(r-a)\right] rR_\ell(r) = 0 \,, \tag{22.67}$$

and, except at $r = a$, the solutions are linear combinations of $j_\ell(kr)$ and $n_\ell(kr)$. The solutions need to be well behaved at the origin, and should be matched to our phase shift expression for $r > a$, c.f. Eq. 22.61

$$R_\ell(r) = \begin{cases} A_\ell j_\ell(kr) & r < a \\ N_\ell\left[h_\ell^{(2)}(kr) + e^{2i\delta(k)}h_\ell^{(1)}(kr)\right] & r > a \end{cases} \,. \tag{22.68}$$

The boundary conditions are that $R_\ell(r)$ is continuous at $r = a$, and integrating across the delta function

$$\left.\frac{d(rR_\ell(r))}{dr}\right|_{r=a^-}^{r=a^+} - \frac{2m\lambda}{\hbar^2}aR_\ell(a) = 0 \,, \tag{22.69}$$

or

$$\frac{R_\ell'(a^+) - R_\ell'(a^-)}{R_\ell(a)} = \frac{2m\lambda}{\hbar^2} \tag{22.70}$$

which becomes

$$\frac{h_\ell'^{(2)}(ka) + e^{2i\delta(k)}h_\ell'^{(1)}(ka)}{h_\ell^{(2)}(ka) + e^{2i\delta(k)}h_\ell^{(1)}(ka)} - \frac{j_\ell'(ka)}{j_\ell(ka)} = \frac{2m\lambda}{\hbar^2 k} \,. \tag{22.71}$$

with solution

$$e^{2i\delta(k)} = \frac{j_\ell(ka)h_\ell'^{(2)}(ka) - j_\ell'(ka)h_\ell^{(2)}(ka) + \frac{2m\lambda}{\hbar^2 k}j_\ell(ka)h_\ell^{(2)}(ka)}{j_\ell(ka)h_\ell'^{(1)}(ka) - j_\ell'(ka)h_\ell^{(1)}(ka) + \frac{2m\lambda}{\hbar^2 k}j_\ell(ka)h_\ell^{(1)}(ka)} \,. \tag{22.72}$$

This can be simplified (but the result is not changed) by using the Wronskians

$$j_\ell(ka)h_\ell'^{(2)}(ka) - j_\ell'(ka)h_\ell^{(2)}(ka) = -\frac{i}{ka^2}$$

$$j_\ell(ka)h_\ell'^{(1)}(ka) - j_\ell'(ka)h_\ell^{(1)}(ka) = \frac{i}{ka^2} \tag{22.73}$$

so that

$$e^{2i\delta(k)} = \frac{1 - i\frac{2m\lambda a^2 k}{\hbar^2} j_\ell(ka) h_\ell^{(2)}(ka)}{1 + i\frac{2m\lambda a^2 k}{\hbar^2} j_\ell(ka) h_\ell^{(1)}(ka)} \tag{22.74}$$

With this form, the numerator is the complex conjugate of the denominator, manifestly showing that the result is a phase factor.

The scattering amplitude is given by expressions Eqs. 22.55, with $|f|^2$ the differential cross section.

22.14   Eikonal approximation

For high energy projectiles (compared to the potential energy), we might expect that the path of the scattered particles would not be deflected very much. In this case, we might expect the projection of the beam on the detector after scattering would be something like a projected image of the target. The name of the eikonal approximation comes from the greek word, for image.

In the eikonal approximation we first write,

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i2\chi(r)} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{22.75}$$

which is completely general, if we allow $\chi(\boldsymbol{r})$ to be an arbitrary complex function. The factor of two multiplying χ is conventional and will later allow us to identify, approximately, χ values with the phase shifts.

Plugging Eq. 22.75 into the Schrödinger equation with $E = \hbar^2 k^2/2m$, the k^2 terms cancel, and dividing by $\psi_{\boldsymbol{k}}(\boldsymbol{r})$, leaves

$$\boldsymbol{k} \cdot \nabla \chi(\boldsymbol{r}) + \nabla^2 \chi(\boldsymbol{r}) = -\frac{m}{2\hbar^2} v(\boldsymbol{r}) \,. \tag{22.76}$$

At this point, we make the approximation that k is large compared to the derivatives of $\chi(\boldsymbol{r})$. That is, with a slowly changing potential, we have the oscillations required by the high energy, and these are modulated by the potential, but the modulation will be on the same length scale as the potential changes. We therefore drop the $\nabla^2 \chi(\boldsymbol{r})$ term compared to $\boldsymbol{k} \cdot \nabla \chi(\boldsymbol{r})$. It is convenient to take the z axis so that $\boldsymbol{k} = k\hat{\boldsymbol{z}}$ and the eikonal approximation to the Schrödinger equation becomes

$$\frac{\partial \chi(\boldsymbol{r})}{\partial z} = -\frac{m}{\hbar^2 k} v(\boldsymbol{r}) \tag{22.77}$$

Writing χ and the potential in Cartesian coordinates we have

$$\frac{\partial \chi(x, y, z)}{\partial z} = -\frac{m}{\hbar^2 k} v(x, y, z) \tag{22.78}$$

we can integrate along z,

$$\chi(x,y,z) = -\frac{m}{2\hbar^2 k} \int_{-\infty}^{z} dz' v(x,y,z') \tag{22.79}$$

where we use that $\chi \to 0$ for large negative z.

The scattering amplitude is

$$
\begin{aligned}
f(\hat{\boldsymbol{r}}) &= -\frac{m}{2\pi\hbar^2} \int d^3r' e^{-ik\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} v(\boldsymbol{r}')\psi_{\boldsymbol{k}}(\boldsymbol{r}') \\
&= -\frac{m}{2\pi\hbar^2} \int d^3r' e^{-i(k\hat{\boldsymbol{r}}-\boldsymbol{k})\cdot\boldsymbol{r}'} v(\boldsymbol{r}')e^{i2\chi(\boldsymbol{r}')} .
\end{aligned} \tag{22.80}
$$

Having used the large k limit in deriving $\chi(\boldsymbol{r})$, we can apply that limit to the scattering amplitude. We neglected the $\nabla^2\chi$ term compared to χ, so we expect the components of both χ and v to correspond to much smaller momentum than $\hbar k$. This means that the transverse components of the momentum transfer will be much smaller than k. Therefore, our eikonal approximation will only apply to small angle scattering. We can then approximate the scattering amplitude for small scattering angles as

$$
\begin{aligned}
f_e(\theta,\phi) &= -\frac{m}{2\pi\hbar^2} \int d^3r' e^{-ik(1-\cos\theta)z' - ik\sin\theta(\cos\phi x' + \sin\phi y')} v(\boldsymbol{r}')e^{i2\chi(\boldsymbol{r}')} \\
&\simeq -\frac{m}{2\pi\hbar^2} \int d^3r' e^{-ik\theta(\cos\phi x' + \sin\phi y')} v(\boldsymbol{r}')e^{i2\chi(\boldsymbol{r}')} \\
&= -\frac{m}{2\pi\hbar^2} \int d^3r' e^{-ik\theta(\cos\phi x' + \sin\phi y')} \frac{d}{dz'}\left[\frac{i\hbar^2 k}{m} e^{2i\chi(x',y,z')}\right] \\
&= -\frac{ik}{2\pi} \int dx'dy' e^{-ik\theta(\cos\phi x' + \sin\phi y')} \left[e^{2i\chi(x',y',\infty)} - 1\right] \tag{22.81}
\end{aligned}
$$

We will now specialize to a central potential. In that case, writing $x' = b\cos\phi'$, $y' = b\sin\phi'$, $\chi(x',y',\infty) = \chi(b)$, and $q = k\theta$, we have

$$
\begin{aligned}
f_e(\theta) &= -ik \int dbb \left(e^{2i\chi(b)} - 1\right) \int_0^{2\pi} \frac{d\phi'}{2\pi} e^{-iqb\cos(\phi'-\phi)} \\
&= -ik \int dbb \left(e^{2i\chi(b)} - 1\right) J_0(qb) \tag{22.82}
\end{aligned}
$$

where $J_0(x)$ is a Bessel function.

22.14.1 Approximate eikonal phase shifts

The scattering amplitude for a central potential in terms of the phase shifts, Eq. 22.55, looks somewhat like our expression for the eikonal scattering amplitude. The eikonal scattering amplitude has an integral over impact parameter instead of a sum over ℓ, a Bessel function instead of Legendre polynomials, various factors different, but $\chi(b)$ looks suspiciously like

275

a phase shift. We can use the small angle or large ℓ, approximation (see appendix E) $J_0[(\ell + \frac{1}{2})\theta] = P_\ell(\cos\theta)$ and approximating the eikonal integral by using the midpoint rule with $b = (\ell + \frac{1}{2})/k$, (the accuracy of the midpoint rule can be checked using the Euler-Maclaurin formula (see appendix F)

$$
\begin{aligned}
f_e(\theta) &= -ik \int_0^\infty dbb \left(e^{2i\chi(b)} - 1\right) J_0(qb) \\
&\simeq \frac{1}{2ik} \sum_{\ell=0}^\infty (2\ell + 1) \left(e^{2i\chi[(\ell+\frac{1}{2})/k]} - 1\right) P_\ell(\cos\theta) \, . \\
&= \frac{1}{k} \sum_{\ell=0}^\infty (2\ell + 1) e^{i\chi[(\ell+\frac{1}{2})/k]} \sin\chi[(\ell + \frac{1}{2})/k] P_\ell(\cos\theta) \, .
\end{aligned}
\tag{22.83}
$$

Comparing with Eq. 22.55 we can approximately identify $\chi[(\ell + \frac{1}{2})/k] \simeq \delta_\ell(k)$. We can understand this identification of the $\chi(b)$ as an eikonal approximation for the phase shifts, by thinking of a wave packet made up of ℓ values peaked around some ℓ_0 value. With a strongly peaked incoming momentum, the packet can be constructed with a peak value of impact parameter around $b = (\ell_0 + \frac{1}{2})/k$. Such a packet will be approximately phase shifted by $e^{2i\chi(b)}$, and this can be identified as the phase shift for the ℓ_0 value.

22.15 Born approximation central potential phase shifts

Since we have an expression for the Eikonal phase shifts for a central potential, it will be useful to derive a Born approximation expression for the phase shifts. Substituting the partial wave expansion for a plane wave for $e^{i\boldsymbol{k}'\cdot\boldsymbol{r}}$ and $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$, with a central potential the angular integrals can be done immediately to give

$$
f_B(\theta) = -\frac{2m}{\hbar^2} \sum_\ell P_\ell(\cos\theta)(2\ell + 1) \int_0^\infty dr r^2 v(r) j_\ell^2(kr) \tag{22.84}
$$

Identifying this with the partial wave expansion, we have

$$
e^{i\delta_\ell^B(k)} \sin\delta_\ell^B(k) = -\frac{2mk}{\hbar^2} \int_0^\infty dr r^2 v(r) j_\ell^2(kr) \, . \tag{22.85}
$$

Note however, that the righthand side is real, so that $\delta_\ell^B(k)$ is not real. This, once again, shows that the Born approximation violates the optical theorem. Since the Born approximation is an expansion in the potential, we require the righthand side to be small. In that case we can expand the left hand side and write

$$
\delta_\ell^B(k) = -\frac{2mk}{\hbar^2} \int_0^\infty dr r^2 v(r) j_\ell^2(kr) \, , \tag{22.86}
$$

as the definition of the Born approximation phase shifts. Since these phase shifts are real, using them in the scattering amplitude expression will automatically satisfy the optical theorem. To the same order of accuracy you can equate this expression for δ_ℓ^B to $\sin\delta_\ell^B$ or $\tan\delta_\ell^B$.

22.16 Problems

Problem 22.1.   a. Calculate an analytic expression for the phase shifts $\delta_\ell(k)$ for the scattering of a particle of mass m and energy $\hbar^2 k^2 / 2m$ for the hard sphere potential

$$v(r) = \begin{cases} \infty & r < a \\ 0 & r > a \end{cases} . \tag{22.87}$$

b. Give a rough estimate of the maximum ℓ value needed in the cross section sums so that higher values of the phase shifts can be ignored. Calculate this value using the properties of the spherical Bessel functions, but also give a sentence or two semiclassical argument involving the maximum classical impact parameter and the corresponding angular momentum.

c. Show using classical mechanics that the classical differential cross section takes the constant value $a^2/4$.

d. Plot the differential cross section divided by the classical differential cross section as a function of the angle between the incident beam and the detector θ for energies 0.01, 1.0, and 10.0 $\hbar^2/2ma^2$. Use a log scale for the cross section axis so that you can plot the results for all three cases on the same graph.

e. Plot the total cross section divided by the projected area as a function of incident energy divided by $\hbar^2/2ma^2$ for the range of energies from 0.01 to 10.0 $\hbar^2/2ma^2$.

f. Explain why the large energy cross section, which corresponds to the classical limit, is twice the projected area. Your plot of the high energy differential cross section may help you understand this.

g. High energy cross sections generally have a large diffraction peak in the forward direction. You should have found such a peak in your plot of the differential cross section. If the solid angle over which the diffraction peak has a value greater than $|f(0)|^2/2$ is called $\Delta\Omega$ (this is essentially the full width at half maximum), then the total cross section must satisfy the inequality

$$\sigma \;\geq\; \int_{\Delta\Omega} d\Omega |f(\Omega)|^2 \geq \frac{1}{2}|f(0)|^2 \Delta\Omega \geq \frac{1}{2}|\mathrm{Im} f(0)|^2 \Delta\Omega = \frac{k^2}{32\pi^2}\sigma^2 \Delta\Omega \tag{22.88}$$

where we have used the optical theorem in the last step. Rearranging, $\Delta\Omega \leq \frac{32\pi^2}{k^2\sigma}$. So that this width of the diffraction peak falls off faster than k^{-2} if σ is relatively constant.

Verify this from your numerical calculations at high energy.

Chapter 23

Lippmann-Schwinger Equation

23.1   Introduction

The Lippmann-Schwinger equation was derived previously for a specific Hamiltonian directly from the differential equation. Here we work out the general method that will work for any a Hamiltonian $H_0 + V$ where H_0 is the Hamiltonian where the projectile and target do not interact, and V is the interaction. The energy eigenstates of H_0 are given by

$$H_0|\Phi_\alpha\rangle = E_\alpha|\Phi_\alpha\rangle \tag{23.1}$$

and are normalized so that a completeness relation for the identity operator can be written as

$$\int d\alpha |\Phi_\alpha\rangle\langle\Phi_\alpha| = 1 \,. \tag{23.2}$$

Here α indicates all quantum numbers and the integral includes a sum over any bound states. Rearranging the Schrödinger equation, the eigenstates of $H_0 + V$ are solutions of

$$(E_\alpha - H_0)|\Psi_\alpha\rangle = V|\Psi_\alpha\rangle \,. \tag{23.3}$$

   Just as in the simpler potential scattering, we want the solutions that look like a noninteracting solution with the addition of outgoing scattered waves. We therefore write

$$|\Psi_\alpha\rangle = |\Phi_\alpha\rangle + \frac{1}{E_\alpha - H_0}\bigg|_{\text{b.c.}} V|\Psi_\alpha\rangle \tag{23.4}$$

where b.c. indicates that we need to specify boundary conditions on the Green's function $(E_\alpha - H_0)^{-1}$. Since energy is conserved, the noninteracting components of the solutions far from the scattering center will have energy E_α, and therefore specifying the boundary conditions means specifying how to handle the singularity when the eigenvalue of H_0 is equal to E_α. The most common choices are to remove the singularities by writing the Lippmann-Schwinger equations as

$$|\Psi_\alpha^\pm\rangle = |\Phi_\alpha\rangle + \frac{1}{E_\alpha - H_0 \pm i\eta}V|\Psi_\alpha^\pm\rangle \tag{23.5}$$

where η is a positive infinitesimal, and $|\Psi_\alpha^+\rangle$ is called an *in* state and $|\Psi_\alpha^-\rangle$ is called an *out* state. We will see that the in state satisfies the boundary conditions that we have the noninteracting state plus outgoing scattered waves.

23.2   T-matrix

The T-matrix (or transition matrix) is defined by the matrix elements

$$T_{\beta\alpha}^\pm \equiv \langle \Phi_\beta | V | \Psi_\alpha^\pm \rangle . \tag{23.6}$$

Inserting the completeness relation for the noninteracting states gives

$$|\Psi_\alpha^\pm\rangle = |\Phi_\alpha\rangle + \int d\beta \frac{1}{E_\alpha - H_0 \pm i\eta} |\Phi_\beta\rangle \langle \Phi_\beta | V | \Psi_\alpha^\pm \rangle \tag{23.7}$$

and since H_0 now can operate on $|\Phi_\beta\rangle$, it can be replaced with E_β. Using the definition of the T matrix for the matrix element then immediately gives the desired result

$$|\Psi_\alpha^\pm\rangle = |\Phi_\alpha\rangle + \int d\beta \frac{|\Phi_\beta\rangle T_{\beta\alpha}^\pm}{E_\alpha - E_\beta \pm i\eta} . \tag{23.8}$$

To get an equation for the T matrix, we multiply this equation on the left by $\langle \Phi_\beta | V$ and change the dummy state index to γ,

$$\langle \Phi_\beta | V | \Psi_\alpha^\pm \rangle = \langle \Phi_\beta | V | \Phi_\alpha \rangle + \int d\gamma \frac{\langle \Phi_\beta | V | \Phi_\gamma \rangle T_{\gamma\alpha}^\pm}{E_\alpha - E_\beta \pm i\eta} . \tag{23.9}$$

Writing the V matrix elements as $V_{\beta\alpha}$ etc., gives

$$T_{\beta\alpha}^\pm = V_{\beta\alpha} + \int d\gamma \frac{V_{\beta\gamma} T_{\gamma\alpha}^\pm}{E_\alpha - E_\gamma \pm i\eta} , \tag{23.10}$$

We have eliminated all kets and operators so that this is now a set of linear equations for the T-matrix elements.

23.2.1   On-shell and Off-shell

Since we began with the Schrödinger equation, the E_α in the T-matrix equation is the H_0 energy of the $|\Phi_\alpha\rangle$ state. An energy that corresponds to the state is called "on energy shell" or just "on shell" (for relativistic calculations, people also refer to "on mass shell").

Sometimes people define the T-matrix equation as a function of energy. E,

$$T_{\beta\alpha}^\pm(E) = V_{\beta\alpha} + \int d\gamma \frac{V_{\beta\gamma} T_{\gamma\alpha}^\pm(E)}{E - E_\gamma \pm i\eta} , \tag{23.11}$$

where E does not have to equal E_α. If E is chosen to be E_α, as we have, it is called the half-on-shell T-matrix. The reason it is called half on shell is because the energy is then equal to the energy of the state α, but not the state β. Since solving for the half-on-shell T-matrix is equivalent to solving the time independent Schrödinger equation, it contains all of the physical results. However, extending the mathematical equation to other E values is a convenient way to analytically continue to unphysically realizable E values. Sometimes this can be helpful in analyzing the results.

We will show that we can calculate the scattering amplitude from the T-matrix. Since scattering only occurs between states that conserve total energy (for example, if the projectile loses energy, the target gains energy), the scattering amplitude will turn out to be proportional to the "on-shell" T-matrix. That is when both E_α and E_β are equal for $T_{\beta\alpha}^{\pm}$.

23.3 Wave Packets

One way to show that the choices of adding $\pm i\eta$ to the energy denominator leads to the desired boundary conditions is to look at wave packets constructed from these solutions. Wave packets formed from the noninteracting states and the interacting states with the same amplitudes are

$$
\begin{aligned}
|\Phi_g(t)\rangle &= \int d\alpha |\Phi_\alpha\rangle g(\alpha) e^{-\frac{i}{\hbar}E_\alpha t}, \\
|\Psi_g^{\pm}(t)\rangle &= \int d\alpha |\Psi_\alpha^{\pm}\rangle g(\alpha) e^{-\frac{i}{\hbar}E_\alpha t},
\end{aligned}
\tag{23.12}
$$

Using the Lippmann-Schwinger equations the wave packets can be written as

$$
|\Psi_g^{\pm}(t)\rangle = \int d\alpha\, g(\alpha) e^{-i\frac{i}{\hbar}E_\alpha t} \left[|\Phi_\alpha\rangle + \int d\beta \frac{|\Phi_\beta\rangle T_{\beta\alpha}^{\pm}}{E_\alpha - E_\beta \pm i\eta} \right].
\tag{23.13}
$$

Interchanging integrations in the second term (this should normally be done carefully, but is OK here), gives

$$
\int d\beta \int d\alpha\, g(\alpha) e^{-i\frac{i}{\hbar}E_\alpha t} \frac{|\Phi_\beta\rangle T_{\beta\alpha}^{\pm}}{E_\alpha - E_\beta \pm i\eta}.
\tag{23.14}
$$

We can view the α integral as an integral along the real axis of the energy where for any degenerate states, the integrand contains an additional integration over over the degenerate states, and for energies without states, the integrand is multiplied by zero. For $t \to \infty$, we can close this contour with a large semicircle in the lower half plane, while for $t \to -\infty$ we can close in the upper half plane. In both cases the exponential in time will give zero contribution for the large semicircle. In general, the T matrix can have singularities in the complex plane, but these will similarly have imaginary parts that give exponentially damped contributions for large magnitude t. The only singularities that contribute are therefore the simple poles at $E_\alpha = E_\beta$, which give $2\pi i$ times the residue for the contour closed in the upper

half plane, and $-2\pi i$ for the contour closed in the lower half plane. The $\pm i\eta$ factors displace the poles slightly so that they only contribute for positive time for Ψ^+ and for negative time for Ψ^-, with the result

$$
\begin{aligned}
|\Psi_g^\pm(t)\rangle &\rightarrow |\Phi_g(t)\rangle \quad \text{for } t \rightarrow \mp\infty \\
|\Psi_g^\pm(t)\rangle &\rightarrow \int d\beta |\Phi_\beta\rangle e^{-\frac{i}{\hbar}E_\beta t} \left[g(\beta) \mp i2\pi \int d\alpha \delta(E_\alpha - E_\beta) g(\alpha) T_{\beta\alpha}^\pm \right] \quad \text{for } t \rightarrow \pm\infty
\end{aligned}
$$

(23.15)

This shows that the in states with the original amplitudes give a wave packet that describes the experiment. It is the original wave packet for large negative times before the scattering takes place, and for large positive times contains additional terms that give the scattering. Notice that with the energy delta function, we can see that the on-shell T-matrix describes the scattered state at large times, and so it will describe the scattering cross section.

23.4 The potential scattering result

Potential scattering corresponds to the case

$$
\begin{aligned}
H_0 &= \frac{p_{\text{op}}^2}{2m} \\
V &= v(\boldsymbol{r}_{\text{op}}) \\
\langle \boldsymbol{r}|\Phi_{\boldsymbol{k}}\rangle &= e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \\
\langle \boldsymbol{r}|\Psi_{\boldsymbol{k}}^+\rangle &= \psi_{\boldsymbol{k}}(\boldsymbol{r}) \,.
\end{aligned}
$$

(23.16)

Let's verify that we recover the potential scattering equations from our more formal expressions.

We begin with the Lippmann-Schwinger equation

$$
|\Psi_{\boldsymbol{k}}\rangle = |\Phi_{\boldsymbol{k}}\rangle + \frac{1}{E_k - H_0 + i\eta} V |\Psi_{\boldsymbol{k}}\rangle
$$

(23.17)

and calculate the matrix elements with $\langle \boldsymbol{r}|$ and insert a complete sets of r and Φ_k states,

$$
\langle \boldsymbol{r}|\Psi_{\boldsymbol{k}}\rangle = \langle \boldsymbol{r}|\Phi_{\boldsymbol{k}}\rangle + \int d^3r' \int \frac{d^3k'}{(2\pi)^3} \langle \boldsymbol{r}|\frac{1}{E_k - H_0 + i\eta}|\Phi_{\boldsymbol{k}'}\rangle \langle \Phi_{\boldsymbol{k}'}|V|\boldsymbol{r}'\rangle \langle \boldsymbol{r}'|\Psi_{\boldsymbol{k}}\rangle \,.
$$

(23.18)

H_0 can now be operated on the $|\Phi_{\boldsymbol{k}'}\rangle$ to give $E_{k'}$, and V can be operated on $|\boldsymbol{r}'\rangle$ to give $v(\boldsymbol{r}')$. Writing in terms of the definitions of the wave functions in the problem we obtain

$$
\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \int d^3r' \int \frac{d^3k'}{(2\pi)^3} \frac{e^{-i\boldsymbol{k}'\cdot(\boldsymbol{r}-\boldsymbol{r}')}}{E_k - E_{k'} + i\eta} v(\boldsymbol{r}') \psi_{\boldsymbol{k}}(\boldsymbol{r}') \,.
$$

(23.19)

Writing $\boldsymbol{R} = \boldsymbol{r} - \boldsymbol{r}'$, the Green's function is

$$
\begin{aligned}
G(\boldsymbol{R}) &= \int \frac{d^3 k'}{(2\pi)^3} \frac{e^{-i\boldsymbol{k}'\cdot\boldsymbol{R}}}{E_k - E_{k'} + i\eta} \\
&= \frac{1}{2\pi^2} \int_0^\infty dk' k'^2 \frac{\sin(k'R)}{k'R} \frac{1}{E_k - E_{k'} + i\eta} \\
&= -\frac{m}{\pi^2\hbar^2 R} \int_0^\infty dk' \frac{k'\sin(k'R)}{k'^2 - (k+i\eta)^2} = -\frac{m}{2\pi^2\hbar^2 R} \int_{-\infty}^\infty dk' \frac{k'\sin(k'R)}{k'^2 - (k+i\eta)^2} \\
&= -\frac{m}{8i\pi^2\hbar^2 R} \int_{-\infty}^\infty dk' \left[e^{ik'R} - e^{-ik'R}\right] \left[\frac{1}{k' - (k+i\eta)} + \frac{1}{k' + (k+i\eta)}\right] .
\end{aligned}
$$
(23.20)

Since R is positive, we can close the contour for the first exponential with a large semicircle in the upper half plane, and only the pole at $k + i\eta$ contributes (recall that k is positive). Similarly, for the second exponential we close in the lower half plane and only the pole at $-k - i\eta$ contributes. The result is

$$
G(\boldsymbol{R}) = -\frac{m}{2\pi\hbar^2 R} e^{ikR}
$$
(23.21)

where we can take η to zero. Plugging this result back in gives

$$
\psi_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{m}{2\pi\hbar^2} \int d^3 r' \frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} v(\boldsymbol{r}')\psi_{\boldsymbol{k}}(\boldsymbol{r}')
$$
(23.22)

which agrees with Baym Eq. 9.9. The scattering amplitude is

$$
\begin{aligned}
f &= -\frac{m}{2\pi\hbar^2} \int d^3 r' e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}'} v(\boldsymbol{r}')\psi_{\boldsymbol{k}}(\boldsymbol{r}') \\
&= -\frac{1}{4\pi} \frac{2m}{\hbar^2} T_{\boldsymbol{k}',\boldsymbol{k}}
\end{aligned}
$$
(23.23)

where, since the magnitudes $k' = k$, the T-matrix is on shell.

23.5 Relationship between the scattering amplitude and T-matrix

More generally, we can think about scattering where the target has structure, that is it is described by its own Hamiltonian when the target and projectile are well separated. Here we will look at the case where the outgoing particle is the same as the incoming one, but the more general case where the scattered particle (or particles) are different can be handled similarly.

We take the scattered particle energy to have the form $\epsilon(p)$, where $\boldsymbol{p} = \hbar\boldsymbol{k}$. We will assume that $\epsilon(p)$ is analytic (at least around the real axis) and $\epsilon(-p) = \epsilon(p)$ so that this covers the usual cases like $\epsilon(p) = p^2/2m$ or $\epsilon(p) = \sqrt{p^2c^2 + m^2c^4}$.

We want to look at the case where the states $|\Phi_\alpha\rangle$ are an outer product (so the wave functions become products) of eigenstates of momentum for the projectile, and eigenstates of the target Hamiltonian. We can write these as $|\Phi_\alpha\rangle = |\boldsymbol{k}n\rangle$. and write

$$H_0|\boldsymbol{k}n\rangle = [E_n + \epsilon(\hbar k)]|\boldsymbol{k}n\rangle \tag{23.24}$$

We can write the Lippmann-Schwinger equation as

$$|\Psi_{\boldsymbol{k}n}\rangle = |\boldsymbol{k}n\rangle + \frac{1}{E_n + \epsilon(\hbar k) - H_0 + i\eta}V|\Psi_{\boldsymbol{k}n}\rangle . \tag{23.25}$$

Operating from the left with $\langle \boldsymbol{r}m|$, (that is with the states that are an outer product of projectile position eigenstates and target energy eigenstates) we have

$$
\begin{aligned}
\langle \boldsymbol{r}m|\Psi_{\boldsymbol{k}n}\rangle &= \langle \boldsymbol{r}m|\boldsymbol{k}n\rangle + \sum_\ell \int \frac{d^3k'}{(2\pi)^3}\langle \boldsymbol{r}m|\frac{1}{E_n + \epsilon(\hbar k) - H_0 + i\eta}|\boldsymbol{k}'l\rangle\langle \boldsymbol{k}'l|V|\Psi_{\boldsymbol{k}n}\rangle . \\
&= \langle \boldsymbol{r}m|\boldsymbol{k}n\rangle + \sum_\ell \int \frac{d^3k'}{(2\pi)^3}\langle \boldsymbol{r}m|\frac{1}{E_n + \epsilon(\hbar k) - E_\ell - \epsilon(\hbar k') + i\eta}|\boldsymbol{k}'l\rangle\langle \boldsymbol{k}'l|V|\Psi_{\boldsymbol{k}n}\rangle .
\end{aligned}
\tag{23.26}
$$

We now need to take r large which means it is large compared to the target position. Therefore we need to insert another complete set of states

$$
\begin{aligned}
\langle \boldsymbol{r}m|\Psi_{\boldsymbol{k}n}\rangle &= \langle \boldsymbol{r}m|\boldsymbol{k}n\rangle + \sum_o \int d^3r' \sum_\ell \int \frac{d^3k'}{(2\pi)^3} \\
&\quad \langle \boldsymbol{r}m|\frac{1}{E_n + \epsilon(\hbar k) - E_\ell - \epsilon(\hbar k') + i\eta}|\boldsymbol{k}'l\rangle\langle \boldsymbol{k}'l|\boldsymbol{r}'o\rangle\langle \boldsymbol{r}'o|V|\Psi_{\boldsymbol{k}n}\rangle .
\end{aligned}
\tag{23.27}
$$

The positions where the matrix elements of V are non zero tell us where the target is. We can evaluate the matrix elements like $\langle \boldsymbol{r}m|\boldsymbol{k}n\rangle = \delta_{mn}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ to obtain

$$\langle \boldsymbol{r}m|\Psi_{\boldsymbol{k}n}\rangle = \delta_{mn}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \int d^3r' \left[\int \frac{d^3k'}{(2\pi)^3}\frac{e^{i\boldsymbol{k}'\cdot(\boldsymbol{r}-\boldsymbol{r}')}}{E_n - E_m + \epsilon(\hbar k) - \epsilon(\hbar k') + i\eta}\right]\langle \boldsymbol{r}'m|V|\Psi_{\boldsymbol{k}n}\rangle . \tag{23.28}$$

The $\boldsymbol{r}$ dependence of the second term is now isolated in the Green's function. We can write

$$G(\boldsymbol{R}) = \int \frac{d^3k'}{(2\pi)^3}\frac{e^{i\boldsymbol{k}'\cdot\boldsymbol{R}}}{E_n - E_m + \epsilon(\hbar k) - \epsilon(\hbar k') + i\eta} \tag{23.29}$$

and we want to calculate its behavior at $R \to \infty$. We begin by evaluating the angular

284

integral as before,

$$
\begin{aligned}
G(\boldsymbol{R}) &= \int \frac{d^3k'}{(2\pi)^3} \frac{e^{i\boldsymbol{k}'\cdot\boldsymbol{R}}}{E_n - E_m + \epsilon(\hbar k) - \epsilon(\hbar k') + i\eta} \\
&= \frac{1}{2\pi^2} \int_0^\infty dk' k' \frac{\sin(k'R)}{R} \frac{1}{E_n - E_m + \epsilon(\hbar k) - \epsilon(\hbar k') + i\eta} \\
&= \frac{1}{4\pi^2} \int_{-\infty}^\infty dk' k' \frac{e^{ik'R} - e^{-ik'R}}{2iR} \frac{1}{E_n - E_m + \epsilon(\hbar k) - \epsilon(\hbar k') + i\eta} .
\end{aligned}
$$
(23.30)

As before, since $R > 0$, for the first exponential, the contour can be closed with a large semicircle in the upper-half plane, and for the second a large semicircle in the lower-half plane[1]. Both of these semicircular paths give zero contribution. The only poles that contribute in the large R limit are the ones infinitesimally shifted away from the real axis where the energy is conserved. Expanding $\epsilon(p)$ around $\pm p_0$, we have

$$
\begin{aligned}
\epsilon(p) &= \epsilon(\pm p_0) + \epsilon'(\pm p_0)(p \mp p_0) + ... \\
&= \epsilon(p_0) \pm \epsilon'(p_0)(p \mp p_0) + ...
\end{aligned}
$$
(23.31)

where in the second line we use the property that $\epsilon(p) = \epsilon(-p)$ which tells us that the derivatives are related by $\epsilon'(p) = -\epsilon'(-p)$. As in the potential scattering problem, both integrations give outgoing spherical waves, and in the large r limit

$$
\begin{aligned}
G(\boldsymbol{R} \to \infty) &= \frac{1}{2\pi} \int dk' k' e^{ik'R} R\delta(\epsilon(\hbar k') + E_m - \epsilon(\hbar k) + E_n) \\
&= -\frac{1}{2\pi} \frac{e^{ik'R}}{R} k' \left[\frac{d}{dk'} \epsilon(\hbar k') \right]^{-1} = -\frac{e^{ik'R}}{R} \frac{k'}{2\pi\hbar^2 v'}
\end{aligned}
$$
(23.32)

where in the last line k' takes the value that makes the argument of the energy conserving delta function zero.

Expanding R we have

$$
\langle \boldsymbol{r} \to \infty m | \Psi^+_{\boldsymbol{k}n} \rangle = \delta_{mn} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{e^{ik'r}}{r} \frac{k'}{2\pi\hbar v'} \int d^3r' e^{-ik'\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'} \langle \boldsymbol{r}'m|V|\Psi^+_{\boldsymbol{k}n} \rangle
$$
(23.33)

To put this in the form of the T-matrix, we can define $\boldsymbol{k}' = k'\hat{\boldsymbol{r}}$, so that $\boldsymbol{k}'$ points in the direction of the detector. We then use the orthogonality of the target states to write

$$
\int d^3r' e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}'} \langle \boldsymbol{r}'m| = \int d^3r' \sum_\ell \langle \boldsymbol{k}'m|\boldsymbol{r}'\ell\rangle\langle \boldsymbol{r}'\ell| = \langle \boldsymbol{k}'m|
$$
(23.34)

---

[1]For a case like $\epsilon(p) = \sqrt{p^2c^2 + m^2c^4}$ there are branch cuts starting at $p = \pm imc$. If we run these out to infinity along the imaginary axis, we can distort the contour to integrate in along one side of the branch cut, around the branch point, and then back out along the other side of the cut and then continue around the semicircle. The contribution of these integrals will be exponentially damped by at least the factor $e^{-mcR/\hbar}$, so for $R \to \infty$ these contributions can be dropped.

so that

$$\langle \boldsymbol{r} \to \infty m | \Psi_{\boldsymbol{k}n}^+ \rangle \;=\; \delta_{mn} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{e^{ik'r}}{r} \frac{k'}{2\pi\hbar v'} \langle \boldsymbol{k}'m | V | \Psi_{\boldsymbol{k}n}^+ \rangle .$$

(23.35)

We can identify the scattering amplitude as

$$f_{m\boldsymbol{k}',n\boldsymbol{k}} = -\frac{p'}{2\pi\hbar^2 v'} T_{m\boldsymbol{k}',n\boldsymbol{k}}^+$$

(23.36)

where $p' = \hbar k'$.

Just as in the 1-dimensional case, when the outgoing particle has less velocity than the incoming particle, we must include the flux factors in the cross section (review the 1-d case of scattering from $v(x) = V_0 \Theta(x)$ if you don't remember this). The cross section for scattering into the solid angle $d\Omega$ given by θ, ϕ where the target is initially in its isolated eigenstate n, and finally in its isolated eigenstate m, is

$$\frac{d\sigma_{n\to m}(\theta,\phi)}{d\Omega} = \frac{v'}{v} |f_{m\boldsymbol{k}',n\boldsymbol{k}}|^2 .$$

(23.37)

Here $\boldsymbol{k}' = k'(\sin\theta\cos\phi\,\hat{\boldsymbol{x}} + \sin\theta\sin\phi\,\hat{\boldsymbol{y}} + \cos\theta\,\hat{\boldsymbol{z}})$, and the magnitude of k' is given so that energy is conserved, $\epsilon(\hbar k) + E_n = \epsilon(\hbar k') + E_m$. That is, the scattering is completely determined by the on-shell T-matrix.

This also tells you that since scattering only determines the on-shell T-matrix, using scattering to probe the target will only give these on-shell matrix elements. To determine the potential we need all the matrix elements. Typically this means that we must have a physical model of our system and use scattering to determine the parameters of our model. In general, uniquely determining V from scattering data alone is not possible. The determination of V from scattering is known as the "inverse scattering problem."

23.6 S-matrix

We see that the on-shell T-matrix describes scattering. A related quantity is the scattering matrix or S-matrix. It was introduce by Heisenberg and like Heisenberg's motivation for matrix mechanics where he tried to use only measurable quantities, it contains just information that you can measure in scattering theory.

The in and out states are given by the Lippmann-Schwinger Equations,

$$|\Psi_\alpha^\pm\rangle = |\Phi_\alpha\rangle + \frac{1}{E_\alpha - H_0 \pm i\eta} V |\Psi_\alpha^\pm\rangle$$

(23.38)

where, as before, $H = H_0 + V$, and $H_0|\Phi_\alpha\rangle = E_\alpha|\Phi_\alpha\rangle$, and we write the orthogonality condition as $\langle \Phi_\beta | \Phi_\alpha \rangle = \delta(\alpha - \beta)$.

The scattering matrix, S, is defined to be the matrix element

$$S_{\beta\alpha} = \left\langle \Psi_\beta^- | \Psi_\alpha^+ \right\rangle. \tag{23.39}$$

Plugging in the Lippmann-Schwinger Equation, this becomes

$$
\begin{aligned}
S_{\beta\alpha} &= \delta(\alpha - \beta) + \langle \Phi_\beta | \frac{1}{E_\alpha - H_0 + i\eta} V | \Psi_\alpha^+ \rangle + \langle \Psi_\beta^- | V \frac{1}{E_\beta - H_0 + i\eta} | \Phi_\alpha \rangle \\
&\quad + \langle \Psi_\beta^- | V \frac{1}{E_\beta - H_0 + i\eta} \frac{1}{E_\alpha - H_0 + i\eta} V | \Psi_\alpha^+ \rangle \\
&= \delta(\alpha - \beta) + \frac{1}{E_\alpha - E_\beta + i\eta} \langle \Phi_\beta | V | \Psi_\alpha^+ \rangle + \frac{1}{E_\beta - E_\alpha + i\eta} \langle \Psi_\beta^- | V | \Phi_\alpha \rangle \\
&\quad + \langle \Psi_\beta^- | V \frac{1}{E_\beta - H_0 + i\eta} \frac{1}{E_\alpha - H_0 + i\eta} V | \Psi_\alpha^+ \rangle
\end{aligned} \tag{23.40}
$$

We can simplify this result by looking at the matrix elements of V,

$$
\begin{aligned}
\langle \Psi_\beta^- | V | \Psi_\alpha^+ \rangle &= \langle \Psi_\beta^- | V | \Phi_\alpha \rangle + \langle \Psi_\beta^- | V \frac{1}{E_\alpha - H_0 + i\eta} V | \Psi_\alpha^+ \rangle \\
&= \langle \Phi_\beta | V | \Psi_\alpha^+ \rangle + \langle \Psi_\beta^- | V \frac{1}{E_\beta - H_0 + i\eta} V | \Psi_\alpha^+ \rangle
\end{aligned} \tag{23.41}
$$

where in the first line we substitute the Lippmann-Schwinger equation for $|\Psi_\alpha^+\rangle$ and in the second for $\langle \Psi_\beta^-|$. Subtracting the two results, we have the identity

$$
\begin{aligned}
\langle \Psi_\beta^- | V | \Phi_\alpha \rangle &= \langle \Phi_\beta | V | \Psi_\alpha^+ \rangle - \langle \Psi_\beta^- | V \left[\frac{1}{E_\alpha - H_0 + i\eta} - \frac{1}{E_\beta - H_0 + i\eta} \right] V | \Psi_\alpha^+ \rangle \\
&= \langle \Phi_\beta | V | \Psi_\alpha^+ \rangle + (E_\alpha - E_\beta) \langle \Psi_\beta^- | V \frac{1}{E_\alpha - H_0 + i\eta} \frac{1}{E_\beta - H_0 + i\eta} V | \Psi_\alpha^+ \rangle
\end{aligned} \tag{23.42}
$$

Substituting into Eq. 23.40,

$$
\begin{aligned}
S_{\beta\alpha} &= \delta(\alpha - \beta) + \left[\frac{1}{E_\alpha - E_\beta + i\eta} + \frac{1}{E_\beta - E_\alpha + i\eta} \right] \langle \Phi_\beta | V | \Psi_\alpha^+ \rangle \\
&\quad + \frac{i\eta}{E_\beta - E_\alpha + i\eta} \langle \Psi_\beta^- | V \frac{1}{E_\beta - H_0 + i\eta} \frac{1}{E_\alpha - H_0 + i\eta} V | \Psi_\alpha^+ \rangle
\end{aligned} \tag{23.43}
$$

Taking the limit $\eta \to 0$, the last term gives 0.[2] We can use the formal identity $(x + i\eta)^{-1} = \mathrm{P}x^{-1} - i\pi\delta(x)$, where P stands for the principal parts, to write

$$S_{\beta\alpha} = \delta(\alpha - \beta) - i2\pi\delta(E_\alpha - E_\beta)\langle \Phi_\beta | V | \Psi_\alpha^+ \rangle. \tag{23.44}$$

---

[2] If you are in doubt, you can take the limits carefully. However, the essential thing to notice is that $\frac{1}{(x+i\eta)^2} = -\frac{d}{dx}\frac{1}{x+i\eta}$ so that this becomes a principal parts piece plus a term proportional to the derivative of a delta function. This derivative of the delta function gives a finite contribution when the intermediate states are summed, and the extra $i\eta$ in the numerator of the part outside the matrix element then makes the whole contribution go to zero.

The matrix element $\langle \Phi_\beta | V | \Psi_\alpha^+ \rangle = T_{\beta\alpha}^+$ is the T matrix, so that

$$S_{\beta\alpha} = \delta(\alpha - \beta) - i2\pi\delta(E_\alpha - E_\beta)T_{\beta\alpha}^+ . \tag{23.45}$$

The energy conserving delta function shows again that scattering depends on the on shell T matrix. Compare the form of Eq. 23.45 to our wave packet expansion.

23.7   Unitarity and the optical theorem

Since the scattering matrix is the transformation matrix between the in and out states, and those states are complete and orthogonal, the scattering matrix defines a unitary transformation and must be unitary[3]. That is $S^{-1} = S^\dagger$, or

$$\delta(\alpha - \gamma) = \int d\beta S_{\beta\alpha} S_{\beta\gamma}^* \tag{23.46}$$

Plugging in the expression from Eq. 23.45 the product of the delta functions gives the left hand side, and the rest must be zero,

$$0 = -i2\pi\delta(E_\alpha - E_\gamma)T_{\gamma\alpha}^+ + i2\pi\delta(E_\alpha - E_\gamma)T_{\alpha\gamma}^{+*} + 4\pi^2 \int d\beta \delta(E_\alpha - E_\beta)\delta(E_\gamma - E_\beta)T_{\beta\alpha}^+ T_{\beta\gamma}^{+*} . \tag{23.47}$$

We can write the product of the delta functions as

$$\delta(E_\alpha - E_\beta)\delta(E_\gamma - E_\beta) = \delta(E_\alpha - E_\gamma)\delta(E_\alpha - E_\beta) \tag{23.48}$$

so that the condition is

$$0 = 2\pi\delta(E_\alpha - E_\gamma)\left[i(T_{\alpha\gamma}^* - T_{\gamma\alpha}) + 2\pi \int d\beta \delta(E_\alpha - E_\beta)T_{\beta\alpha}T_{\beta\gamma}^*\right] \tag{23.49}$$

We see that unitarity of the scattering matrices puts conditions on the on shell T matrix. For the special case, $\alpha = \gamma$, the condition becomes

$$2\mathrm{Im}T_{\alpha\alpha} = 2\pi \int d\beta \delta(E_\alpha - E_\beta)|T_{\beta\alpha}|^2 = v\hbar\sigma \tag{23.50}$$

where the integral is recognized to be the $\hbar$ times the transition rate per unit volume, and the cross section is the transition rate divided by the flux. Writing in terms of the forward scattering amplitude

$$f_{\alpha\alpha} = -\frac{p}{2\pi\hbar^2 v}T_{\alpha\alpha} \tag{23.51}$$

with $p = \hbar k$, the total cross section for scattering from the state α is

$$\sigma = \frac{4\pi}{k}\mathrm{Im}f_{\alpha\alpha} \tag{23.52}$$

which is the optical theorem.

---

[3]The in and out states, as usually defined, do not contain bound states, however this does not change the argument since these same states are left out of both sets and the S-matrix remains a unitary transformation.

23.8 Phase Shifts

Let's look at single particle scattering from a central potential. In that case, the scattering amplitude can be written in terms of the phase shifts,

$$f_k(\gamma) = \frac{1}{2ik} \sum_\ell (2\ell + 1) \left(e^{2i\delta_\ell} - 1\right) P_\ell(\cos\gamma) \tag{23.53}$$

where γ is the angle between the incoming and outgoing directions. This becomes

$$f_{\boldsymbol{k'k}} = \frac{2\pi}{ik} \sum_{\ell m} \left(e^{2i\delta_\ell} - 1\right) Y_{\ell m}(\theta, \phi) Y_{\ell m}^*(\theta', \phi') . \tag{23.54}$$

In terms of the on shell T matrix

$$f_{\boldsymbol{k'k}} = -\frac{p'}{2\pi\hbar^2 v'} T_{\boldsymbol{k'},\boldsymbol{k}} = -\frac{m}{2\pi\hbar^2} T_{\boldsymbol{k'},\boldsymbol{k}} , \tag{23.55}$$

where we assume a nonrelativistic kinetic energy in the last expression, $E_k = \hbar^2 k^2/2m$. We can write

$$\delta^3(\boldsymbol{k} - \boldsymbol{k'}) = \sum_{\ell m} \frac{\delta(k - k')}{k^2} Y_{\ell m}(\theta, \phi) Y_{\ell m}^*(\theta', \phi')$$

$$\delta(E_k - E_k') = \delta(k - k') \frac{dk}{dE} = \frac{m}{\hbar^2 k} \delta(k - k') . \tag{23.56}$$

The T matrix term of the scattering matrix becomes

$$-i2\pi\delta(E_k - E_k') T_{\boldsymbol{k'k}} = i\frac{4\pi^2}{k} \delta(k - k') f_{\boldsymbol{k'},\boldsymbol{k}} \tag{23.57}$$

so that

$$S_{\boldsymbol{k'k}} = (2\pi)^3 \frac{\delta(k - k')}{k^2} \sum_{\ell m} Y_{\ell m}(\theta, \phi) Y_{\ell m}^*(\theta', \phi') e^{2i\delta_\ell} . \tag{23.58}$$

This shows that the phase factors $e^{2i\delta_\ell}$ are the eigenvalues of the scattering matrix with the factor of two the standard convention. Since the scattering matrix is unitary, its eigenvalues are always phase factors, and they are always defined in this way. All scattering in quantum mechanics can therefore be described by real phase shifts and the corresponding eigenvectors.

We can also start from the fact that the scattering matrix is unitary. For central potential scattering, we know that the Hamiltonian is rotationally invariant. We can simultaneously diagonalize H, L^2 and L_z. The S matrix in this case also is rotationally invariant, since rotating the incoming and outgoing wave vectors does not change the scattering amplitude. The angular eigenvectors of the S matrix can therefore be taken to be spherical harmonics, and all of the m values for the same ℓ value will be degenerate, since they are mixed by rotations around axes other than z. The eigenvalues must be phases, and we are lead immediately to the form above for the S matrix. Rotational invariance and unitarity of the scattering matrix is enough to be able to write down the partial wave phase shift expression for the scattering. However, to actually calculate the phase shifts requires us to solve the Schrödinger equation or its equivalent.

23.9 Time dependent derivation of Lippmann-Schwinger Equations

We previously calculated scattering in the Born approximation and outlined how to go to higher order using time-dependent perturbation theory and Fermi's golden rule. Therefore we know that scattering theory can be formulated with time-dependent methods. In this view, since the scattered components of the system do not interact in the incoming and outgoing wave packets, the potential can be turned off there without affecting the physics. The scattering process then consists of the Hamiltonian

$$H = H_0 + V f(t) \tag{23.59}$$

where V is the interaction part of the Hamiltonian, and $f(t)$ is a function of time that goes slowly from zero at $t \to \pm\infty$ to one at finite times when the scattering takes place. The usual choice for $f(t)$ is $\exp(-\eta|t|/\hbar)$ with the limit $\eta \to 0$ taken.

The wave function at $t = 0$ will be given by the free particle wave packet propagated from $t = -\infty$ to $t = 0$. The propagator in the interaction representation is

$$U(0, -\infty) = 1 - \frac{i}{\hbar} \int_{-\infty}^{0} dt_1 V(t_1) + \left(-\frac{i}{\hbar}\right)^2 \int_{-\infty}^{0} dt_1 \int_{-\infty}^{t_1} dt_2 V(t_1) V(t_2) + ... \tag{23.60}$$

Operating on $|\Phi_\alpha\rangle$, inserting complete sets of states, and writing out the interaction representation time dependence,

$$
\begin{aligned}
U(0, -\infty)|\Phi_\alpha\rangle &= |\Phi_\alpha\rangle - \int d\beta |\Phi_\beta\rangle \frac{i}{\hbar} \int_{-\infty}^{0} dt_1 e^{-\frac{i}{\hbar}(E_\alpha - E_\beta + i\eta)t_1} V_{\beta\alpha} \\
&+ \left(-\frac{i}{\hbar}\right)^2 \int d\beta d\gamma |\Phi_\beta\rangle \int_{-\infty}^{0} dt_1 \int_{-\infty}^{t_1} dt_2 e^{-\frac{i}{\hbar}(E_\gamma - E_\beta + i\eta)t_1} V_{\beta\gamma} e^{-\frac{i}{\hbar}(E_\alpha - E_\gamma + i\eta)t_2} V_{\gamma\alpha} + ...
\end{aligned}
\tag{23.61}
$$

The time integrals can now be done to give

$$
\begin{aligned}
U(0, -\infty)|\Phi_\alpha\rangle &= |\Phi_\alpha\rangle + \int d\beta |\Phi_\beta\rangle \frac{1}{E_\alpha - E_\beta + i\eta} V_{\beta\alpha} \\
&+ \int d\beta d\gamma |\Phi_\beta\rangle \frac{1}{E_\alpha - E_\beta + i\eta} V_{\beta\gamma} \frac{1}{E_\alpha - E_\gamma + i\eta} V_{\gamma\alpha} + ...
\end{aligned}
\tag{23.62}
$$

which is the iterated solution to

$$U(0, -\infty)|\Phi_\alpha\rangle = |\Phi_\alpha\rangle + \int d\beta |\Phi_\beta\rangle \frac{1}{E_\alpha - E_\beta + i\eta} \langle\Phi_\beta|V U(0, -\infty)|\Phi_\alpha\rangle. \tag{23.63}$$

We see that

$$|\Psi_\alpha^{(+)}\rangle = U(0, -\infty)|\Phi_\alpha\rangle \tag{23.64}$$

and we have recovered the in state Lippmann-Schwinger Equation using a time-dependent method.

We can now look at the complete matrix element that describes the transition from an eigenstate $|\Phi_\alpha\rangle$ when the interaction has been turned off, to $t = 0$ when the interaction is fully on, to large positive times when the interaction has again been turned off and the state $|\Phi_\beta\rangle$ is an eigenstate

$$\langle\Phi_\beta|U(\infty,0)U(0,-\infty)|\Phi_\alpha\rangle. \tag{23.65}$$

Repeating the analysis above, we find that

$$\langle\Phi_\beta|U(\infty,0) = \langle\Psi_\beta^{(-)}| \tag{23.66}$$

is the out state, and the scattering matrix is

$$S_{\beta\alpha} = \left\langle\Psi_\beta^-|\Psi_\alpha^+\right\rangle = \langle\Phi_\beta|U(\infty,0)U(0,-\infty)|\Phi_\alpha\rangle. \tag{23.67}$$

The scattering operator can be defined to be $U(\infty,0)U(0,-\infty) = U(\infty,-\infty)$.

We will now derive the expression for the cross section in two ways. First, we can write the transition rate exactly as we did for time-dependent perturbation theory. We write for $t < 0$,

$$
\begin{aligned}
U(t,-\infty)|\Phi_\alpha\rangle &= |\Phi_\alpha\rangle - \int d\beta|\Phi_\beta\rangle\frac{i}{\hbar}\int_{-\infty}^t dt_1 e^{-\frac{i}{\hbar}(E_\alpha-E_\beta+i\eta)t_1}V_{\beta\alpha} \\
&\quad + \left(-\frac{i}{\hbar}\right)^2\int d\beta d\gamma|\Phi_\beta\rangle\int_{-\infty}^t dt_1\int_{-\infty}^{t_1} dt_2 e^{-\frac{i}{\hbar}(E_\gamma-E_\beta+i\eta)t_1}V_{\beta\gamma}e^{-\frac{i}{\hbar}(E_\alpha-E_\gamma+i\eta)t_2}V_{\gamma\alpha} + ... \\
&= |\Phi_\alpha\rangle - \int d\beta|\Phi_\beta\rangle\frac{i}{\hbar}\int_{-\infty}^t dt_1 e^{-\frac{i}{\hbar}(E_\alpha-E_\beta+i\eta)t_1}\left[V_{\beta\alpha} + \int d\gamma V_{\beta\gamma}\frac{1}{E_\alpha-E_\gamma+i\eta}V_{\gamma\alpha} + ...\right] \\
&= |\Phi_\alpha\rangle - \int d\beta|\Phi_\beta\rangle\frac{i}{\hbar}\int_{-\infty}^t dt_1 e^{-\frac{i}{\hbar}(E_\alpha-E_\beta+i\eta)t_1}T_{\beta\alpha}^+
\end{aligned}
\tag{23.68}
$$

where we recognize the iterated equation as our T-matrix equation.

The probability of making a transtion from the state α to the state $\beta \neq \alpha$ after time t is then

$$P(t) = \left|-\frac{i}{\hbar}\int_{-\infty}^t dt_1 e^{-\frac{i}{\hbar}(E_\alpha-E_\beta+i\eta)t_1}T_{\beta\alpha}^+\right|^2 = \frac{e^{\frac{2}{\hbar}\eta t}}{(E_\alpha-E_\beta)^2+\eta^2}|T_{\beta\alpha}^+|^2 \tag{23.69}$$

The transition rate when the potential has been fully turned on is $dP(t)/dt$ evaluated at $t = 0$, which is

$$\Gamma_{\alpha\to\beta} = \frac{2}{\hbar}\frac{\eta}{(E_\alpha-E_\beta)^2+\eta^2}|T_{\beta\alpha}^+|^2. \tag{23.70}$$

The function $\eta/(x^2 + \eta^2)$ goes to zero when $\eta \to 0$ except when $x = 0$ where it diverges. This must be constant times a delta function, and integrating over all x gives π, so that

$$\Gamma_{\alpha\to\beta} = \frac{2\pi}{\hbar}|T_{\beta\alpha}^+|^2\delta(E_\alpha-E_\beta). \tag{23.71}$$

The cross section is the transition rate divided by the incoming flux. Let's look at the case we did previously where the projectile has an energy $\epsilon(p)$. The flux factor is v^{-1}, and integrating over the outgoing particles $\boldsymbol{k}'$, that go into the detector gives

$$
\begin{aligned}
\frac{d\sigma_{n\to m}(\theta,\phi)}{d\Omega} &= \frac{1}{v}\frac{1}{(2\pi)^3}\int k'^2 dk'\frac{2\pi}{\hbar}|T^+_{m\boldsymbol{k}',n\boldsymbol{k}}|^2\delta\left(\epsilon(\hbar k')+E_m-\epsilon(\hbar k)-E_n\right) \\
&= \frac{k'^2}{4\pi^2\hbar^2 vv'}|T^+_{m\boldsymbol{k}',n\boldsymbol{k}}|^2 = \frac{p'^2}{4\pi^2\hbar^4 vv'}|T^+_{m\boldsymbol{k}',n\boldsymbol{k}}|^2
\end{aligned} \tag{23.72}
$$

where in the last expression $\boldsymbol{k}'$ must be on shell, and points toward the detector. This agrees with our previous expression Eq. 23.37.

Yet another way to this expression is to use the S-matrix directly. We saw that the S-matrix element $S_{\beta\alpha}$ gives the amplitude to go from the H_0 eigenstate $|\Phi_\alpha\rangle$ to the eigenstate $|\Phi_\beta\rangle$. The probability for this transition is then $|S_{\beta\alpha}|^2$. For $\alpha\neq\beta$, only the T-matrix term in Eq. 23.45 contributes. We need to remember that the delta function squared always means that we need to restrict the integrals that lead to the delta function, and we derived the rule that $\delta^2(E_\alpha-E_\beta)=\frac{T}{2\pi\hbar}\delta(E_\alpha-E_\beta)$. [4] The transition probability for long times T is then

$$
P(T) = \frac{2\pi}{\hbar}\delta(E_\alpha-E_\beta)|T^+_{\beta\alpha}|^2 T \tag{23.74}
$$

and the rate is $dP(T)/dT$ giving

$$
\Gamma_{\alpha\to\beta} = \frac{2\pi}{\hbar}\delta(E_\alpha-E_\beta)|T^+_{\beta\alpha}|^2 \tag{23.75}
$$

directly from the S-matrix.

23.10   More general cases

We have been looking at the case where the outgoing particle is the same as the incoming one. Sometimes other processes occur. For example if you scatter a positron from an atom you can capture (and annihilate) the positron with outgoing electrons and photons. With the appropriate Hamiltonian these processes can be described by the Lippmann-Schwinger equation. Sometimes, however, it is useful in these cases to split H up into different H_0 and V pieces for the incoming and outgoing states. This can be readily handled using the S-matrix formalism since we can calculate $U(0,-\infty)$ using a different split up than we used for $U(\infty,0)$. As we saw, this corresponds to calculating the in states using a different H_0 and V from the outstates and forming the overlaps.

---

[4]Our "sleazy math" way is to write the delta functions as.

$$
\delta(E_\alpha-E_\beta) = \lim_{T\to\infty}\int_{-T/2}^{T/2}\frac{dt}{2\pi\hbar}e^{-\frac{i}{\hbar}(E_\alpha-E_\beta)t}\int_{-T/2}^{T/2}\frac{dt'}{2\pi\hbar}e^{-\frac{i}{\hbar}(E_\alpha-E_\beta)t'} \tag{23.73}
$$

and realize that if one of the integrals becomes a delta function, the other is evaluated with its exponential factor equal to 1 giving our result.

23.11 Problems

Problem 23.1. A single particle with kinetic energy $p^2/2m$ scatters from a rotationally invariant potential that is separable in each angular momentum channel. Its $\boldsymbol{r}$ representation matrix elements are

$$\langle \boldsymbol{r}|V|\boldsymbol{r}'\rangle = \sum_{\ell=0}^{\infty} \lambda_\ell u_\ell(r) u_\ell(r') P_\ell(\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{r}}') \tag{23.76}$$

where $P_\ell(x)$ are the Legendre polynomials and $u_\ell(r)$ are well behaved real functions that go to zero rapidly for large r.

(a) Show that the matrix elements of the potential in the momentum space representation $(\boldsymbol{p}_{\rm op}|\boldsymbol{k}\rangle = \hbar \boldsymbol{k}|\boldsymbol{k}\rangle)$ are

$$\langle \boldsymbol{k}|V|\boldsymbol{k}'\rangle = \sum_{\ell=0}^{\infty} \lambda_\ell \tilde{u}_\ell(k) \tilde{u}_\ell(k') P_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}') \, , \tag{23.77}$$

and write an integral expression for $\tilde{u}_\ell(k)$. Use the normalizations $\langle \boldsymbol{r}|\boldsymbol{r}'\rangle = \delta^3(\boldsymbol{r} - \boldsymbol{r}')$ and $\langle \boldsymbol{k}|\boldsymbol{k}'\rangle = (2\pi)^3 \delta^3(\boldsymbol{k} - \boldsymbol{k}')$.

(b) We showed in class that the Lippmann-Schwinger equation is equivalent to the outgoing T matrix equation (we drop the $+$ superscript for brevity)[5]

$$T = V + V \frac{1}{E - H_0 + i\eta} T \, . \tag{23.78}$$

Solve this equation for the momentum space matrix elements of the T matrix using the potential of part (a). Your result should be

$$\langle \boldsymbol{k}'|T|\boldsymbol{k}\rangle = \sum_{\ell=0}^{\infty} \frac{\lambda_\ell \tilde{u}_\ell(k') \tilde{u}_\ell(k) P_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}')}{1 - \frac{\lambda_\ell}{2\pi^2(2\ell+1)} \int_0^\infty dk'' \frac{k''^2 \tilde{u}_\ell^2(k'')}{E - \frac{\hbar^2 k''^2}{2m} + i\eta}} \tag{23.79}$$

(c) Show that the delta-shell potential $v(r) = \lambda \delta(r - a)$ can be written in the form of V. Calculate its T matrix by performing the integral in Eq. 23.79. Take the value of the energy to be $E = \hbar^2 k_0^2/2m$. Your result should be

$$\langle \boldsymbol{k}'|T|\boldsymbol{k}\rangle = \sum_{\ell} \frac{4\pi \lambda(2\ell+1) a^2 j_\ell(ka) j_\ell(k'a) P_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}')}{1 + \frac{i2m\lambda k_0 a^2}{\hbar^2} j_\ell(k_0 a) h_\ell^{(1)}(k_0 a)} \, . \tag{23.80}$$

(d) Write the scattering amplitude for the delta-shell potential in terms of the on-shell T matrix. Verify your result agrees with the expression we obtained by direct integration of the Schrödinger equation in class.

---

[5]Recall that the operator T is defined so that $\langle \Phi_\alpha|T|\Phi_\beta\rangle = T_{\alpha,\beta} = \langle \Phi_\alpha|V|\Psi_\beta\rangle$.

This problem is an example where the scattering wave function, scattering amplitude, and cross sections can all be calculated by a direct solution of the Lippmann-Schwinger equation in momentum space. In the last part you will demonstrate that the delta-shell potential has this separable form and you can check that your results are correct. Recall that the hard-sphere potential is a special case of the delta shell, so that problem can be done this way.

Problem 23.2. Most targets are composite systems that can be excited by the projectile giving inelastic scattering. We begin with a noninteracting Hamiltonian

$$H_0 = \frac{p^2}{2m} + H_T \tag{23.81}$$

where H_T describes the target. Assume that the target has just two states $|0\rangle$ and $|1\rangle$ with

$$\begin{aligned} H_T|0\rangle &= \epsilon_0|0\rangle \\ H_T|1\rangle &= \epsilon_1|1\rangle, \end{aligned} \tag{23.82}$$

and $\epsilon_1 > \epsilon_0$. The eigenstates of H_0 can then be written as

$$H_0|n\boldsymbol{k}\rangle = \left[\frac{\hbar^2 k^2}{2m} + \epsilon_n\right]|n\boldsymbol{k}\rangle. \tag{23.83}$$

where,

$$\langle n\boldsymbol{r}|m\boldsymbol{k}\rangle = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\delta_{nm}. \tag{23.84}$$

Take the interaction between the projectile and the target to be

$$\langle n\boldsymbol{r}|V|m\boldsymbol{r}'\rangle = \lambda u(r)u(r')\left[\delta_{m1}\delta_{n0} + \delta_{m0}\delta_{n1}\right] \tag{23.85}$$

where $u(r)$ is real with a well behaved Fourier transform,

$$\tilde{u}(k) = \int d^3 r\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} u(r). \tag{23.86}$$

Initially the target is in its ground state, and the projectile has incident energy $\hbar^2 k^2/2m$. We define κ such that

$$\frac{\hbar^2 \kappa^2}{2m} = \epsilon_1 - \epsilon_0 \tag{23.87}$$

a. Solve the Lippmann-Schwinger equation for the matrix elements of the states

$$\langle n\boldsymbol{k}'|\psi_{\boldsymbol{k}}\rangle \equiv \tilde{\psi}_{\boldsymbol{k}}(n, \boldsymbol{k}') \tag{23.88}$$

where the superscript + is dropped for brevity. Your result should be

$$\begin{aligned} \tilde{\psi}_{\boldsymbol{k}}(0, \boldsymbol{k}') &= (2\pi)^3\delta^3(\boldsymbol{k}' - \boldsymbol{k}) + \frac{2m\lambda}{\hbar^2}\frac{\tilde{u}(k)\gamma(k,\kappa)}{1 - \gamma(k,0)\gamma(k,\kappa)}\frac{\tilde{u}(k')}{k^2 - k'^2 + i\eta} \\ \tilde{\psi}_{\boldsymbol{k}}(1, \boldsymbol{k}') &= \frac{2m\lambda}{\hbar^2}\frac{\tilde{u}(k)}{1 - \gamma(k,0)\gamma(k,\kappa)}\frac{\tilde{u}(k')}{k^2 - \kappa^2 - k'^2 + i\eta}. \end{aligned} \tag{23.89}$$

with

$$\gamma(k,\kappa) = \frac{2m\lambda}{\hbar^2}\int\frac{d^3 k'}{(2\pi)^3}\frac{\tilde{u}^2(k')}{k^2 - \kappa^2 - k'^2 + i\eta}. \tag{23.90}$$

294

b. Calculate the cross section, σ_0, for scattering with the target left in its ground state, and the cross section, σ_1, for scattering with the target left in its excited state. Verify that the optical theorem is satisfied, i.e.

$$\frac{4\pi}{k}\mathrm{Im}f_{0\boldsymbol{k},0\boldsymbol{k}} = -\frac{2}{v\hbar}\mathrm{Im}T_{0\boldsymbol{k},0\boldsymbol{k}} = \sigma_0 + \sigma_1 \tag{23.91}$$

here $v = \hbar k/m$.

c. For the case where

$$u(r) = \frac{e^{-\beta r}}{r} \tag{23.92}$$

graph $\sigma_0\beta^2$ and $\sigma_1\beta^2$ as a function of $Em/\hbar^2\beta^2$, where $E = \hbar^2k^2/2m$ is the incident particle's kinetic energy. Take the values of the other parameters to be

$$\begin{aligned}
\kappa^2 &= 2\beta^2 \\
\lambda &= \frac{1}{10}\frac{\hbar^2\beta^3}{m}
\end{aligned} \tag{23.93}$$

Use a log scale for the cross section. This is equivalent to taking natural units with $\hbar = m = \beta = 1$, with numerical values, $\kappa = \sqrt{2}$, and $\lambda = 0.1$.

My graphs are shown in figures 23.1 and 23.2.

As a check, my result for the needed integrations is

$$\begin{aligned}
\tilde{u}(k) &= \frac{4\pi}{k^2 + \beta^2} \\
\gamma(k,\kappa) &= -\frac{4\pi m\lambda}{\beta\hbar^2}\left\{\begin{array}{ll} \left(\beta - i\sqrt{k^2 - \kappa^2}\right)^{-2} & k > \kappa \\ \left(\beta + \sqrt{\kappa^2 - k^2}\right)^{-2} & k < \kappa \end{array}\right. \tag{23.94}
\end{aligned}$$

This problem is very similar to problem 1 but with the possibility of exciting the target and getting inelastic scattering, which has much more experimental relevance than the simpler elastic scattering cases.

Problem 23.3. Calculate the scattering solutions for a system with a separable but nonlocal s-wave potential

$$\begin{aligned}
H_0 &= \frac{p_{\mathrm{op}}^2}{2m} \\
\langle\boldsymbol{r}|V|\boldsymbol{r}'\rangle &= \lambda u(r)u(r')
\end{aligned} \tag{23.95}$$

where λ and $u(r)$ are real. Follow these steps

a. Show that

$$\langle\boldsymbol{k}|V|\boldsymbol{k}'\rangle = \lambda\tilde{u}(k)\tilde{u}(k') \tag{23.96}$$

295
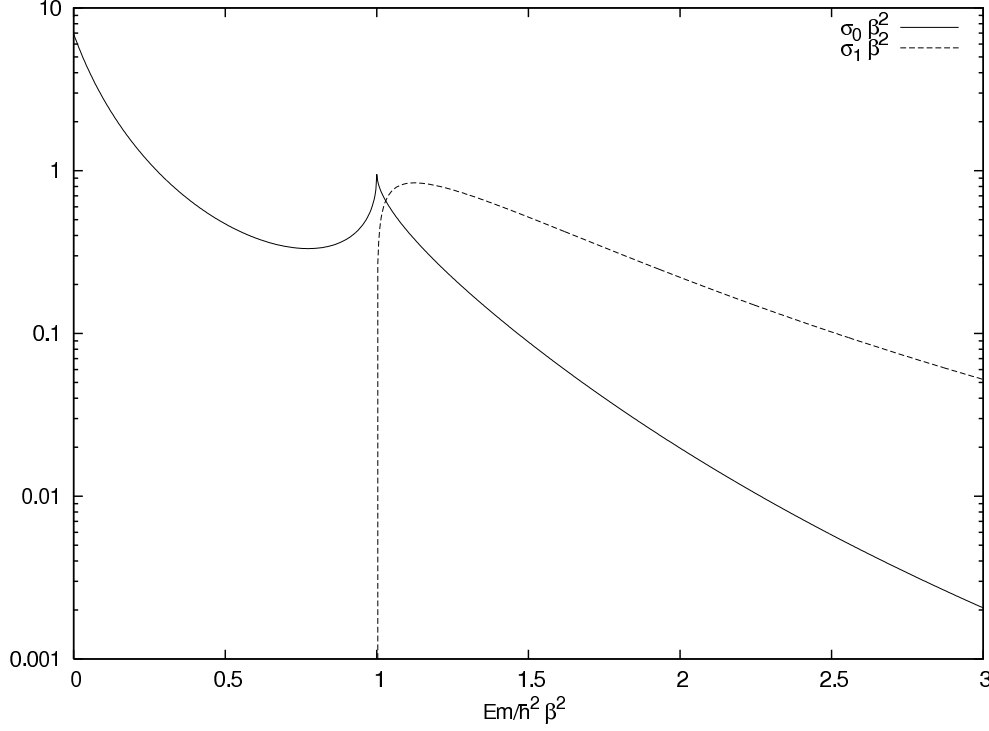
Figure 23.1: The cross sections for the parameters in the problem.

where $|\boldsymbol{k}\rangle = |\Phi_{\boldsymbol{k}}\rangle$ are free particle states, and

$$\tilde{u}(k) = \int d^3r e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} u(r) \,. \tag{23.97}$$

Be sure to show that $\tilde{u}$ does not depend on the direction of $\boldsymbol{k}$. You can assume for the rest of the problem that the Fourier transform is well behaved for real $\boldsymbol{k}$.

b. Solve the Lippmann-Schwinger equation for the matrix elements of the in states

$$\langle \boldsymbol{k}'|\psi_{\boldsymbol{k}}\rangle \equiv \tilde{\psi}_{\boldsymbol{k}}(\boldsymbol{k}') \tag{23.98}$$

where the superscript $+$ is dropped here for brevity. Your result should be

$$\tilde{\psi}_{\boldsymbol{k}}(\boldsymbol{k}') = (2\pi)^3\delta^3(\boldsymbol{k}' - \boldsymbol{k}) + \frac{2m\lambda}{\hbar^2}\left[1 - \frac{2m\lambda}{\hbar^2}\int \frac{d^3k''}{(2\pi)^3}\frac{\tilde{u}^2(k'')}{k^2 - k''^2 + i\eta}\right]^{-1}\frac{\tilde{u}(k)\tilde{u}(k')}{k^2 - k'^2 + i\eta} \tag{23.99}$$

c. Use the result of part b to find an expression for the scattering amplitude. Show that your result satisfies the optical theorem

$$\sigma = \frac{4\pi}{k}\mathrm{Im}f(0) \tag{23.100}$$

where $f(0)$ is the scattering amplitude in the forward direction.

296
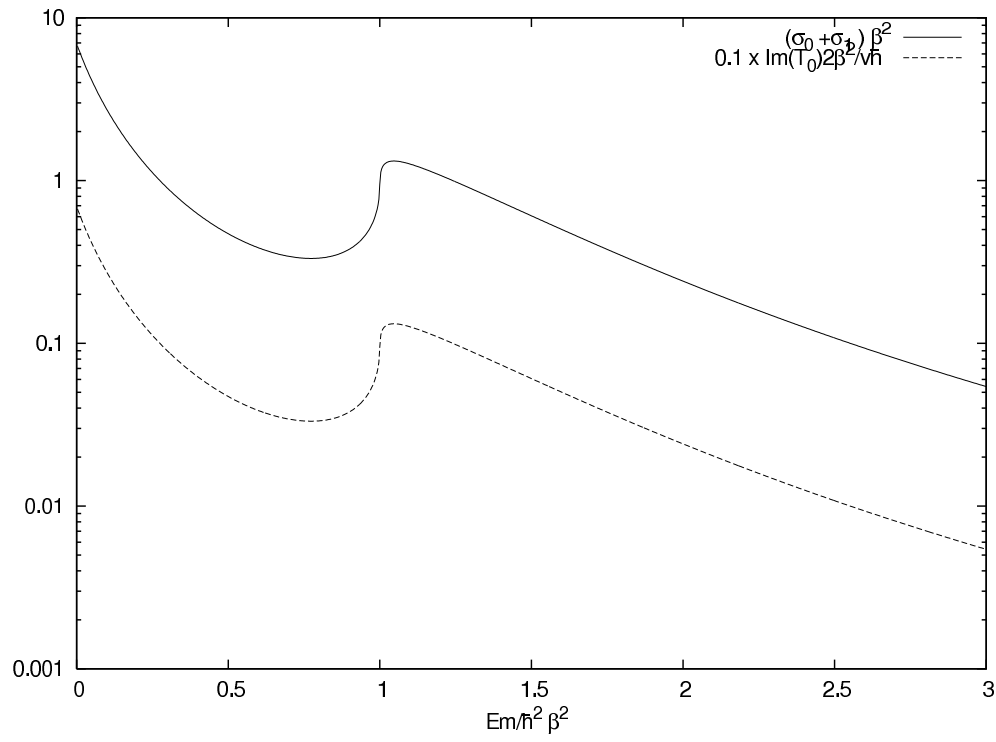
Figure 23.2: A comparison between the total cross section calculated by adding σ_0 and σ_1, and the result from using the optical theorem. The optical theorem values were multiplied by 0.1 to shift the graphs. Otherwise the two curves coincide.

d. For the special case where $u(r)$ has a Yukawa form

$$u(r) = \frac{e^{-\beta r}}{r} \tag{23.101}$$

calculate the scattering amplitude. Your result should be

$$f = \left[-\frac{(k^2 + \beta^2)^2 \hbar^2}{8\pi \lambda m} - ik - \frac{\beta}{2} + \frac{k^2}{2\beta} \right]^{-1} \tag{23.102}$$

e. The effective range expansion for low energy scattering is given by expanding

$$k \cot \delta_0 = -\frac{1}{a} + \frac{1}{2} r_0 k^2 + \ldots, \tag{23.103}$$

where the phase shift is given by $kf = e^{i\delta_0} \sin \delta_0$. Find expressions for the scattering length a and the effective range r_0 of this potential.

Chapter 24

Optical theorem

The name optical theorem comes from the relationship of the forward scattering amplitude to the "index of refraction" which was shown by several people in optics, e.g. Rayleigh, Phil. Mag. **41**, 107-120, 274-279 (1871) in his Rayleigh scattering articles explaining the blue sky.

To calculate the index of refraction of a uniform distribution of scattering centers, we first consider scattering from an infinitesimally thin slab of thickness $h \to 0$. A plane wave incident from negative z will give a wave function,

$$\psi(\boldsymbol{r}) = e^{ikz} + \int d^3r' \rho f(\theta, \phi) \frac{e^{ik|\boldsymbol{r}' - \boldsymbol{r}|}}{|\boldsymbol{r}' - \boldsymbol{r}|} \tag{24.1}$$

where the integral is over the thin slab, ρ is the number density of scatterers, and we have assumed in writing Eq. 24.1 that the scattered wave in the thin slab does not rescatter within the same slab. This places no limit on the scattering amplitude f, but we might have to take the density ρ small to keep multiple scattering from being important. Note also that in quantum mechanics $\psi(\boldsymbol{r})$ is a wave function for the particle traveling through the slab, while in optics we could write the identical equation for the electric field component of a linearly polarized wave.

The effect of the slab at $x = y = 0$ and large positive z is given by expanding the integral with the change of variables, $u^2 = x'^2 + y'^2$

$$e^{ikz} + 2\rho\pi h f(0) z^{-1} \int_0^\infty u \, du e^{ikz(1 + \frac{u^2}{2z^2} + \dots)} \tag{24.2}$$

where we have used the fact that the integrand oscillates wildly for $ku \gg 1$ so that only the forward scattering part of f contributes, and the integral can be made convergent in the usual way by making k have an infinitesimal imaginary part. After integrating, this small imaginary part means we can drop the wildly oscillating upper limit contribution and the result is

$$e^{ikz} \left(1 + i\frac{2\pi}{k} \rho f(0) h \right) \tag{24.3}$$

The effect of the infinitesimally thin slab in the forward direction is to multiply the plane wave by this factor. The effect of a thick slab of thickness $d = Nh$, within the same

299

approximations is

$$e^{ikz}\left(1 + i\frac{2\pi}{k}\rho f(0)\frac{d}{N}\right)^N = e^{ikz}e^{i\frac{2\pi}{k}\rho f(0)d} = e^{ik\left(1 + \frac{2\pi}{k^2}f(0)\rho\right)d}e^{ik(z-d)} \tag{24.4}$$

so that this corresponds to taking the value of k across the slab to be

$$k_{\text{slab}} = k\left(1 + \frac{2\pi}{k^2}f(0)\rho\right) \tag{24.5}$$

which makes the "index of refraction",

$$n = 1 + \frac{2\pi f(0)}{k^2}\rho. \tag{24.6}$$

The only mechanism for attenuation of the incoming beam here is the scattering of particles out of the forward direction. The attenuation per unit length is therefore given by the total cross section σ times the density of scatterers, and equating this to the attenuation from the index of refraction gives

$$e^{-\sigma\rho z} = |e^{iknz}|^2 = e^{ikz(n-n^*)} = e^{-\frac{4\pi Im(f(0))}{k}\rho z}, \tag{24.7}$$

which shows, $\sigma = \frac{4\pi}{k}f(0)$, the optical theorem, once again.

Chapter 25

Analytic properties of the T-matrix

25.1 Analytic properties of the off-shell T-matrix

The off-shell T-matrix was defined for any real energy. We can further analytically continue the T-matrix into the complex plane. The energy E in the off-shell T-matrix equation is then simply a complex parameter. To help make this clear we use the symbol z for this complex parameter, and the symbol E will always be real. We can write

$$T(z) = V + V \frac{1}{z - H_0} T(z). \tag{25.1}$$

Notice that by setting $z = E + i\eta$ we get the equation for T^+, while setting $z = E - i\eta$ we get the equation for T^-. Both equations are then contained in our equation for $T(z)$.

One of the reasons for doing this is to be able to interpret experimental data. If the experimentally measured on-shell T-matrix can be modeled by an function of a complex energy that is analytic except at a few singularities that can be understood physically, often the underlying important degrees of freedom can be understood.

It is helpful to rewrite the T matrix equation by first iterating and then summing

$$\begin{aligned}
T(z) &= V + V(z - H_0)^{-1}V + V(z - H_0)^{-1}V(z - H_0)^{-1}V + ... \\
&= V \left\{ 1 + (z - H_0)^{-1}V + \left[(z - H_0)^{-1}V\right]^2 + ... \right\} \\
&= V \left[1 - (z - H_0)^{-1}V \right]^{-1} \tag{25.2}
\end{aligned}$$

where we have summed using the geometric series $1 + x + x^2 + ... = (1 - x)^{-1}$. We can now factor this recalling that $(AB)^{-1} = B^{-1}A^{-1}$, which can be checked by direct operator multiplication $1 = AB(AB)^{-1} = ABB^{-1}A^{-1}$,

$$\begin{aligned}
T(z) &= V \left[1 - (z - H_0)^{-1}V \right]^{-1} (z - H_0)^{-1}(z - H_0) \\
&= V \left[z - H_0 - V \right]^{-1} (z - H_0 - V + V) \\
&= V \left[z - H \right]^{-1} (z - H + V) \\
&= V + V \frac{1}{z - H} V. \tag{25.3}
\end{aligned}$$

If we insert a complete set of eigenstates of H, we see that there will be poles of z for all of the eigenvalues of H. Therefore there will be isolated poles at the bound state energies (that is along the negative real axis) and a continuum of poles, i.e. a branch cut, for all the continuum energies. For a simple potential model where the target has no structure, we would have poles at the bound state energies and a branch cut running from the origin along the positive real axis. This structure immediately shows us why changing z from $E + i\eta$ to $E - i\eta$ gives us different T-matrices – the two values are on opposite sides of the branch cut so that there is a discontinuity if we cross the cut remaining on the same Riemann sheet.

It turns out that sometimes it is useful to analytically continue through the branch cut to define $T(z)$ on the "second" Riemann sheet.

25.2 Delta-shell potential T-matrix

The T-matrix for the delta-shell potential can be calculated analytically. (See problem 23.1.) The potential is separable and the T-matrix becomes

$$\langle \boldsymbol{k}'|T(z)|\boldsymbol{k}\rangle = 4\pi a^2 \lambda \sum_{\ell=0}^{\infty} \frac{(2\ell+1)j_\ell(ka)j_\ell(k'a)P_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}')}{1 - \frac{2\lambda a^2}{\pi}\int_0^\infty dk'' \frac{k''^2 j_\ell^2(k''a)}{z - \frac{\hbar^2 k''^2}{2m}}}. \tag{25.4}$$

We need to study the analytic properties of the integral in the denominator. Noting that the integrand is an even function of k'', we define, for complex x,

$$I_\ell(x) = \int_{-\infty}^{\infty} du \frac{u^2 j_\ell^2(u)}{u^2 - x} \tag{25.5}$$

so that

$$\langle \boldsymbol{k}'|T(z)|\boldsymbol{k}\rangle = 4\pi a^2 \lambda \sum_{\ell=0}^{\infty} \frac{(2\ell+1)j_\ell(ka)j_\ell(k'a)P_\ell(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{k}}')}{1 + \frac{2m\lambda a}{\hbar^2 \pi}I_\ell(z\frac{2ma^2}{\hbar^2})}. \tag{25.6}$$

First notice that the integral $I(x)$ is well defined for any x not on the positive real axis, and therefore $T(z)$ is well defined for any z not on the positive real axis. You plug in the value of x and perform the integral over real values of u. This integral can be done numerically. The Bessel functions are always evaluated on the real axis.

We can also use contour integration to perform these integrals. Rewriting one of the $j_\ell(u)$ as $\frac{1}{2}[h_\ell^{(1)}(u) + h_\ell^{(2)}(u)]$, we get two integrals. The contour for the integral with $h_\ell^{(1)}(u)$ can be closed with a large semicircle in the upper half plane, while the one with $h_\ell^{(2)}(u)$ can be closed with a large semicircle in the lower half plane. The denominator shows that there are poles at $u = \pm\sqrt{x}$. We therefore need to decide our rule for calculating $\sqrt{x}$, i.e. where to put the branch cut for $\sqrt{x}$. Putting the branch cut on the positive real axis will correspond to our original integral, which then corresponds to taking $x = \rho e^{i\phi}$ with $0 < \phi < 2\pi$, $\rho \geq 0$, and $\sqrt{x} = \sqrt{\rho}e^{i\phi/2}$, with $\sqrt{\rho} \geq 0$. Since $0 < \phi/2 < \pi$, $\sqrt{x}$ will be in the upper half plane.

The contour integration then gives

$$I_\ell(x) = i\pi\sqrt{x}j_\ell(\sqrt{x})h_\ell^{(1)}(\sqrt{x}) \tag{25.7}$$

It is helpful to define $z = \frac{\hbar^2 k_0^2}{2m}$, so that $\sqrt{x} = k_0 a$, and k_0 is required to be in the upper half plane to enforce the branch cut position for z. The delta-shell potential T-matrix becomes

$$\langle \boldsymbol{k}'|T|\boldsymbol{k}\rangle = \sum_\ell \frac{4\pi\lambda(2\ell+1)a^2 j_\ell(ka)j_\ell(k'a)P_\ell(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}')}{1 + \frac{i2m\lambda k_0 a^2}{\hbar^2}j_\ell(k_0 a)h_\ell^{(1)}(k_0 a)} \, . \tag{25.8}$$

From our general analysis, we expect poles corresponding to the bound-state energies. These would correspond to $k_0 = i\kappa$, and the bound-state energies $E_b = -\frac{\hbar^2\kappa^2}{2m}$. We have the identities

$$
\begin{aligned}
j_\ell(i\kappa a) &= (i)^\ell i_\ell(\kappa a) \\
h_\ell(i\kappa) &= -\frac{2}{\pi}(-i)^\ell k_\ell(\kappa a)
\end{aligned}
\tag{25.9}
$$

so that the T-matrix will have poles when

$$\frac{4m\lambda\kappa a^2}{\pi\hbar^2}i_\ell(\kappa a)k_\ell(\kappa a) = -1 \, . \tag{25.10}$$

We can check that this is the requirement for a bound state by integrating the radial Schrödinger equation. For $E = -\frac{\hbar^2\kappa^2}{2m}$, the solutions that are well behaved at the origin are $i_\ell(\kappa r)$, and the solutions that go to zero and are normalizable for large r are $k_\ell(\kappa r)$. The bound state radial solution must be

$$R_\ell(r) = \begin{cases} Ai_\ell(\kappa r) & r < a \\ Bk_\ell(\kappa r) & r > a \end{cases} \, . \tag{25.11}$$

Integrating the radial Schrödinger equation across the delta function gives the requirement

$$-\frac{\hbar^2 a}{2m}[R_\ell'(a^+) - R_\ell'(a^-)] + \lambda a R_\ell(a) = 0 \tag{25.12}$$

Integrating again gives the requirement that R_ℓ is continuous. Substituting the Bessel function forms, we have

$$\frac{k_\ell'(\kappa a)}{k_\ell(\kappa a)} - \frac{i_\ell'(\kappa a)}{i_\ell(\kappa a)} = \frac{2m\lambda}{\hbar^2\kappa} \, . \tag{25.13}$$

Using the Wronskian

$$k_\ell'(x)i_\ell(x) - k_\ell(x)i_\ell'(x) = -\frac{\pi}{2x^2} \tag{25.14}$$

this becomes

$$\frac{4m\lambda\kappa a^2}{\pi\hbar^2}i_\ell(\kappa a)k_\ell(\kappa a) = -1 \tag{25.15}$$

which are the locations of the poles of the T-matrix.

Since $i_\ell(\kappa a)$ is a monotonically increasing positive function, and $k_\ell(\kappa a)$ is a monotonically decreasing positive function, we will only have poles if λ is negative (i.e. the potential is

attractive). Further, we can have at most one bound state for each ℓ value. Taking the limit $\kappa \to 0$,

$$
\begin{aligned}
i_\ell(x) &= \frac{x^\ell}{(2\ell + 1)!!} + \dots \\
k_\ell(x) &= \frac{\pi(2\ell - 1)!!}{2x^{\ell+1}} + \dots
\end{aligned}
\tag{25.16}
$$

Substituting the these expansions we get the requirement that

$$
-\lambda > \frac{(2\ell + 1)\hbar^2}{2ma}
\tag{25.17}
$$

for a bound state of angular momentum ℓ. If we imagine slowly increasing the magnitude of λ, we find that when we pass through the magnitude $\frac{(2\ell+1)\hbar^2}{2ma}$ a bound state appears and the corresponding pole in the T-matrix. The bound state pole will appear on the negative real axis at the origin and move to more negative energies as the magnitude of λ is increase.

You might ask, where was this pole just before it appeared? We can try to answer that question by taking the next term in the expansions. For $\ell = 0$, we find that

$$
-1 = \frac{2m\lambda a}{\hbar^2}(1 - \kappa a + \dots)
\tag{25.18}
$$

which shows the pole when $|\lambda|$ is large enough to have a bound state. It also shows that we can solve this equation for $|\lambda|$ somewhat smaller than the value that gives a bound state, but κ will be negative. This would correspond to k_0 being on the negative imaginary axis, which corresponds to z not having a phase between 0 and 2π, but a phase of $-i\pi$, that is it corresponds to an energy on the negative real axis, but not on the physical Riemann sheet. We need to analytically continue z through the branch cut onto the second sheet, and then to the negative real axis. The corresponding κ will then be negative and the continued scattering amplitude will have a pole. As $|\lambda|$ increases this pole eventually moves to the branch point at $k_0 = 0$, and then onto the physical sheet and along the negative real axis of the physical sheet. A pole corresponding to k_0 on the negative imaginary axis is sometimes called a virtual state – becoming a real bound state if the potential's attraction is increased sufficiently. The presence of a virtual state's pole near $k_0 = 0$ can affect the physical low energy scattering.

The delta-shell potential also has resonances when λ is large and positive. In this case, we can think of a particle with $r < a$ as almost trapped in a spherical cavity. If the the projectile particle's energy is close to the corresponding spherical cavity energy, the particle can tunnel into this region and be trapped for a relatively long time if λ is large. This gives rise to a large cross section. Resonances occur for a variety of reasons, but the behavior of the cross section is much the same.

In order for the cross section to become large, we expect the denominator of $T(z)$ to be small when $z = E + i\eta$. In order for the denominator to be small, both its real and imaginary

parts must be small. We begin by looking for energies where the real part is small. The denominator writing $h_\ell^{(1)}(x) = j_\ell(x) + in_\ell(x)$ is

$$1 - \frac{2m\lambda k_0 a^2}{\hbar^2} j_\ell(k_0 a) n_\ell(k_0 a) + i\frac{2m\lambda k_0 a^2}{\hbar^2} j_\ell^2(k_0 a) . \tag{25.19}$$

For λ large, the real part will be zero only when either $j_\ell(k_0 a)$ or $n_\ell(k_0 a)$ is small. If $n_\ell(k_0 a)$ is near a zero, $j_\ell(k_0 a)$ will not be, so that the imaginary part of the denominator will be large. The cross section will not be particularly large. These are sometimes called broad resonances. The more interesting terms are near the zeroes of $j_\ell(k_0 a)$. In that case $j_\ell(k_0 a)$ will be small, and therefore $j_\ell^2(k_0 a)$ will be very small. These are sometimes called sharp resonances. We can first expand around a zero of $j_\ell(k_0 a)$ for real positive k_0, $j_\ell(k_z a) = 0$, using the Wronskian we get

$$k_0 a j_\ell(k_0 a) n_\ell(k_0 a) = -\frac{1}{k_z}(k_0 - k_z) + \dots \tag{25.20}$$

The real part of the denominator will then be zero when $k_0 = k_r$ with

$$1 + \frac{2m\lambda a}{\hbar^2 k_z}(k_r - k_z) = 0 \tag{25.21}$$

or

$$k_r = \left[1 - \frac{\hbar^2}{2m\lambda a}\right] k_z . \tag{25.22}$$

Therefore, for k_0 near k_r and $|\lambda$ large the denominator is

$$\frac{2m\lambda a}{\hbar^2 k_r} \left[(k_0 - k_r) + ik_r^2 a j_\ell^2(k_r a)\right] . \tag{25.23}$$

Taking k_0 to be complex, this would have a pole for k_0 in the lower half plane. Again, since the physical Riemann sheet corresponds to k_0 only in the upper half plane, to get k_0 in the lower half plane, we must analytically continue to the second Riemann energy sheet. Sharp resonances correspond to poles with small negative imaginary parts which are on this second sheet.

Chapter 26

K-matrix

26.1 The K-matrix

When we calculate using approximate methods, including numerical methods, sometimes the resulting approximate S-matrix will not be unitary. Since this can lead to completely unphysical results, approximate methods that guarantee a unitary S-matrix are often preferred. One way to formulate the scattering problem that guarantees a unitary S-matrix is to use the K-matrix, which is also known as the reaction or sometimes the reactance matrix. In the analogous electromagnetic scattering, the on shell K-matrix corresponds to the reactance matrix which relates the voltage across the "ports" to the current through them.

The nonunitarity problem usually arises because we approximately solve the T-matrix equations, and while the S-matrix is unitary, the T matrix has no easily enforced special properties. An alternative way of describing scattering is to look at the off shell T-matrix equations for $T^{\pm}(E)$

$$T^{\pm}(E) = V + V \frac{1}{E - H_0 \pm i\eta} T^{\pm}(E) \tag{26.1}$$

and ask how we could modify the Green's function to obtain a Hermitian result. If we iterate the T-matrix, the terms that contain a Green's function are not Hermitian because the Green's function is not Hermitian. We can find a Hermitian Green's function by taking the average of the T-matrix Green's function with its Hermitian conjugate. That is

$$\frac{1}{2}\left(\frac{1}{E - H_0 + i\eta} + \frac{1}{E - H_0 - i\eta} \right) = \mathcal{P} \frac{1}{E - H_0} \, . \tag{26.2}$$

where $\mathcal{P}$ means the principle parts integration. The K-matrix equation is then

$$K(E) = V + V\mathcal{P} \frac{1}{E - H_0} K(E) \, . \tag{26.3}$$

Since the principal parts Green's function is half the sum of the retarded and advanced Green's functions, the scattering wave packets will not be so easily constructed as with the retarded solution, however, we know that since these form a complete set, all of the scattering information is contained in their solution. We will now work on extracting that information.

Iterating Eq. 26.3, every term is hermitian, so the resulting $K(E)$ is hermitian. The corresponding integral equation for the state is

$$|\Psi_\alpha\rangle = |\Phi_\alpha\rangle + \mathcal{P}\frac{1}{E - H_0}V|\Psi_\alpha\rangle, \tag{26.4}$$

which are sometimes called the standing-wave solutions. The K-matrix in terms of these solutions is

$$K_{\beta\alpha}(E) = \langle\Phi_\beta|K(E)|\Phi_\alpha\rangle = \langle\Phi_\beta|V|\Psi_\alpha\rangle \tag{26.5}$$

Notice that while the K-matrix with a particular fixed E is hermitian, the half on shell K-matrix, where $E = E_\alpha$ changes with the state α, is not hermitian. The fully on shell K-matrix where we are restricted to states where both α and β have the same energy $E = E_\alpha = E_\beta$ is hermitian. This means that, like the unitary S-matrix, the eigenstates of the on-shell K-matrix form an orthogonal set. In addition, its eigenvalues will be real.

We can write the T-matrix equation as

$$T^+(E) = V + V\frac{1}{E - H_0 + i\eta}T^+(E) = V + V\left[\mathcal{P}\frac{1}{E - H_0} - i\pi\delta(E - H_0)\right]T^+(E). \tag{26.6}$$

Iterating the last form, we can write out all of the terms and there are terms that have no delta functions and other terms which contain at least one delta function and may also have principal parts Green's function contributions. If we gather together all of the terms that have no delta functions, these give $K(E)$. All other terms have at least one delta function in the product, and therefore when looking from the left have a first delta function. To the left of this first delta function, we can have any number (including zero) of principal parts Green's functions. The sum of all possible such terms is $K(E)$. To the right of the first delta function we can have any number of either principal parts or delta function contributions – but the sum of these is $T^+(E)$. Therefore we can write

$$T^+(E) = K(E) + K(E)[-i\pi\delta(E - H_0)]T^+(E). \tag{26.7}$$

Operating on the left with $\langle\Phi_\beta|$, on the right with $|\Phi_\alpha\rangle$ and inserting a complete set of $|\Phi_\gamma\rangle$ states we can write

$$\langle\Phi_\beta|T^+(E)|\Phi_\alpha\rangle = \langle\Phi_\beta|K(E)|\Phi_\alpha\rangle - i\pi\int d\gamma\delta(E - E_\gamma)\langle\Phi_\beta|K(E)|\Phi_\gamma\rangle\langle\Phi_\gamma|T^+(E)|\Phi_\alpha\rangle \tag{26.8}$$

and now setting $E = E_\alpha$ and multiplying by $-i2\pi\delta(E_\alpha - E_\beta)$ to select just the on-shell matrix elements, we get

$$
\begin{aligned}
-i2\pi\delta(E_\alpha - E_\beta)T^+_{\beta\alpha} &= -i2\pi\delta(E_\alpha - E_\beta)K_{\beta\alpha} \\
&+ \int d\gamma[-i\pi\delta(E_\beta - E_\gamma)K_{\beta\gamma}][-i2\pi\delta(E_\gamma - E_\alpha)T^+_{\gamma\alpha}]
\end{aligned}
\tag{26.9}
$$

where we drop the energy arguments to indicate that these are all on shell. Since the S-matrix is

$$S_{\beta\alpha} = \delta(\alpha - \beta) - i2\pi\delta(E_\alpha - E_\beta)T^+_{\beta\alpha} \tag{26.10}$$

this becomes

$$\int d\gamma \left[\delta(\beta - \gamma) + i\pi\delta(E_\beta - E_\gamma)K_{\beta\gamma}\right] S_{\gamma\alpha} = \delta(\beta - \alpha) - i\pi\delta(E_\beta - E_\alpha)K_{\beta\alpha}. \tag{26.11}$$

Defining the on-shell K-matrix multiplied by $\pi\delta(E_\beta - E_\alpha)$ as

$$\mathcal{K}_{\beta\alpha} \equiv \pi\delta(E_\beta - E_\alpha)K_{\beta\alpha} \tag{26.12}$$

and the corresponding operator

$$\mathcal{K} = \int d\beta \int d\alpha |\Phi_\beta\rangle \mathcal{K}_{\beta\alpha} \langle\Phi_\alpha| \tag{26.13}$$

This is

$$[1 + i\mathcal{K}] S = [1 - i\mathcal{K}] \tag{26.14}$$

or

$$\mathcal{K} = -i\frac{1 - S}{1 + S}. \tag{26.15}$$

$\mathcal{K}$ therefore has the same eigenvectors as S, and if the eigenvector of S has eigenvalue $e^{2i\delta}$, the corresponding eigenvalue of $\mathcal{K}$ will be

$$-i\frac{1 - e^{2i\delta}}{1 + e^{2i\delta}} = -\tan\delta. \tag{26.16}$$

An alternative method to get to this same place is to look at the wave packet as we did previously for the T-matrix equations. In this way, we see that a state with amplitude $g(\gamma) = \delta(\gamma - \alpha)$ has amplitude of state β at large negative times of $\delta(\alpha - \beta) + i\pi\delta(E_\alpha - E_\beta)K_{\beta\alpha}$, and at large positive times it becomes $\delta(\alpha - \beta) - i\pi\delta(E_\alpha - E_\beta)K_{\beta\alpha}$. So multiplying the first expression which represents the large negative time solution of the K-matrix state, by the S-matrix must give the corresponding K-matrix state at large positive times, which is the second expression. Writing the $\delta(\alpha - \beta)$ as the unit operator, and $\pi\delta(E_\alpha - E_\beta)K_{\beta\alpha}$ as $\mathcal{K}$, we have, $S(1 + i\mathcal{K}) = (1 - i\mathcal{K})$, or

$$\begin{aligned} S &= \frac{1 - i\mathcal{K}}{1 + i\mathcal{K}} \\ \mathcal{K} &= -i\frac{1 - S}{1 + S} \end{aligned} \tag{26.17}$$

as before.

As an aside, we can also define the on-shell quantity

$$\mathcal{T}_{\beta\alpha} = -\pi\delta(E_\alpha - E_\beta)T_{\beta\alpha} \tag{26.18}$$

We then have $S = 1 + 2i\mathcal{T}$, so that S and $\mathcal{T}$ share the same eigenvectors, and the eigenvalues of $\mathcal{T}$ are $\sin\delta e^{i\delta}$ as in the scattering amplitude. We have the relationship

$$\mathcal{K}\mathcal{T} = 1 + i\mathcal{T} \tag{26.19}$$

or equivalently

$$\mathcal{K} = i + \mathcal{T}^{-1}. \tag{26.20}$$

To solve for the scattering, we can solve the K-matrix equations. Any standard approximate technique that gives an approximation to the K-matrix that is hermitian, will give orthogonal eigenvectors of the on-shell K-matrix, with real eigenvalues. These real eigenvalues tell us $-\tan\delta$ which then uniquely gives the phases $e^{2i\delta}$ which are the corresponding eigenvalues of the S-matrix, which in turn gives us the cross sections, etc.

26.2   Central potential Example

Let's look at the case where the potential is rotationally invariant but nonlocal. We take the kinetic energy to be $H_0 = \frac{p^2}{2m}$. Since the full Hamiltonian commutes with angular momentum we can first find its eigenvectors and the Hamiltonian will be block diagonal in these states – the K-matrix will also have this symmetry. We write a basis given by the magnitude of the wave vector and the angular momentum states as $|k\ell m\rangle$. A consistent set of normalizations and completeness relations are

$$\langle \boldsymbol{k}'|\boldsymbol{k}\rangle = (2\pi)^3 \delta^3(\boldsymbol{k} - \boldsymbol{k}')$$

$$1 = \int \frac{d^3k}{(2\pi)^3}|\boldsymbol{k}\rangle\langle\boldsymbol{k}|$$

$$\langle k'\ell'm'|k\ell m\rangle = \frac{(2\pi)^3\delta(k-k')}{k^2}\delta_{\ell\ell'}\delta_{mm'}$$

$$1 = \sum_{\ell m}\frac{1}{(2\pi)^3}\int_0^\infty dk\, k^2 |k\ell m\rangle\langle k\ell m|$$

$$\langle \boldsymbol{k}'|k\ell m\rangle = \frac{(2\pi)^3\delta(k-k')}{k^2}Y_{\ell m}(\hat{\boldsymbol{k}}') \tag{26.21}$$

Since the potential is rotationally invariant, it cannot change the angular momentum states. Furthermore rotating around x or y will change the m value but not the matrix element, so the matrix elements of V will be independent of m,

$$\langle k'\ell'm'|V|k\ell m\rangle = V_\ell(k', k)\delta_{\ell\ell'}\delta_{mm'}. \tag{26.22}$$

We can specify the potential by giving the functions $V_\ell(k', k)$. To demonstrate the rotational symmetry of the K-matrix, we can either write out the iterated equation and realize that each term is rotationally invariant or write

$$\langle k'\ell'm'|K|k\ell m\rangle = \langle k'\ell'm'|V|k\ell m\rangle + \langle k'\ell'm'|V\mathcal{P}\frac{1}{E - H_0}K|k\ell m\rangle$$

$$= \delta_{\ell\ell'}\delta_{mm'}V_\ell(k', k) + \mathcal{P}\int\frac{dk''}{(2\pi)^3}k''^2\frac{V_\ell'(k', k'')}{E - E_{k''}}\langle k''\ell'm'|K|k\ell m\rangle \tag{26.23}$$

310

If ℓ', m' are not equal to ℓm, there is no source term and the solution is $\langle k'\ell'm'|K|k\ell m\rangle = 0$
Therefore we can write

$$\langle k'\ell'm'|K|k\ell m\rangle = \delta_{\ell\ell'}\delta_{mm'}K_\ell(k',k)\,. \tag{26.24}$$

The channel equations are then

$$K_\ell(k',k) = V_\ell(k',k) - \frac{m}{4\pi^3\hbar^2}\mathcal{P}\int_0^\infty dk''\frac{k''^2 V_\ell(k',k'')K_\ell(k'',k)}{k''^2 - k_0^2} \tag{26.25}$$

where we write $E = \frac{\hbar^2 k_0^2}{2m}$.

Before going on, let's extract the phase shifts. We know that the on shell K-matrix has $|k\ell m\rangle$ eigenfunctions, so

$$\langle k\ell m|\mathcal{K}|k'\ell'm'\rangle = -(2\pi)^3\frac{\delta(k-k')}{k^2}\delta_{\ell\ell'}\delta_{mm'}\tan\delta_\ell(k) = \pi\delta(E_k - E'_k)\,K_\ell(k,k')|_{k_0=k}\,. \tag{26.26}$$

Multiplying by $k^2/(2\pi)^3$ and integrating over k', we get

$$\tan\delta_l(k) = -\frac{km}{8\pi^2\hbar^2}K_\ell(k,k)\,, \tag{26.27}$$

where $k_0 = k$.

26.3   Numerical solution method

To make the principal parts integration a little easier numerically, we can calculate directly that[1]

$$\mathcal{P}\int_0^\infty dx\frac{1}{x^2 - x_0^2} = 0 \tag{26.29}$$

so that we can write

$$
\begin{aligned}
K_\ell(k',k) &= V_\ell(k',k) - \frac{m}{4\pi^3\hbar^2}\mathcal{P}\int_0^\infty dk''\frac{k''^2 V_\ell(k',k'')K_\ell(k'',k) - k_0^2 V_\ell(k',k_0)K_\ell(k_0,k)}{k''^2 - k_0^2}\\
&= V_\ell(k',k) - \frac{m}{4\pi^3\hbar^2}\int_0^\infty dk''\frac{k''^2 V_\ell(k',k'')K_\ell(k'',k) - k_0^2 V_\ell(k',k_0)K_\ell(k_0,k)}{k''^2 - k_0^2}
\end{aligned} \tag{26.30}
$$

---

[1]Taking x_0 positive,

$$
\begin{aligned}
\mathcal{P}\int_0^L dx\frac{1}{x^2 - x_0^2} &= \mathcal{P}\frac{1}{2x_0}\int_0^L dx\left[\frac{1}{x-x_0} - \frac{1}{x+x_0}\right]\\
&= \frac{1}{2x_0}\lim_{\eta\to 0}\left[\int_0^{x_0-\eta}dx\frac{1}{x-x_0} + \int_{x_0+\eta}^L dx\frac{1}{x-x_0} - \ln(L+x_0) + \ln(x_0)\right]\\
&= \frac{1}{2x_0}\ln\left(\frac{L-x_0}{L+x_0}\right) \tag{26.28}
\end{aligned}
$$

and taking the limit $L\to\infty$ gives 0.

and we drop the principal parts since there is now no singularity at $k'' = k_0$.

If we now replace the integration with any standard numerical integration formula, this will become a set of linear equations which we can solve. To be specific, let's use the trapezoidal rule for the integration. We define

$$\Delta k = \frac{k_{\max}}{N_k} \tag{26.31}$$

where we choose a maximum k value, $k_{\max}$, such that the potential (and therefore K) is zero or small enough to ignore. We choose N_k large enough to give a step size that captures the physics. We can increase these to check numerically that we have chosen adequately large values. Since we don't need to use the $k = k_{\max}$ point we can label the k points as

$$k_i = \begin{cases} i\Delta k & 0 \le i < N_k \\ k_0 & i = N_k \end{cases} \tag{26.32}$$

Plugging in the trapezoidal rule and taking $k = k_0$ we get

$$K_\ell(k_i, k_{N_k}) = V_\ell(k_i, k_{N_k}) - \frac{m\Delta k}{4\pi^3\hbar^2} \sum_{j=0}^{N_k-1} [1 - \tfrac{1}{2}\delta_{j0}] \frac{k_j^2 V_\ell(k_i, k_j) K_\ell(k_j, k_{N_k}) - k_{N_k}^2 V_\ell(k_i, k_{N_k}) K_\ell(k_{N_k}, k_{N_k})}{k_j^2 - k_0^2}$$

$$\tag{26.33}$$

where the kronecker delta enforces the trapezoidal rule at the endpoint $k_j = 0$. The other endpoint has been dropped since we assume $V = 0$ there. We can write this in matrix notation as

$$v_i = \sum_{j=1, N_k} A_{ij} K_j \tag{26.34}$$

where

$$\begin{aligned} v_i &= V_\ell(k_i, k_{N_k}) \\ K_i &= K_\ell(k_i, k_{N_k}) \end{aligned} \tag{26.35}$$

and

$$A_{ij} = \begin{cases} \delta_{ij} + \frac{m\Delta k}{4\pi^3\hbar^2} \frac{k_i^2 V_\ell(k_i, k_j)}{k_j^2 - k_{N_k}^2} & j \ne N_k \\ -\sum_{n=0}^{N_k-1} [1 - \tfrac{1}{2}\delta_{n,0}] \frac{m\Delta k}{4\pi^3\hbar^2} \frac{k_{N_k}^2 V_\ell(k_i, k_{N_k})}{k_n^2 - k_{N_k}^2} & j = N_k \end{cases} \tag{26.36}$$

Since calculating the integrand for k_0 equal to an integration point would require taking a limit, which would complicate the numerics, we always choose k_0 so that it is not one of the integration points. Taking the k_0 points to be, in turn, the midpoints of the intervals i.e. at $(i - \tfrac{1}{2})\Delta k$ is simplest.

26.4   Problems

Problem 26.1. Study the separable potential of problem 23.1 using the K-matrix formalism.

(a) Solve the K-matrix Lippmann-Schwinger equation with the separable potential for the matrix elements $K_\ell(k', k)$. Your result should be

$$K_\ell(k', k) = \frac{4\pi\lambda_\ell}{2\ell + 1} \frac{\tilde{u}_\ell(k')\tilde{u}_\ell(k)}{1 - \frac{\lambda_\ell m}{\pi^2(2\ell+1)}\mathcal{P}\int_0^\infty dk'' \frac{k''^2 \tilde{u}_\ell^2(k'')}{k_0^2 - k''^2}} \qquad (26.37)$$

where $\frac{\hbar^2 k_0^2}{2m}$ is the energy E.

(b) From your results of part (a), calculate an expression for the phase shifts for energy E. Your result should be

$$\tan\delta_\ell = -\frac{km\lambda_\ell}{2\pi(2\ell + 1)\hbar^2} \frac{\tilde{u}_\ell^2(k_0 a)}{1 - \frac{\lambda_\ell m}{\hbar^2\pi^2(2\ell+1)}\mathcal{P}\int_0^\infty dk'' \frac{k''^2 \tilde{u}_\ell^2(k'')}{k_0^2 - k''^2}} \qquad (26.38)$$

(c) Evaluate the integral in the case of the delta shell potential where $v(r) = \lambda\delta(r - a)$, and verify that your phase shifts agree with those obtained by direct integration

$$\tan\delta_\ell = -\frac{\frac{2m\lambda a^2 k_0}{\hbar^2} j_\ell^2(k_0 a)}{1 - \frac{2m\lambda a^2 k_0}{\hbar^2} j_\ell(k_0 a)n_\ell(k_0 a)} . \qquad (26.39)$$

Chapter 27

Quantizing the electromagnetic field

27.1 Lorentz Force Law

We begin with a review of the classical electromagnetism we will need. The Lorentz force on a particle of charge q_k at the position $\boldsymbol{r}_k(t)$, interacting with electric and magnetic fields[1] $\boldsymbol{E}$ and $\boldsymbol{B}$, is

$$\boldsymbol{F}_k(t) = q_k \left[\boldsymbol{E}(\boldsymbol{r}_k(t), t) + \frac{1}{c} \boldsymbol{v}_k(t) \times \boldsymbol{B}(\boldsymbol{r}_k(t), t) \right] \tag{27.1}$$

where the velocity is

$$\boldsymbol{v}_k(t) = \frac{d\boldsymbol{r}_k(t)}{dt} . \tag{27.2}$$

27.2 Interaction Lagrangian

A typical classical procedure uses experiments to find the form of the force, and then a Lagrangian is developed to give that force. The charge and current densities of a set of point charges are

$$\rho(\boldsymbol{r}, t) = \sum_j q_j \delta^3(\boldsymbol{r} - \boldsymbol{r}_j(t))$$

$$\boldsymbol{J}(\boldsymbol{r}, t) = \sum_j q_j \boldsymbol{v}_j \delta^3(\boldsymbol{r} - \boldsymbol{r}_j(t)) \tag{27.3}$$

and the interaction Lagrangian found is

$$\begin{aligned} \mathcal{L}_{\text{int}} &= \frac{1}{c} \int d^3 r \, \boldsymbol{J}(\boldsymbol{r}, t) \cdot \boldsymbol{A}(\boldsymbol{r}, t) - \int d^3 r \, \rho(\boldsymbol{r}, t) \Phi(\boldsymbol{r}, t) \\ &= \frac{1}{c} \sum_j q_j \boldsymbol{v}_j(t) \cdot \boldsymbol{A}(\boldsymbol{r}_j(t), t) - \sum_j q_j \Phi(\boldsymbol{r}_j(t), t) \end{aligned} \tag{27.4}$$

---

[1]Here $\boldsymbol{B}$ is the fundamental field. Some people call it the magnetic flux density and call the derived field, $\boldsymbol{H}$, the magnetic field.

where $\boldsymbol{A}$ and Φ are the scalar and vector potentials.

We can verify that this is the correct interaction Lagrangian by direct calculation. Taking a nonrelativistic kinetic energy, the Lagrangian is

$$\mathcal{L} = \sum_j \left[\frac{1}{2} m_j v_j^2 + \frac{1}{c} q_j \boldsymbol{v}_j \cdot \boldsymbol{A}(\boldsymbol{r}_j, t) - q_j \Phi(\boldsymbol{r}_j, t) \right] \tag{27.5}$$

where we drop writing the time dependence of $\boldsymbol{r}_j$ and $\boldsymbol{v}_j$ to simplify the notation, but remember that that time dependence is still there. Extremizing the action gives the Lagrange equations

$$
\begin{aligned}
0 &= \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial v_{\alpha k}} - \frac{\partial \mathcal{L}}{\partial r_{\alpha k}} \\
&= \frac{d}{dt} \left[m_k v_{k\alpha} + q_k \frac{1}{c} A_\alpha(\boldsymbol{r}_k, t) \right] + q_k \frac{\partial \Phi(\boldsymbol{r}_k, t)}{\partial r_{\alpha k}} - q_k \frac{1}{c} \boldsymbol{v}_k \cdot \frac{\partial \boldsymbol{A}(\boldsymbol{r}_k, t)}{\partial r_{k\alpha}},
\end{aligned}
\tag{27.6}
$$

where α indicates the component (i.e. x, y, or z). Notice that the momentum conjugate to $\boldsymbol{r}_k$ is not $m\boldsymbol{v}_k$. The momentum is given by the derivative of the Lagrangian with respect to the velocity and is therefore the quantity in the brackets

$$\boldsymbol{p}_k = m_k \boldsymbol{v}_k + \frac{q_k}{c} \boldsymbol{A}(\boldsymbol{r}_k, t). \tag{27.7}$$

Expanding the time derivative using the chain rule, gives in vector notation

$$0 = m_k \frac{d\boldsymbol{v}_k}{dt} + \frac{q_k}{c} \frac{\partial \boldsymbol{A}(\boldsymbol{r}_k, t)}{\partial t} + \frac{q_k}{c} (\boldsymbol{v}_k \cdot \boldsymbol{\nabla}_k) \boldsymbol{A}(\boldsymbol{r}_k, t) + q_k \boldsymbol{\nabla}_k \Phi(\boldsymbol{r}_k, t) - \frac{q_k}{c} \boldsymbol{\nabla}_k [\boldsymbol{v}_k \cdot \boldsymbol{A}(\boldsymbol{r}_k, t)]. \tag{27.8}$$

We can now use the identity

$$\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = \boldsymbol{B}(\boldsymbol{A} \cdot \boldsymbol{C}) - \boldsymbol{C}(\boldsymbol{A} \cdot \boldsymbol{B}) \tag{27.9}$$

being careful to continue operating the derivatives on $\boldsymbol{A}$ to write

$$\boldsymbol{\nabla}(\boldsymbol{v} \cdot \boldsymbol{A}) - (\boldsymbol{v} \cdot \boldsymbol{\nabla})\boldsymbol{A} = \boldsymbol{v} \times (\boldsymbol{\nabla} \times \boldsymbol{A}) \tag{27.10}$$

so that

$$m_k \frac{d\boldsymbol{v}_k}{dt} = q_k \left[-\boldsymbol{\nabla}_k \Phi(\boldsymbol{r}_k(t), t) - \frac{1}{c} \frac{\partial \boldsymbol{A}(\boldsymbol{r}_k, t)}{\partial t} \right] + \frac{q_k}{c} \boldsymbol{v}_k \times [\boldsymbol{\nabla}_k \times \boldsymbol{A}(\boldsymbol{r}_k(t), t)] \tag{27.11}$$

which is the Lorentz Force law, since the electric and magnetic fields are

$$
\begin{aligned}
\boldsymbol{E} &= -\boldsymbol{\nabla}\Phi - \frac{1}{c} \frac{\partial \boldsymbol{A}}{\partial t} \\
\boldsymbol{B} &= \boldsymbol{\nabla} \times \boldsymbol{A}.
\end{aligned}
\tag{27.12}
$$

27.3 The Hamiltonian of the particles and interaction

To use the Schrödinger equation, we want the Hamiltonian. We Legendre transform from the Lagrangian which is a function of the position and velocity, to the variables position and the derivative of the Lagrangian with respect to velocity which we defined as the conjugate momentum,

$$H = \sum_j \boldsymbol{v}_j \cdot \boldsymbol{p}_j - \mathcal{L} \tag{27.13}$$

and H must be written in terms of position and momentum. From the equation for the momentum, the velocity is

$$\boldsymbol{v}_k = \frac{\boldsymbol{p}_k - \frac{q_k}{c}\boldsymbol{A}(\boldsymbol{r}_k,t)}{m_k}. \tag{27.14}$$

Substituting this into the expression for H to eliminate the velocity, after a bit of algebra, the Hamiltonian becomes

$$H = \sum_j \left[\frac{[\boldsymbol{p}_j - \frac{q_j}{c}\boldsymbol{A}(\boldsymbol{r}_j,t)]^2}{2m_k} + q_j \Phi(\boldsymbol{r}_j,t) \right]. \tag{27.15}$$

You should verify that Hamilton's equations

$$\begin{aligned}
\dot{\boldsymbol{r}}_k &= \boldsymbol{\nabla}_{p_k} H \\
\dot{\boldsymbol{v}}_k &= -\boldsymbol{\nabla}_k H
\end{aligned} \tag{27.16}$$

gives the Lorentz force law, where $\boldsymbol{\nabla}_{p_k}$ means taking the derivatives with respect to the components of $\boldsymbol{p}_k$.

Notice the simplicity of the inclusion of the coupling in the Hamiltonian formalism. We make the substitutions

$$\boldsymbol{p}_j \rightarrow \boldsymbol{p}_j - \frac{q_j}{c}\boldsymbol{A}(\boldsymbol{r}_j,t)$$
$$H \rightarrow H + \sum_j q_j \Phi(\boldsymbol{r}_j,t) \tag{27.17}$$

Essentially, the energy and momentum are modified by simply subtracting the scalar and vector potential terms. The modern view point is that this form is dictated by gauge invariance and the requirements of quantum mechanics.

27.4 Quantum charge density and current density operators

The classical charge density depends only on the position of the particles, so changing $\boldsymbol{r}_j$ to the corresponding operator $\boldsymbol{r}_{j\ \mathrm{op}}$ gives the quantum charge density,

$$\rho_{\mathrm{op}}(\boldsymbol{r},t) = \sum_j q_j \delta^3(\boldsymbol{r} - \boldsymbol{r}_{j\ \mathrm{op}}). \tag{27.18}$$

The classical velocity is given by Eq. 27.14. If we substitute these operators into the classical current density expressions with the momentum to the left of the position operators the result is not Hermitian. To obtain a Hermitian operator, we can symmetrize the order and our proposed current density operator is

$$\boldsymbol{J}(\boldsymbol{r},t) = \sum_j \left\{ \frac{q_j}{2m_j} \left[\boldsymbol{p}_{j\ \text{op}} \delta^3(\boldsymbol{r} - \boldsymbol{r}_{j\ \text{op}}) + \delta^3(\boldsymbol{r} - \boldsymbol{r}_{j\ \text{op}}) \boldsymbol{p}_{j\ \text{op}} \right] - \frac{q_j^2}{m_j c^2} \boldsymbol{A}(\boldsymbol{r}_{j\ \text{op}}, t) \delta^3(\boldsymbol{r} - \boldsymbol{r}_{j\ \text{op}}) \right\}$$

(27.19)

As an exercise, you should verify that the continuity equation

$$\boldsymbol{\nabla} \cdot \langle \psi(t) | \boldsymbol{J}(\boldsymbol{r},t) | \phi(t) \rangle + \frac{\partial}{\partial t} \langle \psi(t) | \rho(\boldsymbol{r}) | \phi(t) \rangle = 0 \tag{27.20}$$

is satisfied by all matrix elements when $|\psi(t)\rangle$ and $|\phi(t)\rangle$ are solutions of the Schrödinger equation $H|\phi(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\phi(t)\rangle$, $H|\psi(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi(t)\rangle$. We therefore take Eq. 27.19 to be the quantum current density operator.

It is sometimes convenient to write the quantum Hamiltonian Eq. 27.15 as

$$
\begin{aligned}
H_0 &= \sum_j \frac{p_j^2}{2m_j} \\
H &= H_0 + \int d^3r \, \boldsymbol{J}(\boldsymbol{r},t) \cdot \boldsymbol{A}(\boldsymbol{r},t) + \int d^3r \, \rho(\boldsymbol{r}) \Phi(\boldsymbol{r},t) \, .
\end{aligned}
\tag{27.21}
$$

27.5 Classical Gauge Invariance

Classically we can modify the potentials by adding the derivatives of a scalar function

$$
\begin{aligned}
\boldsymbol{A}'(\boldsymbol{r},t) &= \boldsymbol{A}(\boldsymbol{r},t) + \boldsymbol{\nabla}\chi(\boldsymbol{r},t) \\
\Phi'(\boldsymbol{r},t) &= \Phi(\boldsymbol{r},t) - \frac{1}{c}\frac{\partial \chi(\boldsymbol{r},t)}{\partial t} \, .
\end{aligned}
\tag{27.22}
$$

Since the curl of a gradient is zero, the curl of $\boldsymbol{A}'$ will be equal to the curl of $\boldsymbol{A}$ and the magnetic field will be unchanged. Similarly, gradient of the additional term in the scalar potential will cancel the time derivative of the additional term in the vector potential so that the electric field is unchanged. The Lorentz force law will therefore remain the same, and the classical motion will be unchanged.

27.6 Quantum Gauge Invariance

Let's begin with a single free particle of charge q. The Hamiltonian is

$$H_0 = \frac{p^2}{2m} \tag{27.23}$$

and the energy eigenfunctions are plane waves $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$, with eigenvalues $\hbar^2 k^2/2m$. We can view this as the solution with the scalar and vector potentials equal to zero. Let's make a gauge transformation using the function $\chi(\boldsymbol{r}, t) = \boldsymbol{\alpha} \cdot \boldsymbol{r}$ with $\boldsymbol{\alpha}$ a constant vector. The scalar potential is still zero, while the transformed vector potential is $\boldsymbol{A}' = \boldsymbol{\alpha}$. The transformed Hamiltonian is

$$H = \frac{(\boldsymbol{p} - \frac{q}{c}\boldsymbol{\alpha})^2}{2m} \tag{27.24}$$

and we can integrate the Schrödinger equation to obtain the energy eigenvectors

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r} + i\frac{q\boldsymbol{\alpha}\cdot\boldsymbol{r}}{\hbar c}} = e^{\frac{iq}{\hbar c}\chi(\boldsymbol{r},t)} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{27.25}$$

with energies $\hbar^2 k^2/2m$. Notice that the physically measurable energies do not change, but the phase of the wave function has changed to compensate the change in the Hamiltonian.

Notice that the gauge function χ gave an additional phase to the wave function when evaluated at the particle position $\boldsymbol{r}$. Multiplying all wave functions by an overall phase has no affect on the physical matrix elements. Here we see that a gauge transformation is a method of changing the phase of the wave function depending on the particle position (and its charge). A modern view point is that this invariance under a local change of phase is the property that is fundamental and therefore should be built into our quantum mechanical theories. The steps are as follows (here we will take a single particle, and the generalization to many particles is left as an exercise, see Baym chapter 13, problem 1).

First we operate with $\boldsymbol{p} - \frac{q}{c}A'(\boldsymbol{r}, t)$ on $e^{\frac{iq}{\hbar c}\chi(\boldsymbol{r},t)}\psi(\boldsymbol{r}, t)$. In the r representation, the derivative from the $\boldsymbol{p}$ operator brings down a gradient of χ giving

$$\begin{aligned}
[\boldsymbol{p} - \frac{q}{c}A'(\boldsymbol{r}, t)]e^{\frac{iq}{\hbar c}\chi(\boldsymbol{r},t)}\psi(\boldsymbol{r}, t) &= e^{\frac{iq}{\hbar c}\chi(\boldsymbol{r},t)}\left[\boldsymbol{p} - \frac{q}{c}\left(\boldsymbol{A}'(\boldsymbol{r}, t) - \boldsymbol{\nabla}\chi(\boldsymbol{r}, t)\right)\right]\psi(\boldsymbol{r}, t) \\
&= e^{\frac{iq}{\hbar c}\chi(\boldsymbol{r},t)}\left[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r}, t)\right]\psi(\boldsymbol{r}, t). \tag{27.26}
\end{aligned}$$

Similarly operating with $H' + \frac{\hbar}{i}\frac{\partial}{\partial t}$ shows that

$$\begin{aligned}
\left[H' + \frac{\hbar}{i}\frac{\partial}{\partial t}\right]e^{\frac{iq}{\hbar c}\chi(\boldsymbol{r},t)}\psi(\boldsymbol{r}, t) &= e^{\frac{iq}{\hbar c}\chi(\boldsymbol{r},t)}\left[H' + \frac{q}{c}\frac{\partial\chi(\boldsymbol{r}, t)}{\partial t} + \frac{\hbar}{i}\frac{\partial}{\partial t}\right]\psi(\boldsymbol{r}, t) \\
&= e^{\frac{iq}{\hbar c}\chi(\boldsymbol{r},t)}\left[H + \frac{\hbar}{i}\frac{\partial}{\partial t}\right]\psi(\boldsymbol{r}, t). \tag{27.27}
\end{aligned}$$

We see that the gauge terms are exactly those needed to compensate for the derivatives of the arbitrary phase of the wave function. All matrix elements will be invariant if we simultaneously gauge transform the momentum and Hamiltonian operators along with a change of phase as a function of position of the wave function.[2]

---

[2]The modern philosophy is that these gauge transformation functions must represent dynamical degrees of freedom, and therefore we need to add electromagnetism as a massless vector field to our quantum theory as a gauge field. Similar arguments about the change of phase of, for example, the two component spinors leads to nonabelian gauge fields, but this is all well beyond the scope of this course.

27.7 Gauge Choice

Since we have shown that all physical matrix elements will be independent of the gauge, we can pick the gauge that makes our problems simplest. For nonrelativistic quantum mechanics, this is usually Coulomb gauge. Another popular choice is the Lorentz gauge where the potentials become a four vector,

$$\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0 \qquad \text{Coulomb gauge}$$
$$\boldsymbol{\nabla} \cdot \boldsymbol{A} + \frac{1}{c}\frac{\partial \Phi}{\partial t} = 0 \qquad \text{Lorentz gauge}. \tag{27.28}$$

There are many other gauge choices, e.g. $A_z = 0$ or $\Phi = 0$ are two other possible gauge choices. If we calculate accurately in any gauge, we must get the same results.

27.8 Interacting charges in Coulomb gauge

By writing the fields in terms of the vector and scalar potentials, the two Maxwell equations

$$\begin{aligned}
\boldsymbol{\nabla} \cdot \boldsymbol{B} &= 0 \\
\boldsymbol{\nabla} \times \boldsymbol{E} &= -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}
\end{aligned} \tag{27.29}$$

are automatically satisfied. The divergence of $\boldsymbol{E}$ equation becomes in Coulomb gauge

$$\begin{aligned}
\boldsymbol{\nabla} \cdot \boldsymbol{E} &= 4\pi\rho \\
\boldsymbol{\nabla} \cdot \left[-\boldsymbol{\nabla}\Phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} \right] &= 4\pi\rho
\end{aligned} \tag{27.30}$$

or

$$-\nabla^2\Phi = 4\pi\rho \tag{27.31}$$

which is Poisson's equation. The solution for a set of point charges of charge q_j at positions $\boldsymbol{r}_j$ is therefore instantaneous Coulomb's law,

$$\Phi(\boldsymbol{r}) = \sum_j \frac{q_j}{|\boldsymbol{r} - \boldsymbol{r}_j|}. \tag{27.32}$$

If we have interacting charges, we need to include all of the energy in the Hamiltonian. That means that we need to include any associated field energy as well. When written in terms of the vector and scalar potentials, the Lagrangian for the fields is from Jackson's book, Classical Electrodynamics[3]

$$\mathcal{L}_{\text{em}} = \frac{1}{8\pi}\int d^3r(|\boldsymbol{E}(\boldsymbol{r})|^2 - |\boldsymbol{B}(\boldsymbol{r})|^2) \tag{27.33}$$

---

[3]Alternatively, the only Lorentz invariants that you can form from $\boldsymbol{E}$ and $\boldsymbol{B}$ are $|\boldsymbol{E}|^2 - |\boldsymbol{B}|^2$ which is a scalar, and $\boldsymbol{E} \cdot \boldsymbol{B}$ which is a pseudoscalar. Even if you include an $\boldsymbol{E} \cdot \boldsymbol{B}$ term in the free field Lagrangian, it does not contribute to the equation of motion and can be dropped. The constant in front of the $|\boldsymbol{E}|^2 - |\boldsymbol{B}|^2$ is a field normalization; changing it can always be compensated by changing the electric charge – in fact just this difference happens when we change from Gaussian to Heaviside-Lorentz or to SI units.

and plugging in the potentials, we get

$$\mathcal{L}_{\text{em}}\left(\boldsymbol{A}(\boldsymbol{r},t),\frac{\partial \boldsymbol{A}(\boldsymbol{r},t)}{\partial t}\right) = \frac{1}{8\pi}\int d^3 r \, |\boldsymbol{\nabla}\Phi(\boldsymbol{r},t)|^2 + \frac{1}{8\pi}\int d^3 r \left(\left|\frac{1}{c}\frac{\partial \boldsymbol{A}(\boldsymbol{r},t)}{\partial t}\right|^2 - |\boldsymbol{\nabla}\times\boldsymbol{A}(\boldsymbol{r},t)|^2\right), \tag{27.34}$$

where we drop the electric field cross terms that contain $\boldsymbol{\nabla}\Phi \cdot \frac{\partial \boldsymbol{A}}{\partial t}$ since integrating them by parts (i.e. using the divergence theorem) gives a divergence of the vector potential which is zero in Coulomb gauge. Integrating the $|\boldsymbol{\nabla}\Phi|^2$ term by parts (i.e. using the divergence theorem) gives

$$\frac{1}{8\pi}\int d^3 r \, |\boldsymbol{\nabla}\Phi(\boldsymbol{r},t)|^2 = -\frac{1}{8\pi}\int d^3 r \Phi(\boldsymbol{r},t)\nabla^2\Phi(\boldsymbol{r},t) = \frac{1}{2}\int d^3 r \rho(\boldsymbol{r},t)\Phi(\boldsymbol{r},t). \tag{27.35}$$

this term cancels half of the $\int d^3 r \rho\Phi$ term in Eq. 27.21. The resulting term is the usual instantaneous Coulomb energy for a set of interacting particles

$$V = \sum_{j<k}\frac{q_j q_k}{|\boldsymbol{r}_j - \boldsymbol{r}_k|} \tag{27.36}$$

where the infinite constant self energy is dropped since it has no physical significance.

When the vector potential components are added, you get retardation effects. From relativistic theory we know that these are order v^2/c^2 effects, and for nonrelativistic theories they will be small and can be ignored. The usual hydrogen atom solutions were given in Coulomb gauge with the vector potential dropped for this reason. Notice that if you calculated in Lorentz Gauge and dropped the vector potential components you would make mistakes of order v/c. In hydrogen like atoms with nuclear charge Z the electron speed is roughly $Z(e^2/\hbar c)c$ so that the corrections would be of order $Z/137$ rather than that squared. Choosing a gauge that is consistent with your other approximations is important.

Another advantage of the Coulomb gauge is that Eq. 27.36 is given completely by the charged particle coordinates. It is simply Coulomb's law. Notice that if there are no charges, it is zero. That is, for a vacuum, only $\boldsymbol{A}$ appears. Another way of seeing this is that there is no momentum variable conjugate to Φ when we construct the field Hamiltonian.

In classical physics, when we have a holonomic constraint, we can either substitute the constraint equation into our Lagrangian and then vary the action, or we can vary the action subject to the constraints applied through, for example, Lagrange multipliers. Choosing Coulomb gauge and enforcing it by writing the scalar potential part in terms of the particle coordinate is like enforcing the constraint directly on our Lagrangian. This simplifies the equations. One disadvantage is that Coulomb gauge is not manifestly Lorentz invariant. That is, if you Lorentz transform Coulomb gauge $\boldsymbol{A}$, Φ as a Lorentz 4-vector, the resulting $\boldsymbol{A}'$ and Φ' will not be in Coulomb gauge. This is not a big problem for us since we are using a nonrelativistic Hamiltonian for the charged particles anyway. Enforcing a gauge of course means that we will lose manifest gauge invariance. The underlying electrodynamics is both Lorentz and gauge invariant which can be shown after quantization.

27.9 Quantizing the Electromagnetic Field

We have solved for the electronic states of atoms ignoring the interaction with the radiation field. To include it perturbatively, we need to quantize the radiation field also ignoring this interaction. We can then use perturbation theory to calculate the effects of the interaction of the radiation field and the particles. Since we have included the Coulomb energy in the particle Hamiltonian, the remaining interaction term is $\frac{1}{c}\int d^3 r\, \boldsymbol{J}(\boldsymbol{r},t)\cdot \boldsymbol{A}(\boldsymbol{r},t)$ as in classical electromagnetism.

We therefore quantize the radiation field with no particles. The free fields are given by Eq. 27.12 with $\Phi = 0$. The Lagrangian is

$$\mathcal{L}_{\text{em}}\left(\boldsymbol{A}(\boldsymbol{r},t), \frac{\partial \boldsymbol{A}(\boldsymbol{r},t)}{\partial t}\right) = \frac{1}{8\pi}\int d^3 r \left(\left|\frac{1}{c}\frac{\partial \boldsymbol{A}(\boldsymbol{r},t)}{\partial t}\right|^2 - |\nabla \times \boldsymbol{A}(\boldsymbol{r},t)|^2\right). \qquad (27.37)$$

It is helpful to work in periodic boundary conditions in a cube of volume L^3. We can verify that $\boldsymbol{A}(\boldsymbol{r})$ are the canonical coordinates by performing the usual variation of the action integral

$$\delta \int_{t_1}^{t_2} \mathcal{L}_{\text{em}} dt = 0. \qquad (27.38)$$

To do this substitute $\boldsymbol{A}(\boldsymbol{r},t)+\delta\boldsymbol{A}(\boldsymbol{r},t)$ into the action integral and integrate by parts terms which have derivatives of the variation so that all derivatives appear on the $\boldsymbol{A}$ terms. The surface terms vanish from either the rule that the variation is zero at t_1 and t_2 or from the periodic boundary conditions. The coefficient of the variation will be the wave equation

$$\nabla^2 \boldsymbol{A}(\boldsymbol{r},t) - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}(\boldsymbol{r},t)}{\partial t^2} = 0 \qquad (27.39)$$

This is identical to the solution to Maxwell's equations for a free field in coulomb gauge. The Maxwell divergence equations are satisfied automatically for the coulomb gauge potentials, and the curl equations are

$$\begin{aligned}
\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r},t) &= -\frac{1}{c}\frac{\partial \boldsymbol{B}(\boldsymbol{r},t)}{\partial t} \\
\boldsymbol{\nabla} \times \boldsymbol{B}(\boldsymbol{r},t) &= +\frac{1}{c}\frac{\partial \boldsymbol{E}(\boldsymbol{r},t)}{\partial t}
\end{aligned} \qquad (27.40)$$

Substituting in the vector potential for the free field, the first equation is automatically satisfied for coulomb gauge potentials and the second is the wave equation above.

Classically, the wave equation can be separated in space and time and solved for the modes consistent with the boundary conditions. All solutions are then linear combinations of these separated solutions. We transform the Lagrangian by writing the solution of the wave equation in terms of its classical modes. This transforms the Lagrangian into a set of classical harmonic oscillator terms. These are easy to quantize. Since the vector potential is real, we write

$$\boldsymbol{A}(\boldsymbol{r},t) = \sum_{\boldsymbol{k},\alpha} \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha}\left[q_{\boldsymbol{k},\alpha}(t)\sin(\boldsymbol{k}\cdot\boldsymbol{r}) + Q_{\boldsymbol{k},\alpha}(t)\cos(\boldsymbol{k}\cdot\boldsymbol{r})\right] \qquad (27.41)$$

where the allowed values of $\boldsymbol{k}$ are the usual

$$\boldsymbol{k} = \frac{2\pi}{L}(m\hat{\boldsymbol{x}} + n\hat{\boldsymbol{y}} + o\hat{\boldsymbol{z}}) \tag{27.42}$$

and since changing the sign of $\boldsymbol{k}$ does not produce different sine and cosine basis functions, if we include only one of the pair m, n, o and $-m, -n, -o$. The Q and q variables are the real coefficients needed to produce the classical fields. The $m = n = o = 0$ term is a constant that has no physical significance and can be dropped. The polarization vectors $\lambda_{\boldsymbol{k}\alpha}$ must be chosen to be two orthonormal directions so that $\hat{\boldsymbol{k}}$, and the two polarizations form a three dimensional orthogonal set. The exclusion of a polarization along $\boldsymbol{k}$ enforces the coulomb gauge condition.

The curl and the time derivative of this expansion are easy to calculate. We substitute into the expression for the Lagrangian and integrate over $\boldsymbol{r}$. The orthogonality of the $\hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha}$ and the sin and cos then reduce the double sums to

$$\mathcal{L}_{\text{em}} = \sum_{\boldsymbol{k},\alpha} \frac{L^3}{16\pi} \left[q_{\boldsymbol{k},\alpha}^2(t)|\boldsymbol{k}|^2 + \frac{1}{c^2}\dot{q}_{\boldsymbol{k},\alpha}^2(t) + Q_{\boldsymbol{k},\alpha}^2(t)|\boldsymbol{k}|^2 + \frac{1}{c^2}\dot{Q}_{\boldsymbol{k},\alpha}^2(t) \right] \tag{27.43}$$

The Momenta $p_{\boldsymbol{k},\alpha}$, $P_{\boldsymbol{k},\alpha}$ conjugate to the $q_{\boldsymbol{k},\alpha}$ and $Q_{\boldsymbol{k},\alpha}$ variables are given by the derivatives with respect to $\dot{q}_{\boldsymbol{k},\alpha}$ and $\dot{Q}_{\boldsymbol{k},\alpha}$. The standard rule for obtaining the Hamiltonian is

$$H = \sum_{\boldsymbol{k},\alpha} \left[P_{\boldsymbol{k},\alpha}\dot{Q}_{\boldsymbol{k},\alpha} + p_{\boldsymbol{k},\alpha}\dot{q}_{\boldsymbol{k},\alpha} \right] - \mathcal{L}_{\text{em}} \tag{27.44}$$

where we need to substitute so that H is written in terms of q, Q and p, P,

$$H = \sum_{\boldsymbol{k},\alpha} \frac{1}{2} \left[\frac{8\pi c^2}{L^3} p_{\boldsymbol{k},\alpha}^2 + \frac{L^3}{8\pi c^2}\omega_k^2 q_{\boldsymbol{k},\alpha}^2 + \frac{8\pi c^2}{L^3} P_{\boldsymbol{k},\alpha}^2 + \frac{L^3}{8\pi c^2}\omega_k^2 Q_{\boldsymbol{k},\alpha}^2 \right] \tag{27.45}$$

where I have defined

$$\omega_k \equiv |\boldsymbol{k}|c. \tag{27.46}$$

We can now quantize in the usual way. One way would be to work in the q and Q representation. In this case we would write

$$\begin{aligned} p_{\boldsymbol{k},\alpha} &= -i\hbar\frac{\partial}{\partial q_{\boldsymbol{k},\alpha}} \\ P_{\boldsymbol{k},\alpha} &= -i\hbar\frac{\partial}{\partial Q_{\boldsymbol{k},\alpha}} \end{aligned} \tag{27.47}$$

and we can then write the wave function for the free radiation field in terms of a product of harmonic oscillator functions: an oscillator function for each wave vector and polarization. This would be the first quantized form of the theory. However, this is generally a notational nightmare. An exactly equivalent formulation is to use the second quantized formulation for the harmonic oscillator wave functions. The advantage is notational convenience. Using

second quantized notation does not change any of the physics. Note that we can do quantum field theory in first quantized notation using explicit wave functions if we want.

Using second quantized notation, we define a set of creation and destruction operators exactly as for the harmonic oscillator. To bring our terms into the standard form for the Hamiltonian for the harmonic oscillator, we can use the identification

$$M = \frac{L^3}{8\pi c^2}$$
$$\Omega = \omega_k \tag{27.48}$$

so that all our terms are in the form,

$$\frac{p^2}{2M} + \frac{1}{2}M\Omega^2 q^2. \tag{27.49}$$

The usual definition of the creation operator a^+ is

$$a^+ = \sqrt{\frac{M\Omega}{2\hbar}}\left(q - i\frac{p}{M\Omega}\right). \tag{27.50}$$

Using this and defining $b_{\boldsymbol{k},\alpha}^+$ and $B_{\boldsymbol{k},\alpha}^+$ as the creation operators for the small and capital p,q oscillators, we get

$$b_{\boldsymbol{k},\alpha}^+ = \sqrt{\frac{\omega_k L^3}{16\pi c^2 \hbar}}\left(q_{\boldsymbol{k},\alpha} - ip_{\boldsymbol{k},\alpha}\frac{8\pi c^2}{L^3 \omega_k}\right)$$
$$B_{\boldsymbol{k},\alpha}^+ = \sqrt{\frac{\omega_k L^3}{16\pi c^2 \hbar}}\left(Q_{\boldsymbol{k},\alpha} - iP_{\boldsymbol{k},\alpha}\frac{8\pi c^2}{L^3 \omega_k}\right) \tag{27.51}$$

and the Hamiltonian is

$$H = \sum_{\boldsymbol{k},\alpha}(b_{\boldsymbol{k},\alpha}^+ b_{\boldsymbol{k},\alpha} + B_{\boldsymbol{k},\alpha}^+ B_{\boldsymbol{k},\alpha} + 1)\hbar\omega_k \tag{27.52}$$

We can write the vector potential in terms of these operators too as,

$$\boldsymbol{A}_{op}(\boldsymbol{r}) = \sum_{\boldsymbol{k},\alpha}\sqrt{\frac{4\pi c^2 \hbar}{L^3 \omega_k}}\hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha}\left[(b_{\boldsymbol{k},\alpha}^+ + b_{\boldsymbol{k},\alpha})\sin(\boldsymbol{k}\cdot\boldsymbol{r}) + (B_{\boldsymbol{k},\alpha}^+ + B_{\boldsymbol{k},\alpha})\cos(\boldsymbol{k}\cdot\boldsymbol{r})\right] \tag{27.53}$$

and then use that to do perturbation theory. However, $\boldsymbol{A}$ will then be in terms of sines and cosines. Often calculations are easier using a plane wave basis. Now that we have quantized the hamiltonian, we can easily transform to these coordinates. Writing a set of new creation operators as

$$a_{\boldsymbol{k},\alpha}^+ = \frac{1}{\sqrt{2}}(B_{\boldsymbol{k},\alpha}^+ + ib_{\boldsymbol{k},\alpha}^+)$$
$$a_{-\boldsymbol{k},\alpha}^+ = \frac{1}{\sqrt{2}}(B_{\boldsymbol{k},\alpha}^+ - ib_{\boldsymbol{k},\alpha}^+) \tag{27.54}$$

recalling that our original set was restricted such that we included only one of $\boldsymbol{k}$ and $-\boldsymbol{k}$. It is straightforward to show that these new operators have the same commutation relations as the old set, and therefore the transformation is canonical.

Substituting back we get the hamiltonian

$$H = \sum_{\boldsymbol{k},\alpha}(a^+_{\boldsymbol{k},\alpha}a_{\boldsymbol{k},\alpha} + \tfrac{1}{2})\hbar\omega_k \tag{27.55}$$

where now the $\boldsymbol{k}$ sum is over all values of m,n,o in Eq. 27.42 The vector potential is

$$\boldsymbol{A}_{op}(\boldsymbol{r}) = \sqrt{\frac{2\pi c^2\hbar}{L^3}}\sum_{\boldsymbol{k},\alpha}\frac{1}{\sqrt{\omega_k}}\hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha}\left[a_{\boldsymbol{k},\alpha}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + a^+_{\boldsymbol{k},\alpha}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{27.56}$$

Throughout these calculations, I have assumed that the two independent polarizations for each $\boldsymbol{k}$ value are linear, so the polarization unit vectors have real coefficients in cartesian coordinates. In the general case the polarization vectors can be complex and therefore can give elliptic and circular polarizations. These complex unit vectors can be produced by a canonical transformation. For example

$$\begin{aligned}
c_{\boldsymbol{k},1} &= \cos(\theta)a_{\boldsymbol{k},1} + \sin(\theta)e^{i\phi}a_{\boldsymbol{k},2}\\
c_{\boldsymbol{k},2} &= -\sin(\theta)e^{-i\phi}a_{\boldsymbol{k},1} + \cos(\theta)a_{\boldsymbol{k},2}
\end{aligned} \tag{27.57}$$

and defining the new polarization directions to be

$$\begin{aligned}
\hat{\boldsymbol{\lambda}}_{new\ 1} &= \cos(\theta)\hat{\boldsymbol{\lambda}}_1 + \sin(\theta)e^{-i\phi}\hat{\boldsymbol{\lambda}}_2\\
\hat{\boldsymbol{\lambda}}_{new\ 2} &= \cos(\theta)\hat{\boldsymbol{\lambda}}_2 - \sin(\theta)e^{i\phi}\hat{\boldsymbol{\lambda}}_2
\end{aligned} \tag{27.58}$$

will produce the vector potential,

$$\boldsymbol{A}_{op}(\boldsymbol{r}) = \sqrt{\frac{2\pi c^2\hbar}{L^3}}\sum_{\boldsymbol{k},\alpha}\frac{1}{\sqrt{\omega_k}}\left[c_{\boldsymbol{k},\alpha}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} + c^+_{\boldsymbol{k},\alpha}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\hat{\boldsymbol{\lambda}}^*_{\boldsymbol{k}\alpha}\right] \tag{27.59}$$

In Eq. 27.57, the angles θ and ϕ can be different for different $\boldsymbol{k}$ values.

I have tried to be very plodding in deriving the quantized vector potential operator. I have kept all the classical coordinates real as they should be, and only changed to plane waves and complex polarizations after quantization. A faster way is to use the complex notation from the beginning, but make sure that your Lagrangian and the coordinates you use are real (that is you can use the real and imaginary parts of a complex variable as two real variables).

Note that while the vector potential is written here in the Schrödinger representation, it is easy to change to the interaction or Heisenberg representations to obtain the time dependent fields.

27.10 The Hamiltonian for nonrelativistic charged particles with electromagnetism

Dropping unimportant constants and combining our results we get the Hamiltonian for nonrelativistic charged particles and the electromagnetic field to be

$$
\begin{aligned}
H &= \sum_j \frac{[\boldsymbol{p}_j - \frac{q_j}{c}\boldsymbol{A}_{\mathrm{op}}(\boldsymbol{r}_j)]^2}{2m_j} + \sum_{j<n} \frac{q_j q_n}{|\boldsymbol{r}_j - \boldsymbol{r}_n|} + \sum_{\boldsymbol{k},\alpha} \hbar k c\, a^+_{\boldsymbol{k}\alpha} a_{\boldsymbol{k}\alpha} \\
\boldsymbol{A}_{\mathrm{op}}(\boldsymbol{r}) &= \sqrt{\frac{2\pi c^2 \hbar}{L^3}} \sum_{\boldsymbol{k},\alpha} \frac{1}{\sqrt{kc}} \left[a_{\boldsymbol{k},\alpha} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} + a^+_{\boldsymbol{k},\alpha} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \hat{\boldsymbol{\lambda}}^*_{\boldsymbol{k}\alpha} \right] \qquad (27.60)
\end{aligned}
$$

Most electronic structure calculations begin with this Hamiltonian, with the electrons and nuclei being the particles, and neglecting the $\boldsymbol{p}_j \cdot \boldsymbol{A}$ and A^2 terms that couple the photons to the charged particles. The particle interactions are then just the instantaneous coulomb interaction. After solving this problem (often approximately), perturbation theory can be used to calculate electromagnetic transitions. Spin and other neglected terms that come from relativistic corrections can also be included if needed. Often these are included perturbatively as well.

At low orders we get accurate results using this Hamiltonian along with perturbation theory. However, at higher orders, most results become infinite when the photon intermediate states are summed over – the upper limit of the sum gives an infinite result. These infinities become less strong (e.g. instead of a result proportional to the photon momentum, they become proportional to its logarithm) when relativistic Dirac electrons (and positrons) are used. These results become finite when the theory is renormalized. This means that parameters, like m and the coupling between the electrons and photons q_j that appear in the Hamiltonian are adjusted along with the high energy cutoffs to give the experimental mass and charge, etc. We will see the beginnings of the renormalization methods when we discuss the Lamb shift.

27.11 An electromagnetic transition rate

Let's write the Hamiltonian for a hydrogen atom and the electromagnetic field, taking the proton to be fixed in space,

$$
\begin{aligned}
H &= H_0 + V \\
H_0 &= H_{\text{hyd}} + \sum_{\boldsymbol{k}\alpha} \hbar\omega_k a^+_{\boldsymbol{k},\alpha} a_{\boldsymbol{k},\alpha} \\
H_{\text{hyd}} &= \frac{p^2}{2m} - \frac{e^2}{r} \\
V &= \frac{-e\boldsymbol{p}\cdot\boldsymbol{A}_{\text{op}}(\boldsymbol{r})}{mc} + \frac{e^2}{2mc^2}\left|\boldsymbol{A}_{\text{op}}(\boldsymbol{r})\right|^2 \\
\boldsymbol{A}_{op}(\boldsymbol{r}) &= \sqrt{\frac{2\pi c^2\hbar}{L^3}} \sum_{\boldsymbol{k},\alpha} \frac{1}{\sqrt{\omega_k}} \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} \left[a_{\boldsymbol{k},\alpha}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + a^+_{\boldsymbol{k},\alpha}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \right] ,
\end{aligned}
\qquad (27.61)
$$

where we drop the constant zero point energy from the electromagnetic Hamiltonian.

What is the transition rate for an electron in a $2p_z$ state and the electromagnetic field in its ground state, to make a transition to the $1s$ state emitting a photon? The initial eigenstate of H_0 is $|2p_z\ 0\rangle$ where the 0 represents the electromagnetic vacuum. The energy of this state is E_{2p_z}. The final state has one photon of wave vector $\boldsymbol{k}$ and polarization $\hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha}$ and can be written as $|1s\ \boldsymbol{k}\hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha}\rangle$. The energy of this state is $E_{1s} + \hbar kc$. Fermi's golden rule gives the transition rate, and this needs to be summed over the photon quantum states

$$
\Gamma = \sum_{\boldsymbol{k}\alpha} \frac{2\pi}{\hbar} \left| \langle 1s\ \boldsymbol{k}\hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha}|V|2p_z\ 0\rangle \right|^2 \delta(E_{1s} + \hbar kc - E_{2p_z}) . \qquad (27.62)
$$

The only term in V that couples a state with no photon to one with 1 photon is the term linear in A. The term quadratic in A couples states with the same number of photons and states with the number differing by 2. The electromagnetic part of the matrix element can be evaluated immediately and it becomes

$$
\langle 1s\ \boldsymbol{k}\hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha}|V|2p_z\ 0\rangle = -\frac{e}{mc}\sqrt{\frac{2\pi c\hbar}{L^3 k}} \langle 1s|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}|2p_z\rangle \qquad (27.63)
$$

The energy conserving delta function will set the value of k, and in this case we find that the wave length is much larger than the size of the atom. That means that $\boldsymbol{k}\cdot\boldsymbol{r}$ will be very small, and it is a good approximation to set $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ equal to unity. This is called the electric dipole approximation.[4] An analogous approximation is made in classical electromagnetism.

---

[4]The photon energy will be order the atomic unit of energy (Since the $2p$ state has energy $-\frac{1}{8}$ of an atomic unit, and the $1s$ state has energy $-\frac{1}{2}$ of an atomic unit, the photon will have energy of $\frac{3}{8}$ of an atomic unit, but here we just need a rough estimate.) so $\hbar kc \sim \frac{e^4 m}{\hbar^2}$, and the matrix elements of r will be of order the Bohr radius $\frac{\hbar^2}{me^2}$ – see Eq. 27.72 – so that kr in a hydrogenic matrix element will be of order $\frac{e^2}{\hbar c} \sim \frac{1}{137}$.

We now can calculate the matrix elements of the momentum operator in the hydrogenic states. Our matrix element is therefore

$$\langle 1s | \boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} | 2p_z \rangle = \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} \cdot \int d^3 r \psi_{1s}^*(\boldsymbol{r}) \frac{\hbar}{i} \boldsymbol{\nabla} \psi_{2p_z}(\boldsymbol{r}) \,. \tag{27.64}$$

There is no particular difficulty in this calculation, but often we want to have an expression that looks more like the electric dipole operator. The commutator

$$[p^2, x] = [p_x^2, x] = p_x[p_x, x] + [p_x, x]p_x = -2i\hbar p_x \,. \tag{27.65}$$

Since positions commute with other positions, we can write this as

$$[H_{\text{hyd}}, x] = -\frac{i\hbar}{m} p_x \,. \tag{27.66}$$

The momentum matrix element can therefore be calculated as

$$\langle 1s | \boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} | 2p_z \rangle = \frac{im}{\hbar} \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} \cdot \langle 1s | [H_{\text{hyd}}, \boldsymbol{r}] | 2p_z \rangle = \frac{im}{\hbar} (E_{1s} - E_{2p_z}) \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} \cdot \langle 1s | \boldsymbol{r} | 2p_z \rangle \,. \tag{27.67}$$

Our Fermi's golden rule expression becomes

$$
\begin{aligned}
\Gamma &= \frac{4\pi^2 e^2}{\hbar^2 c L^3} \sum_{\boldsymbol{k}\alpha} k^{-1} (E_{1s} - E_{2p_z})^2 \left| \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} \cdot \langle 1s | \boldsymbol{r} | 2p_z \rangle \right|^2 \delta(E_{1s} + \hbar k c - E_{2p_z}) \,. \\
&= \frac{e^2}{2\pi\hbar^2 c} \sum_{\alpha} \int d^3 k \, k^{-1} (E_{1s} - E_{2p_z})^2 \left| \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} \cdot \langle 1s | \boldsymbol{r} | 2p_z \rangle \right|^2 \delta(E_{1s} + \hbar k c - E_{2p_z})
\end{aligned}
\tag{27.68}
$$

where we change the k sum to an integral. Notice that the volume terms have all canceled as they must. We need to deal with the polarization sum. One way is to explicitly construct two polarization unit vectors orthogonal to $\boldsymbol{k}$ and each other. However since here we are summing over both polarizations, we can use a trick. If we imagine summing over both polarizations and the direction $\hat{\boldsymbol{k}}$, we would be summing over all three components of the matrix element of $\boldsymbol{r}$ and summing their squares. That would just be the magnitude squared of the matrix element of the vector $\boldsymbol{r}$. If we then subtract the square of the components along $\hat{\boldsymbol{k}}$ we have the polarization sum. That is

$$\sum_{\alpha} (\boldsymbol{d}_1 \cdot \hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha})(\hat{\boldsymbol{\lambda}}_{\boldsymbol{k}\alpha} \cdot \boldsymbol{d}_2) = \boldsymbol{d}_1 \cdot \boldsymbol{d}_2 - (\boldsymbol{d}_1 \cdot \hat{\boldsymbol{k}})(\hat{\boldsymbol{k}} \cdot \boldsymbol{d}_2) \,. \tag{27.69}$$

Performing the polarization sum and integrating over the magnitude of k, gives

$$\Gamma = \frac{e^2}{2\pi c^3 \hbar^4} (E_{2p_z} - E_{1s})^3 \int d\cos\theta_k d\phi_k \left[|\boldsymbol{d}|^2 - (\hat{\boldsymbol{k}} \cdot \boldsymbol{d})^2 \right] \tag{27.70}$$

328

where $\boldsymbol{d} = \langle 1s|\boldsymbol{r}|2p_z\rangle$, and $\hat{\boldsymbol{k}} = \hat{\boldsymbol{x}}\sin\theta_k\cos\phi_k + \hat{\boldsymbol{y}}\sin\theta_k\sin\phi_k + \hat{\boldsymbol{z}}\cos\theta_k$. If we define $\hbar\omega \equiv E_{2p_z} - E_{1s}$ this can be written as

$$\Gamma = \frac{e^2\omega^3}{2\pi c^3\hbar}\int d\cos\theta_k d\phi_k\left[|\boldsymbol{d}|^2 - (\hat{\boldsymbol{k}}\cdot\boldsymbol{d})^2\right] \tag{27.71}$$

We have performed the matrix element integration before. The only component of r that survives is the z component. The result after integation is

$$\boldsymbol{d} = \hat{\boldsymbol{z}}\frac{8}{\sqrt{2}}\left(\frac{2}{3}\right)^5 a_0 , \tag{27.72}$$

and the transition rate becomes

$$\Gamma = \frac{e^2\omega^3}{2\pi c^3\hbar}|d|^2\int d\cos\theta_k d\phi_k(1 - \cos^2\theta_k) = \frac{4e^2\omega^3}{3c^3\hbar}|d|^2$$
$$= \frac{2^{17}e^2\omega^3 a_0^2}{3^{11}\hbar c^3} \tag{27.73}$$

The $1/e$ lifetime is given by

$$\tau = \frac{1}{\Gamma} = 1.59\times 10^{-9}\text{seconds} \tag{27.74}$$

in excellent agreement with experiment.

27.12   Transitions without fixing the proton

The approximation in the previous section that the proton was fixed in space simplifies the calculation, but fixing the proton destroys the translational invariance of the Hamiltonian so that momentum is not conserved. After the transition, the photon has momentum, but the atom has not recoiled. Let's now go back and include the proton degree of freedom. The Hamiltonian is

$$H = H_0 + V$$
$$H_0 = \frac{p_e^2}{2m_e} + \frac{p_p^2}{2m_p} - \frac{e^2}{|\boldsymbol{r}_e - \boldsymbol{r}_p|} + \sum_{\boldsymbol{k}\alpha}\hbar\omega_k a_{\boldsymbol{k}\alpha}^+ a_{\boldsymbol{k}\alpha}$$
$$V = \frac{e}{m_e c}\boldsymbol{p}_e\cdot\boldsymbol{A}_{\text{op}}(\boldsymbol{r}_e) - \frac{e}{m_p c}\boldsymbol{p}_p\cdot\boldsymbol{A}_{\text{op}}(\boldsymbol{r}_p) + \frac{e^2}{m_e c^2}A_{\text{op}}^2(\boldsymbol{r}_e) + \frac{e^2}{m_p c^2}A_{\text{op}}^2(\boldsymbol{r}_p)$$
$$\boldsymbol{A}_{\text{op}}(\boldsymbol{r}) = \sum_{\boldsymbol{k}\alpha}\sqrt{\frac{2\pi\hbar c^2}{L^3\omega_k}}\left[a_{\boldsymbol{k},\alpha}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\hat{\boldsymbol{\lambda}}_\alpha + a_{\boldsymbol{k},\alpha}^+ e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\hat{\boldsymbol{\lambda}}_\alpha^*\right] . \tag{27.75}$$

As before, the single photon transition dominates, and only the linear terms in the vector potential will contribute substantially. Since the electromagnetic field couples to the charged

particles, writing down the Hamiltonian is simplest in electron and proton coordinates. The calculation of the matrix elements is easiest in relative and center of mass coordinates, (See chapter 18 for a review)

$$
\begin{aligned}
\boldsymbol{r} &= \boldsymbol{r}_e - \boldsymbol{r}_p \\
\boldsymbol{R} &= \frac{m_e \boldsymbol{r}_e - m_p \boldsymbol{r}_p}{m_e + m_p}
\end{aligned}
\tag{27.76}
$$

or

$$
\begin{aligned}
\boldsymbol{r}_e &= \boldsymbol{R} + \frac{m_p}{m_p + m_e}\boldsymbol{r} \\
\boldsymbol{p}_e &= \frac{m_e}{m_e + m_p}\boldsymbol{P} + \boldsymbol{p} \\
\boldsymbol{r}_p &= \boldsymbol{R} - \frac{m_e}{m_p + m_e}\boldsymbol{r} \\
\boldsymbol{p}_p &= \frac{m_p}{m_e + m_p}\boldsymbol{P} - \boldsymbol{p} \,.
\end{aligned}
\tag{27.77}
$$

Writing the total and reduced masses as

$$
\begin{aligned}
M &= m_e + m_p \\
\frac{1}{m} &= \frac{1}{m_e} + \frac{1}{m_p}
\end{aligned}
\tag{27.78}
$$

the unperturbed Hamiltonian becomes

$$
H_0 = \frac{p^2}{2m} + \frac{P^2}{2M} - \frac{e^2}{r} + \sum_{\boldsymbol{k}\alpha} \hbar\omega_k a_{\boldsymbol{k}\alpha}^+ a_{\boldsymbol{k}\alpha} \,.
\tag{27.79}
$$

As we showed in the previous section, the electric dipole approximation makes errors in the matrix elements for this transition proportional to the fine structure constant, so making the electric dipole approximation has the same order error as ignoring the A^2 terms.

In the electric dipole approximation here, it is the relative coordinate times k that will be small. The center of mass coordinate can be anything so that $\boldsymbol{k} \cdot \boldsymbol{R}$ cannot be dropped. The plane waves in the vector potential will become

$$
\begin{aligned}
e^{i\boldsymbol{k}\cdot\boldsymbol{r}_e} &= e^{i\boldsymbol{k}\cdot\boldsymbol{R}} e^{i\frac{m_p}{M}\boldsymbol{k}\cdot\boldsymbol{r}} \\
e^{i\boldsymbol{k}\cdot\boldsymbol{r}_p} &= e^{i\boldsymbol{k}\cdot\boldsymbol{R}} e^{-i\frac{m_e}{M}\boldsymbol{k}\cdot\boldsymbol{r}}
\end{aligned}
\tag{27.80}
$$

The perturbation that survives in the electric dipole approximation is therefore

$$
\begin{aligned}
V &= \left[\frac{e}{m_e c}\boldsymbol{p}_e - \frac{e}{m_p c}\boldsymbol{p}_p\right] \cdot \boldsymbol{A}_{\mathrm{op}}(\boldsymbol{R}) \\
&= \frac{e}{mc}\boldsymbol{p} \cdot \boldsymbol{A}_{\mathrm{op}}(\boldsymbol{R}) \,.
\end{aligned}
\tag{27.81}
$$

330

The initial state can be written as

$$|2p_z \ \boldsymbol{K} \ 0\rangle \tag{27.82}$$

where $\boldsymbol{K}$ stands for the center of mass state. If the atom is initially at rest, we will take $\boldsymbol{K} = 0$, and 0 is the electromagnetic vacuum. The final state is $|1s \ \boldsymbol{K}' \ \boldsymbol{k}\alpha\rangle$, where $\boldsymbol{k}$ and α stand for the final state photon's wave vector and polarization. Fermi's golden rule summed over the photon energy, its polarization, and the center of mass final momentum gives

$$\begin{aligned}
d\Gamma &= \sum_{\boldsymbol{K}'}\sum_{\alpha} \frac{L^3}{(2\pi)^3} \int \frac{k^2 dk d\Omega_k}{(2\pi)^3} \left| \langle 1s \ \boldsymbol{K}'\boldsymbol{k}\alpha| \frac{e}{mc}\boldsymbol{p}\cdot\boldsymbol{A}_{\text{op}}(\boldsymbol{R})|2p_z \ \boldsymbol{K} \ 0\rangle \right|^2 \\
&\quad \delta\left(\frac{\hbar^2 K^2}{2M} + E_{2p} - \frac{\hbar^2 K'^2}{2M} - E_{1s} - \hbar k c \right).
\end{aligned} \tag{27.83}$$

Evaluating the center of mass matrix element, it will be zero unless $\boldsymbol{K}' + \boldsymbol{k} = \boldsymbol{K}$ giving total momentum conservation. Similarly, the delta function enforces total energy conservation. If we now take $K = 0$, we can evaluate the electromagnetic matrix element and integrate over the delta functions to give

$$d\Gamma = d\Omega_k \frac{\omega_k e^3}{2\pi\hbar m^2 c^3 (1 + \frac{\hbar\omega_k}{Mc^2})} \left[|\langle 1s|\boldsymbol{p}|2p_z\rangle|^2 - |\langle 1s|\boldsymbol{p}\cdot\hat{\boldsymbol{k}}|2p_z\rangle|^2 \right] \tag{27.84}$$

where ω_k must satisfy energy and momentum conservation. We can now use the commutation relation for the relative coordinates, exactly as we did with for the electron coordinates with the proton fixed, to write

$$\boldsymbol{p} = \frac{im}{\hbar} \left[\frac{p^2}{2m} - \frac{e^2}{r}, \boldsymbol{r} \right] \tag{27.85}$$

so that

$$d\Gamma = d\Omega_k \frac{\omega_k (E_{2p} - E_{1s})^2 e^3}{2\pi\hbar^3 c^3 (1 + \frac{\hbar\omega_k}{Mc^2})} \left[|\langle 1s|\boldsymbol{r}|2p_z\rangle|^2 - |\langle 1s|\boldsymbol{r}\cdot\hat{\boldsymbol{k}}|2p_z\rangle|^2 \right]. \tag{27.86}$$

Substituting the matrix element Eq. 27.72 with the reduced mass a_0 into the transition rate

$$d\Gamma = d\Omega_k \frac{k(E_{2p} - E_{1s})^2 e^3}{2\pi\hbar^3 c^2 (1 + \frac{\hbar\omega_k}{Mc^2})} \frac{2^{14}}{3^{10}} \left[1 + \cos^2\theta_k \right] \tag{27.87}$$

we see the angular distribution of photons.

For the case $K = 0$, $\boldsymbol{k} = -\boldsymbol{K}'$, and the photon energy is independent of direction. We can therefore perform the angular integration easily to get

$$\Gamma = \frac{4k(E_{2p} - E_{1s})^2 e^3}{3\hbar^3 c^2 (1 + \frac{\hbar\omega_k}{Mc^2})} \frac{2^{15}}{3^{10}} \tag{27.88}$$

The energy difference is

$$E_{2p} - E_{1s} = \frac{3e^4 m}{8\hbar^2} \tag{27.89}$$

and the energy and momentum conservation equations require that

$$\hbar k c = E_{2p} - E_{1s} - \frac{\hbar^2 k^2}{2M} \tag{27.90}$$

or

$$\hbar k c = \sqrt{M^2 c^4 + 2Mc^2(E_{2p} - E_{1s})} - Mc^2 \simeq (E_{2p} - E_{1s})\left[1 + O((E_{2p} - E_{1s})/Mc^2\right] \tag{27.91}$$

so in the nonrelativistic limit, the photon carries away nearly all the energy. $E_{2p} - E_{1s}$ is about 10.2 eV, and Mc^2 is 9.388×10^8 eV, so the correction to the energy is negligible. The only remaining difference is the difference between the electron and reduced masses. The result is proportional to m^3, so the main difference between this result and the result where the proton is fixed in space is that the transition rate here is slightly reduced by a factor of $(m/m_e)^3$, or about 0.1 percent, and the hydrogen atom will recoil properly to conserve momentum.

The calculation goes through exactly as above if the hydrogen atom has an initial velocity V in the laboratory frame, so that $\hbar K = MV$. The momentum conservation comes from the center of mass matrix element and gives

$$\hbar K' = MV - \hbar k \tag{27.92}$$

Fermi's golden rule gives energy conservation and requires

$$E_{2p} + \frac{1}{2}MV^2 = E_{1s} + \frac{(MV - \hbar k)^2}{2M} + \hbar k c \tag{27.93}$$

or

$$V \cdot k\hbar + \frac{\hbar^2 k^2}{2M} + \hbar k c = E_{2p} - E_{1s} \,. \tag{27.94}$$

The k value therefore depends on the direction of emission of the photon. For nonrelativistic velocities we can expand the result to obtain

$$\hbar k c \simeq \frac{E_{2p} - E_{1s}}{1 - \frac{V}{c}\cos\theta} \simeq (E_{2p} - E_{1s})\left[1 + \frac{V}{c}\cos\theta\right] \tag{27.95}$$

where $\cos\theta$ is the cosine of the angle between V and k. The factor in brackets is the lowest order Doppler shift.

The angular integration above to get the total transition rate is somewhat more complicated. However, keeping just the linear term in V/c, the additional term integrates to zero and we have the same result as before; there are corrections at next order in V/c which would give the correct time dilation effect if the rest of our theory were relativistic. Of course the simplest way to do the calculation would be to work in the rest frame and then Lorentz transform the results to the moving frame. In the rest frame the nonrelativistic theory is entirely adequate. To see all of the relativistic effects drop out of the calculation done in any frame, we need to make the whole theory relativistic – not just the electromagnetic part as was done here.

27.13 Problems

Problem 27.1. A classical string with tension T and mass per unit length ρ stretched along x vibrates transversely along the z direction only. The shape of the string at time t is given by $z(x,t)$. Assume that the vibrations are small.

a. Analyze the force on a small element of the string when it is displaced, and use this and Newton's law to derive the equation of motion for the string. That is find the differential equation satisfied by $z(x,t)$.

b. Again analyze the displacement and velocity of a small element of the string to calculate its kinetic and potential energies. Write the Lagrangian in terms of the variables $z(x,t)$, and $\dot{z}(x,t)$. Show that setting the variation of the action with respect to $z(x,t)$ to zero gives the differential equation of part a. Assume that either the ends of the string are fixed so that $z(-L/2,t) = z(L/2,t) = \dot{z}(-L/2,t) = \dot{z}(L/2,t) = 0$ or that the string is periodic $z(-L/2,t) = z(L/2,t)$, $\dot{z}(-L/2,t) = \dot{z}(L/2,t)$.

c. Assume periodic boundary conditions so that $z(x+L,t) = z(x,t)$ and L goes to infinity at the end of the calculations. Expand $z(x,t)$ in terms of the normal modes (eigenfunction) of the differential equation. As in electromagnetism, using sines and cosines with real coefficients automatically enforces the constraint that the classical $z(x,t)$ must be real, but you can use exponentials if you are careful. Calculate the Hamiltonian and identify the canonical coordinates. Show that the Hamiltonian separates into a set of harmonic oscillators, one for each mode.

   Note, since we will be looking at localized systems compared to $L \to \infty$, the analysis and the boundary conditions at infinity cannot matter, the calculation is somewhat simplified if the mode, which is not harmonic, corresponding to a translation of the whole string is dropped.

d. Quantize the field by writing $z(x,t)$ in terms of the appropriate phonon creation and destruction operators. Write the Hamiltonian of the string H_s and $z_{\rm op}(x)$ in second quantized notation in the Schrödinger representation. Transform your results into

$$H_s = \sum_k \left[a_k^+ a_k + \frac{1}{2} \right] \hbar \omega_k$$

$$z_{\rm op}(x) = \sum_k \sqrt{\frac{\hbar}{2\rho L \omega_k}} \left[a_k e^{ikx} + a_k^+ e^{-ikx} \right] \tag{27.96}$$

   if they are not already in that form.

e. A particle (identified as particle 1) in a harmonic oscillator well has the Hamiltonian, in the Schrödinger representation,

$$H_p = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x_1^2} + \frac{1}{2} m \omega_0^2 x_1^2. \tag{27.97}$$

The particle also interacts weakly with the string through the interaction

$$V = \alpha z(x_1) \tag{27.98}$$

where α is a small coupling constant. The full Hamiltonian of the system is then

$$H = H_p + H_s + V. \tag{27.99}$$

Calculate the transition rate for the particle to decay from its first excited state to its ground state while emitting 1 phonon. Assume that the string is initially in its ground state.

Problem 27.2. The string in problem 27.1 is modified so that along with the restoring force from the tension, each element of the string also has a linear restoring force that pulls it back to z = 0.

a. Show that the classical differential equation that the string now solves can be written in the form

$$\left(\frac{\partial^2}{\partial x^2} - \mu^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) z(x, t) = 0. \tag{27.100}$$

b. Show that the quantization goes through as before except that the frequency, ω_k, corresponding to a wave vector k, is modified to

$$\omega_k = c\sqrt{k^2 + \mu^2}. \tag{27.101}$$

Notice that this is the energy momentum relationship for a relativistic massive particle. The phonons have acquired a "mass"; this form is also similar to the dispersion of optical phonons in a diatomic chain.

c. The Born-Oppenheimer approximation is often used when there are heavy particles (like the nuclei of atoms) and light particles (like their electrons). We used the Born-Oppenheimer approximation when we calculated the van der Waals interaction. The wave function of the total system is approximated by first calculating the ground state energy of the light particles when the heavy particles are fixed. This ground state energy as a function of the positions of the heavy particles is then used as the potential between the heavy particles, and their Schrödinger equation is then solved. In exactly this way, calculate, using second order second order perturbation theory, the potential *between* two heavy particles located at x_1 and x_2 if they couple to the string via the interaction,

$$V = \alpha z(x_1) + \alpha z(x_2). \tag{27.102}$$

The matrix elements you use can be interpreted as corresponding to the exchange of one virtual phonon. You should get a 1-dimensional version of a Yukawa potential. You can think of the excitations of the string as being mesons or optical phonons. The differential equation of the string is a scalar wave equation, so these would be called scalar particles. The electromagnetic case is an example of a massless vector wave equation and photons are vector particles.

Problem 27.3. An electron when it passes through matter or passes by another particle can scatter and emit photons called bremsstrahlung. We can model this using the unperturbed Hamiltonian for a nonrelativistic free electron and the electromagnetic field,

$$H_0 = \frac{p^2}{2m} + \sum_{\mathbf{k}\alpha} \hbar\omega_k a^+_{\mathbf{k}\alpha} a_{\mathbf{k}\alpha} \,. \tag{27.103}$$

The interaction models the matter or the scattering center as a potential, and includes the usual electromagnetic coupling. In the Schrödinger representation it is

$$V = v(\mathbf{r}) + \frac{e}{mc}\mathbf{p} \cdot \mathbf{A}(\mathbf{r}) + \frac{e^2}{2mc^2}|\mathbf{A}(\mathbf{r})|^2 \tag{27.104}$$

where the vector potential operator was derived in class

$$\mathbf{A}(\mathbf{r}) = \sqrt{\frac{2\pi c^2 \hbar}{L^3}} \sum_{\mathbf{k},\alpha} \frac{1}{\sqrt{\omega_k}} \left[a_{\mathbf{k},\alpha} e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\boldsymbol{\lambda}}_\alpha + a^+_{\mathbf{k},\alpha} e^{-i\mathbf{k}\cdot\mathbf{r}} \hat{\boldsymbol{\lambda}}^*_\alpha \right] \,. \tag{27.105}$$

To simplify the calculations, assume that $v(\mathbf{r})$ is small so that only one power of $v(\mathbf{r})$ needs to be kept in the perturbed wave function. Also assume that since $e^2/\hbar c \approx 1/137$ is small, that we only need to keep terms of order e in the perturbed wave function. You can therefore drop the A^2 term in V. Notice that Fermi's golden rule using the term $\mathbf{p} \cdot \mathbf{A}$ leads to energy and momentum conserving delta functions and will give zero since we cannot conserve energy and momentum in this case. Therefore to emit a photon, the second order terms containing both time orderings of the $\mathbf{p} \cdot \mathbf{A}$ and $v(\mathbf{r})$ terms must be included.

The system begins with the electron in a momentum eigenstate of momentum $\mathbf{p}_i$ and no photons. Calculate in terms of the Fourier transform

$$\tilde{v}(\mathbf{k}) = \int d^3r \, e^{-i\mathbf{k}\cdot\mathbf{r}} v(\mathbf{r}) \tag{27.106}$$

the lowest order differential scattering cross section

$$\frac{d\sigma_0}{d\Omega_e} \tag{27.107}$$

for an electron of momentum $\mathbf{p}_i$ to scatter into solid angle $d\Omega_e$ without emitting a photon. This will be the usual Born approximation result. Next calculate the lowest order differential scattering cross section for the electron to scatter into a solid angle $d\Omega_e$, along with the emission of a photon with angular frequency ω within $d\omega$ in the direction defined by the solid angle element $d\Omega_\gamma$ with polarization λ_γ,

$$\frac{d^3\sigma_1}{d\omega d\Omega_\gamma d\Omega_e} \tag{27.108}$$

In the limit $\omega \to 0$ show that

$$\frac{d^3\sigma_1}{d\Omega_\gamma d\omega d\Omega_e} = \frac{d\sigma_0}{d\Omega_e}\frac{dN}{d\omega d\Omega_\gamma} \tag{27.109}$$

where $dN/d\omega d\Omega_\gamma$ agrees with the classical result given in Jackson Eq. 15.3.

The classical result is only correct in this limit, since it does not conserve energy and momentum. The quantum result gets these kinematic factors correctly.

Problem 27.4. The $2S$ state of hydrogen if it is not perturbed (which can mix it with the $2P$ state which then decays rapidly as we calculated in class) can only decay to the $1S$ ground state and, when spin orbit and Lamb shift corrections are included, to the $2P_{\frac{1}{2}}$ state. The tiny energy difference between these latter states makes the lifetime for transition to the $2P_{\frac{1}{2}}$ state about 20 years so it can be ignored. Electric dipole transitions are forbidden by parity. Magnetic dipole and electric quadrupole transitions are forbidden for our nonrelativistic Hamiltonian, but the magnetic dipole transition can occur if the electron is treated relativistically with the Dirac equation but the lifetime is about two days and again can be ignored. The dominant process is the emission of two photons, both of which can be treated in the electric dipole approximation.

Calculate an expression for the transition rate. You can assume that the proton is infinitely massive. Manipulate your result until it is in the form:

$$
\begin{aligned}
\Gamma \;=\; & \frac{4e^4}{3\pi\hbar^7 c^6} \int_0^{E_{2S}-E_{1S}} d\epsilon\,\epsilon^3 (E_{2s} - E_{1s} - \epsilon)^3 \left| \sum_{n\geq 2} \langle 1S|z|nP_z\rangle\langle nP_z|z|2S\rangle \right. \\
& \left. \cdot \left[\frac{1}{E_{1S} - E_n + \epsilon} + \frac{1}{E_{2S} - E_n - \epsilon} \right] \right|^2
\end{aligned}
\tag{27.110}
$$

where the sum includes an integration over the continuum states.[5]

Problem 27.5. As we discussed when looking at the perturbation calculation of the polarizability of hydrogen, the reason for the necessity to include the continuum states is because the eigenstates of H_0 do not describe the perturbed state very well. To obtain faster convergence, I change the hydrogen atom Hamiltonian to

$$
H_0 = \frac{p^2}{2m} - \frac{e^2}{r} + v(r)
\tag{27.111}
$$

where

$$
v(r) = \begin{cases} 0 & r < R \\ \infty & r \geq R \end{cases} .
\tag{27.112}
$$

For R much larger than the size of the $1S$ and $2S$ orbitals, the results will be independent of R, but the spectrum will be discrete and rapidly convergent. Using the applet on the class web page, I obtained radial wave functions and energies. From the radial wave functions I numerically integrated to get the needed matrix elements as shown in table 27.5 for $R = 30a_0$ (I checked that the calculation had converged by recalculating with $R = 50a_0$). A file containing these numbers is also on the class web site.

---

[5]This is identical to the expressions Eqs. 6.1 and 6.2 in G. Breit and E. Teller, *Metastability of hydrogen and helium levels*, Astrophysical Journal, **91**, 215 (1940).

Table 27.1: The energies of the nP states and their matrix elements with $1S$ and $2S$ states for a Hydrogen atom in a spherical cavity of radius $R = 30a_0$.

n	$a_0^{-1}\langle 1S|z|nP_z\rangle$	$a_0^{-1}\langle 2S|z|nP_z\rangle$	$E_{nP}\frac{a_0}{e^2}$	n	$a_0^{-1}\langle 1S|z|nP_z\rangle$	$a_0^{-1}\langle 2S|z|nP_z\rangle$	$E_{nP}\frac{a_0}{e^2}$
2	0.7449356	2.9999988	-0.1250000	17	0.0431602	-0.0224093	1.3497274
3	0.3004253	-1.7796378	-0.0554713	18	0.0368013	-0.0179944	1.5342806
4	0.2342117	-0.9261387	-0.0259140	19	0.0315067	-0.0153115	1.7298999
5	0.2471151	-0.7015368	0.0092110	20	0.0270852	-0.0124264	1.9365750
6	0.2387975	-0.4921913	0.0573932	21	0.0233801	-0.0108259	2.1542973
7	0.2179167	-0.3418603	0.1178545	22	0.0202639	-0.0088435	2.3830592
8	0.1921311	-0.2387594	0.1901200	23	0.0176328	-0.0078790	2.6228542
9	0.1657865	-0.1714346	0.2739397	24	0.0154025	-0.0064550	2.8736767
10	0.1411739	-0.1244378	0.3691657	25	0.0135045	-0.0058781	3.1355218
11	0.1193074	-0.0933488	0.4757023	26	0.0118830	-0.0048142	3.4083850
12	0.1004587	-0.0702375	0.5934834	27	0.0104924	-0.0044806	3.6922625
13	0.0845068	-0.0547723	0.7224610	28	0.0092953	-0.0036574	3.9871508
14	0.0711526	-0.0423950	0.8625994	29	0.0082610	-0.0034803	4.2930468
15	0.0600391	-0.0341708	1.0138706	30	0.0073642	-0.0028232	4.6099478
16	0.0508144	-0.0270179	1.1762525				

Using these matrix elements, perform the sum and integral to find the numerical value of the two-photon transition rate. Your result should be 8.23 second^{-1} in agreement with previous numerical calculations.[6]

Problem 27.6. Show that when $\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{\lambda}} = 0$,

$$\hat{\boldsymbol{k}} \cdot \boldsymbol{r}\hat{\boldsymbol{\lambda}} \cdot \boldsymbol{p} = -\frac{1}{2}\hat{\boldsymbol{\lambda}} \cdot (\hat{\boldsymbol{k}} \times \boldsymbol{L}) + \frac{im}{6\hbar}[H, \hat{\boldsymbol{\lambda}} \cdot \overset{\leftrightarrow}{Q} \cdot \hat{\boldsymbol{k}}] \qquad (27.113)$$

where the quadrupole moment operator, hamiltonian and angular momentum are

$$\begin{aligned} Q_{ij} &= 3x_i x_j - \delta_{ij} r^2 \\ H &= \frac{p^2}{2m} + v(\boldsymbol{r}) \\ \boldsymbol{L} &= \boldsymbol{r} \times \boldsymbol{p}. \end{aligned} \qquad (27.114)$$

Use these forms to derive the ℓ and m selection rules for magnetic dipole (the angular momentum term) and the electric quadrupole transitions for a hydrogen atom. You can use the properties of the angular momentum operators, parity, the properties of spherical harmonics, or any other technique.

[6]L. Spitzer, Jr. and J. L. Greenstein, *Continuous emission from planetary nebulae*, Astrophysical Journal **114**, 407 (1951); J. Shapiro and G. Breit, *Metastability of 2s states of hydrogenic atoms*, Phys. Rev. **113**, 179 (1959).

Problem 27.7. Calculate the photoelectric cross setion of a hydrogen atom in its ground state. That is calculate the differential cross section for a photon of energy $\hbar\omega$ incoming along the z axis with polarization along x, to be absorbed by a hydrogen atom in its ground state and an electron emitted along the directions given by θ ϕ, the usual spherical coordinate angles. Do not use the dipole approximation, but do take the photon energy to be large compared to the electron binding energy, but small enough that relativistic effects are not important. Approximate the outgoing wave function of the electron by a plane wave. Show that the differntial cross section can be written as

$$\frac{d\sigma}{d\Omega} = r_0^2 \left(\frac{e^2}{\hbar c}\right)^4 \left(\frac{mc^2}{\hbar\omega}\right)^{7/2} \frac{\sqrt{32}\sin^2\theta\cos^2\phi}{(1 - \frac{v}{c}\cos\theta)^4} \tag{27.115}$$

where r_0 is the classical electron radius e^2/mc^2, and v is the speed of the outgoing electron.

Problem 27.8. Verify the Schrödinger representation continuity equation equation Eq. 27.20, for matrix elements whose states satisfy the Schrödinger equation.

Chapter 28

Classical and quantum fields

28.1 Introduction

We have, until now, used two kinds of electromagnetic couplings to our quantum mechanical particle Hamiltonian. For systems where we have a classical electromagnetic field, we have simply substituted the scalar and vector potentials into our Hamiltonian and viewed the classical position $\boldsymbol{r}$ as the quantum position operator. On the other hand, when dealing with transition rates and photon emission, we have used the fully quantized electromagnetic field Hamiltonian. For the quantum case we have used Coulomb gauge exclusively, and the scalar potential is the same as in the classical case, but the vector potential has photon oscillators that we quantized and wrote using second quantized notation.

Clearly, the fully quantum case we developed can always be used, but will be difficult to apply in many cases that are easy using the classical vector potential. An example is an atom in a constant classical magnetic field.

28.2 Coherent state representation

One way to separate the classical contribution while still using the full quantum mechanical Hamiltonian is to use coherent states. I'll first describe the properties of the coherent states and then show how to use them for our calculations. If you don't want to review the coherent state development you can skip to Eq. 28.29 and see how they will be used. A coherent state is often defined as the eigenstate of the destruction operator.

$$a|\alpha\rangle = \alpha|\alpha\rangle \tag{28.1}$$

Since a is not hermitian, α is not necessarily real, and in fact it is easy to construct coherent states with α being any complex number. Let's expand $|\alpha\rangle$ as a linear combination of our good number operator states,

$$a^+a|n\rangle = n|n\rangle \,, \tag{28.2}$$

$$|\alpha\rangle = \sum_n c_n|n\rangle \,. \tag{28.3}$$

Applying the a operator we get

$$a|\alpha\rangle = \sum_n c_n \sqrt{n}|n-1\rangle = \alpha \sum_n c_n|n\rangle \tag{28.4}$$

and applying $\langle m-1|$, gives us

$$c_m \sqrt{m} = \alpha c_{m-1} \tag{28.5}$$

or

$$c_m = \frac{\alpha}{\sqrt{m}} c_{m-1} = \frac{\alpha^m}{\sqrt{m!}} c_0 . \tag{28.6}$$

Normalizing and picking the phase so that c_0 is real gives

$$|\alpha\rangle = \sum_n e^{-\frac{|\alpha|^2}{2}} \frac{\alpha^n}{\sqrt{n!}} |n\rangle . \tag{28.7}$$

There is no restriction on α, it can be any complex number.

Recall that

$$|n\rangle = \frac{(a^+)^n}{\sqrt{n!}} |0\rangle . \tag{28.8}$$

We can use this result to write

$$
\begin{aligned}
|\alpha\rangle &= \sum_n e^{-\frac{|\alpha|^2}{2}} \frac{(\alpha \, a^+)^n}{n!} |0\rangle \\
&= e^{-\frac{|\alpha|^2}{2} + \alpha a^+} |0\rangle .
\end{aligned}
\tag{28.9}
$$

which is a convenient operator to use to construct the coherent states.

Coherent states with different eigenvalues are not orthogonal. The overlap is

$$
\begin{aligned}
\langle \beta | \alpha \rangle &= \langle 0| e^{-\frac{|\beta|^2}{2} + \beta^* a} e^{-\frac{|\alpha|^2}{2} + \alpha a^+} |0\rangle \\
&= e^{-\frac{|\beta|^2}{2} - \frac{|\alpha|^2}{2}} \langle 0| e^{\beta^* a} e^{\alpha a^+} |0\rangle .
\end{aligned}
\tag{28.10}
$$

The exponentials can be expanded to

$$
\begin{aligned}
\langle 0| e^{\beta^* a} e^{\alpha a^+} |0\rangle &= \sum_n \langle 0| \frac{(\beta^* \, a)^n}{n!} \frac{(\alpha \, a^+)^n}{n!} |0\rangle \\
&= \sum_n \frac{\beta^* \alpha}{n!} = e^{\beta^* \, \alpha}
\end{aligned}
\tag{28.11}
$$

The overlap is therefore

$$\langle \beta | \alpha \rangle = e^{-\frac{|\beta|^2}{2} - \frac{|\alpha|^2}{2} + \beta^* \, \alpha} \tag{28.12}$$

and the magnitude squared is

$$|\langle \beta | \alpha \rangle|^2 = e^{-|\beta - \alpha|^2} \tag{28.13}$$

So the "distance" between β and α in the complex plane tells how orthogonal the states are.

The expectation value of a and a^+ are

$$
\begin{aligned}
\langle\alpha|a|\alpha\rangle &= \alpha \\
\langle\alpha|a^+|\alpha\rangle &= \alpha^*
\end{aligned}
\tag{28.14}
$$

since $|\alpha\rangle$ is the right eigenvector of a and the left eigenvector of a^+ by construction.

Recall that the x and p variables of the harmonic oscillator are

$$
\begin{aligned}
x &= \sqrt{\frac{\hbar}{2m\omega}}(a + a^+), \\
p &= i\sqrt{\frac{\hbar m\omega}{2}}(a^+ - a),
\end{aligned}
\tag{28.15}
$$

so we get the expectation values of p and x are

$$
\begin{aligned}
\langle\alpha|p|\alpha\rangle &= \sqrt{2\hbar m\omega}\ Im(\alpha) \\
\langle\alpha|x|\alpha\rangle &= \sqrt{\frac{2\hbar}{m\omega}}\ Re(\alpha)
\end{aligned}
\tag{28.16}
$$

Notice that we can therefore specify the coherent state by giving the expectation values of x and p since these then fully specify the α value

You can also calculate the expectation value of x^2 and p^2. Simply write the operators in terms of a and a^+, commute the operators using

$$
[a, a^+] = 1
\tag{28.17}
$$

to get all the a operators to the right of the a^+ operators. You can then replace a with α and a^+ with α^*. The result is

$$
\begin{aligned}
\langle\alpha|p^2|\alpha\rangle &= -\frac{\hbar m\omega}{2}\left[(\alpha - \alpha^*)^2 - 1\right] \\
\langle\alpha|x^2|\alpha\rangle &= \frac{\hbar}{2m\omega}\left[(\alpha + \alpha^*)^2 + 1\right]
\end{aligned}
\tag{28.18}
$$

and the fluctuations are then

$$
\begin{aligned}
(\Delta p)^2 \equiv \langle\alpha|p^2|\alpha\rangle - (\langle\alpha|p|\alpha\rangle)^2 &= \frac{\hbar m\omega}{2} \\
(\Delta x)^2 \equiv \langle\alpha|x^2|\alpha\rangle - (\langle\alpha|x|\alpha\rangle)^2 &= \frac{\hbar}{m\omega 2}.
\end{aligned}
\tag{28.19}
$$

So that

$$
\Delta x \Delta p = \frac{\hbar}{2}.
\tag{28.20}
$$

Which shows that the coherent states are the minimim uncertainty wave packets.

341

Before returning to electrodynamics lets see how we can use these to better understand statistical mechanics. The usual completeness relation does not work for these states since they are not orthogonal, but there is a nice result because the coherent states form an overcomplete set. Let's integrate over all values of α, which is like integrating over all values of the expectation value of x and p.

$$\int d^2\alpha |\alpha\rangle\langle\alpha| = \int d^2\alpha \left[\sum_n e^{-\frac{|\alpha|^2}{2}} \frac{\alpha^n}{\sqrt{n!}}|n\rangle\right]\left[\sum_m\langle m|e^{-\frac{|\alpha|^2}{2}}\frac{(\alpha^*)^m}{\sqrt{m!}}\right] \tag{28.21}$$

Changing to cylindrical coordinates and writing $d^2\alpha = |\alpha|d|\alpha|d\phi$, the ϕ integration is immediate and gives 2π if $m = n$ and 0 otherwise. The integration over $|\alpha|$ is the Gamma function form, and just gives a factorial. The result is

$$\int \frac{d^2\alpha}{\pi}|\alpha\rangle\langle\alpha| = 1 \tag{28.22}$$

which is the over-completeness relation for the coherent states. Recall that we can write α in terms of the expectation values of p and x. Let's call those $\bar{p}$ and $\bar{x}$, and changing variables, the integration becomes

$$d^2\alpha = \frac{d\bar{x}d\bar{p}}{2\hbar} \tag{28.23}$$

and the overcompleteness relation becomes

$$\int \frac{d\bar{x}d\bar{p}}{2\pi\hbar}|\bar{x}\ \bar{p}\rangle\langle\bar{x}\ \bar{p}| \tag{28.24}$$

Notice this is the usual phase space integral that we do in classical statistical mechanics. If the fluctuations are small, the operators p and x in the Boltzmann factor can be replaced with their expectation and we recover the classical statistical mechanics result. For example to calculate the partition function for N particles, we can write

$$\begin{aligned}
Z &= \sum_n e^{-\beta E_n} = \sum_n \langle n|e^{-\beta H}|n\rangle \\
&= \int \prod_{i=1}^N \frac{d\bar{x}_i d\bar{p}_i}{2\pi\hbar} \sum_n \langle n|e^{-\beta H}|\bar{x}_1\bar{p}_1, ..., \bar{x}_N, \bar{p}_N\rangle\langle\bar{x}_1\bar{p}_1, ..., \bar{x}_N, \bar{p}_N|n\rangle \\
&= \sum_n \int \prod_{i=1}^N \frac{d\bar{x}_i d\bar{p}_i}{2\pi\hbar} \langle\bar{x}_1\bar{p}_1, ..., \bar{x}_N, \bar{p}_N|n\rangle\langle n|e^{-\beta H}|\bar{x}_1\bar{p}_1, ..., \bar{x}_N, \bar{p}_N\rangle \\
&= \int \prod_{i=1}^N \frac{d\bar{x}_i d\bar{p}_i}{2\pi\hbar} \langle\bar{x}_1\bar{p}_1, ..., \bar{x}_N, \bar{p}_N|e^{-\beta H}|\bar{x}_1\bar{p}_1, ..., \bar{x}_N, \bar{p}_N\rangle. \tag{28.25}
\end{aligned}$$

Ignoring the commutators of x_i and p_i, we can replace the operators by their expectation values and we get the standard classical partition function. By keeping and expanding in

342

powers of the commutators we get an expansion in powers of $\hbar$, and can get the quantum corrections.

The probability of finding n excitations (photons or phonons) in a coherent state is

$$P_n = |\langle n|\alpha\rangle|^2 = \frac{|\alpha|^{2n}}{n!}e^{-|\alpha|^2} \tag{28.26}$$

which is the usual Poisson distribution

$$P_n(x) = \frac{x^n}{n!}e^{-x^2} \tag{28.27}$$

For large $|\alpha|$, we get a well defined number of excitations. It is straightforward to show that the fluctuation in photon number goes like $\sqrt{N}$ where N is the average number.

The time dependence of the coherent states with the oscillator Hamiltonian is straightforward to calculate. In the number representation, we propagate forward in time by multiplying each state $|n\rangle$ by $\exp(-i\omega nt)$. Plugging this into Eq. 28.7, we see that a coherent state remains a coherent state,

$$\alpha(t) = \alpha(0)\exp(-i\omega t)\,. \tag{28.28}$$

Notice that the expectation values $\bar{x}$ and $\bar{p}$, simply satisfy the classical equations of motion (as is always true), but because the coherent state is a minimum uncertainty wave packet, the packet does not spread. This is a special property that is true only for the harmonic oscillator Hamiltonian.

28.3   Coherent states in electrodynamics

Now let's apply the coherent state formulations to electrodynamics. We defined the coherent states as eigenstates of the destruction operator, but another way to motivate a single coherent state is to view it as the new vacuum after a canonical transformation. Let's define a new destruction operator,

$$b = a - \beta \tag{28.29}$$

By direct substitution, you can verify that

$$[b, b^+] = [a, a^+] = 1 \tag{28.30}$$

so this is a canonical transformation. The new vaccuum in the b representation is clearly the coherent state $|\beta\rangle$. For our electrodynamics, we have a different oscillator for each of the modes. Let's make a possibly different canonical transformation like that above for each of the modes. First let's see what happens to the vector potential operator in the new representation,

The original operator is

$$\boldsymbol{A}_{op}(\boldsymbol{r}) = \sqrt{\frac{2\pi c^2\hbar}{L^3}} \sum_{\boldsymbol{k},\alpha} \frac{1}{\sqrt{\omega_k}} \left[a_{\boldsymbol{k},\alpha}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\hat{\boldsymbol{\lambda}}_\alpha + a_{\boldsymbol{k},\alpha}^+ e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\hat{\boldsymbol{\lambda}}_\alpha^* \right] \tag{28.31}$$

and substituting $b_{\boldsymbol{k},\alpha} = a_{\boldsymbol{k},\alpha} - \beta_{\boldsymbol{k},\alpha}$, we get

$$
\begin{aligned}
\boldsymbol{A}_{op}(\boldsymbol{r}) \;=\;\;& \sqrt{\frac{2\pi c^2 \hbar}{L^3}} \sum_{\boldsymbol{k},\alpha} \frac{1}{\sqrt{\omega_k}} \left[b_{\boldsymbol{k},\alpha} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \hat{\boldsymbol{\lambda}}_\alpha + b^+_{\boldsymbol{k},\alpha} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \hat{\boldsymbol{\lambda}}^*_\alpha \right] \\
& + \sqrt{\frac{2\pi c^2 \hbar}{L^3}} \sum_{\boldsymbol{k},\alpha} \frac{1}{\sqrt{\omega_k}} \left[\beta_{\boldsymbol{k},\alpha} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \hat{\boldsymbol{\lambda}}_\alpha + \beta^*_{\boldsymbol{k},\alpha} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \hat{\boldsymbol{\lambda}}^*_\alpha \right] .
\end{aligned}
\tag{28.32}
$$

Notice that the second term is nothing but a general solution to the classical Free Maxwell's equations. We can rewrite this as

$$
\boldsymbol{A}_{op}(\boldsymbol{r}) = \sqrt{\frac{2\pi c^2 \hbar}{L^3}} \sum_{\boldsymbol{k},\alpha} \frac{1}{\sqrt{\omega_k}} \left[b_{\boldsymbol{k},\alpha} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \hat{\boldsymbol{\lambda}}_\alpha + b^+_{\boldsymbol{k},\alpha} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \hat{\boldsymbol{\lambda}}^*_\alpha \right] + \boldsymbol{A}_{\text{classical}}(\boldsymbol{r})
\tag{28.33}
$$

In addition, the free electromagnetic Hamiltonian must be transformed too.

$$
\begin{aligned}
H_{em} \;=\;\;& \sum_{\boldsymbol{k},\alpha} (a^+_{\boldsymbol{k},\alpha} a_{\boldsymbol{k},\alpha} + \frac{1}{2}) \hbar\omega_k \\
=\;\;& \sum_{\boldsymbol{k},\alpha} (b^+_{\boldsymbol{k},\alpha} b_{\boldsymbol{k},\alpha} + \frac{1}{2} + b_{\boldsymbol{k},\alpha} \beta^*_{\boldsymbol{k},\alpha} + b^+_{\boldsymbol{k},\alpha} \beta_{\boldsymbol{k},\alpha} + \beta^2_{\boldsymbol{k},\alpha}) \hbar\omega_k
\end{aligned}
\tag{28.34}
$$

To do a calculation such as a transition rate when a classical field is present, or to assess the validity of a classical representation, you would include your classical $\boldsymbol{A}_{\text{classical}}$ vector potential. You can evaluate the $\beta_{\boldsymbol{k},\alpha}$ by Fourier transform. The initial state of the system assuming that you want to calculate with just the classical field, would be the new vacuum state (which is our coherent state in the old representation). If you think you can ignore the quantum fluctuations, put all the terms which contain b and b^+ contributions into a perturbation and solve the rest of the Schroedinger equation. You can now do perturbation theory with the terms containing b and b^+ to see if the classical result is valid. Notice that besides the terms that we had before where the particle emits or absorbs a photon, we also have the $b^+_{\boldsymbol{k},\alpha} \beta_{\boldsymbol{k},\alpha}$ and $b_{\boldsymbol{k},\alpha} \beta^*_{\boldsymbol{k},\alpha}$ terms that, when included, give the effects of the quantum fluctuations of the electromagnetic field.

Chapter 29

Light scattering

29.1   Introduction

The Kramers-Heisenberg formula gives the lowest order quantum mechanical cross section for light scattering from an atom. They derived its form using ideas from old quantum theory and the correspondence principle[1]

We start with a photon in state given by a wavevector $\boldsymbol{k}$ and polarization $\hat{\boldsymbol{\lambda}}$ along with an atom in state A. After scattering we have a photon in state $\boldsymbol{k}'$, $\hat{\boldsymbol{\lambda}}'$, and the atom in state B (where A and B could be the same state).

The Hamiltonian is H_0+V. H_0 is the sum of the atom Hamiltonian with the instantaneous coulomb interaction along with the free electromagnetic Hamiltonian. V contains the $\boldsymbol{J} \cdot \boldsymbol{A}$ interaction terms.

To simplify the discussion let's consider scattering from a Hydrogen atom. It will then be easy to generalize to more complicated systems at the end. For the hydrogen atom

$$H_0 = \frac{p_p^2}{2m_p} + \frac{p_e^2}{2m_e} - \frac{e^2}{|r_e - r_p|} + \sum_{\boldsymbol{k},\alpha} \hbar k c a_{\boldsymbol{k}\alpha}^+ a_{\boldsymbol{k}\alpha} \tag{29.1}$$

and the interaction is

$$\begin{aligned}
V &= \frac{e}{m_e c} \boldsymbol{p}_e \cdot \boldsymbol{A}_{\mathrm{op}}(\boldsymbol{r}_e) + \frac{e^2}{2m_e c^2} \boldsymbol{A}_{op}(\boldsymbol{r}_e) \cdot \boldsymbol{A}_{op}(\boldsymbol{r}_e) \\
&\quad - \frac{e}{m_p c} \boldsymbol{p}_p \cdot \boldsymbol{A}_{\mathrm{op}}(\boldsymbol{r}_p) + \frac{Z^2 e^2}{2m_p c^2} \boldsymbol{A}_{op}(\boldsymbol{r}_p) \cdot \boldsymbol{A}_{op}(\boldsymbol{r}_p)
\end{aligned} \tag{29.2}$$

with

$$\boldsymbol{A}_{\mathrm{op}}(\boldsymbol{r}) = \sum_{\boldsymbol{k}\alpha} \sqrt{\frac{2\pi\hbar c^2}{L^3 \omega_k}} \left[a_{\boldsymbol{k},\alpha} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \hat{\boldsymbol{\lambda}}_\alpha + a_{\boldsymbol{k},\alpha}^+ e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \hat{\boldsymbol{\lambda}}_\alpha^* \right] \tag{29.3}$$

and $\boldsymbol{A}_{\mathrm{op}}$ in Coulomb gauge as usual.

---

[1]H.A. Kramers and W. Heisenberg, Zs. f. Phys., **31**, 681 (1925). Translated into english as "On the dispersion of radiation by atoms," and reprinted in *Sources of Quantum Mechanics*, Ed. B. L. van der Waerden, (Dover, Mineola, 1967).

29.2   Photon Scattering

To scatter, we must destroy the original photon and create the scattered photon. The A^2 terms in the perturbation can do this at first order. The $\boldsymbol{p} \cdot \boldsymbol{A}$ terms contribute at second order in perturbation theory producing the same order in $e^2/\hbar c$. Therefore the lowest order contributions will have the A^2 part of V contributing at first order with the $\boldsymbol{p} \cdot \boldsymbol{A}$ terms contributing at second order.

We can approach this directly with time dependent perturbation theory. We saw generally that the result for the transition rate becomes

$$\Gamma_{i \to j} = \frac{2\pi}{\hbar} |T_{fi}|^2 \delta(E_f - E_i) \tag{29.4}$$

where we can write out the T matrix to second order

$$T = V + V \frac{1}{E - H_0 + i\eta} V + ... \tag{29.5}$$

and keep terms of order e^2. The incident flux for one photon in a box of volume L^3 is c/L^3, so the differential cross section for scattering a photon with the atom initially in state A and finally in state B, possibly equal to state A, is

$$
\begin{aligned}
\frac{d\sigma(A, \lambda \to B, \lambda')}{d\Omega_{k'}} &= \frac{L^6}{(2\pi)^3 c} \int dk' k'^2 \Gamma_{A,\boldsymbol{k},\hat{\boldsymbol{\lambda}} \to B,\boldsymbol{k}',\lambda'} \\
&= \frac{L^6}{(2\pi)^2 \hbar c} \int dk' k'^2 |T_{B,\boldsymbol{k}',\hat{\boldsymbol{\lambda}}';A,\boldsymbol{k},\lambda}|^2 \delta(E_A + \hbar k c - E_B - \hbar k' c)
\end{aligned}
\tag{29.6}
$$

29.3   Calculation of the T-matrix

At this point we have all the pieces and just have to plug in the values, calculate any matrix elements and do the final state summations.

The first simplification we will make is to drop the nucleus terms. If we compare to the electron terms, the nuclear interaction terms will be down by a factor m_e/m_p, compared to the electron terms. Note that there is no fundamental difficulty in keeping the nuclei terms; it just doubles the number of terms in our equations and obscures the main physics. From here on I will write m instead of m_e for convenience since there will not be any other masses in our problem.

We can now write out the T matrix. Figure 29.1 shows diagrammatically the terms that involve a single electron. We can use the diagrams to help write down the expression. The vertices give V terms, and the energy of the intermediate states between the vertices tells us
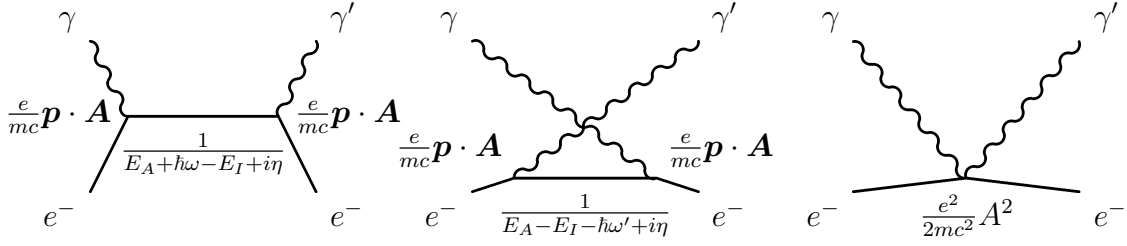
Figure 29.1: The 3 processes that contribute to photon scattering at lowest order. Here the initial state is a photon of energy $\hbar\omega$ and an electronic state of energy E_A. The final state is a photon of energy $\hbar\omega'$ and an electronic state of energy E_B. For the second order terms, the electronic energy is E_I. Notice that cutting across the second order diagrams shows that the energy of the intermediate state is E_I for the first diagram and $E_I + \hbar\omega + \hbar\omega'$ for the second diagram. Subtracting these energies from $E_A + \hbar\omega$ gives the energy denominators shown.

.

the energy denominator.

$$
\begin{aligned}
T_{B,\boldsymbol{k}',\hat{\boldsymbol{\lambda}}';A,\boldsymbol{k},\lambda} &= \frac{e^2}{mc^2}\hat{\boldsymbol{\lambda}}\cdot\hat{\boldsymbol{\lambda}}'^*\frac{2\pi\hbar c^2}{L^3\sqrt{\omega\omega'}}\langle B|e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}_e}|A\rangle \\
&+ \frac{e^2}{m^2c^2}\frac{2\pi\hbar c^2}{L^3\sqrt{\omega\omega'}}\sum_I\left[\frac{\langle B|\boldsymbol{p}_e\cdot\hat{\boldsymbol{\lambda}}'^*e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}_e}|I\rangle\langle I|\boldsymbol{p}_e\cdot\hat{\boldsymbol{\lambda}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}_e}|A\rangle}{E_A - E_I + \hbar\omega + i\eta}\right. \\
&\left.+ \frac{\langle B|\boldsymbol{p}_e\cdot\hat{\boldsymbol{\lambda}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}_e}|I\rangle\langle I|\boldsymbol{p}_e\cdot\hat{\boldsymbol{\lambda}}'^*e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}_e}|A\rangle}{E_A - E_I - \hbar\omega' + i\eta}\right] ,
\end{aligned}
\tag{29.7}
$$

where the factor of 2 in the first term comes because the first $\boldsymbol{A}$ can destroy the incoming photon and the second can create the outgoing photon or the first $\boldsymbol{A}$ can create the outgoing photon and the second one destroy the incoming photon. All of the photon matrix elements have been evaluated leaving just the atomic matrix elements.

Before continuing on, it is now easy to see what we need to do if we have many particles. The first order term is summed over all the particles with their correct masses and charges. The second order term has each of the terms summed over all the particles. That is for N charged particles

$$
\begin{aligned}
T^{\text{general}}_{B,\boldsymbol{k}',\hat{\boldsymbol{\lambda}}';A,\boldsymbol{k},\lambda} &= \sum_{j=1}^{N}\frac{q_j^2}{m_jc^2}\hat{\boldsymbol{\lambda}}\cdot\hat{\boldsymbol{\lambda}}'^*\frac{2\pi\hbar c^2}{L^3\sqrt{\omega\omega'}}\langle B|e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}_j}|A\rangle \\
&+ \sum_{j=1}^{N}\sum_{j'=1}^{N}\frac{q_jq_{j'}}{m_jm_{j'}c^2}\frac{2\pi\hbar c^2}{L^3\sqrt{\omega\omega'}}\sum_I\left[\frac{\langle B|\boldsymbol{p}_j\cdot\hat{\boldsymbol{\lambda}}'^*e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}_j}|I\rangle\langle I|\boldsymbol{p}_{j'}\cdot\hat{\boldsymbol{\lambda}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}_{j'}}|A\rangle}{E_A - E_I + \hbar\omega + i\eta}\right. \\
&\left.+ \frac{\langle B|\boldsymbol{p}_j\cdot\hat{\boldsymbol{\lambda}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}_j}|I\rangle\langle I|\boldsymbol{p}_{j'}\cdot\hat{\boldsymbol{\lambda}}'^*e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}_{j'}}|A\rangle}{E_A - E_I - \hbar\omega' + i\eta}\right] ,
\end{aligned}
\tag{29.8}
$$

347

and the matrix elements are between the many-particle H_0 eigenstates. So if we understand how to calculate for the hydrogen atom, the only complication for more complex systems is the need to calculate all of the needed matrix elements. Note also that if we include all of the particles, the center of mass momentum will take up the momentum difference of the two photons. In our transition rate we need to sum or integrate over the final state center of mass quantum numbers.

29.4   The Dipole Approximation

Going back to our hydrogen atom example, let's look more closely at the case where the photon wavelength is much longer than the atomic size. In that case, we can write the exponentials as

$$
e^{i\boldsymbol{k}\cdot\boldsymbol{r}_e} = e^{i\boldsymbol{k}\cdot[\boldsymbol{R}_{cm}+\frac{m_p}{m_p+m_e}\boldsymbol{r}]} = e^{i\boldsymbol{k}\cdot\boldsymbol{R}_{cm}}[1+O(kr)]. \tag{29.9}
$$

Dropping the order kr term gives the dipole approximation. With this approximation all of the terms contain a center of mass part that simply multiplies the center of mass wave function by $e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_{cm}}$ so that the center of mass takes up the extra momentum. This, of course, changes the energy of the final center of mass state, and that energy needs to be included in the delta function in Fermi's golden rule. Here we will again make the approximation that the proton mass is so large so that we can ignore the change in the center of mass energy. Again, note that it is not substantially more difficult to include it, but it does make the expressions a little longer. We will also make the same approximation for the momenta and replace $\boldsymbol{p}_e$ by $\boldsymbol{p}$, the relative momentum. The states $|A\rangle$, $|B\rangle$, and $|I\rangle$ then refer to just the relative coordinates in this approximation.

The T matrix for the hydrogen atom assuming an infinitely massive proton and in the dipole approximation is then

$$
\begin{aligned}
T_{B,\boldsymbol{k}',\hat{\boldsymbol{\lambda}}';A,\boldsymbol{k},\lambda} \;=\; & \frac{2\pi e^2\hbar}{L^3 m\sqrt{\omega\omega'}}\left\{\, \hat{\boldsymbol{\lambda}}\cdot\hat{\boldsymbol{\lambda}}'^{*}\langle B|A\rangle \right. \\
& \left. +\frac{1}{m}\sum_I\left[\frac{\langle B|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}'^{*}|I\rangle\langle I|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}|A\rangle}{E_A-E_I+\hbar\omega+i\eta} + \frac{\langle B|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}|I\rangle\langle I|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}'^{*}|A\rangle}{E_A-E_I-\hbar\omega'+i\eta}\right]\right\}.
\end{aligned}
$$
$$\tag{29.10}$$

Plugging this into our expression for the cross section, we have

$$
\begin{aligned}
\frac{d\sigma(A,\lambda \to B,\lambda')}{d\Omega_{k'}} &= \frac{L^6}{(2\pi)^2 \hbar c} \int dk' k'^2 |T_{B,\boldsymbol{k}',\hat{\boldsymbol{\lambda}}';A,\boldsymbol{k},\lambda}|^2 \delta(E_A + \hbar k c - E_B - \hbar k' c) \\
&= \int dk' k'^2 \frac{e^4 \hbar}{m^2 c \omega \omega'} \delta(E_A + \hbar\omega - E_B - \hbar\omega') \Bigg| \hat{\boldsymbol{\lambda}} \cdot \hat{\boldsymbol{\lambda}}'^* \langle B|A\rangle \\
&\quad + \frac{1}{m} \sum_I \left[\frac{\langle B|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}'^*|I\rangle\langle I|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}|A\rangle}{E_A - E_I + \hbar\omega + i\eta} + \frac{\langle B|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}|I\rangle\langle I|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}'^*|A\rangle}{E_A - E_I - \hbar\omega' + i\eta} \right] \Bigg|^2 \\
&= \left(\frac{e^2}{mc^2}\right)^2 \frac{\omega'}{\omega} \Bigg| \hat{\boldsymbol{\lambda}} \cdot \hat{\boldsymbol{\lambda}}'^* \langle B|A\rangle \\
&\quad + \frac{1}{m} \sum_I \left[\frac{\langle B|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}'^*|I\rangle\langle I|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}|A\rangle}{E_A - E_I + \hbar\omega + i\eta} + \frac{\langle B|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}|I\rangle\langle I|\boldsymbol{p}\cdot\hat{\boldsymbol{\lambda}}'^*|A\rangle}{E_A - E_I - \hbar\omega' + i\eta} \right] \Bigg|^2
\end{aligned}
\tag{29.11}
$$

where, because of the delta function integral, we must set $\hbar\omega' = E_A - E_B + \hbar\omega$. The factor $r_0 = e^2/mc^2 \simeq 2.82 \times 10^{-13}$cm is called the classical radius of the electron since it is the radius where e^2/r_0 gives the electron rest energy.

29.5 Limiting Cases

Let's look at some interesting limits. First let's explore elastic scattering, so that $\omega = \omega'$. We will look in the two limits of low and high energy photons.

29.5.1 Thomson scattering

If the photon energy is much larger than $|E_A - E_I|$, the energy difference of the atomic states that are coupled by the momentum operator, then we can ignore the $E_A - E_I$ in the energy denominators. The intermediate states then give a completeness relation which can be summed. Since all components of $\boldsymbol{p}$ commute, the matrix elements of the last two terms are then equal and the denominators are equal and opposite. These two terms cancel leaving just the A_{op}^2 contribution. The atomic state is then unchanged and the cross section is the Thomson cross section.

$$
\frac{d\sigma}{d\Omega_{k'}} = r_0^2 |\hat{\boldsymbol{\lambda}} \cdot \hat{\boldsymbol{\lambda}}'^*|^2 .
\tag{29.12}
$$

Let's look at this in more detail. If we do not measure the outgoing polarization we must add the cross sections for the two polarizations. We can use the polarization summation

trick to write

$$\frac{d\sigma}{d\Omega_{k'}}\bigg|_{\text{polarization not measured}} = \sum_{\lambda'} r_0^2 \hat{\boldsymbol{\lambda}} \cdot \hat{\boldsymbol{\lambda}}'^* \hat{\boldsymbol{\lambda}}' \cdot \hat{\boldsymbol{\lambda}}^* =$$

$$= r_0^2 \left[1 - \left| \hat{\boldsymbol{\lambda}} \cdot \hat{\boldsymbol{k}}' \right|^2 \right] . \tag{29.13}$$

If further we use unpolarized light for our beam, then it is like doing two kinds of experiments, half the time we use one of the polarizations and half the other. The cross section becomes

$$\frac{d\sigma}{d\Omega_{k'}}\bigg|_{\text{unpolarized polarization not measured}} = \frac{r_0^2}{2} \sum_{\lambda} \left[1 - \left| \hat{\boldsymbol{\lambda}} \cdot \hat{\boldsymbol{k}}' \right|^2 \right]$$

$$= \frac{r_0^2}{2} \left\{ 2 - \left[1 - \left(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}' \right)^2 \right] \right\} = \frac{r_0^2}{2} \left[1 + \left(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{k}}' \right)^2 \right]$$

$$= \frac{r_0^2}{2} \left[1 + \cos^2 \theta_{k'} \right] \tag{29.14}$$

where $\theta_{k'}$ is the angle between the incident photon direction and the outgoing photon direction.

Let's choose the incident photon to be propagating along $\hat{\boldsymbol{z}}$, with linear polarization along $\hat{\boldsymbol{x}}$. The outgoing photon's direction is given by $\theta_{k'}$ and $\phi_{k'}$ defined in the usual way according to the chosen axes so that

$$\boldsymbol{k}' = k \left[\sin\theta_{k'} \cos\phi_{k'} \hat{\boldsymbol{x}} + \sin\theta_{k'} \sin\phi_{k'} \hat{\boldsymbol{y}} + \cos\theta_{k'} \hat{\boldsymbol{z}} \right] . \tag{29.15}$$

We then have

$$\frac{d\sigma}{d\Omega_{k'}}\bigg|_{\text{polarization not measured}} = r_0^2 \left[1 - \sin^2 \theta_{k'} \cos^2 \phi_{k'} \right] . \tag{29.16}$$

Notice that picking the incoming polarization along $\hat{\boldsymbol{y}}$ would simply replace $\cos\phi_{k'}$ with $\sin\phi_{k'}$ and averaging these would give the unpolarized result of Eq. 29.14.

To analyze further, we can select two orthogonal polarizations for the outgoing photon. It is convenient to pick one polarization vector λ_1' in the $x-y$ plane. Orthogonalizing to $\hat{\boldsymbol{k}}'$ and the then taking $\lambda_2' = \hat{\boldsymbol{k}}' \times \lambda_1'$ gives

$$\begin{aligned} \lambda_1' &= \sin\phi_{k'} \hat{\boldsymbol{x}} - \cos\phi_{k'} \hat{\boldsymbol{y}} \\ \lambda_2' &= \cos\theta_{k'} \cos\phi_{k'} \hat{\boldsymbol{x}} + \cos\theta_{k'} \sin\phi_{k'} \hat{\boldsymbol{y}} - \sin\theta_{k'} \hat{\boldsymbol{z}} . \end{aligned} \tag{29.17}$$

The cross section for measuring these two polarizations for an x polarized incoming photon would be

$$\begin{aligned} \frac{d\sigma(\hat{\boldsymbol{x}} \to \lambda_1')}{d\Omega_{k'}} &= \sin^2 \phi_{k'} \\ \frac{d\sigma(\hat{\boldsymbol{x}} \to \lambda_2')}{d\Omega_{k'}} &= \cos^2 \theta_{k'} \cos^2 \phi_{k'} . \end{aligned} \tag{29.18}$$

We can also ask what is the amount of polarization of the outgoing photon if the incoming photon (still along $\hat{\bm{z}}$) is unpolarized. In that case, again using our polarization summation trick, we have

$$\frac{d\sigma}{d\Omega_{k'}}\bigg|_{\text{unpolarized}} = r_0^2 \left[1 - \left|\lambda' \cdot \hat{\bm{k}}\right|^2\right]. \tag{29.19}$$

This shows that the density matrix in our $\hat{\bm{\lambda}}_1'$, $\hat{\bm{\lambda}}_2'$ basis is proportional to

$$\rho \propto \begin{pmatrix} 1 - (\hat{\bm{\lambda}}_1' \cdot \hat{\bm{k}})^2 & -\hat{\bm{\lambda}}_1' \cdot \hat{\bm{k}}\hat{\bm{\lambda}}_2' \cdot \hat{\bm{k}} \\ -\hat{\bm{\lambda}}_2' \cdot \hat{\bm{k}}\hat{\bm{\lambda}}_1' \cdot \hat{\bm{k}} & 1 - (\hat{\bm{\lambda}}_2' \cdot \hat{\bm{k}})^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 - \sin^2\theta_{k'} \end{pmatrix} \tag{29.20}$$

so that by amazing luck, we have chosen the polarization states that diagonalize the density matrix. If we had chosen different ones, we could now diagonalize the density matrix to find these polarizations.

Since the trace of ρ must be 1, the density matrix is

$$\rho = \begin{pmatrix} \frac{1}{1+\cos^2\theta_{k'}} & 0 \\ 0 & \frac{\cos^2\theta_{k'}}{1+\cos^2\theta_{k'}} \end{pmatrix}. \tag{29.21}$$

We see that the polarization of the outgoing photon can be viewed as a fraction of $2\cos^2\theta_{k'}/(1+\cos^2\theta_{k'})$ unpolarized light and a fraction $(1 - \cos^2\theta_{k'})/(1 + \cos^2\theta_k')$ polarized along λ_1', that is, in the $x - y$ plane. The scattering is fully polarized at $\theta_{k'} = \pi/2$.

29.5.2   Rayleigh Scattering

Rayleigh scattering is the low frequency limit for elastic scattering. Since the classical Rayleigh theory explained the blue sky and its polarization, we should expect quantum theory to give the same result. However, now all three terms must be kept and at first glance they do not look very much like Rayleigh's classical result.

To get something that looks like the dipole operator, we can replace the p_α by the commutator $p_\alpha = \frac{im}{\hbar}[H_{\text{hydrogen}}, r_\alpha]$, and expand the energy denominators in a power series in ω. In order to combine the $\hat{\bm{\lambda}} \cdot \hat{\bm{\lambda}}'^*$ term we insert a complete set of states, and write the dot product in terms of momenta and position commutators,

$$\begin{aligned} \hat{\bm{\lambda}} \cdot \hat{\bm{\lambda}}'^* &= \sum_{\alpha,\beta} \hat{\lambda}_\alpha \delta_{\alpha,\beta} \hat{\lambda}_\beta'^* = \frac{1}{i\hbar}\hat{\lambda}_\alpha[r_\alpha, p_\beta]\hat{\lambda}_\beta'^* \\ &= \frac{1}{i\hbar}\left(\bm{r} \cdot \hat{\bm{\lambda}} \bm{p} \cdot \hat{\bm{\lambda}}'^* - \bm{p} \cdot \hat{\bm{\lambda}}'^* \bm{r} \cdot \hat{\bm{\lambda}}\right) \end{aligned} \tag{29.22}$$

Using our commutator above we can write

$$\langle I|[H_{\text{hydrogen}}, \bm{r}]|A\rangle = (E_I - E_A)\langle I|\bm{r}|A\rangle \tag{29.23}$$

or

$$\langle I|\boldsymbol{r}|A\rangle = \frac{i\hbar}{m(E_I - E_A)}\langle I|\boldsymbol{p}|A\rangle. \tag{29.24}$$

We then can write

$$\langle A|\hat{\boldsymbol{\lambda}} \cdot \hat{\boldsymbol{\lambda}}'^*|A\rangle = -\frac{2}{m}\sum_I \frac{\langle A|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}|I\rangle\langle I|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}'^*|A\rangle}{E_A - E_I} \tag{29.25}$$

Repeating the calculation for $\hat{\boldsymbol{\lambda}}'^* \cdot \hat{\boldsymbol{\lambda}}$ and and averaging the result we have

$$\langle A|\hat{\boldsymbol{\lambda}} \cdot \hat{\boldsymbol{\lambda}}'^*|A\rangle = -\frac{1}{m}\sum_I \frac{\langle A|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}|I\rangle\langle I|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}'^*|A\rangle + \langle A|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}'^*|I\rangle\langle I|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}|A\rangle}{E_A - E_I}. \tag{29.26}$$

The Kramers-Heisenberg result for elastic scattering can now be written as

$$\frac{d\sigma}{d\Omega_{k'}} = \frac{r_0^2}{m^2}\left|\sum_I -\frac{\hbar\omega\langle A|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}'^*|I\rangle\langle I|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}|A\rangle}{(E_A - E_I)(E_A - E_I + \hbar\omega + i\eta)} + \frac{\hbar\omega\langle A|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}|I\rangle\langle I|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}'^*|A\rangle}{(E_A - E_I)(E_A - E_I - \hbar\omega + i\eta)}\right|^2. \tag{29.27}$$

Expanding in powers of ω, we get

$$\frac{d\sigma}{d\Omega_{k'}} = \frac{r_0^2\hbar^4\omega^4}{m^2}\left|\sum_I \frac{\langle A|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}'^*|I\rangle\langle I|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}|A\rangle + \langle A|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}|I\rangle\langle I|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}'^*|A\rangle}{(E_I - E_A)^3}\right|^2. \tag{29.28}$$

Using our commutator expression for $\boldsymbol{p}$ once again and rearranging the factors, this becomes

$$\frac{d\sigma}{d\Omega_{k'}} = \frac{\omega^4}{c^4}\left|\sum_I \frac{\langle A|e\boldsymbol{r} \cdot \hat{\boldsymbol{\lambda}}'^*|I\rangle\langle I|e\boldsymbol{r} \cdot \hat{\boldsymbol{\lambda}}|A\rangle + \langle A|e\boldsymbol{r} \cdot \hat{\boldsymbol{\lambda}}|I\rangle\langle I|e\boldsymbol{r} \cdot \hat{\boldsymbol{\lambda}}'^*|A\rangle}{E_I - E_A}\right|^2. \tag{29.29}$$

The summation is the expression we found last semester for the static polarizability tensor, which gives the induced dipole in terms of the applied field

$$\boldsymbol{d} = \overset{\leftrightarrow}{\alpha} \cdot \boldsymbol{E}_{\text{applied}}. \tag{29.30}$$

Recalling that calculation, applying a constant in space, static, field $\boldsymbol{E}_{\text{applied}}$, the perturbing potential is

$$V = e\boldsymbol{r} \cdot \boldsymbol{E}_{\text{applied}} \tag{29.31}$$

The perturbation to the state $|A\rangle$, to first order in the perturbation, is

$$|A\rangle + \sum_{I \neq A} \frac{|I\rangle\langle I|e\boldsymbol{r} \cdot \boldsymbol{E}_{\text{applied}}|A\rangle}{E_A - E_I} \tag{29.32}$$

and the dipole moment is then[2]

$$\boldsymbol{d} = \sum_{I \neq A} \frac{\langle A|er|I\rangle \langle I|er \cdot \boldsymbol{E}_{\text{applied}}|A\rangle + \langle A|er \cdot \boldsymbol{E}_{\text{applied}}|I\rangle \langle I|er|A\rangle}{E_I - E_A} \tag{29.33}$$

Since the matrix elements with $I = A$ do not contribute, Eq. 29.29 is

$$\frac{d\sigma}{d\Omega_{k'}} = \frac{\omega^4}{c^4} \left| \hat{\boldsymbol{\lambda}} \cdot \overset{\leftrightarrow}{\alpha} \cdot \hat{\boldsymbol{\lambda}}'^* \right|^2 . \tag{29.34}$$

Rayleigh didn't know how to calculate the polarizability, but he assumed that Eq. 29.30 was correct. The classical vector potential in Lorentz gauge with a harmonic current $\boldsymbol{J}(\boldsymbol{r})e^{-i\omega t}$ (as usual, we must take the real part of all quantities with an $e^{-i\omega t}$ factor to get the physical quantity) is

$$\boldsymbol{A}(\boldsymbol{r}, t) = \frac{1}{c} \int d^3r' dt' \frac{\boldsymbol{J}(\boldsymbol{r}')e^{-i\omega t'}}{|\boldsymbol{r} - \boldsymbol{r}'|} \delta(t' - t + c^{-1}|\boldsymbol{r} - \boldsymbol{r}'|) = \frac{1}{c} \int d^3r' \frac{\boldsymbol{J}(\boldsymbol{r})e^{i\frac{\omega}{c}|\boldsymbol{r} - \boldsymbol{r}'| - i\omega t}}{|\boldsymbol{r} - \boldsymbol{r}'|} . \tag{29.35}$$

Taking r large to look at the radiation fields, we expand $|\boldsymbol{r} - \boldsymbol{r}'|$ and

$$\boldsymbol{A}(\boldsymbol{r} \to \infty, t) = \frac{e^{i\frac{\omega}{c}r - i\omega t}}{rc} \int d^3r' \boldsymbol{J}(\boldsymbol{r}')e^{-i\frac{\omega}{c}\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} . \tag{29.36}$$

If the wavelength is long compared to the size of the current distribution, we can set the exponential to 1 just as in our quantum calculation. We then use the identity

$$\begin{aligned}
\hat{\boldsymbol{\alpha}} \cdot J(\boldsymbol{r}) &= \boldsymbol{\nabla} \cdot [\hat{\boldsymbol{\alpha}} \cdot \boldsymbol{r} \boldsymbol{J}(\boldsymbol{r})] - \hat{\boldsymbol{\alpha}} \cdot \boldsymbol{r} \boldsymbol{\nabla} \cdot \boldsymbol{J}(\boldsymbol{r}) \\
&= \boldsymbol{\nabla} \cdot [\hat{\boldsymbol{\alpha}} \cdot \boldsymbol{r} \boldsymbol{J}(\boldsymbol{r})] - i\omega \hat{\boldsymbol{\alpha}} \cdot \boldsymbol{r} \rho(\boldsymbol{r}) .
\end{aligned} \tag{29.37}$$

Substituting the divergence theorem gives zero for the integral of a divergence of the localized current source, and the second term gives the usual dipole approximation. Writing

$$\boldsymbol{d} = \int d^3r \boldsymbol{r} \rho(\boldsymbol{r}) \tag{29.38}$$

The vector potential becomes

$$\begin{aligned}
\boldsymbol{A}(\boldsymbol{r} \to \infty, t) &= -\frac{i\omega e^{i\frac{\omega}{c}r - i\omega t}}{rc} \boldsymbol{d} \\
&= -\frac{i\omega e^{i\frac{\omega}{c}r - i\omega t}}{rc} \overset{\leftrightarrow}{\alpha} \cdot \boldsymbol{E}_{\text{applied}} .
\end{aligned} \tag{29.39}$$

Since the scalar potential only gives a longitudinal part, we can calculate the radiation electric field along the polarization $\hat{\boldsymbol{\lambda}}'$ by taking the negative time derivative divided by c

---

[2]If the Hamiltonian and the state $|A\rangle$ are invariant under time reversal, the two terms are equal and can be combined.
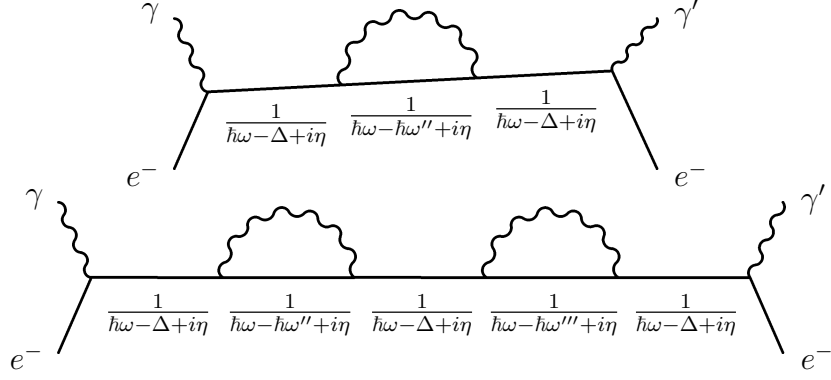
Figure 29.2: The first two additional diagrams that contribute to the $1s$-$2p$ resonance scattering. The electron changes from $1s$ to $2p$ or $2p$ to $1s$ at every photon vertex. The vertices all have $\frac{e}{mc}\boldsymbol{p}\cdot\boldsymbol{A}$ factors. The intermediate loop photon states must be summed over, and $\Delta \equiv E_{2p} - E_{1s}$

and dotting it into the polarization direction. The classical Rayleigh cross section is then the power radiated into a solid angle with polarization $\hat{\boldsymbol{\lambda}}'$ divided by the incident power per unit area from the applied field. The applied electric field is along $\hat{\boldsymbol{\lambda}}$ so that

$$\frac{d\sigma_{\text{classical}}}{d\Omega_{k'}} = \frac{\frac{c}{4\pi}|\hat{\boldsymbol{\lambda}}'^* \cdot \boldsymbol{E}(\boldsymbol{r}\to\infty)|^2 r^2}{\frac{c}{4\pi}|\boldsymbol{E}_{\text{applied}}|^2} = \frac{\omega^4}{c^4}|\hat{\boldsymbol{\lambda}}^* \cdot \overset{\leftrightarrow}{\alpha} \cdot \hat{\boldsymbol{\lambda}}'|^2 \tag{29.40}$$

in agreement with the quantum mechanical result. But be sure to realize that the classical result does not give a value for $\overset{\leftrightarrow}{\alpha}$ while quantum mechanics does. Quantum mechanics also does not need to assume elastic scattering and can describe photons of any frequency.

29.6 Resonance

The energy denominators in the T-matrix Eq. 29.10 can become small. Let's sketch out an explicit case where $|A\rangle$ is the hydrogen $1s$ ground state and the incident photon has an energy very close to $E_{2p} - E_{1s}$. In that case, when the state $|I\rangle$, is the appropriate $2p$ state corresponding to the incoming photon polarization, the energy denominator $E_A - E_I + \hbar\omega + i\eta$ becomes small and the corresponding term in the T-matrix becomes large. In fact it becomes infinite if the photon energy is exactly $E_{2p} - E_{1s}$. Of course if a term in a perturbation expansion becomes large it indicates that the expansion may not be converging. One way to deal with this is to try to sum more terms in the T-matrix equation to get a convergent result. If we look at higher order terms, we need to address those that are also getting large. This will occur when they have the same energy denominator.

Diagrammatically, the relevant terms at higher order that are all diverging are like those shown in fig. 29.2. Here the electronic state changes from $1s$ to $2p$ or $2p$ to $1s$ at each vertex and the internal photon loop will give a diverging contribution when that photon also has an

energy of $E_{2p} - E_{1s}$. Each internal photon gives the same contribution and summing up all of these types of diagrams multiplies the diverging diagram at lowest order by a geometric series

$$S = 1 + x + x^2 + x^3 + ... = \frac{1}{1-x} \tag{29.41}$$

where (using the dipole approximation)

$$x = \frac{1}{E_{1s} - E_{2p} + \hbar\omega + i\eta} \sum_{\lambda''} \int \frac{d^3k''}{(2\pi)^3} \frac{2\pi e^2 \hbar c^2}{m^2 c^2} \frac{|\langle 2p|\boldsymbol{p} \cdot \hat{\boldsymbol{\lambda}}''|1s\rangle|^2}{\hbar\omega - \hbar\omega'' + i\eta}. \tag{29.42}$$

Under the integral sign the denominator can be replaced by a principal parts integration which will give a small real term, and $-i\pi\delta(\hbar\omega - \hbar\omega')$ which will give an imaginary part. Notice that if $\hbar\omega = E_{2p} - E_{1s}$, that the imaginary integral is $-\frac{i}{2}\hbar\Gamma_{2p\to 1s}$ where $\Gamma_{2p\to 1s}$ is the transition rate we calculated for the $2p$ to $1s$ transition. Writing the sum and integral as $\Sigma(\omega)$ we have

$$x = \frac{\text{Re}\Sigma(\omega) + i\text{Im}\Sigma(\omega)}{E_{1s} - E_{2p} + \hbar\omega + i\eta} \tag{29.43}$$

the diverging energy denominator is replaced by

$$\frac{S}{E_{1s} - E_{2p} + \hbar\omega + i\eta} = \frac{1}{E_{1s} - E_{2p} + \hbar\omega - \text{Re}\Sigma(\omega) - i\text{Im}\Sigma(\omega)} \tag{29.44}$$

where we can drop the $i\eta$ since we have added the finite imaginary part of Σ. Near the resonance (where I assume the frequency shift from the real part of Σ is small and can be ignored) this becomes

$$\frac{1}{E_{1s} - E_{2p} + \hbar\omega - \text{Re}\Sigma(\omega) - i\text{Im}\Sigma(\omega)} \to \frac{1}{E_{1s} - E_{2p} + \hbar\omega + i\frac{\hbar}{2}\Gamma_{2p\to 1s}} \tag{29.45}$$

and we get the usual Lorentzian resonant shape. Quite generally you get a resonance whenever there is a long lived intermediate state. Mathematically the T-matrix develops a pole near the real axis which tends to dominate the scattering at that energy.

29.7   Inelastic Scattering

Inelastic photon scattering is usually called Raman scattering. Here $\omega \neq \omega'$ and $|A\rangle \neq |B\rangle$. If $|A\rangle$ is the ground state, then we must have $\omega' < \omega$. This is called the Stokes line which is lower frequency than the incident radiation. If the system is in an excited state, then we can also have $\omega' > \omega$. These are called anti-Stokes lines. Notice that the states $|A\rangle$ and $|B\rangle$ are coupled by two dipole operators. Since the dipole operator changes parity, these two states must have the same parity. On the other hand, electric dipole decays and therefore the reverse process which is absorption and there is no outgoing photon, the states must be coupled by a dipole operator.

- Transitions that have a large absorption cross section in the dipole approximation change the parity of the electronic state.

- Transitions that have a large Raman cross section in the dipole approximation do not change the parity of the electronic state.

Therefore these two kinds of experiments probe different sets of states of the system.

Chapter 30

Lorentz QED

In physical matrix elements for measurable quantities we can always arrange things so that we first create and then detect all the photons so the initial and final states can be the vacuum. In perturbation theory we calculate time-ordered products of the perturbing potential, and this leads electromagnetic matrix elements like

$$\langle 0|A_\nu(x')A_\mu(x)|0\rangle\Theta(t'-t) \tag{30.1}$$

or

$$\langle 0|A_\nu(x)A_\mu(x')|0\rangle\Theta(t-t') \tag{30.2}$$

where $|0\rangle$ is the vacuum state, μ and ν are indices indicating four component objects. Here I take units where $\hbar$ and c are unity. Relativistic notation is used, so that in terms of the usual space and spin positions

$$x_\mu \to (x_0, x_1, x_2, x_3) \equiv (t, -\boldsymbol{r})$$
$$x^\mu \to (x^0, x^1, x^2, x^3) \equiv (t, \boldsymbol{r}) \tag{30.3}$$

and for the four vectors,

$$x^2 = x_\nu x^\mu = t^2 - r^2$$
$$p^\mu \to (E, \boldsymbol{p})$$
$$p_1 \cdot p_2 = E_1 E_2 - \boldsymbol{p}_1 \cdot \boldsymbol{p}_2$$
$$\boldsymbol{p} \cdot \boldsymbol{x} = tE - \boldsymbol{r} \cdot \boldsymbol{p} \tag{30.4}$$

and the metric tensor $g_{\mu\nu}$ is diagonal with diagonal elements $(1, -1, -1, -1)$.

Our vector potential in Coulomb gauge has only components along the transverse directions to the $\boldsymbol{k}$ components. The electromagnetic amplitudes we calculate will be given by the time ordered products of matrix elements with only these transverse components. These can be written in terms of the Feynman propagator with only transverse components

$$
\begin{aligned}
iD_F^{TR}(x', x)_{\mu\nu} &= \langle 0|A_\nu(x')A_\mu(x)|0\rangle\Theta(t'-t) + \langle 0|A_\mu(x)A_\nu(x')|0\rangle\Theta(t-t') \\
&= \langle 0|T[A_\mu(x')A_\mu(x)]|0\rangle \tag{30.5}
\end{aligned}
$$

where T is the time ordering operator. We can evaluate the transverse propagator by plugging in our expression for the Coulomb gauge vector potential operator. Only terms where we create then destroy the same photon will contribute. We get

$$iD_F^{TR}(x', x)_{\mu\nu} \;=\; -i4\pi \int \frac{d^3k}{(2\pi)^3} \frac{1}{2|\boldsymbol{k}|} \sum_{\alpha=1,2} \hat{\boldsymbol{\lambda}}_\mu^{(\alpha)}(\boldsymbol{k}) \hat{\boldsymbol{\lambda}}_\nu^{(\alpha)}(\boldsymbol{k}) \left[\Theta(t'-t)e^{i\boldsymbol{k}\cdot(\boldsymbol{r}'-\boldsymbol{r})-i|\boldsymbol{k}|(t'-t)}\right.$$
$$\left. +\Theta(t-t')e^{-i\boldsymbol{k}\cdot(\boldsymbol{r}'-\boldsymbol{r})+i|\boldsymbol{k}|(t'-t)}\right] . \tag{30.6}$$

It is convenient to change variables from $-\boldsymbol{k}$ to $\boldsymbol{k}$ in the second term to give

$$iD_F^{TR}(x', x)_{\mu\nu} \;=\; -i4\pi \int \frac{d^3k}{(2\pi)^3} \frac{1}{2|\boldsymbol{k}|} \sum_{\alpha=1,2} \hat{\boldsymbol{\lambda}}_\mu^{(\alpha)}(\boldsymbol{k}) \hat{\boldsymbol{\lambda}}_\nu^{(\alpha)}(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot(\boldsymbol{r}'-\boldsymbol{r})}$$
$$\left[\Theta(t'-t)e^{i|\boldsymbol{k}|(t'-t)} + \Theta(t-t')e^{i|\boldsymbol{k}|(t'-t)}\right] . \tag{30.7}$$

To put this into a more Lorentz covariant formalism, we write

$$\int_{-\infty}^\infty \frac{d\omega}{2\pi} \frac{e^{-i\omega(t'-t)}}{\omega^2 - |\boldsymbol{k}|^2 + i\eta} = \frac{1}{2|\boldsymbol{k}|} \left[\Theta(t'-t)e^{-i|\boldsymbol{k}|(t'-t)} + \Theta(t-t')e^{i|\boldsymbol{k}|(t'-t)}\right] \tag{30.8}$$

which can be checked by contour integration. Combining this gives the more covariant looking result,

$$D_F^{TR}(x', x)_{\mu\nu} = -4\pi \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot(x'-x)}}{k^2 + i\eta} \sum_{\alpha=1,2} \hat{\boldsymbol{\lambda}}_\mu^{(\alpha)}(k) \hat{\boldsymbol{\lambda}}_\nu^{(\alpha)}(k) . \tag{30.9}$$

Now the only part of the matrix element that is not manifestly Lorentz covariant is the polarization sum. We will now pick two more polarization vectors to make a complete set, and add and subtract these from the propagator. The term with the full set of polarization vectors will be manifestly covariant. We will see that the extra terms cancel the instantaneous Coulomb interaction leaving terms that will not couple to a conserved current and can be dropped for physical calculations. The result is electrodynamics described by photon exchange with photons of 4 polarizations.

We pick one new polarzation in the time direction $\hat{\boldsymbol{\eta}} = (1, 0, 0, 0)$ and one along $\hat{\boldsymbol{k}}$. The latter can be written in terms of the four vector k^μ as

$$\hat{\boldsymbol{k}}^\mu = \tilde{k}^\mu = \frac{k^\mu - (k\cdot\eta)\eta^\mu}{\sqrt{(k\cdot\eta)^2 - k^2}} \tag{30.10}$$

$\tilde{k}$ is space like, and we can write the metric tensor as

$$g_{\mu\nu} = \eta_\mu\eta_\nu - \tilde{k}_\mu\tilde{k}_\nu - \sum_{\alpha=1,2} \hat{\boldsymbol{\lambda}}_\mu^{(\alpha)}(k) \hat{\boldsymbol{\lambda}}_\nu^{(\alpha)}(k) \tag{30.11}$$

Therefore we can write our transverse polarization sum as

$$\sum_{\alpha=1,2} \hat{\boldsymbol{\lambda}}_\mu^{(\alpha)}(k)\hat{\boldsymbol{\lambda}}_\nu^{(\alpha)}(k) = -g_{\mu\nu} + \eta_\mu\eta_\nu - \tilde{k}_\mu\tilde{k}_\nu . \tag{30.12}$$

Plugging into the transverse propagator gives

$$
\begin{aligned}
D_F^{TR}(x',x)_{\mu\nu} &= D_F(x',x)_{\mu\nu} \\
&\quad -4\pi \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot(x'-x)}}{k^2+i\eta} \left[\frac{k^2\hat{\boldsymbol{\eta}}_\mu\hat{\boldsymbol{\eta}}_\nu - (k\cdot\hat{\boldsymbol{\eta}})(k_\nu\hat{\boldsymbol{\eta}}_\mu + \hat{\boldsymbol{\eta}}_\nu k_\mu) + k_\mu k_\nu}{(k\cdot\eta)^2 - k^2} \right]
\end{aligned}
\tag{30.13}
$$

where the Feynman propagator is

$$D_F(x',x)_{\mu\nu} = g_{\mu\nu}4\pi \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot(x'-x)}}{k^2+i\eta} . \tag{30.14}$$

The matrix elements we calculate always have the potential coupled to a conserved current. That is they have forms like $j^\mu A_\mu$, where the currents satisfy the continuity equation. In 3-vector notation this is

$$\boldsymbol{\nabla}\cdot\boldsymbol{J} + \frac{\partial\rho}{\partial t} = 0 \tag{30.15}$$

and Fourier transforming will give $k_\mu j^\mu(k) = 0$ where $j^\mu(k)$ is the Fourier transform of the current. The form above shows that integrating over the particle part of the wave function will give a fourier transform from the $e^{-ik\cdot(x'-x)}$ and therefore all of the terms proportional to k_μ or k_ν will give zero from current conservation. If we drop these we get the Feynman propagator and the term

$$
\begin{aligned}
-4\pi \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot(x'-x)}}{k^2+i\eta} \frac{k^2\hat{\boldsymbol{\eta}}_\mu\hat{\boldsymbol{\eta}}_\nu}{(k\cdot\eta)^2 - k^2} &= -\hat{\boldsymbol{\eta}}_\mu\hat{\boldsymbol{\eta}}_\nu 4\pi \int \frac{d\omega}{2\pi} e^{-i\omega(t'-t)} \int \frac{d^3k}{(2\pi)^3} \frac{e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')}}{|\boldsymbol{k}|^2} \\
&= -g_{\nu 0}g_{\mu 0}\delta(t'-t)\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}
\end{aligned}
\tag{30.16}
$$

which cancels the instantaneous Coulomb interaction term in Coulomb gauge. The $g_{\mu 0}g_{\nu 0}$ terms pick out just the charge density terms from the four-current operator.

The result is that we can calculate all matrix elements using the manifestly Lorentz covariant Feynman propagator which exchanges photons of four polarizations and without any direct interaction between the particles. The corresponding vector potential is in Lorentz Gauge.

Alternatively, you can view the Coulomb gauge result as coming from integrating out the photons with longitudinal and time like polarizations from the Lorentz Gauge result to give the instantaneous Coulomb interaction.

Chapter 31

Lamb shift and renormalization ideas

31.1 Infinities and self energy

We have managed to calculate the lowest order transition rates using quantum electrodynamics and a nonrelativistic Hamiltonian for the electrons. Unfortunately, when we go to higher order, infinities appear which tells us that perturbation theory breaks down. Fortunately, these infinities can be renormalized away when using relativistic quantum mechanics.

Classical electrodynamics also gives infinities for many quantities, and there is no consistent solution for charged point particles obeying causality and classical mechanics. Classically, the self energy of an electron for a charge at rest

$$E_{\text{self}} = \frac{1}{8\pi} \int d^3r \rho(\boldsymbol{r}) \Phi(\boldsymbol{r}) \tag{31.1}$$

diverges as $\rho \to \delta^3(\boldsymbol{r})$. Lorentz, and others thought that a sperical charge density localized in a small volume would bound this energy, and if you assume that $E_{\text{self}} \sim mc^2$ then the radius of the charge distributions will be roughly $r_0 = mc^2/e^2$, which is called the classical electron radius for this reason.

You should realize that we cannot just say that an electron does not interact with itself at all. The reason is that we know that an accelerating electron radiates and it must lose energy and momentum to the electromagnetic field. For an isolated electron, the only field that can do work on the electron is its own field. Feynman and Wheeler developed one method that is consistent with classical physics by using both the retarded and advanced potentials (they replaced the self interaction with the advanced reaction to the motion of other charges, and assume that all radiation is eventually absorbed), but their theory violates causality in a small region of size r_0 around the electron.

In quantum mechanics, we still have divergences of the self energy, but if we use a relativistic Hamiltonian for the electron (i.e. a Dirac field), the divergence is rather weak. Weiskopf showed that if we cut off the sum over field oscillators at some energy Λ, so that $\hbar\omega_k < \Lambda$, the electromagnetic contribution to the electron mass is

$$\delta m = \frac{3m}{2\pi} \left(\frac{e^2}{\hbar c} \right) \ln \left(\frac{\Lambda}{mc^2} \right) + \text{Constant} \, . \tag{31.2}$$

If we calculate the value of Λ which gives $\delta m = m$, we find

$$\Lambda \sim mc^2 e^{\frac{2\pi}{3}\frac{\hbar c}{e^2}} \sim 10^{100} mc^2 \,. \tag{31.3}$$

Since the mass of the visible universe is estimated to be $10^{80} mc^2$, and the Planck mass where gravity and quantum mechanics are equally important is $10^{22} m$, we expect unknown physics to be important long before photon energies of this scale contribute. Perhaps a reasonable guess is that while this expression is divergent, the "real" correction to the mass of the electron from electromagnetism is order $\frac{e^2}{\hbar c} m$ or a few percent. A somewhat oversimplified idea of renormalization is that since we cannot turn off electromagnetism and measure the "bare" mass of the electron we instead cut off the theory at high energy to make the corrections finite (for example, take a finite Λ). We then adjust the bare mass of the electron such that the corrected mass for a free electron is the experimentally measured mass. We do the same sort of procedure for other diverging experimentally measured quantities (for example the coupling constants). We can then use this theory to calculate other quantities at higher order. If the measured quantities are independent of the cutoff as the cutoff goes to infinity, the theory is renormalizable which then means that our results are independent of the details of whatever new physics really supplies the cutoff. You can also have non-renormalizable theories – these indicate that the details of the physical cutoff will matter. Quantum electrodynamics is a renormalizable theory.

31.2 Bethe's calculation of the Lamb shift

The nonrelativistic hydrogen atom has all of the $n = 2$ levels degenerate. These have $\ell = 0$ and $\ell = 1$, with spin up or down, so there are 8, $n = 2$ levels. In the Dirac equation solution, the spin-orbit interaction mixes the spin and orbital angular momentum states, so that with $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, we have $|jm\ell s\rangle$ states. The $J = 3/2$, $\ell = 1$ states are split, but the $J = 1/2, \ell = 1$ states are still degenerate with the $J = 1/2$ $\ell = 0$ states. Lamb was able to drive transitions between these two states using microwaves with a frequency of about 1060 MHz. Indicating that these levels are not degenerate. This splitting is called the Lamb shift.

If we assume that the Dirac theory is correct, then this extra splitting must come from coupling the electromagnetic field to the atom. The problem was that this coupling as described above gave an infinite correction. Let's look at how this works for our nonrelativistic theory for a free electron. First notice that in Coulomb gauge, the static interaction would never change, so it cannot give any physical effects. The self interaction from the $\boldsymbol{A}$ coupling would give terms that are pictured diagramatically in fig. 31.1

The A^2 term creates and then immediately destroys a photon. Calculating the change in energy from first-order perturbation theory, we see that the plane-wave components in the vector potential cancel and there is no dependence on the electronic state. That is, with a cutoff to make this term finite, the energy of all electronic states is changed by the same constant value. Since adding a constant to the Hamiltonian has no physical consequences, we can drop this term.
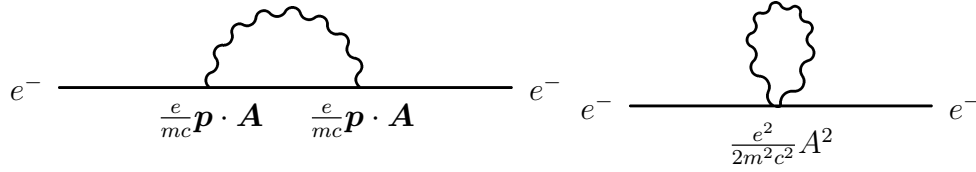
Figure 31.1: The lowest order self energy diagrams.

The term second order in $\boldsymbol{p} \cdot \boldsymbol{A}$, when acting on the electronic free-particle state and electromagnetic vacuum, $|\boldsymbol{q}0\rangle$ has one photon intermediate states gives the perturbed energy

$$\Delta E(q) = \frac{e^2}{m^2 c^2} \sum_{\boldsymbol{q}', \hat{\boldsymbol{\lambda}} \boldsymbol{k}} \frac{|\langle \boldsymbol{q}0|\boldsymbol{p} \cdot \boldsymbol{A}|\boldsymbol{q}'\hat{\boldsymbol{\lambda}}\boldsymbol{k}\rangle|^2}{\frac{\hbar^2 q^2}{2m} - \frac{\hbar^2 q'^2}{2m} - \hbar k c} \tag{31.4}$$

Only the destruction operator $a_{\boldsymbol{k}\lambda}$ in $\boldsymbol{A}$ gives a nonzero matrix element, $e^{i\boldsymbol{k}\cdot\boldsymbol{r}_{\mathrm{op}}}$ in $\boldsymbol{A}$ is the momentum translation operator which, along with the states, requires that momentum is conserved. The $\boldsymbol{p}$ operator gives gives $\hbar\boldsymbol{q}$ when operating on the state $|\boldsymbol{q}0\rangle$. The result is

$$\Delta E(q) = \frac{e^2 2\pi\hbar^3}{m^2} \int \frac{d^3 k}{(2\pi)^3} \sum_{\lambda} (\boldsymbol{q} \cdot \hat{\boldsymbol{\lambda}})^2 \frac{1}{kc} \frac{1}{\frac{\hbar^2 \boldsymbol{k}\cdot\boldsymbol{q}}{m} - \frac{\hbar^2 k^2}{2m} - \hbar k c} . \tag{31.5}$$

Since the electron velocity $\frac{\hbar k}{m}$ must be much smaller than c for the relativistic equation to be valid, the $\frac{\hbar^2 \boldsymbol{k}\cdot\boldsymbol{q}}{m}$ term must be much smaller than $\hbar k c$, and we drop it. We next calculate that $\frac{\hbar^2 k^2}{2m}$ will will be the same order as $\hbar k c$ when $\hbar k c \sim mc^2$, so we drop it as well and realize that a cut off above mc^2 would again require a relativistic theory.

Obviously with all of these approximations, we can only hope to get a rough idea of the magnitudes of the effects, not a precise value that can be compared in detail with experiments.

Dropping the electronic energy terms in the denominator we have

$$\begin{aligned}
\Delta E(q) &= -\frac{e^2 2\pi\hbar^3}{m^2} \int \frac{d^3 k}{(2\pi)^3} \sum_{\lambda} (\boldsymbol{q} \cdot \hat{\boldsymbol{\lambda}})^2 \frac{1}{kc} \frac{1}{\hbar k c} \\
&= -\frac{2e^2\hbar^2}{3\pi m^2 c^2} q^2 \int_0^\Lambda dk = -\frac{2e^2\hbar^2}{3\pi m^2 c^2} q^2 \Lambda = -\frac{\hbar^2 q^2}{2m} \cdot \frac{e^2}{\hbar c} \cdot \frac{4}{3\pi} \cdot \frac{\hbar c \Lambda}{mc^2} .
\end{aligned} \tag{31.6}$$

In the last expression, we can see that the effect of the electromagnetic coupling is to change the factor multiplying q^2 in the kinetic energy. This factor is what we experimentally call $\frac{\hbar^2}{2m}$, that is this is a way to measure the electron mass. When we couple to the electromagnetic field, the calculated mass is no longer the value of m that we used in the Hamiltonian. Rather the m in the Hamiltonian is the "bare" mass we would measure if we could somehow turn off the electromagnetic degrees of freedom and just have an electron by itself. The "experimental" mass is then changed at second order in the field coupling to the expression here. We get a factor of the fine structure constant, $\frac{e^2}{\hbar c} \sim \frac{1}{137}$, a factor of order one, and

the ratio of the photon energy at the cutoff to the electron rest energy. Of course, our nonrelativistic theory makes no sense if $\hbar c \Lambda$ is allowed to be bigger than mc^2, so we can hope that the "true" correction to the bare mass is of order the fine structure constant. This energy diverges linearly with the cutoff Λ.

At this order, with a reasonable cutoff, this result says that we should change our bare Hamiltonian. We should modify the bare mass so that after adding this electromagnetic contribution, we get the experimental mass.

Therefore we use the free particle Hamiltonian,

$$H_0 = \frac{p^2}{2m} + C(\Lambda)p^2 \tag{31.7}$$

where $C(\Lambda) = \frac{2e^2}{3\pi m^2 c^2}\Lambda$, and $C(\Lambda)p^2$ is called a counter term. m in the Hamiltonian is the experimental electron mass, and when we go through our calculation above with this Hamiltonian, we get the correct experimental momentum energy relation to first order in the fine structure constant, since the counter term cancels the second order energy to that order.

Having done this, we get no observable change in the free particle result. The question, of course, is what happens to a bound electron in a hydrogen atom, and, in particular, can radiative corrections explain the Lamb shift?

Chapter 32

Spin half particles

32.1 Introduction

Spin one-half is a special case of the angular momentum described previously in chapter 15. It simplifies some expressions to define $\boldsymbol{\sigma}$ as

$$\boldsymbol{S}_{\text{op}} = \frac{\hbar}{2}\boldsymbol{\sigma}_{\text{op}} = \frac{\hbar}{2}\left(\hat{\boldsymbol{x}}\sigma_{x\ \text{op}} + \hat{\boldsymbol{y}}\sigma_{y\ \text{op}} + \hat{\boldsymbol{z}}\sigma_{z\ \text{op}}\right) \tag{32.1}$$

For spin one-half, the angular momentum operator $\boldsymbol{S}_{\text{op}}$ in the eigenbasis of S_z and S^2 is given by the general expressions specialized to $L^2 = \hbar^2\frac{3}{4}$, we write

$$\begin{aligned}|s = \tfrac{1}{2}\ m_s = \tfrac{1}{2}\rangle &= |\uparrow\rangle \\ |s = \tfrac{1}{2}\ m_s = -\tfrac{1}{2}\rangle &= |\downarrow\rangle\end{aligned} \tag{32.2}$$

with the matrix represenations

$$\begin{aligned}
\sigma_x &= \begin{pmatrix} \langle\uparrow|\sigma_{x\ \text{op}}|\uparrow\rangle & \langle\uparrow|\sigma_{x\ \text{op}}|\downarrow\rangle \\ \langle\downarrow|\sigma_{x\ \text{op}}|\uparrow\rangle & \langle\downarrow|\sigma_{x\ \text{op}}|\downarrow\rangle \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\sigma_y &= \begin{pmatrix} \langle\uparrow|\sigma_{y\ \text{op}}|\uparrow\rangle & \langle\uparrow|\sigma_{y\ \text{op}}|\downarrow\rangle \\ \langle\downarrow|\sigma_{y\ \text{op}}|\uparrow\rangle & \langle\downarrow|\sigma_{y\ \text{op}}|\downarrow\rangle \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
\sigma_z &= \begin{pmatrix} \langle\uparrow|\sigma_{z\ \text{op}}|\uparrow\rangle & \langle\uparrow|\sigma_{z\ \text{op}}|\downarrow\rangle \\ \langle\downarrow|\sigma_{z\ \text{op}}|\uparrow\rangle & \langle\downarrow|\sigma_{z\ \text{op}}|\downarrow\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{aligned} \tag{32.3}$$

These are called the Pauli matrices in the standard representation. Since they are the matrix representations of Hermitian operators they are Hermitian. The eigenvalues of σ_z are obviously ± 1, and by rotational symmetry you would expect the eigenvalues of σ_x and σ_y to also be ± 1, which you can check by direct calculation in the standard representation if the general symmetry argument is not convincing.

Squaring a Hermitian matrix with eigenvalues ± 1 must give a Hermitian matrix with all eigenvalues 1, and this matrix therefore must be the identity. Therefore

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1. \tag{32.4}$$

Again, you can check this by direct multiplication of the matrices in the standard representation if you are not convinced.

The angular momentum commutation relations are

$$[S_\alpha, S_\beta] = i\hbar \sum_\gamma \epsilon_{\alpha\beta\gamma} S_\gamma \tag{32.5}$$

which become

$$[\sigma_\alpha, \sigma_\beta] = 2i \sum_\gamma \epsilon_{\alpha\beta\gamma} \sigma_\gamma . \tag{32.6}$$

Looking at the specific case for $\alpha = x$ and $\beta = y$, we have

$$\sigma_x \sigma_y - \sigma_y \sigma_x = 2i\sigma_z . \tag{32.7}$$

If we multiply this equation from the left or from the right with σ_x and use $\sigma_x^2 = 1$ we get the two equations

$$\sigma_y - \sigma_x \sigma_y \sigma_x = 2i\sigma_x \sigma_z$$
$$\sigma_x \sigma_y \sigma_x - \sigma_y = 2i\sigma_z \sigma_x . \tag{32.8}$$

Adding them gives

$$(\sigma_x \sigma_y + \sigma_y \sigma_x) = 0 \tag{32.9}$$

Repeating for the other pairs (or using cyclic permutations of the coordinates), and combining with $\sigma_\alpha^2 = 1$, we can get the anticommutation relations for the Pauli matrices

$$\{\sigma_\alpha, \sigma_\beta\} = \sigma_\alpha \sigma_\beta + \sigma_\beta \sigma_\alpha = 2\delta_{\alpha\beta} . \tag{32.10}$$

Again, these can be checked by direct multiplication in the standard representation if desired.

Adding the commutation and anticommutation relations, we have

$$\sigma_\alpha \sigma_\beta = \tfrac{1}{2}\{\sigma_\alpha, \sigma_\beta\} + \tfrac{1}{2}[\sigma_\alpha, \sigma_\beta] = \delta_{\alpha,\beta} + i \sum_\gamma \epsilon_{\alpha\beta\gamma} \sigma_\gamma \tag{32.11}$$

which shows that products of Pauli matrices can always be reduced to a linear combination of the Pauli matrices and the identity. This should not be too surprising since in the standard representation we can write, with I representing the identity matrix,

$$aI + \boldsymbol{b} \cdot \boldsymbol{\sigma} = \begin{pmatrix} a + b_z & b_x - ib_y \\ b_x + ib_y & a - b_z \end{pmatrix} \tag{32.12}$$

and for any complex 2×2 matrix, we can solve for the needed a and $\boldsymbol{b}$. This shows that we can write any 2×2 matrix as a linear combination of the Pauli matrices and the identity. Often it is useful to use the Pauli matrices for general two-state systems since we can simplify the algebra using the relationships above.

A particularly useful result is, for operators that commute with $\boldsymbol{\sigma}$ (and if they do not, they can be rewritten in terms of $\boldsymbol{\sigma}$ and operators that commute with $\boldsymbol{\sigma}$), we have

$$
\begin{aligned}
\boldsymbol{\sigma} \cdot \boldsymbol{A}\, \boldsymbol{\sigma} \cdot \boldsymbol{B} &= \sum_{\alpha\beta} A_\alpha B_\beta \sigma_\alpha \sigma_\beta = \sum_{\alpha\beta} A_\alpha B_\beta \left[\delta_{\alpha,\beta} + i \sum_\gamma \epsilon_{\alpha\beta\gamma} \sigma_\gamma \right] \\
&= \boldsymbol{A} \cdot \boldsymbol{B} + i \boldsymbol{\sigma} \cdot (\boldsymbol{A} \times \boldsymbol{B})
\end{aligned}
\tag{32.13}
$$

where, since we have kept the same order of $\boldsymbol{A}$ and $\boldsymbol{B}$ throughout, $\boldsymbol{A}$ and $\boldsymbol{B}$ do not need to commute with each other.

32.2 Spin one-half rotations

The operator that rotates a state about the unit vector $\hat{\boldsymbol{n}}$ by an angle θ is $e^{-\frac{i}{\hbar} \boldsymbol{L}_{\mathrm{op}} \cdot \hat{\boldsymbol{n}} \theta}$, where $\boldsymbol{L}$ is the angular momentum operator. For a spin one-half state, we can write

$$
e^{-\frac{i}{\hbar} \boldsymbol{S}_{\mathrm{op}} \cdot \hat{\boldsymbol{n}} \theta} = e^{-i \boldsymbol{\sigma}_{\mathrm{op}} \cdot \hat{\boldsymbol{n}} \frac{\theta}{2}} .
\tag{32.14}
$$

Working in the standard basis (although again all of the things we show are operator indentities true in any basis) we can Taylor series expand this

$$
e^{-i \boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}} \frac{\theta}{2}} = 1 - \left(i \boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}} \frac{\theta}{2} \right) + \frac{1}{2!} \left(i \boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}} \frac{\theta}{2} \right)^2 + \ldots
\tag{32.15}
$$

and using our identity Eq. 32.13 above, $(\boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}})^2 = 1$, and therefore all even powers of $\boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}}$ are unity. We have

$$
\begin{aligned}
e^{-i \boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}} \frac{\theta}{2}} &= \left[1 - \frac{1}{2!} \left(\frac{\theta}{2} \right)^2 + \frac{1}{4!} \left(\frac{\theta}{2} \right)^4 + \ldots \right] - i \boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}} \left[\frac{\theta}{2} + \frac{1}{3!} \left(\frac{\theta}{2} \right)^3 + \ldots \right] \\
&= \cos \left(\frac{\theta}{2} \right) - i \boldsymbol{\sigma} \cdot \hat{\boldsymbol{n}} \sin \left(\frac{\theta}{2} \right) .
\end{aligned}
\tag{32.16}
$$

Let's rotate the states $|\uparrow\rangle$ and $|\downarrow\rangle$ around y by $\theta = \frac{\pi}{2}$. The rotation matrix is then

$$
R = \frac{\sqrt{2}}{2} - i \frac{\sqrt{2}}{2} \sigma_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} .
\tag{32.17}
$$

and operating on the up and down vectors

$$
\begin{aligned}
R \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
R \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}
\end{aligned}
\tag{32.18}
$$

we get the eigenvectors of σ_x as we might have expected.

Of course we can also rotate operators in the usual way by realizing that rotating states and operators together will leave matrix elements invariant

$$\langle\psi_1|O|\psi_2\rangle = \langle\psi_1|R^+ROR^+R|\psi_2\rangle == \langle\psi_1^R|ROR^+|\psi_2^R\rangle \tag{32.19}$$

so that the rotated operator is

$$O^R = ROR^+ . \tag{32.20}$$

For example, rotating σ_z with our 90° rotation around y gives

$$
\begin{aligned}
R\sigma_z R^+ &= \frac{1}{\sqrt{2}}(1 - i\sigma_y)\sigma_z \frac{1}{\sqrt{2}}(1 + i\sigma_y) \\
&= \frac{1}{2}(\sigma_z - i\sigma_y\sigma_z)(1 + i\sigma_y) \\
&= \frac{1}{2}(\sigma_z + \sigma_x)(1 + i\sigma_y) \\
&= \frac{1}{2}(\sigma_z + \sigma_x + i\sigma_z\sigma_y + i\sigma_x\sigma_y) \\
&= \frac{1}{2}(\sigma_z + \sigma_x - \sigma_x + \sigma_z) = \sigma_z \tag{32.21}
\end{aligned}
$$

again as we should expect.

Spin one-half has the special property that there is a direction $\hat{\boldsymbol{n}}$ such that

$$\boldsymbol{S}\cdot\hat{\boldsymbol{n}}|\psi\rangle = \frac{\hbar}{2}|\psi\rangle \tag{32.22}$$

for any $|\psi\rangle$. You can see this because rotating the state $|\uparrow\rangle$ around z by an angle γ will give it an arbitrary phase. Then rotating it around y by an angle θ and around z again by ϕ will point the quantization axis along $\hat{\boldsymbol{n}} = \sin\theta(\cos\phi\hat{\boldsymbol{x}} + \sin\phi\hat{\boldsymbol{y}}) + \cos\theta\hat{\boldsymbol{z}}$. The rotation around y gives any relative magnitudes for the upper and lower components by choosing the angle. The rotation around around z by ϕ gives any relative phase for the components, and the first rotation around z by gamma gives any overall phase. We can therefore specify the spin state up to an overall phase by giving its quantization axis. Often we call this the "direction" of the spin.

There is one oddity about spin one-half rotations. If we rotate by 2π, we have

$$e^{-i\boldsymbol{\sigma}\cdot\hat{\boldsymbol{n}}\pi} = \cos\pi - i\sin(\pi)\boldsymbol{\sigma}\cdot\hat{\boldsymbol{n}} = -1 . \tag{32.23}$$

Spin rotations are double valued. If you rotate by 2π you get -1 times the original state, not the original state. Since the spin is attached to a particle, this gives no physical effects, however, there are some consequences – for example, the time reversal operator squared is not 1 if there are an odd number of spins.

32.3 Spin effects and interactions

Spin has two physical consequences. The first is that physical states are symmetric or antisymmetric under interchange of identical particles. Even if the spins do not interact, the symmetry of under interchange will dictate the kinds of states that are allowed. For example, if we have two identical fermions, interchanging them interchanges their positions and spins. If they have the same spin, their state must be antisymmetric under spatial interchange alone, since interchanging their spins will not change anything. However, if they have opposite spins, their state will be required to have any particular spatial symmetry. The lowest energy state can then have different energies for different spin configurations even if the spins themselves don't interact. The second consequence is that spins can correspond, for example, to magnetic moments and give directly interacting terms in the Hamiltonian.

Perhaps the most straightforward method for obtaining the consistent spin interactions for our nonrelativistic Hamiltonian is to start with the relativistic Dirac equation and take its nonrelativistic limit. Before doing that, we will write down the lowest order Pauli equation that has some of this physics. For an electromagnetic field and a spin one-half particle of mass m and charge q, the Pauli equation Hamiltonian is

$$H = \frac{\left\{ \boldsymbol{\sigma} \cdot [\boldsymbol{p} - \frac{q}{c} \boldsymbol{A}(\boldsymbol{r}, t)] \right\}^2}{2m} + q\Phi(\boldsymbol{r}, t) + ... \tag{32.24}$$

where the terms correct to order m^{-1} are shown. We can expand

$$\boldsymbol{\sigma} \cdot [\boldsymbol{p} - \frac{q}{c} \boldsymbol{A}(\boldsymbol{r}, t)] \boldsymbol{\sigma} \cdot [\boldsymbol{p} - \frac{q}{c} \boldsymbol{A}(\boldsymbol{r}, t)] = \left| \boldsymbol{p} - \frac{q}{c} \boldsymbol{A}(\boldsymbol{r}, t) \right|^2 - \frac{iq}{c} \boldsymbol{\sigma} \cdot (\boldsymbol{p} \times \boldsymbol{A}(\boldsymbol{r}, t) + \boldsymbol{A}(\boldsymbol{r}, t) \times \boldsymbol{p}) \tag{32.25}$$

We can write in components, using the summation convention

$$\begin{aligned}
(\boldsymbol{p} \times \boldsymbol{A}(\boldsymbol{r}, t) + \boldsymbol{A}(\boldsymbol{r}, t) \times \boldsymbol{p})_\gamma &= p_\alpha A_\beta \epsilon_{\alpha\beta\gamma} + A_\alpha p_\beta \epsilon_{\alpha\beta\gamma} \\
&= p_\alpha A_\beta \epsilon_{\alpha\beta\gamma} - A_\beta p_\alpha \epsilon_{\alpha\beta\gamma} \\
&= [p_\alpha, A_\beta] \epsilon_{\alpha\beta\gamma} \\
&= \frac{\hbar}{i} (\partial_\alpha A_\beta) \epsilon_{\alpha\beta\gamma} \tag{32.26}
\end{aligned}$$

so that

$$\boldsymbol{p} \times \boldsymbol{A}(\boldsymbol{r}, t) + \boldsymbol{A}(\boldsymbol{r}, t) \times \boldsymbol{p} = \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r}, t) \equiv \boldsymbol{B}(\boldsymbol{r}, t) \tag{32.27}$$

and the Pauli equation to this order becomes

$$H = \frac{[\boldsymbol{p} - \frac{q}{c} \boldsymbol{A}(\boldsymbol{r}, t)]^2}{2m} - \frac{q\hbar}{2mc} \boldsymbol{\sigma} \cdot \boldsymbol{B}(\boldsymbol{r}, t) + q\Phi(\boldsymbol{r}, t) \tag{32.28}$$

which shows that the Pauli equation gives a spin interaction with the magnetic field.

To compare with the orbital interaction, we take a static magnetic field and choose the gauge

$$\boldsymbol{A}(\boldsymbol{r}) = -\tfrac{1}{2} \boldsymbol{r} \times \boldsymbol{B} \tag{32.29}$$

which you can check gives the static field $\boldsymbol{B}$. Since in our gauge $\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$, we can write

$$\frac{1}{2m}[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}]^2 = \frac{p^2}{2m} + \frac{q^2 A^2}{2mc^2} - \frac{q}{mc}\boldsymbol{p} \cdot \boldsymbol{A} \tag{32.30}$$

this last term is

$$-\frac{q}{mc}\boldsymbol{p} \cdot \boldsymbol{A} = -\frac{q}{2mc}\boldsymbol{p} \cdot (\boldsymbol{B} \times \boldsymbol{r}) = -\frac{q}{2mc}\boldsymbol{B} \cdot (\boldsymbol{r} \times \boldsymbol{p}) = -\frac{q}{2mc}\boldsymbol{L} \cdot \boldsymbol{B} \,. \tag{32.31}$$

So in a constant magnetic field we have

$$H = \frac{p^2}{2m} + \frac{q^2}{8mc^2}|\boldsymbol{r} \times \boldsymbol{B}|^2 - \frac{q}{2mc}\left[\boldsymbol{L} + 2\boldsymbol{S}\right] \cdot \boldsymbol{B} \tag{32.32}$$

where $\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\sigma}$. Taking the z axis along $\boldsymbol{B}$, the last term is

$$-\frac{q}{2mc}\left[L_z + 2S_z\right]|B| = -\frac{q\hbar}{2mc}\left[m_L + 2m_S\right]|B| \tag{32.33}$$

and for electrons the factor $-\frac{q\hbar}{2mc}$ is $\frac{e\hbar}{2mc}$ which is called the Bohr magneton. For magnetic fields larger than the neglected terms in the Pauli equation, this term predicts the Zeeman splittings of atoms.

The factor of 2 in front of $\boldsymbol{S}$ is called the Landé g factor, or gyromagnetic ratio. g is almost 2 for electrons. When we couple to the full quantized electromagnetic field there are small corrections and Schwinger's result that agrees with experiment is[1]

$$g = 2\left[1 + \frac{e^2}{\hbar c}\frac{1}{2\pi} + ...\right] . \tag{32.34}$$

For composite spin one-half particles the magnetic moment can be substantially different. For example, the proton has a moment that is about 2.8 times larger than the Dirac or Pauli theories would predict for a point particle, while the neutron has a moment that is about -2/3 of the proton. These can be roughly predicted by simple quark models, where the quarks are Dirac particles.

The biggest of the neglected terms that are order m^{-2}, is the spin-orbit interaction. Again, this drops out immediately when we take the nonrelativistic limit of the Dirac equation. Einstein gave a quick and dirty (and off by a factor of 2) argument for the spin-orbit interaction in a Bohr orbit. Since in the Bohr theory the electron orbits the proton, Einstein imagined looking in the rest frame of the electron. In this case you would see the proton circling the electron, and this circling charge would produce a magnetic field

$$\boldsymbol{B} = \frac{e\boldsymbol{v} \times \hat{\boldsymbol{r}}}{cr^2} = e\frac{\boldsymbol{v}}{c} \times \boldsymbol{E} \tag{32.35}$$

where $\boldsymbol{E}$ is the electric field from the proton. This is also the result you would get from Lorentz transforming an electric field $\boldsymbol{E}$ to a frame moving at velocity $\boldsymbol{v}$ to lowest order in

---

[1]I have been told that $\frac{\alpha}{2\pi}$ is engraved on Schwinger's tombstone.

v/c. Of course, the problem here is that the elecron is not in an inertial frame since it is accelerating. However, if we plug this magnetic field into our Pauli equation result anyway we get

$$\Delta H_{\mathrm{wrong}} = \frac{q}{mc} \boldsymbol{S} \cdot \left(\frac{\boldsymbol{v}}{c} \times \boldsymbol{E} \right) . \tag{32.36}$$

As noted, this wrong, but the correct result differs only by a factor of 2. Classically, this comes from correctly going into the rotating frame. The correct result (along with some other terms) that drops out of the Dirac equation for the hydrogen atom is

$$\frac{1}{2mc^2} \boldsymbol{S} \cdot (\boldsymbol{r} \times \boldsymbol{v}) \frac{1}{r} \frac{\partial V}{\partial r} = \frac{1}{2m^2 c^2} \frac{1}{r} \frac{\partial V}{\partial r} \boldsymbol{L} \cdot \boldsymbol{S} . \tag{32.37}$$

As we will see when we discuss angular momentum addition, the simplest way to deal with terms like this is to use states with good J^2, J_z, L^2, and S^2, with $\boldsymbol{J} = \boldsymbol{L} \cdot \boldsymbol{S}$. Since $J^2 = L^2 + S^2 + 2\boldsymbol{L} \cdot \boldsymbol{S}$, these states are eigenstates of $\boldsymbol{L} \cdot \boldsymbol{S}$.

Chapter 33

Quantum teleportation

Quantum teleportation is essentially the ability to transfer a quantum state from one place to another by a set of measurements on different parts of the system. It is analogous to sending a classical fax with a fax machine that shreds your original document.

For simplicity we will work with spin-1/2 systems. Each spin has the usual $|\uparrow\rangle$ and $|\downarrow\rangle$ states quantized along $\hat{z}$. The usual Pauli matrices give our spin operators

$$\frac{\hbar}{2}\boldsymbol{\sigma} = \boldsymbol{S} \tag{33.1}$$

and in the standard representation,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & \text{-i} \\ \text{i} & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & \text{-1} \end{pmatrix} \tag{33.2}$$

You should verify that for 2 spins, that $\sigma_{1x}\sigma_{2x}$ and $\sigma_{1z}\sigma_{2z}$ commute, so we can find simultaneous eigenstates of these. Squaring these Hermitian operators gives 1, so their eigenvalues are ± 1. Looking at the Pauli matrices tells us that $\sigma_{1z}\sigma_{2z}$ is 1 for like spin states and -1 for unlike spins, while $\sigma_{1x}\sigma_{2x}$ flips both spins. Either by direct diagonalization or from the above, We can construct the simultaneous eigenstates

$$\frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right)$$

$$\frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\rangle - |\downarrow\downarrow\rangle\right)$$

$$\frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right)$$

$$\frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) \tag{33.3}$$

with eigenvalues of $\sigma_{1z}\sigma_{2z}$ and $\sigma_{1x}\sigma_{2x}$, 1,1, 1,-1, -1,1, -1,-1, respectively. These are the so called "entangled" states.

Let's assume we have prepared an "entangled" state of two particles which we will call particles 2 and 3, with eigenvalues 1,1, sometime in the past. Our partner carries particle 3 to the destination and awaits further instructions. A customer arrives with a particle which we call particle 1 in an unknown to us spin state,

$$\alpha|\uparrow\rangle + \beta|\downarrow\rangle. \tag{33.4}$$

Of course $|\alpha|^2 + |\beta|^2 = 1$, and only the relative magnitude and phase of α and β are important. The customer wants that state to be at the destination. Now if the customer tells us what α and β are, we can just tell our partner to prepare a quantum state, but if the customer doesn't know α and β, and can't prepare more states like this for us to measure, there is no way for us to determine them, since for example measuring the z component of the spin will give us a single value of $\pm\hbar/2$, and thereafter the state will be either up or down. However, we don't need to know α and β to make particle 3 have this same wave function. Of course in producing particle 3 in the

$$\alpha|\uparrow\rangle + \beta|\downarrow\rangle \tag{33.5}$$

wave function, we will no longer have particle 1 in that wave function. This is what is meant by quantum teleportation.

To accomplish the teleportation, we measure $\sigma_{1z}\sigma_{2z}$ and $\sigma_{1x}\sigma_{2x}$. That is we make these measurements between the customer's particle 1 and our half of the entangled pair, particle 2. We know that this will pick out the part of the total wave function with the eigenvectors of the measured eigenvalues of $\sigma_{1z}\sigma_{2z}$ and $\sigma_{1x}\sigma_{2x}$.

Expanding the full wave function (written in the form particle 1, particle 2, particle 3),

$$(\alpha|\uparrow\rangle + \beta|\downarrow\rangle)\left[\frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)\right]$$

$$= \frac{1}{2}\left[\frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)\right](\alpha|\uparrow\rangle + \beta|\downarrow\rangle)$$

$$+ \frac{1}{2}\left[\frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle - |\downarrow\downarrow\rangle)\right](\alpha|\uparrow\rangle - \beta|\downarrow\rangle)$$

$$+ \frac{1}{2}\left[\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)\right](\beta|\uparrow\rangle + \alpha|\downarrow\rangle)$$

$$+ \frac{1}{2}\left[\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)\right](-\beta|\uparrow\rangle + \alpha|\downarrow\rangle) \tag{33.6}$$

We make the measurement, and with with probability 1/4, we will get one of the 4 terms above. Since we know which of the terms we measure, we now know how the state of particle 3 is related to the original state of particle 1. For example if we measure 1,1, particle 3 is in the same state as particle 1 was originally, and we have teleported the state. If we measure -1,1, we get the second term. In this case the down state of particle 3 has the wrong sign. We can correct this by a unitary transformation. For example we can rotate the state by an angle of 180 degrees around the z axis. If we measure 1,-1, we can get the original state

by rotating by 180 degrees around the x axis. Finally if we measure -1,-1, we can rotate particle 3 around x and around z. One way to accomplish these rotations would be to apply a magnetic field for a prescribed length of time in the appropriate directions.

Chapter 34

Dirac Equation

34.1   Many roads to the Dirac equation

The Dirac equation is the wave equation that is consistent with quantum mechanics, spin one-half, and special relativity. As you might expect, you can arrive at the Dirac equation by starting with nonrelativistic spin one-half quantum theory and trying to make it consistent with relativity. This is the route that Baym takes. You can also try to derive the general form of a consistent quantum field theory invariant under Lorentz transformations.

34.2   Dirac's argument

Perhaps the fastest way is to follow Dirac. The Klein-Gordon equation had apparently been invented and abandoned by Schrödinger. We know now that it is perfectly good wave equation for spin-zero bosons. Not surprsingly, it does not give the correct fine structure for the electronic states of hydrogen. In addition, it is second-order in time, and, when interpreted like the Schrödinger equation, a complete set of solutions requires both positive and negative energy eigenstates. We have already quantized the electromagnetic field without any problem with negative energy states. That is, we worked out the appropriate way to do quantum mechanics for relativistic bosons, and were lead directly to writing a quantum field. For photons, the "negative energies" corresponded to the $e^{i\omega t}$ terms, which were always associated with the creation operators. Exactly the same thing happens for all relativistic bosons, so, in fact, the second-order in time derivatives, wave equation, is not the problem. However, Dirac did not know this and decided to look for a relativistic wave equation that is first order in the time derivative. Since Lorentz invariance treats space and time together, to be Lorentz invariant, it would have to be first order in the spatial derivatives too. Dirac therefore wrote a Schrödinger equation

$$H|\psi(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi(t)\rangle \tag{34.1}$$

with the Hamiltonian

$$H = \boldsymbol{\alpha} \cdot \boldsymbol{p}c + \beta mc^2 \, . \tag{34.2}$$

Realize that $\boldsymbol{\alpha}$ cannot be a normal vector since it would pick out a prefered direction in space. A little reflection will tell you that $\boldsymbol{\alpha}$ and therefore β will have to be related to the electron spin operators if this is to work out. We therefore assume that they can be represented by matrices just as our $\boldsymbol{\sigma}$ matrices represent the spin one-half operators. Since the Hamiltonian is the generator of time translations, and these must be represented by unitary transformations, the Hamiltonian must be Hermitian. Therefore, the matrices $\boldsymbol{\alpha}$ and β must be Hermitian.

In order to describe a particle of mass m, the relativistic relationship

$$p^2c^2 + m^2c^4 = E^2 \tag{34.3}$$

should still hold. That is, if we have a state with one particle at rest, its energy will be mc^2 and it will have zero momentum. Lorentz transforming this to a frame where it has momentum $\boldsymbol{p}$ requires that that energy be given by Eq. 34.3. Therefore, if we solve the Schrödinger equation above in the usual way, we would have energy eigenstates of H

$$H|\psi_n\rangle = E_n|\psi_n\rangle \tag{34.4}$$

so that

$$H^2|\psi_n\rangle = E_n^2|\psi_n\rangle = (p^2c^2 + m^2c^4)|\psi_n\rangle \tag{34.5}$$

Expanding H^2 we get

$$
\begin{aligned}
H^2 &= \left(\boldsymbol{\alpha}\cdot\boldsymbol{p}c + \beta mc^2\right)\left(\boldsymbol{\alpha}\cdot\boldsymbol{p}c + \beta mc^2\right) \\
&= \alpha_i p_i \alpha_j p_j c^2 + \alpha_i \beta m p_i c^3 + \beta\alpha_j m p_j c^3 + \beta^2 m^2 c^4
\end{aligned} \tag{34.6}
$$

where i and j indicate the x, y, and z components. In order for this to equal $p^2c^2 + m^2c^4$, we must have $\alpha_i^2 = \beta^2 = 1$ and all the anticommutators between different α_i and β zero,

$$
\begin{aligned}
\{\alpha_i, \alpha_j\} &= 2\delta_{ij} \\
\{\alpha_i, \beta\} &= 0\,.
\end{aligned} \tag{34.7}
$$

Let's try to find the properties of such Hermitian matrices. If we find such matrices, we could take one of them, and transform into its eigenbasis. The eigenvalues are real, and squaring the matrix must give the identity matrix. Therefore all of the eigenvalues of the α_i and β matrices must be either 1 or -1.

The anticommutator between two of these matrices shows that, for example, $\alpha_x\beta = -\beta\alpha_x$. Multiplying on the left by α_x and using $\alpha_x^2 = 1$, this becomes $\beta = -\alpha_x\beta\alpha_x$. We can now take the trace, and using the cyclic property of the trace we have $\text{tr}\alpha_x\beta\alpha_x = \text{tr}\alpha_x^2\beta = \text{tr}\beta$. So that $\text{tr}\beta = 0$. Repeating, we find that the traces of all of the α_i and β matrices are zero. Since the trace is the sum of the eigenvalues, these matrices must have even rank, i.e. they must be 2×2, or 4×4, etc.

We can eliminate the possibility of the 2×2 case for three spatial dimensions immediately. We know that a general 2×2 Hermitian matrix can be written as a real linear combination

of the three Pauli matrices and the identity. The three Pauli matrices anticommute with each other and their squares are the identity, but the identity does not have trace 0, so there is no fourth matrix available. We cannot construct the necessary four 2×2 matrices. This shows that α_i, and β must be at least 4×4 matrices. We will see that the 4×4 case describes massive relativistic spin-half particles.

Before continuing, let's think about the momentum eigenstates. Since we construct a translationally invariant Hamiltonian, it commutes with the momentum and momentum eigenstates will be energy eigenstates. The zero momentum eigenstates will have $\boldsymbol{p} = 0$, and the Hamiltonian in this subset of states is then just $\beta m c^2$. Diagonalizing this Hamiltonian, we will have a diagonal β. We know that β has two $+1$ eigenvalues and two -1 eigenvalues. Its standard form is

$$
\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} , \tag{34.8}
$$

where in the second form, I and 0 represent the 2×2 identity and zero matrices. The first thing you might notice is that while we have two eigenstates with the rest energies mc^2, we also have two with negative energies. The negative energy "problem" is still here, however, the two states with energies mc^2 look like they might describe the two spin states.[1]

Having chosen the diagonal form for β, let's find the possible forms for the α_i. Taking a general 4×4 Hermitian matrix

$$
\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{12}^* & a_{22} & a_{23} & a_{24} \\ a_{13}^* & a_{23}^* & a_{33} & a_{24} \\ a_{14}^* & a_{24}^* & a_{34}^* & a_{44} \end{pmatrix} \tag{34.9}
$$

and setting its anticommutator with β to zero shows it must have the form

$$
\begin{pmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ a_{13}^* & a_{23}^* & 0 & 0 \\ a_{14}^* & a_{24}^* & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & A \\ A^+ & 0 \end{pmatrix} \tag{34.10}
$$

where again in the last expression A represents a 2×2 matrix and A^+ is its Hermitian conjugate. The squares of the α_i matrices must be the identity, so $AA^+ = 1$, so the A matrices must be unitary. The anticommutator of two such matrices shows that we require $\{A_i, A_j\} = 2\delta_{ij}$. We already know a set of three 2×2 matrices which satisfy these relationships: the

---

[1] As in the boson case, the negative energy "problem" will go away if we realize that we have to view the relativistic theory as a many-particle problem. Dirac's original solution, which is mathematicaly equivalent to later more elegant techniques is to assume that the vacuum state has all of the negative energy states filled with fermions. An excitation of one of these negative energy fermions to a positive energy state gives a physical fermion and a hole in the Dirac "sea." This hole is the antiparticle of the fermion.

Pauli matrices. We can therefore choose the α_i to be

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}. \tag{34.11}$$

Other choices correspond to other choices for the combinations of the degenerate eigenvectors of β and their phases. The standard form is then

$$\alpha_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\alpha_y = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$

$$\alpha_z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \tag{34.12}$$

The four coupled first-order differential equations that these give, collectively, are called the Dirac equation.

34.3   Nonrelativistic limit

For the states we are interested in, in the nonrelativistic limit, the energy eigenvalues can be written as $E = \epsilon + mc^2$, with $\epsilon \ll mc^2$. Let's write this as the equation for a particle of charge q coupled to static electromagnetic fields represented by the static potentials $\Phi(\boldsymbol{r})$ and $\boldsymbol{A}(\boldsymbol{r})$. In the usual way, we make the substitution $H \to H + q\Phi(\boldsymbol{r})$, and $\boldsymbol{p} \to \boldsymbol{p} - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r})$, and the equations become

$$\begin{pmatrix} mc^2 + q\Phi(\boldsymbol{r}) & \boldsymbol{\sigma} \cdot [\boldsymbol{p}c - q\boldsymbol{A}(\boldsymbol{r})] \\ \boldsymbol{\sigma} \cdot [\boldsymbol{p}c - q\boldsymbol{A}(\boldsymbol{r})] & -mc^2 + q\Phi(\boldsymbol{r}) \end{pmatrix} \begin{pmatrix} |\psi_u\rangle \\ |\psi_l\rangle \end{pmatrix} = (\epsilon + mc^2) \begin{pmatrix} |\psi_u\rangle \\ |\psi_l\rangle \end{pmatrix}. \tag{34.13}$$

where $|\psi_u\rangle$ and $|\psi_l\rangle$ are called the "upper" and "lower" components and are each two component spinors. This becomes the two equations

$$\begin{aligned} \boldsymbol{\sigma} \cdot [\boldsymbol{p}c - q\boldsymbol{A}(\boldsymbol{r})] |\psi_l\rangle &= [\epsilon - q\Phi(\boldsymbol{r})] |\psi_u\rangle \\ \boldsymbol{\sigma} \cdot [\boldsymbol{p}c - q\boldsymbol{A}(\boldsymbol{r})] |\psi_u\rangle &= [(2mc^2 + \epsilon) - q\Phi(\boldsymbol{r})] |\psi_l\rangle \end{aligned} \tag{34.14}$$

In the nonrelativistic limit, we can drop $\epsilon - q\Phi(\boldsymbol{r})$ in comparison to $2mc^2$ in the second term. (Notice that $2mc^2$ is the minimum energy to create an electron-positron pair.) We can then

substitute the resulting second equation into the first, eliminating the lower components, to give

$$\left\{ \frac{\left(\boldsymbol{\sigma}\cdot\left[\boldsymbol{p}-\frac{q}{c}\boldsymbol{A}(\boldsymbol{r})\right]\right)^2}{2m} + q\Phi(\boldsymbol{r})\right\}|\psi_u\rangle = \left[\frac{\left(p-\frac{q}{c}\boldsymbol{A}(\boldsymbol{r})\right)^2}{2m} - \frac{q\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{B}(\boldsymbol{r}) + q\Phi(\boldsymbol{r})\right]|\psi_u\rangle = \epsilon|\psi_u\rangle,$$

(34.15)

where $B(\boldsymbol{r}) = \boldsymbol{\nabla}\times\boldsymbol{A}(\boldsymbol{r})$, which is the nonrelativistic time independent Schrödinger equation coupled to static potentials with the correct spin-magnetic field interaction. We used the Pauli matrix identity $\boldsymbol{\sigma}\cdot\boldsymbol{A}\boldsymbol{\sigma}\cdot\boldsymbol{B} = \boldsymbol{A}\cdot\boldsymbol{B} + i\boldsymbol{\sigma}\cdot(\boldsymbol{A}\times\boldsymbol{B})$ to simplify the spin matrices.

34.4   Foldy-Wouthuysen Transformation

This gets much more interesting if we couple the particle to a time dependent electromagnetic field and take the next order terms which includes the spin-orbit coupling. While the same substitution and expansion can be used as above, a more systematic method is the Foldy-Wouthuysen transformation[2]. We will follow their notation. Their idea is to try to find a canonical transformation of the Dirac Hamiltonian that has no operators like $\boldsymbol{\alpha}$ that couple upper and lower components. Foldy and Wouthuysen call operators that couple only upper to lower and lower to upper odd operators, while operators that couple upper to upper or lower to lower are called even operators. If such a transformation can be found, the Hamiltonian will be block diagonal in the upper and lower components, and the upper components will correspond to the positive energy solutions i.e. the sector where for weak potentials, no positrons are created.

One advantage of the Foldy-Wouthuysen method is that it also makes clear some conceptual difficulties with our method above. For example, for the free Dirac Hamiltonian, if we measure the velocity along x using the Heisenberg representation, we would use the Heisenberg velocity operator $\frac{dx}{dt}$ which is

$$\frac{dx}{dt} = \frac{i}{\hbar}[H_{\mathrm{Dirac}}, x] = \frac{i}{\hbar}[\boldsymbol{\alpha}\cdot\boldsymbol{p}c + \beta mc^2, x] = \alpha_x c \tag{34.16}$$

and has eigenvalues $\pm c$. However, if we measure the nonrelativistic velocity we have

$$\frac{dx}{dt} = \frac{i}{\hbar}[H_{\mathrm{nr}}, x] = \frac{i}{\hbar}\left[\frac{p^2}{2m}, x\right] = \frac{p}{m} \tag{34.17}$$

which has continuous eigenvalues running from $-\infty$ to ∞. How did taking the nonrelativistic limit completely change the character of the velocity operator? The Foldy-Wouthuysen method tells us this. We eliminate the coupling between the upper and lower components with a canonical transformation and all other operators must be transformed as well.

To find the nonrelativistic limit, we will perform a series of canonical transformations to eliminate the odd operators in a power series in m^{-1}. The nonrelativistic limit consists of

---

[2]L.L. Foldy and S.A. Wouthuysen, *On the Dirac theory of spin 1/2 particles and its non-relativistic limit*, Phys. Rev. **78**, 29 (1950).

neglecting higher order terms in this expansion, so the Foldy-Wouthuysen transformations give the nonrelativistic result, and a systematic way to obtain the low order relativistic corrections.

The Dirac Hamiltonian can always be put into the form

$$H = \beta mc^2 + \mathcal{E} + \mathcal{O} \tag{34.18}$$

where $\mathcal{E}$ is an even operator and $\mathcal{O}$ is an odd operator in the Foldy-Wouthuysen sense. For all physical cases $\mathcal{E}$ and $\mathcal{O}$ will be order m^{-n} with $n \geq 0$. Since $\mathcal{E}$ operators only couple upper to upper and lower to lower, $[\beta, \mathcal{E}] = 0$, while since $\mathcal{O}$ operators only couple upper to lower and lower to upper, $\{\beta, \mathcal{O}\} = 0$. That is

$$\begin{aligned} \beta\mathcal{E} &= \mathcal{E}\beta \\ \beta\mathcal{O} &= -\mathcal{O}\beta \end{aligned} \tag{34.19}$$

We transform the Schrödinger equation

$$H|\psi(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi(t)\rangle \tag{34.20}$$

with a unitary transformation $e^{\frac{i}{\hbar}S}$ where S is Hermitian and can be time dependent

$$\begin{aligned} e^{\frac{i}{\hbar}S}He^{-\frac{i}{\hbar}S}e^{\frac{i}{\hbar}S}|\psi(t)\rangle &= -\frac{\hbar}{i}e^{\frac{i}{\hbar}S}\frac{\partial}{\partial t}|\psi(t)\rangle \\ &= -\frac{\hbar}{i}\frac{\partial}{\partial t}e^{\frac{i}{\hbar}S}|\psi(t)\rangle + \frac{\hbar}{i}\left(\frac{\partial}{\partial t}e^{\frac{i}{\hbar}S}\right)e^{-\frac{i}{\hbar}S}e^{\frac{i}{\hbar}S}|\psi(t)\rangle \end{aligned} \tag{34.21}$$

where in the last term the time derivative only acts on $e^{\frac{i}{\hbar}S}$. Taking the transformed state to be $|\psi'(t)\rangle = e^{\frac{i}{\hbar}S}|\psi(t)\rangle$, The Schrödinger equation becomes

$$H'|\psi'(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi'(t)\rangle \tag{34.22}$$

with the transformed Hamiltonian

$$H' = e^{\frac{i}{\hbar}S}He^{-\frac{i}{\hbar}S} - \frac{\hbar}{i}\left(\frac{\partial}{\partial t}e^{\frac{i}{\hbar}S}\right)e^{-\frac{i}{\hbar}S} = e^{\frac{i}{\hbar}S}He^{-\frac{i}{\hbar}S} + \frac{\hbar}{i}e^{\frac{i}{\hbar}S}\left(\frac{\partial}{\partial t}e^{-\frac{i}{\hbar}S}\right). \tag{34.23}$$

Since we will be expanding in powers of m^{-1}, and we will see that this corresponds to expanding in powers of S, we can expand the transformed Hamiltonian as

$$H' = H + \frac{\partial S}{\partial t} + \frac{i}{\hbar}\left[S, H + \frac{1}{2}\frac{\partial S}{\partial t}\right] + \left(\frac{i}{\hbar}\right)^2\frac{1}{2!}\left[S, \left[S, H + \frac{1}{3}\frac{\partial S}{\partial t}\right]\right] + \dots \tag{34.24}$$

To develop our series, we want an S that, when commuting with the βmc^2 term of the Hamiltonian, cancels the $\mathcal{O}$ term. That is, we want

$$\frac{i}{\hbar}[S, \beta mc^2] = -\mathcal{O} \tag{34.25}$$

382

From Eq. 34.19, even operators commute with β so they will not eliminate the odd operators – they only mix the even operators. We choose S odd and therefore it anticommutes with β,

$$\frac{i}{\hbar}[S, \beta mc^2] = \frac{imc^2}{\hbar}(S\beta - \beta S) = -\frac{i2mc^2}{\hbar}\beta S = -\mathcal{O}. \tag{34.26}$$

Multiplying through by β using $\beta^2 = 1$, gives

$$S = \frac{\hbar}{i2mc^2}\beta\mathcal{O}. \tag{34.27}$$

Substituting this form, and using the relations of Eq. 34.19 the transformed Hamiltonian correct to order m^{-2} is

$$
\begin{aligned}
H' &= \beta mc^2 + \mathcal{E} + \frac{\beta}{2mc^2}\mathcal{O}^2 - \frac{1}{8m^2c^4}\left[\mathcal{O}, [\mathcal{O}, \mathcal{E}] + \hbar\frac{\partial \mathcal{O}}{\partial t}\right] \\
&\quad -\frac{\hbar}{2mc^2}\beta\frac{\partial \mathcal{O}}{\partial t} + \frac{\beta}{2mc^2}[\mathcal{O}, \mathcal{E}] - \frac{1}{3m^2c^4}\mathcal{O}^3 + ...
\end{aligned} \tag{34.28}
$$

The first line of Eq. 34.28 contains only even operators. The second contains odd operators, but now at order m^{-1}. If we call these a new $\mathcal{O}$, we can repeat the transformation to eliminate them. By continuing this, we can eliminate the odd operators to any desired order m^{-n}.

34.4.1 Free particles

Before working out the nonrelativistic limit for an electron interacting with the electromagnetic field, let's look at the transformation at lowest order for the free particle Hamiltonian so that we can answer the question raised previously of how to interpret the velocity. With $H = \boldsymbol{\alpha} \cdot \boldsymbol{p}c + \beta mc^2$, we have $\mathcal{E} = 0$, and $\mathcal{O} = \boldsymbol{\alpha} \cdot \boldsymbol{p}c$, the lowest order transformation has

$$S = \frac{\hbar}{i2mc^2}\beta\boldsymbol{\alpha} \cdot \boldsymbol{p}c \tag{34.29}$$

and the Hamiltonian becomes

$$H' = \beta mc^2 + \beta\frac{p^2}{2m} - \frac{p^2}{3m^2c^2}\boldsymbol{\alpha} \cdot \boldsymbol{p}c + ... \tag{34.30}$$

Stopping here, which is appropriate for our usual nonrelativistic limit since applying a transformation to eliminate the last term shown will produce the term needed in the expansion of the energy $\beta\sqrt{p^2c^2 + m^2c^4}$. The momentum operator $\boldsymbol{p}$ commutes with S, so $\boldsymbol{p}' = \boldsymbol{p}$, and the momentum is identical in the new and old representations. Keeping order m^{-1} terms, the transformed position operator x' will be

$$
\begin{aligned}
x' &= x + \frac{i}{\hbar}[S, x] + ... \\
&= x + \frac{1}{mc}\beta\boldsymbol{\alpha} \cdot [\boldsymbol{p}, x] + ... = x - i\frac{\hbar}{mc}\beta\alpha_x + ...
\end{aligned} \tag{34.31}
$$

383

The difference between these is the odd operator of order m^{-1}. We can see its effect by looking at x'^2,

$$x'^2 = \left(x - i\frac{\hbar}{mc}\beta\alpha_x \right) \left(x + i\frac{\hbar}{mc}\alpha_x\beta \right) = x^2 + \frac{\hbar^2}{m^2c^2} + i\frac{2\hbar}{mc}\alpha_x\beta x \qquad (34.32)$$

A particle sharply localized around the origin in x, will only be localized within about a Compton wave length of the origin in x'. This is the source of the smoothing that gets done on the Compton wave length scale. In other words, we will calculate the solution of H' using the original operators x and p, not the transformed operators x' and p'. This form gives our nonrelativistic equations and shows that our nonrelativistic operator x is related to the full relativistic calculation averaged over a Compton wave length. The only way to have results sensitive to this averaging is to have momentum components $p \sim mc$, which would make the nonrelativistic limit inappropriate.

34.4.2 Coupling to general classical electromagnetic field

We now take the Hamiltonian to be

$$H = \boldsymbol{\alpha} \cdot (\boldsymbol{p}c - q\boldsymbol{A}(\boldsymbol{r},t)) + \beta mc^2 + q\Phi(\boldsymbol{r},t) \qquad (34.33)$$

Our first transformation is then

$$S = \frac{\hbar}{i2mc^2}\boldsymbol{\alpha} \cdot (\boldsymbol{p}c - q\boldsymbol{A}(\boldsymbol{r},t)) \qquad (34.34)$$

Substituting, the calculation is straightforward but tedious,

$$
\begin{aligned}
H' &= \beta mc^2 + q\Phi(\boldsymbol{r},t) + \beta\frac{(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t))^2}{2m} - \frac{q\hbar}{2mc}\beta\boldsymbol{\sigma}\cdot\boldsymbol{B}(\boldsymbol{r},t) \\
&\quad - \frac{q\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot\{\boldsymbol{E}(\boldsymbol{r},t)\times[\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)]\} + \frac{q\hbar^2}{8m^2c^2}\boldsymbol{\nabla}\cdot\boldsymbol{E}(\boldsymbol{r},t) \\
&\quad + \text{ odd terms } O(m^{-1}) + \text{ odd terms } O(m^{-2}) + ... \qquad (34.35)
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{B}(\boldsymbol{r},t) &= \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r},t) \\
\boldsymbol{E}(\boldsymbol{r},t) &= -\boldsymbol{\nabla}\Phi(\boldsymbol{r},t) - \frac{1}{c}\frac{\partial\boldsymbol{A}(\boldsymbol{r},t)}{\partial t} \;. \qquad (34.36)
\end{aligned}
$$

We next eliminate the odd term of order m^{-1}, but this produces new even terms that are order m^{-3}, and odd terms of order m^{-2}, so the even terms at order m^{-2} are unchanged. Eliminating the odd term of order m^{-2}, will not produce any lower order terms, so our approximation is given by the terms shown in Eq. 34.35. We take our nonrelativistic

384

Hamiltonian coupled to a classical electromagnetic field to be the positive energy solutions, i.e. when $\beta = 1$, or

$$
\begin{aligned}
H_{nr} \; = \; & \frac{(\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t))^2}{2m} + q\boldsymbol{\Phi}(\boldsymbol{r},t) - \frac{q\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{B}(\boldsymbol{r},t) \\
& - \frac{q\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot\{\boldsymbol{E}(\boldsymbol{r},t) \times [\boldsymbol{p} - \frac{q}{c}\boldsymbol{A}(\boldsymbol{r},t)]\} + \frac{q\hbar^2}{8m^2c^2}\boldsymbol{\nabla}\cdot\boldsymbol{E}(\boldsymbol{r},t)\,. \quad (34.37)
\end{aligned}
$$

The first line of terms are the same as we found in the static case. The nonradiatively corrected fermion g factor of 2 drops out. The last term with the cross product contains the spin orbit interaction as well as a nonlinear in the field spin interaction. You can verify that the spin orbit interaction drops out by taking $\boldsymbol{E}$ along $\hat{\boldsymbol{r}}$ as the nuclear field would be in an atom. The term is then proportional to $\boldsymbol{\sigma}\cdot(\boldsymbol{r}\times\boldsymbol{p})$ or proportional to $\boldsymbol{L}\cdot\boldsymbol{S}$. Since the original Dirac theory is properly relativistic, the correct factor accounting for the Thomas precession is automatic. The last term (recalling that $\boldsymbol{\nabla}\cdot\boldsymbol{E} = 4\pi\rho$) gives a zero range force from other charges – it is called the Darwin term, and causes a small splitting of the $2S$ level of hydrogen in addition to the bigger Lamb shift.

34.5   Coupling to quantized electromagnetism

Previously, in chapter 27, we quantized the electromagnetic field, and coupled it to the non-relativistic equations. In most of that work, we used the Schrödinger representation for the field operators. If we continue to work in the Schrödinger representation, the field operators will be independent of time, so none of the time derivatives we had in our calculation above will occur. However, since the state of our system will now include the electromagnetic field, its time development equation, the Schrödinger equation, will necessarily need to add the free electromagnetic field Hamiltonian

$$
H_{em} = \sum_{\boldsymbol{k}\alpha} \hbar\omega_k a^+_{\boldsymbol{k}\alpha} a_{\boldsymbol{k}\alpha} \tag{34.38}
$$

to the particle Hamiltonian we used. The commutators in Eq. 34.28 will need to include the commutators of H_{em} and $\boldsymbol{A}(\boldsymbol{r})$.

We can of course repeat our calculation being careful to keep these commutator terms, however, since we have already included the time derivatives for the classical field, it is simpler to change our quantized field to the interaction representation. In that case, the time development operator, our effective Hamiltonian, will no longer contain H_{em}, and we will not have to calculate its commutators. Instead, we will have time derivatives of the

quantized vector potential in the interaction representation

$$
\begin{aligned}
\boldsymbol{A}_{\text{int}}(\boldsymbol{r},t) &= \sum_{\boldsymbol{k}\alpha} \sqrt{\frac{2\pi\hbar c^2}{L^3\omega_k}} \left[a_{\boldsymbol{k},\alpha} e^{i\boldsymbol{k}\cdot\boldsymbol{r}-i\omega_k t}\hat{\boldsymbol{\lambda}}_\alpha + a_{\boldsymbol{k},\alpha}^+ e^{-i\boldsymbol{k}\cdot\boldsymbol{r}+i\omega_k t}\hat{\boldsymbol{\lambda}}_\alpha^* \right] \\
\frac{\partial \boldsymbol{A}_{\text{int}}(\boldsymbol{r},t)}{\partial t} &= -i\sum_{\boldsymbol{k}\alpha} \sqrt{\frac{2\pi\hbar c^2\omega_k}{L^3}} \left[a_{\boldsymbol{k},\alpha} e^{i\boldsymbol{k}\cdot\boldsymbol{r}-i\omega_k t}\hat{\boldsymbol{\lambda}}_\alpha - a_{\boldsymbol{k},\alpha}^+ e^{-i\boldsymbol{k}\cdot\boldsymbol{r}+i\omega_k t}\hat{\boldsymbol{\lambda}}_\alpha^* \right] \\
&= \frac{i}{\hbar}\left[H_{em}, \boldsymbol{A}_{\text{int}}(\boldsymbol{r},t) \right] .
\end{aligned}
\tag{34.39}
$$

If we wish to use the Schrödinger representation, we simply need to evaluate the interaction Hamiltonian at $t = 0$, this will give the same result as if we had calculated the commutators of the Foldy-Wouthuysen transformation with the full Schrödinger Hamiltonian.

Chapter 35

Angular momentum addition

35.1 Introduction

Since spin 1/2 particles interact with a magnetic field, they must produce a magnetic field, and therefore, two of them will interact. This interaction has a dipole-dipole form and is the hyperfine interaction. A simpler version is the Heisenberg form

$$H_{\text{spin}} = A\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 = \frac{4A}{\hbar^2}\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 \tag{35.1}$$

where A is a coefficient with units of energy. If these were classical spin vectors, the lowest energy state would have $\boldsymbol{S}_1 = -\boldsymbol{S}_2$ for $A > 0$ and $\boldsymbol{S}_1 = \boldsymbol{S}_2$ for $A < 0$. Quantum mechanically we need to diagonalize the Hamiltonian. We could do this without thinking, by writing a matrix representation for the spin operators, or we can use symmetry and angular momentum addition to find a more general, and simpler, solution.

35.1.1 Diagonalizing without thinking

Each spin has two states, so the system with two particles has four states which we can write as $|\uparrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, $|\downarrow\uparrow\rangle$, $|\downarrow\downarrow\rangle$. Where the first up or down specifies the state of spin 1 and the second the state of spin 2. The Pauli matrices when operating on a single particle basis $|\uparrow\rangle$, $|\downarrow\rangle$ are

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{35.2}$$

We can therefore immediately write down their representation in our 2 particle basis

$$
\boldsymbol{\sigma}_{1x} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad
\boldsymbol{\sigma}_{1y} = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \quad
\boldsymbol{\sigma}_{1z} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}
$$

$$
\boldsymbol{\sigma}_{2x} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad
\boldsymbol{\sigma}_{2y} = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \quad
\boldsymbol{\sigma}_{2z} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (35.3)
$$

We now can form the dot product by matrix multiplication and addition

$$
\begin{aligned}
\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \\
&\quad + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (35.4)
\end{aligned}
$$

Our Hamiltonian is block diagonal, with $| \uparrow\uparrow \rangle$, and $| \downarrow\downarrow \rangle$ eigenstates with eiegenvalue 1. Diagonalizing the central block, we find the additional eigenstates and $(| \uparrow\downarrow \rangle + | \downarrow\uparrow \rangle)/\sqrt{2}$ with eigenvalue 1 and $(| \uparrow\downarrow \rangle - | \downarrow\uparrow \rangle)/\sqrt{2}$ with eigenvalue -3.

35.1.2 Rotational invariance

Using the rotational symmetry of our spin Hamiltonian gives a simpler and more general way to treat systems like this. Since the Hamiltonian is a dot product of two vectors, we should expect that it is rotationally invariant. In order to rotate the system, we must rotate both $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$. To rotate $\boldsymbol{S}_1$ around the axis $\hat{\boldsymbol{\alpha}}$ by the angle $|\boldsymbol{\alpha}|$, we use the rotation operator $e^{\frac{i}{\hbar} \boldsymbol{S}_1 \cdot \boldsymbol{\alpha}}$, and to rotate $\boldsymbol{S}_2$, we use $e^{\frac{i}{\hbar} \boldsymbol{S}_2 \cdot \boldsymbol{\alpha}}$. Since $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ operate on different spaces, they commute, and the combined rotation operator is $e^{\frac{i}{\hbar} (\boldsymbol{S}_1 + \boldsymbol{S}_2) \cdot \boldsymbol{\alpha}}$. This is the rotation operator

for the angular momentum $\boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2$. We can check directly that our Hamiltonian commutes with $\boldsymbol{S}$. Using the summation convention

$$
\begin{aligned}
[S_\alpha, S_{1\beta}S_{2\beta}] &= [S_{1\alpha}, S_{1\beta}S_{2\beta}] + [S_{2\alpha}, S_{1\beta}S_{2\beta}] \\
&= [S_{1\alpha}, S_{1\beta}]S_{2\beta} + [S_{2\alpha}, S_{2\beta}]S_{1\beta} \\
&= i\hbar\epsilon_{\alpha\beta\gamma}S_{1\gamma}S_{2\beta} + i\hbar\epsilon_{\alpha\beta\gamma}S_{2\gamma}S_{1\beta} = 0
\end{aligned}
\tag{35.5}
$$

$\boldsymbol{S}$ satisfies the angular momentum commutation relations

$$
\begin{aligned}
[S_\alpha, S_\beta] &= [S_{1\alpha} + S_{2\alpha}, S_{1\beta} + S_{2\beta}] \\
&= [S_{1\alpha}, S_{1\beta}] + [S_{2\alpha}, S_{2\beta}] = i\hbar\epsilon_{\alpha\beta\gamma}(S_{1\gamma} + S_{2\gamma}) \\
&= i\hbar\epsilon_{\alpha\beta\gamma}S_\gamma
\end{aligned}
\tag{35.6}
$$

We can find simultaneous eigenvectors of S^2 and S_z. Since $\boldsymbol{S}$ commutes with H, we can go through all of our previous arguments we used for central potentials. The Hamiltonian must be block diagonal in this spin basis with the blocks corresponding to specific values of S^2 and S_z. Also, since H commutes with S_x and S_y and therefore $S_\pm$, the states with different S_z for the same S^2 will be degenerate.

We now simply have to find the possible values for S^2 and how many times each S^2 value occurs. The argument we use here will be generalized when we add arbitrary angular momenta, so it is important to understand the steps here. We write our original basis as the outer product of S_1^2, S_{1z} and S_2^2, S_{2z} states. These are our four $|\uparrow\uparrow\rangle$, $|\downarrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, $|\downarrow\downarrow\rangle$, states. These states are eigenstates of $S_z = S_{1z} + S_{2z}$. The maximum M value is 1 and occurs for just the $|\uparrow\uparrow\rangle$, state where $m_1 = m_2 = 1/2$. This state, since it is the only state with $M = 1$ must be an eigenstate of S^2. To have $m = 1$, we must have $s \geq 1$. If, however, $s = 2$, then there would have to be a state with $m = 2$. There is no such state, so that $|\uparrow\uparrow\rangle = e^{i\phi}|s = 1, m = 1, s_1 = 1/2, s_2 = 1/2\rangle$ where $e^{i\phi}$ is an arbitrary phase factor. The standard convention is to take this phase factor to be 1, so that

$$
|\uparrow\uparrow\rangle = |s = 1, m = 1, s_1 = 1/2, s_2 = 1/2\rangle .
\tag{35.7}
$$

The $s = 1$ states must include $m = 0$ and $m = -1$ states. Using our angular momentum algebra, and the usual phase convention, we know

$$
L_\pm|\ell m\rangle = \hbar\sqrt{\ell(\ell + 1) - m(m \pm 1)}|\ell m \pm 1\rangle \equiv \hbar A(\ell, \pm m)|\ell m \pm 1\rangle
\tag{35.8}
$$

where it is useful to define

$$
A(\ell, m) \equiv \sqrt{\ell(\ell + 1) - m(m + 1)}
\tag{35.9}
$$

so that

$$
\begin{aligned}
\hbar^{-1}S_-|s = 1, m = 1, s_1 = 1/2, s_2 = 1/2\rangle &= \sqrt{2}|s = 1, m = 0, s_1 = 1/2, s_2 = 1/2\rangle \\
&= \hbar^{-1}(S_{1-} + S_{2-})|\uparrow\uparrow\rangle = |\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle
\end{aligned}
\tag{35.10}
$$

or

$$|s = 1, m = 0, s_1 = 1/2, s_2 = 1/2\rangle = \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}} \tag{35.11}$$

Repeating this lowering on the $M = 0$ state gives

$$|s = 1, m = -1, s_1 = 1/2, s_2 = 1/2\rangle = |\downarrow\downarrow\rangle. \tag{35.12}$$

Notice that there were two original states that had $m = 0$. Having found the $s = 1, m = 0$ state, we can orthogonalize to find the orthogonal $m = 0$ state which will be any arbitrary phase multiplied by $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$. We need to choose a phase convention, and one standard way is to pick the term with the largest m_1 value to have a real positive coefficient. This second $m = 0$ zero state must have $s \geq 0$, but there are no other $m > 0$ states unassigned. Therefore this must be an $s = 0$ state.

These states must block diagonalize H, but since each of these states has a unique set of angular momentum quantum numbers, we get 1×1 blocks, i.e. these must be the eigenstates of our H, and in fact, the eigenstates of any rotationally invariant H that is a function of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$. For our H, we can see that

$$S^2 = (\boldsymbol{S}_1 + \boldsymbol{S}_2) \cdot (\boldsymbol{S}_1 + \boldsymbol{S}_2) = S_1^2 + S_2^2 + 2\boldsymbol{S}_1 \cdot \boldsymbol{S}_2. \tag{35.13}$$

We then have

$$\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 = \frac{S^2 - S_1^2 - S_2^2}{2} \tag{35.14}$$

when operating on our states these become

$$\frac{4A}{\hbar^2} \boldsymbol{S}_1 \cdot \boldsymbol{S}_2 |s, m, s_1 = 1/2, s_2 = 1/2\rangle = A[2s(s+1) - 3]|s, m, s_1 = 1/2, s_2 = 1/2\rangle \tag{35.15}$$

with eigenvalues A for $s = 1$ and $-3A$ for $s = 0$ as before.

35.2 Clebsch-Gordan Coefficients

The general problem, analogous to our spin problem, is to write the eigenstates of J^2, J_z, J_1^2 and J_2^2 in terms of the eigenstates of J_1^2, J_{1z}, J_2^2, and J_{2z}. That is we want to find the coefficients C so that

$$|jmj_1j_2\rangle = \sum_{m_1,m_2} C_{j_1,m_1,j_2,m_2}^{jm} |j_1m_1j_2m_2\rangle$$

$$= \sum_{m_1,m_2} |j_1m_1j_2m_2\rangle\langle j_1m_1j_2m_2|jmj_1j_2\rangle \tag{35.16}$$

the Clebsch-Gordan coeffiecients, C_{j_1,m_1,j_2,m_2}^{jm} are therefore

$$C_{j_1,m_1,j_2,m_2}^{jm} = \langle j_1m_1j_2m_2|jmj_1j_2\rangle. \tag{35.17}$$

We can follow the same arguments that we used for the spin problem. Just like that problem, we could simply write down the matrix representation of $\boldsymbol{J}_1 \cdot \boldsymbol{J}_2$ in the basis of good J_1^2, J_{1z}, J_2^2, J_{2z} which will be block diagonal in the J_z blocks, and then diagonalize the blocks. After choosing a phase convention for the eigenvectors, we would have the Clebsch-Gordan coefficients. However, like the spin problem, it is easier to use the angular momentum algebra directly to find the eigenvectors of J^2 and J_z.

1. The states $|j_1 m_1 j_2 m_2\rangle$ are J_z eigenstates with eigenvalues $m_1 + m_2$. Therefore, $m_1 + m_2 = m$ or the Clebsch-Gordan coefficient is zero.

2. The state with both the largest m_1, $m_1 = j_1$ and largest m_2, $m_2 = j_2$ has $m = j_1 + j_2$. Therefore this state must have $j \geq j_1 + j_2$. Since there are no larger m values, there cannot be any components with $j > j_1 + j_2$, and this state is an eigenstate of J^2 with $j = j_1 + j_2$. We choose the phase so that these states are equal

$$|j = j_1 + j_2, m = j_1 + j_2, j_1 j_2\rangle = |j_1, m_1 = j_1, j_2, m_2 = j_2\rangle . \qquad (35.18)$$

That is

$$C_{j_1,j_1,j_2,j_2}^{j_1+j_2,j_1+j_2} = 1 . \qquad (35.19)$$

3. We use the lowering operator to produce the $j = j_1 + j_2$ states for other m values. For example

$$\hbar^{-1} J_- |j = j_1 + j_2, m = j_1 + j_2, j_1 j_2\rangle = \hbar^{-1}(J_{1-} + J_{2-})|j_1, m_1 = j_1, j_2, m_2 = j_2\rangle$$
$$A(j_1 + j_2, -(j_1 + j_2))|j = j_1 + j_2, m = j_1 + j_2 - 1, j_1 j_2\rangle =$$
$$\sqrt{2(j_1 + j_2)}|j = j_1 + j_2, m = j_1 + j_2 - 1, j_1 j_2\rangle =$$
$$A(j_1, -j_1)|j_1, m_1 = j_1 - 1, j_2, m_2 = j_2\rangle + A(j_2, -j_2)|j_1, m_1 = j_1, j_2, m_2 = j_2 - 1\rangle =$$
$$\sqrt{2j_1}|j_1, m_1 = j_1 - 1, j_2, m_2 = j_2\rangle + \sqrt{2j_2}|j_1, m_1 = j_1, j_2, m_2 = j_2 - 1\rangle$$
$$(35.20)$$

This tells us that

$$C_{j_1,j_1-1,j_2,j_2}^{j_1+j_2,j_1+j_2-1} = \sqrt{\frac{j_1}{j_1 + j_2}}$$

$$C_{j_1,j_1,j_2,j_2-1}^{j_1+j_2,j_1+j_2-1} = \sqrt{\frac{j_2}{j_1 + j_2}} .$$
$$(35.21)$$

All of the $j = j_1 + j_2$ coefficients can be produced by repeated application of the lowering operator. Recall that we chose the phase relationship so that the coefficient $A(j, -m)$ for the lowering operator was real and positive. Using this lowering method keeps our phase choices consistent.

4. If $j_1 = 0$ is zero, $j = j_2$, and $m = m_2$, and similarly for $j_2 = 0$. In all other cases, the example in the last step produces one combination of the states with $m = j_1 + j_2 - 1$. We can form an orthogonal combination[1]. Using our previous argument, since this is the only remaining $m = j_1 + j_2 - 1$ state, it must be a $j = j_1 + j_2 - 1$, $m = j_1 + j_2 - 1$ state.

5. We now lower this state to get more Clebsch-Gordan coefficients. If this does not exhaust the states, we will find an orthogonal combination of $m = j_1 + j_2 - 2$ states which must have $j = j_1 + j_2 - 2$.

6. repeat until you have all of the Clebsch-Gordan coefficients.

The total number of original states is $(2j_1 + 1)(2j_2 + 1)$. As we proceed, we first produce $2(j_1 + j_2) + 1$ states with $j = j_1 + j_2$ by application of the lowering operator. We then orthogonalize to get a $j = j_1 + j_2 - 1$ state, and lowering produces $2(j_1 + j_2 - 1) + 1$ of these states. We repeat until there are no more states left. That will occur when

$$
(2j_1 + 1)(2j_2 + 1) = \sum_{j=j_{\min}}^{j_1+j_2} (2j + 1) = \left[2\frac{j_1 + j_2 + j_{\min}}{2} + 1 \right] (j_1 + j_2 - j_{\min} + 1)
$$
$$
= (j_1 + j_2 + 1)^2 - j_{\min}^2 \tag{35.22}
$$

or

$$
j_{\min}^2 = j_1^2 - 2j_1 j_2 + j_2^2 = (j_1 - j_2)^2 \tag{35.23}
$$

or $j_{\min} = |j_1 - j_2|$. The possible values for j are then given by the triangle inequality

$$
|j_1 - j_2| \leq j \leq j_1 + j_2 . \tag{35.24}
$$

The algorithm we used can be written as a set of recursion relations between the Clebsch-Gordan coefficients. A relation that we will find useful later is that given by operating with $J^2 = J_1^2 + J_2^2 + 2\boldsymbol{J}_1 \cdot \boldsymbol{J}_2$, with $2\boldsymbol{J}_1 \cdot \boldsymbol{J}_2 = J_{1+}J_{2-} + J_{1-}J_{2+} + 2J_{1z}J_{2z}$ on our states,

$$
j(j + 1) \sum_{m_1 m_2} C_{j_1 m_1 j_2 m_2}^{jm} |j_1 m_1 j_2 m_2\rangle
$$
$$
= \sum_{m_1 m_2} C_{j_1 m_1 j_2 m_2}^{jm} \left\{ [j_1(j_1 + 1) + j_2(j_1 + 1) + 2m_1 m_2]|j_1 m_1 j_2 m_2\rangle \right.
$$
$$
+ A(j_1, m_1)A(j_2, -m_2)|j_1, m_1 + 1, j_2, m_2 - 1\rangle
$$
$$
\left. + A(j_1, -m_1)A(j_2, m_2)|j_1, m_1 - 1, j_2, m_2 + 1\rangle \right\} \tag{35.25}
$$

---

[1]We must choose a phase convention here. Everyone chooses the coefficients to be real, but the sign is up to you, and therefore there are several choices in the literature. The most common is the Condon and Shortley choice[3]. In that convention, the sign is chosen so that $C_{j_1, j_1, j_2, -j_2}^{j, j_1 - j_2} > 0$.

Equating coefficients of the states, (i.e. we can operate from the left with $\langle j_1 m_1' j_2 m_2'|$ which picks those terms out of the sum – since this eliminates m_1 and m_2 from the expressions, we can then drop the primes from m_1' and m_2')

$$
\begin{aligned}
[j(j+1) &- j_1(j_1+1) - j_2(j_2+1) - 2m_1 m_2]\, C^{jm}_{j_1 m_1 j_2 m_2} \\
&= A(j_1, m_1 - 1)A(j_2, -m_2 - 1))C^{jm}_{j_1, m_1-1, j_2, m_2+1} \\
&+ A(j_1, -m_1 - 1)A(j_2, m_2 - 1)C^{jm}_{j_1, m_1+1, j_2, m_2-1}
\end{aligned}
$$

$$(35.26)$$

35.3   Closed form

The algorithm for finding the Clebsch-Gordan coefficients, if written down in general, leads to a closed form expression. This can be easily programmed on a computer to calculate any Clebsch-Gordan coefficients that are needed and tables are available. You need to know that the triangle inequality must hold and that $m_1 + m_2 = m$ so that you only calculate nonzero coefficients. One form is

$$
\begin{aligned}
C^{jm}_{j_1 m_1 j_2 m_2} = \delta_{m, m_1+m_2} & \left[\frac{(2j+1)(-j+j_1+j_2)!(j-j_1+j_2)!(j+j_1-j_2)!}{(j_1+j_2+j+1)!} \right]^{1/2} \\
& \sqrt{(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!(j+m)!(j-m)!} \sum_x (-1)^{x+j_1-m_1}
\end{aligned}
$$

$$
\frac{1}{x!(j_1+j_2-j-x)!(j_1-m_1-x)!(j_2+m_2-x)!(j-j_2+m_1+x)!(j-j_1-m_2+x)!}
$$

$$(35.27)$$

where the x sum is over all integer values of x that have all of the factorial arguments in the denominator greater than or equal to 0.

35.4   Symmetry properties

As you should expect, the Clebsch-Gordan coefficients have many symmetry properties. For example, reversing time reverses the direction of angular momenta, so we would expect changing the signs of all of m_1, m_2 and m would at most change the sign of the coefficient.

Symmetries must be contained in the general expression for the Clebsch-Gordan coefficients. If we change the signs of all the m values in Eq. 35.27 the factors outside of the sum do not change. To bring the sum into the same form, we can make the substitution $x = j_1 + j_2 - j - x'$, and sum over x' instead, we find that each of the pairs of factors maps

into the corresponding pair with the negative m values:

$$
\begin{aligned}
x!(j_1 + j_2 - j - x)! &= (j_1 + j_2 - j - x')!x'! \\
(j_1 + m_1 - x)!(j - j_2 - m_1 + x)! &= (j - j_2 + m_1 + x')!(j_1 - m_1 - x')! \\
(j_2 - m_2 - x)!(j - j_1 + m_2 + x)! &= (j - j_1 - m_2 + x')!(j_2 + m_2 - x')! \\
(-1)^x &= (-1)^{j_1 + j_2 - j - x'} = (-1)^{j - j_1 - j_2}(-1)^{x'}. \quad (35.28)
\end{aligned}
$$

Comparing with the original equation we see that

$$
C^{j,-m}_{j_1,-m_1,j_2,-m_2} = (-1)^{j - j_1 - j_2} C^{j,m}_{j_1,m_1,j_2,m_2}. \quad (35.29)
$$

Another obvious symmetry is if we change the angular momentum we call j_1 with the one we call j_2. This should again, at most change the sign of the coefficient. Again, the factors in front of the sum in Eq. 35.27 do not change. Making the same substitution as before $x = j_1 + j_2 - j - x'$, and sum over x' instead, we find that

$$
\begin{aligned}
x!(j_1 + j_2 - j - x)! &= (j_1 + j_2 - j - x')!x'! \\
(j_2 - m_2 - x)!(j - j_1 + m_2 + x)! &= (j - j_1 - m_2 + x')!(j_2 + m_2 - x')! \\
(j_1 + m_1 - x)!(j - j_2 - m_1 + x)! &= (j - j_2 + m_1 + x')!(j_1 - m_1 - x')! \\
(-1)^x &= (-1)^{j_1 + j_2 - j - x'} = (-1)^{j - j_1 - j_2}(-1)^{x'}. \quad (35.30)
\end{aligned}
$$

and again comparing the equations shows

$$
C^{j,m}_{j_2,m_2,j_1,m_1} = (-1)^{j - j_1 - j_2} C^{j,m}_{j_1,m_1,j_2,m_2} \quad (35.31)
$$

There are also simple relationships between adding j_1 and j_2 to get j and adding j and j_1 to get j_2, etc. These are more simply written in terms of the 3-j symbol described below, but can again be derived from Eq. 35.27.

35.5 Wigner 3-j symbols

An alternative way to write the Clebsch-Gordan coefficients has the advantage of treating all three angular momenta j_1, j_2 and j on the same footing, and therefore has simpler looking symmetries. It is called the Wigner 3-j symbol

$$
C^{j_3 m_3}_{j_1 m_1 j_2 m_2} = (-1)^{j_1 - j_2 + m_3}\sqrt{2j_3 + 1}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} \quad (35.32)
$$

where we write $j = j_3$ here to emphasize that we don't pick out any of the three as special.

Even permutations of the columns leave the 3-j symbol invariant

$$
\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} j_2 & j_3 & j_1 \\ m_2 & m_3 & m_1 \end{pmatrix} = \begin{pmatrix} j_3 & j_1 & j_2 \\ m_3 & m_1 & m_2 \end{pmatrix} \quad (35.33)
$$

While odd permutations multiply it by $(-1)^{j_1+j_2+j_3}$,

$$(-1)^{j_1+j_2+j_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} = \begin{pmatrix} j_1 & j_3 & j_2 \\ m_1 & m_3 & m_2 \end{pmatrix} = \begin{pmatrix} j_3 & j_2 & j_1 \\ m_3 & m_2 & m_1 \end{pmatrix} .$$
(35.34)

Changing the signs of the m values gives the same factor

$$(-1)^{j_1+j_2+j_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix}$$
(35.35)

We can think of the 3-j symbols as the amplitude of adding three angular momenta to get a total angular momentum 0. Imagine first adding j_1, m_1 and j_2, m_2 to get j', m', and then adding j_3, m_3 to j', m' to get 0. We will need to make a phase choice eventually. Here we write

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = e^{i\phi} \sum_{j'm'} C^{00}_{j'm'j_3m_3} C^{j'm'}_{j_1m_1j_2m_2}$$
(35.36)

with $e^{i\phi}$ the phase choice. The Clebsch-Gordan coefficient $C^{00}_{j'm'j_3m_3}$ is zero except when $j' = j_3$, and $m' = -m_3$. Evaluating the general expression Eq. 35.27 gives

$$C^{00}_{j_3,-m_3,j_3,m_3} = \frac{(-1)^{j_3+m_3}}{\sqrt{2j_3+1}}$$
(35.37)

and

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = e^{i\phi} \frac{(-1)^{j_3+m_3}}{\sqrt{2j_3+1}} C^{j_3,-m_3}_{j_1m_1j_2m_2}$$
(35.38)

changing $m_3 \to -m_3$ we see that to agree with the definition of the 3-j symbol, the phase factor is $(-1)^{-j_1+j_2-j_3}$, so that

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} = \frac{(-1)^{-j_1+j_2-m_3}}{\sqrt{2j_3+1}} C^{j_3,m_3}_{j_1m_1j_2m_2}$$
(35.39)

35.6 Irreducible spherical tensor operators and the Wigner-Eckart theorem

Let's imagine that you decide to calculate the dipole transition matrix elements for an atom. For our purposes, we can just take the hydrogen atom as our example. Imagine that you want to calculate all of the possible dipole transition matrix elements from the 5g to the 4f states, so that you can determine their transition rates. the g states have $\ell = 4$, so there are 9 of them. The f states have $\ell = 3$, so there are 7 of them. We need to calculate the matrix element of x, y and z between them, so if we don't know the selection rules, we would have to calculate $3 \cdot 7 \cdot 9 = 189$ integrals for the matrix elements. We are already smarter

in that we know that the z matrix element only couples equal m values, and the x and y only couple $m' = m \pm 1$ values. So this would reduce our calculation to 7 integrals for the z matrix elements and 14 each for the x and y matrix elements. So we would have 35 integrals to do. If you think about it, you might reduce this number further. For example, if instead of calculating x and y matrix elements you calculated $x + iy$ and $x - iy$ matrix elements, $x + iy$ would only couple $m' = m + 1$ terms, while $x - iy$ would just couple $m' = m - 1$ terms. This would further reduce the number of integrals to 21. This last simplication, is the idea behind the Wigner-Eckart theorem, and it leads to the result that we only have to calculate 1 integral to get all of the matrix elements.

The key is to write operators that have the same algebra as an angular momentum state. These are called irreducible spherical tensor operators. Here $\boldsymbol{r}$ is a cartesian vector operator with usual $r_x = x$, $r_y = y$ and $r_z = z$ components, so that

$$[J_\alpha, r_\beta] = i\hbar\epsilon_{\alpha\beta\gamma}r_\gamma \,. \tag{35.40}$$

If we calculate the commutator of $J_\pm$ and J_z with z, we get

$$
\begin{aligned}
{[J_\pm, z]} &= [J_x \pm iJ_y, z] = \hbar(-iy \mp x) = \hbar\sqrt{2}\frac{\mp(x \pm iy)}{\sqrt{2}} = A(1, \pm 1)\frac{\mp(x \pm iy)}{\sqrt{2}} \,. \\
{[J_z, z]} &= 0
\end{aligned}
\tag{35.41}
$$

If we define

$$
\begin{aligned}
r_1 &= -\frac{x + iy}{\sqrt{2}} \\
r_0 &= z \\
r_{-1} &= \frac{x - iy}{\sqrt{2}}
\end{aligned}
\tag{35.42}
$$

and calculate all of the commutation relations with $J_\pm$ and J_z, they become

$$
\begin{aligned}
{[J_\pm, r_m]} &= \hbar A(1, \pm m)r_{m \pm 1} \\
{[J_z, r_m]} &= \hbar m r_m \,.
\end{aligned}
\tag{35.43}
$$

Notice that the angular momentum commutators give the same coefficients and m values that these same angular momentum operators would give when operating on a $j = 1$ m state. We call the set of r_m, an irreducible spherical tensor of rank 1, since it has the same coefficients as a $j = 1$ angular momentum.

Let's look at how we would apply this to our example problem. Operating on $r_{m_1}|n\ell m_2\rangle$, where $n = 4$ for the f states, with J_z and $J_\pm$ we get

$$
\begin{aligned}
\hbar^{-1}J_z r_{m_1}|n\ell m_2\rangle &= \hbar^{-1}[J_z, r_{m_1}]|n\ell m_2\rangle + \hbar^{-1}r_{m_1}J_z|n\ell m_2\rangle = (m_1 + m_2)r_{m_1}|n\ell m_2\rangle \\
\hbar^{-1}J_\pm r_{m_1}|n\ell m_2\rangle &= \hbar^{-1}[J_\pm, r_{m_1}]|n\ell m_2\rangle + r_{m_1}\hbar^{-1}J_\pm|n\ell m_2\rangle = \\
&= A(1, \pm m_1)r_{m_1 \pm 1}|n\ell m_2\rangle + A(\ell, \pm m_2)r_{m_1}|n\ell m_2 \pm 1\rangle \,. \tag{35.44}
\end{aligned}
$$

This is the same algebra that we have when we operate on two angular momenta. That is, compare this to

$$
\begin{aligned}
\hbar^{-1} J_z |\ell_1 m_1 \ell_2 m_2\rangle &= (m_1 + m_2)|\ell_1 m_1 \ell_2 m_2\rangle \\
\hbar^{-1} J_\pm |\ell_1 m_1 \ell_2 m_2\rangle &= A(\ell_1, \pm m_1)|\ell_1 m_1 \pm 1 \ell_2 m_2\rangle + A(\ell_2, \pm m_2)|\ell_1 m_1 \ell_2 m_2 \pm 1\rangle
\end{aligned}
$$
$$(35.45)$$

Using these operators we will be able to use the algebra of angular momentum addition to simplify matrix elements.

To exploit this, in general, we require a set of irreducible spherical tensor operators to have the commutation relations

$$
\begin{aligned}
\left[J_\pm, T_{\ell m}\right] &= \hbar A(\ell, \pm m) T_{\ell m \pm 1} \\
\left[J_z, T_{\ell m}\right] &= \hbar m T_{\ell m} .
\end{aligned}
$$
$$(35.46)$$

This may seem like a difficult set of commutation relations to come up with, but it is not. Any vector can be combined into a spherical tensor of rank 1 exactly as we did for $\boldsymbol{r}$ above. For example, $\boldsymbol{p}$ can be written as a spherical tensor, and since it is $-i\hbar\boldsymbol{\nabla}$ in the r representation, the gradient can be written as a spherical tensor. The spherical harmonics satisfy the irreducible spherical tensor operator commutation relations for the corresponding ℓ and m values. Any combination of components $x^{n_x} y^{n_y} z^{n_z}$ can be written as a linear combination of spherical harmonics and powers of r, so it can be written as a linear combination of irreducible spherical tensors. Similarly with the components of any other vector.

We will need the commutator with J^2 later. It can be calculated from[2]

$$
\left[J^2, T_{\ell m}\right] = \hbar^2 \ell(\ell+1) T_{\ell m} + [J_-, T_{\ell,m}]J_+ + [J_+, T_{\ell,m}]J_- + 2[J_z, T_{\ell,m}]J_z .
$$
$$(35.48)$$

The analog of adding two angular momenta is to combine the operators times states to produce what we conjecture to be good angular momentum combinations,

$$
|\Psi_{jm}\rangle = \sum_{m_1,m_2} C^{jm}_{\ell_2 m_2 \ell_1 m_1} T_{\ell_2 m_2} |n\ell_1 m_1\rangle .
$$
$$(35.49)$$

---

[2]For any three operators

$$
[O_1 O_2, T] = [O_1, [O_2, T]] + [O_1, T]O_2 + [O_2, T]O_1
$$
$$(35.47)$$

which can be verified by expanding the commutators. Applying this to $[J_\pm J_\mp, T_{\ell m}]$ and $[J_z^2, T_{\ell m}]$ gives Eq. 35.48.

We can verify that these really are good angular momentum states

$$
\begin{aligned}
J_z|\Psi_{jm}\rangle &= \sum_{m_1,m_2} C^{jm}_{\ell_2 m_2 \ell_1 m_1} \left([J_z, T_{\ell_2 m_2}] + T_{\ell_2,m_2} J_z\right) |n\ell_1 m_1\rangle \\
&= \sum_{m_1,m_2} C^{jm}_{\ell_2 m_2 \ell_1 m_1} \hbar(m_1 + m_2) T_{\ell_2 m_2} |n\ell_1 m_1\rangle \\
&= \hbar m \sum_{m_1,m_2} C^{jm}_{\ell_2 m_2 \ell_1 m_1} T_{\ell_2 m_2} |n\ell_1 m_1\rangle |n\ell_1 m_1\rangle = \hbar m |\Psi_{jm}\rangle \qquad (35.50)
\end{aligned}
$$

where in the last line we used the property of the Clebsch-Gordan coefficients that $m_1 + m_2 = m$ or the coefficients are zero. So they are good J_z states with eigenvalue $m\hbar$. Similarly we can write

$$
\begin{aligned}
\hbar^{-2} J^2 |\Psi_{jm}\rangle &= \hbar^{-2} \sum_{m_1,m_2} C^{jm}_{\ell_2 m_2 \ell_1 m_1} \left([J^2, T_{\ell_2 m_2}] + T_{\ell_2 m_2} J^2\right) |n\ell_1 m_1\rangle \\
&= \sum_{m_1,m_2} C^{jm}_{\ell_2 m_2 \ell_1 m_1} \left\{ [\ell_1(\ell_1+1) + \ell_2(\ell_2+1) + 2m_1 m_2] T_{\ell_2 m_2}|n\ell_1 m_1\rangle \right. \\
&\quad + A(\ell_2, m_2) A(\ell_1, -m_1) T_{\ell_2, m_2+1}|n\ell_1 m_1 - 1\rangle \\
&\quad \left. + A(\ell_2, -m_2) A(\ell_1, m_1+1) T_{\ell_2, m_2-1}|n\ell_1 m_1 + 1\rangle \right\} \\
&= \sum_{m_1,m_2} \left\{ [\ell_1(\ell_1+1) + \ell_2(\ell_2+1) + 2m_1 m_2] C^{jm}_{\ell_2 m_2 \ell_1 m_1} \right. \\
&\quad + C^{jm}_{\ell_2 m_2 - 1 \ell_1 m_1 + 1} A(\ell_2, m_2 - 1) A(\ell_1, -m_1 - 1) \\
&\quad \left. + C^{jm}_{\ell_2 m_2 + 1 \ell_1 m_1 - 1} A(\ell_2, -m_2 - 1) A(\ell_1, m_1 - 1) \right\} T_{\ell_2, m_2}|n\ell_1 m_1\rangle .
\end{aligned}
$$

$$(35.51)$$

Using the relationship between the Clebsch-Gordan coefficients, Eq. 35.26, that we derived for the good j combinations, this becomes

$$
\hbar^{-2} J^2 |\Psi_{jm}\rangle = j(j+1) \sum_{m_1,m_2} C^{jm}_{\ell_2 m_2 \ell_1 m_1} T_{\ell_2 m_2}|n\ell_1 m_1\rangle = j(j+1)|\Psi_{jm}\rangle, \qquad (35.52)
$$

and the $|\Psi_{jm}\rangle$ are eigenstates of J^2 and J_z as expected.

Since these are good angular momentum states in the original space, we can expand them in terms of the complete set in our original basis $|njm\rangle$,

$$
|\Psi_{jm}\rangle = \sum_{n'} |n'jm\rangle\langle n'jm|\Psi_{jm}\rangle \qquad (35.53)
$$

If we operate on this with $J_\pm A^{-1}(j, \pm m)$, we will raise or lower the m index of the kets, but the coefficients do not change. This shows that the coefficients are independent of m,

$$
\langle n'jm|\Psi_{jm}\rangle \equiv \langle n'j|\Psi_j\rangle \qquad (35.54)
$$

where the second form is written with the m value dropped to show this independence explicitly.

Since

$$\delta_{m_1,m_1'}\delta_{m_2,m_2'} = \sum_{jm}\langle\ell_2 m_2\ell_1 m_1|jm\ell_1\ell_2\rangle\langle jm\ell_1\ell_2|\ell_2 m_2'\ell_1 m_1'\rangle = \sum_{jm}C^{jm}_{\ell_2 m_2\ell_1 m_1}C^{jm}_{\ell_2 m_2'\ell_1 m_1'} \quad (35.55)$$

$$\sum_{jm}C^{jm}_{\ell_2 m_2\ell_1 m_1}|\Psi_{jm}\rangle = \sum_{jm,m_1',m_2'}C^{jm}_{\ell_2 m_2\ell_1 m_1}\left[C^{jm}_{\ell_2 m_2'\ell_1 m_1'}T_{\ell_2 m_2'}|n\ell_1 m_1'\rangle\right]$$
$$= T_{\ell_2 m_2}|n\ell_1 m_1\rangle. \quad (35.56)$$

Therefore

$$\langle n'\ell_3 m_3|T_{\ell_2 m_2}|n\ell_1 m_1\rangle = \sum_{jm}C^{jm}_{\ell_2 m_2\ell_1 m_1}\langle n'\ell_3 m_3|\Psi_{jm}\rangle$$
$$= C^{\ell_3 m_3}_{\ell_2 m_2\ell_1 m_1}\langle n'\ell_3|\Psi_{\ell_3}\rangle \quad (35.57)$$

For historical reasons, and to simplify some expressions the quantity $\langle n'\ell_3|\Psi_{\ell_3}\rangle$ is written as

$$\langle n'\ell_3|\Psi_{\ell_3}\rangle = (-1)^{-\ell_1+\ell_2+\ell_3}\frac{\langle n'\ell_3||T_{\ell_2}||n\ell_1\rangle}{\sqrt{2\ell_3+1}} \quad (35.58)$$

where $\langle n'\ell_3||T_{\ell_2}||n\ell_1\rangle$ is called the reduced matrix element. The Wigner-Eckart theorem is then

$$\langle n'\ell_3 m_3|T_{\ell_2 m_2}|n\ell_1 m_1\rangle = (-1)^{-\ell_1+\ell_2+\ell_3}\frac{C^{\ell_3 m_3}_{\ell_2 m_2\ell_1 m_1}}{\sqrt{2\ell_3+1}}\langle n'\ell_3||T_{\ell_2}||n\ell_1\rangle$$
$$= (-1)^{2\ell_2}\frac{C^{\ell_3 m_3}_{\ell_1 m_1\ell_2 m_2}}{\sqrt{2\ell_3+1}}\langle n'\ell_3||T_{\ell_2}||n\ell_1\rangle \quad (35.59)$$

The key result is that all of the m_1, m_2 and m_3 dependence is in the Clebsch-Gordan coefficient. The typical way you use this is to calculate the left-hand side in the usual way for *one* particular m_1, m_2 m_3 value where the Clebsch-Gordan coefficient is nonzero. This tells you what the value of the reduced matrix element is. Given the reduced matrix elements, you only need to plug in the values of the Clebsch-Gordan coefficients to get all of the other matrix elements. For our 5g to 4f transition, you need to calculate one of the nonzero matrix elements. Clebsch-Gordan coefficients give you the rest.

In terms of the 3-j symbols we get the expression

$$\langle n'\ell_3 m_3|T_{\ell_2 m_2}|n\ell_1 m_1\rangle = (-1)^{\ell_3-m_3}\begin{pmatrix}\ell_3 & \ell_2 & \ell_1 \\ -m_3 & m_2 & m_1\end{pmatrix}\langle n'\ell_3||T_{\ell_2}||n\ell_1\rangle. \quad (35.60)$$

35.7 Products of irreducible spherical tensor operators

Using the commutation relations, and the recursion formula's for the Clebsch-Gordan coefficients, it is straightforward to show that coupling together irreducible spherical tensor operators like angular momenta also gives irreducible spherical tensor operators. That is for irreducible spherical tensors $U_{\ell_1 m_1}$ and $V_{\ell_2 m_2}$, we can construct

$$W_{\ell_3 m_3} = \sum_{m_1 m_2} C^{\ell_3 m_3}_{\ell_1 m_1 \ell_2 m_2} U_{\ell_1 m_1} V_{\ell_2 m_2} . \tag{35.61}$$

This is sometimes written (with the m indices suppressed) as

$$W_{\ell_3} = \{U_{\ell_1} \otimes V_{\ell_2}\}_{\ell_3} \tag{35.62}$$

We often couple two irreducible spherical tensors of the same rank to form a scalar,

$$W_0 = \{U_\ell \otimes V_\ell\}_0 = \sum_m C^{00}_{\ell,-m,\ell,m} U_{\ell,-m} V_{\ell,m} = \frac{1}{\sqrt{2\ell+1}} \sum_m (-1)^{\ell+m} U_{\ell,-m} V_{\ell,m} \tag{35.63}$$

Sometimes, for integer ℓ, the scalar product is defined to be $\sum_m (-1)^m U_{\ell,-m} V_{\ell,m} = U_\ell \cdot V_\ell$, but I will not use that notation here.

We can multiply Eq. 35.61 by $C^{\ell_3 m_3}_{\ell_1 m_1' \ell_2 m_2'}$ and sum over ℓ_3 and m_3 using the orthogonality of the Clebsch-Gordan coefficients (and then drop the primes on m_1' and m_2' since m_1 and m_2 have been summed over and are no longer used) to express the original product as a linear combination of tensor operators

$$U_{\ell_1 m_1} V_{\ell_2 m_2} = \sum_{\ell_3 m_3} C^{\ell_3 m_3}_{\ell_1 m_1 \ell_2 m_2} \{U_{\ell_1} \otimes V_{\ell_2}\}_{\ell_3 m_3} \tag{35.64}$$

35.7.1 Reduced matrix element of the spherical harmonics

Let's apply this to calculating the reduced matrix element for the spherical harmonics. It is easy to verify that the spherical harmonics $Y_{\ell m}(\theta_{\mathrm{op}}, \phi_{\mathrm{op}})$ are irreducible spherical tensors of rank ℓ. This means that while they are the r space wave functions, they can also be an irreducible spherical tensor operators. The matrix element of a spherical harmonic between two angular momentum states is

$$
\begin{aligned}
\langle \ell_3 m_3 | Y_{\ell_2 m_2}(\theta_{\mathrm{op}}, \phi_{\mathrm{op}}) | \ell_1 m_1 \rangle &= \int_{-1}^1 d\cos\theta \int_0^{2\pi} d\phi \langle \ell_3 m_3 | Y_{\ell_2 m_2}(\theta_{\mathrm{op}}, \phi_{\mathrm{op}}) | \theta\phi \rangle \langle \theta\phi | \ell_1 m_1 \rangle \\
&= \int_{-1}^1 d\cos\theta \int_0^{2\pi} d\phi \, Y^*_{\ell_3 m_3}(\theta, \phi) Y_{\ell_2 m_2}(\theta, \phi) Y_{\ell_1 m_1}(\theta, \phi) .
\end{aligned}
\tag{35.65}
$$

The Wigner-Eckart theorem tells us

$$\langle \ell_3 m_3 | Y_{\ell_2 m_2}(\theta_{\mathrm{op}}, \phi_{\mathrm{op}}) | \ell_1 m_1 \rangle = (-1)^{\ell_1 - \ell_2 + \ell_3} \frac{C^{\ell_3 m_3}_{\ell_1 m_1 \ell_2 m_2}}{\sqrt{2\ell_3 + 1}} \langle \ell_3 || Y_{\ell_2} || \ell_1 \rangle . \tag{35.66}$$

The straightforward way would be to pick one matrix element for each set of ℓ values and perform the integral. We shall instead use a trick. Notice that the integral that we need to do is the same one we would have to do if we were calculating the expectation value of three spherical harmonics in an $\ell = 0$ state,

$$4\pi \langle 00|(-1)^{m_3} Y_{\ell_3,-m_3}(\theta_{\rm op}, \phi_{\rm op}) Y_{\ell_2 m_2}(\theta_{\rm op}, \phi_{\rm op}) Y_{\ell_1 m_1}(\theta_{\rm op}, \phi_{\rm op})|00\rangle = \langle \ell_3 m_3 | Y_{\ell_2 m_2}(\theta_{\rm op}, \phi_{\rm op})|\ell_1 m_1 \rangle . \tag{35.67}$$

We can write the product of three-spherical tensors by using Eq. 35.64 twice

$$Y_{\ell_3,-m_3}(\theta_{\rm op}, \phi_{\rm op}) Y_{\ell_2 m_2}(\theta_{\rm op}, \phi_{\rm op}) Y_{\ell_1 m_1}(\theta_{\rm op}, \phi_{\rm op}) = \sum_{\ell m \ell' m'} C^{\ell m}_{\ell_3,-m_3,\ell',m'} C^{\ell' m'}_{\ell_2 m_2 \ell_1 m_1} \{Y_{\ell_3} \otimes \{Y_{\ell_2} \otimes Y_{\ell_1}\}_{\ell'}\}_{\ell m} \tag{35.68}$$

The Wigner-Eckart theorem (or just thinking about the rotational symmetry), tells us that the expectation value of a tensor operator in an angular momentum zero state is proportional to $C^{00}_{\ell m 00}$ which is zero unless $\ell = m = 0$. Therefore only the scalar or $\ell = m = 0$ component of Eq. 35.68 will be nonzero and this scalar component will be independent of θ and ϕ. Therefore

$$\langle \ell_3 m_3 | Y_{\ell_2 m_2}(\theta_{\rm op}, \phi_{\rm op})|\ell_1 m_1 \rangle = 4\pi (-1)^{m_3} C^{00}_{\ell,-m_3,\ell_3,m_3} C^{\ell_3 m_3}_{\ell_1 m_1 \ell_2 m_2} \{Y_{\ell_3} \otimes \{Y_{\ell_2} \otimes Y_{\ell_1}\}_{\ell_3}\}_0 . \tag{35.69}$$

Comparing with the Wigner-Eckart theorem, and substituting Eq. 35.37 for $C^{00}_{\ell,-m_3,\ell_3,m_3}$ the reduced matrix element is

$$\langle \ell_3 || Y_{\ell_2} || \ell_1 \rangle = 4\pi (-1)^{\ell_3} \{Y_{\ell_3} \otimes \{Y_{\ell_2} \otimes Y_{\ell_1}\}_{\ell_3}\}_0 . \tag{35.70}$$

Evaluating this scalar

$$\{Y_{\ell_3} \otimes \{Y_{\ell_2} \otimes Y_{\ell_1}\}_{\ell_3}\}_0 = \sum_{m_1,m_2,m_3} Y_{\ell_3,-m_3}(\theta, \phi) C^{00}_{\ell_3 -m_3 \ell_3 m_3} C^{\ell_3 m_3}_{\ell_2 m_2 \ell_1 m_1} Y_{\ell_2 m_2}(\theta, \phi) Y_{\ell_1 m_1}(\theta, \phi) . \tag{35.71}$$

Since the scalar commutes with $\boldsymbol{J}$ it is independent of θ and ϕ as noted above. Therefore we can evaluate it at any convenient values for θ and ϕ. The easiest place is $\theta = 0$ where

$$Y_{\ell m}(0,0) = \sqrt{\frac{2\ell+1}{4\pi}} \delta_{m,0} \tag{35.72}$$

We find

$$\begin{aligned} \{Y_{\ell_3} \otimes \{Y_{\ell_2} \otimes Y_{\ell_1}\}_{\ell_3}\}_0 &= \sqrt{\frac{2\ell_1+1}{4\pi}} \sqrt{\frac{2\ell_2+1}{4\pi}} \sqrt{\frac{2\ell_3+1}{4\pi}} C^{00}_{\ell_3 0 \ell_3 0} C^{\ell_3 0}_{\ell_1 0 \ell_2 0} \\ &= \sqrt{\frac{2\ell_1+1}{4\pi}} \sqrt{\frac{2\ell_2+1}{4\pi}} \sqrt{\frac{1}{4\pi}} (-1)^{\ell_3} C^{\ell_3 0}_{\ell_1 0 \ell_2 0} \end{aligned} \tag{35.73}$$

401

where I have used Eq. 35.37 to evaluate $C^{00}_{\ell_3 0 \ell_3 0}$.

Substituting we get the final expression for the reduced matrix element

$$\langle \ell_3 || Y_{\ell_2} || \ell_1 \rangle = \sqrt{\frac{(2\ell_1 + 1)(2\ell_2 + 1)}{4\pi}} C^{\ell_3 0}_{\ell_1 0 \ell_2 0}. \tag{35.74}$$

The Clebsch-Gordan coefficient is

$$C^{\ell_3 0}_{\ell_1 0 \ell_2 0} = \begin{cases} 0 & \ell_1 + \ell_2 + \ell_3 \text{ odd} \\ \frac{(-1)^{s-\ell_3}\sqrt{2\ell_3+1}s!}{(s-\ell_1)!(s-\ell_2)!(s-\ell_3)!} \left[\frac{(2s-2\ell_1)!(2s-2\ell_2)!(2s-2\ell_3)!}{(2s+1)!} \right]^{1/2} & \ell_1 + \ell_2 + \ell_3 = 2s \text{ even} \end{cases}. \tag{35.75}$$

You probably already realize that since the spherical harmonics have parity $(-1)^\ell$, if the sum of the three ℓ values is odd, the integrand is odd under inversion and the integral is zero. This drops out of the calculation automatically.

Our complete result is summarized as

$$\langle \ell_3 || Y_{\ell_2} || \ell_1 \rangle = \sqrt{\frac{(2\ell_1 + 1)(2\ell_2 + 1)}{4\pi}} C^{\ell_3 0}_{\ell_1 0 \ell_2 0}$$

$$\int_{-1}^{1} d\cos\theta \int_0^{2\pi} d\phi Y^*_{\ell_3 m_3}(\theta,\phi) Y_{\ell_2 m_2}(\theta,\phi) Y_{\ell_1 m_1}(\theta,\phi) = \sqrt{\frac{(2\ell_1 + 1)(2\ell_2 + 1)}{4\pi(2\ell_3 + 1)}} C^{\ell_3 0}_{\ell_1 0 \ell_2 0} C^{\ell_3 m_3}_{\ell_1 m_1 \ell_2 m_2}$$

$$\tag{35.76}$$

35.7.2 Vector couplings

As mentioned above, we can form an irreducible spherical tensor of rank 1 from any vector

$$A_{1,1} = -\frac{A_x + iA_y}{\sqrt{2}}$$

$$A_{1,0} = A_z$$

$$A_{1,-1} = -\frac{A_x - iA_y}{\sqrt{2}}. \tag{35.77}$$

Using two vectors, we can form three irreducible tensor products with $\ell = 0, 1, 2$. Forming the $\ell = 0$ scalar and $\ell = 1$ vector couplings and evaluating the Clebsch-Gordan coefficients gives results proportional to the familiar dot and cross products:

$$C = \{A_1 \otimes B_1\}_{00} \quad \text{becomes} \quad C = -\frac{1}{\sqrt{3}} \boldsymbol{A} \cdot \boldsymbol{B}$$

$$C_{1m} = \{A_1 \otimes B_1\}_{1m} \quad \text{becomes} \quad \boldsymbol{C} = \frac{i}{\sqrt{2}} \boldsymbol{A} \times \boldsymbol{B}$$

$$\tag{35.78}$$

In the next section we will look at combinations of couplings that lead to scalars $\ell = 0$. For three vectors $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, we have

$$D = \{\{A_1 \otimes B_1\}_1 \otimes C_1\}_0 \quad \text{becomes} \quad D = -\frac{i}{\sqrt{6}} (\boldsymbol{A} \times \boldsymbol{B}) \cdot \boldsymbol{C}$$

(35.79)

while for four vectors $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$

$$E = \{\{A_1 \otimes B_1\}_0 \otimes \{C_1 \otimes D_1\}_0\}_0 \quad \text{becomes} \quad E = \frac{1}{3} (\boldsymbol{A} \cdot \boldsymbol{B}) (\boldsymbol{C} \cdot \boldsymbol{D})$$

$$E = \{\{A_1 \otimes B_1\}_1 \otimes \{C_1 \otimes D_1\}_1\}_0 \quad \text{becomes} \quad E = \frac{1}{2\sqrt{3}} [(\boldsymbol{A} \cdot \boldsymbol{C})(\boldsymbol{B} \cdot \boldsymbol{D}) - (\boldsymbol{A} \cdot \boldsymbol{D})(\boldsymbol{B} \cdot \boldsymbol{C})]$$

$$E = \{\{A_1 \otimes B_1\}_2 \otimes \{C_1 \otimes D_1\}_2\}_0 \quad \text{becomes} \quad \frac{1}{6\sqrt{5}} [3 (\boldsymbol{A} \cdot \boldsymbol{C}) (\boldsymbol{B} \cdot \boldsymbol{D}) - 2 (\boldsymbol{A} \cdot \boldsymbol{B}) (\boldsymbol{C} \cdot \boldsymbol{D})$$
$$+ 3 (\boldsymbol{A} \cdot \boldsymbol{D}) (\boldsymbol{B} \cdot \boldsymbol{C})]$$

(35.80)

In the last two equations, I have assumed that the operators commute. If that is not so, then the dot products must be written out in component sums and the components put in the order A,B,C,D as in the original expression.

A vector like $\boldsymbol{r}$ that has no derivatives and also does not simplify like a spin variable so that products of components do not simplify, can be written as spherical harmonics

$$\{...\{\{A_1 \otimes A_1\}_{\ell_2} \otimes A_1\}_{\ell_3}... \otimes A_1\}_{\ell_n m_n} = \sqrt{\frac{4\pi}{2\ell_n + 1}} |\boldsymbol{A}|^n Y_{\ell_n m_n}(\theta, \phi) \prod_{j=2}^{n} C_{10\ell_{j-1}0}^{\ell_n 0}$$

(35.81)

where $\ell_1 = 1$ and $|A|$, θ and ϕ are given by $A_x = |A| \sin \theta \cos \phi$, $A_y = |A| \sin \theta \sin \phi$, $A_z = |A| \cos \theta$.

35.8   Possible forms for an interaction

Since we believe that our physical laws are independent of orientation, we expect the Hamiltonian to be a scalar under rotations. That is, it is an irreducible spherical tensor of rank 0. We might ask what are the possible kinds of interactions we could have between various kinds of particles that are consistent with this. For example, you might want to measure the interaction between protons by scattering experiments[3], and then fit the results to a two-body potential. What are the possible forms that you could take for this proton-proton interaction? We go to the center of mass frame, and the interaction could obviously contain a central potential $v_c(r)$, where $r = |\boldsymbol{r}|$ and $\boldsymbol{r}$ is the relative coordinate between the protons.

---

[3]These are done by accelerating protons and scattering a beam of protons off hydrogen – either hydrogen gas or liquid. At the energies of interest the scattering from electrons is negligible.

When you scatter protons, you find that the interaction can flip the spin of the protons, so there must be interactions with the spins. We already saw that fermions have a spin-orbit interaction, so we might expect that. Spin-orbit interactions contain a momentum as well as position. To simplify our initial analysis, let's see what sort of possible forms we get if we have no momentum terms.

Protons are spin $1/2$ particles. The spin operators $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are vectors, so we can form them into spherical tensors of rank 1,

$$
\begin{aligned}
S_{m=1} &= -\frac{S_x + iS_y}{\sqrt{2}} \\
S_{m=0} &= S_z \\
S_{m=-1} &= \frac{S_x - iS_y}{\sqrt{2}} \, .
\end{aligned}
\tag{35.82}
$$

Spatial vector can be formed into r_m as discussed above, and all other combinations of x, y, and z can be expanded as functions of r times $Y_{\ell m}(\theta, \phi)$. You should easily see that we cannot do this for spin $1/2$, since $\sigma_z^2 = 1$, any combinations of products of spin components can be reduced to a scalar that is independent of spin and a vector. Therefore spin $1/2$ has only a rank 0 and rank 1 possibilities.

We could couple a rank 1 spin operator with a rank 1 space operator. Coupling two rank 1 irreducible spherical tensors is equivalent to forming the dot product of the vectors so this gives a term $\boldsymbol{S}_1 \cdot \boldsymbol{r}$. While allowed by rotational symmetry, it is not invariant under parity. The weak interaction does violate parity conservation, however, the proton-proton interaction is dominated by the strong and electromagnetic interactions which conserve parity. We will therefore not include parity violating terms.

Since one spin operator cannot do anything, we must couple two of them together. We have two spin $1/2$ particles, so we have their two rank 1 irreducible spherical tensors. These can be coupled to form a rank 0, a rank 1 or a rank 2 tensor. Coupling each of these to the corresponding spatial irreducible spherical tensor, again the rank 1 tensor violates parity. The rank 0 tensor is again just the dot product, so we can have a term like $v_s(r)\boldsymbol{S}_1 \cdot \boldsymbol{S}_2$. The rank two spatial tensor is proportional $Y_{2m}(\hat{\boldsymbol{r}})$, which is proportional to $\{r \otimes r\}_2$ so this last term has the form

$$
\begin{aligned}
\{\{r \otimes r\}_2 \otimes \{S_1 \otimes S_2\}_2\}_0 &= \frac{1}{6\sqrt{5}} \left[3\left(\boldsymbol{r} \cdot \boldsymbol{S}_1\right)\left(\boldsymbol{r} \cdot \boldsymbol{S}_2\right) - 2\left(\boldsymbol{r} \cdot \boldsymbol{r}\right)\left(\boldsymbol{S}_1 \cdot \boldsymbol{S}_2\right) + 3\left(\boldsymbol{r} \cdot \boldsymbol{S}_2\right)\left(\boldsymbol{r} \cdot \boldsymbol{S}_1\right) \right] \\
&= \frac{r^2}{3\sqrt{5}} \left[3\boldsymbol{S}_1 \cdot \hat{\boldsymbol{r}}\, \boldsymbol{S}_2 \cdot \hat{\boldsymbol{r}} - \boldsymbol{S}_1 \cdot \boldsymbol{S}_2 \right]
\end{aligned}
\tag{35.83}
$$

and the interaction must have the form

$$
v_t(r) \left[3\boldsymbol{S}_1 \cdot \hat{\boldsymbol{r}}\, \boldsymbol{S}_2 \cdot \hat{\boldsymbol{r}} - \boldsymbol{S}_1 \cdot \boldsymbol{S}_2 \right] \, .
\tag{35.84}
$$

This is called the tensor interaction, and you should recognize it as the form you get from the interaction of two dipoles $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$.

In fact, these three terms are all present and large in the proton proton interaction. In addition, there are terms that depend on $\boldsymbol{p}$ the relative momentum. The first of these you would expect is given by coupling the spin to $\boldsymbol{r}$ to $\boldsymbol{p}$ which gives, from Eq. 35.79 a term proportional to $\boldsymbol{S}_1 \cdot (\boldsymbol{r} \times \boldsymbol{p})$ and a similar term with $\boldsymbol{S}_2$. These do not violate parity, and give a spin orbit interaction. Continuing on adding additional powers of p gives other possibilties. Currently, keeping two powers of $\boldsymbol{p}$ in the interactions allows a fit to the nucleon-nucleion interaction for laboratory kinetic energies below 350 MeV with $\chi^2 \simeq 1$ per data point[4].

35.9 Coupling three (or more) angular momenta

Often we want to couple 3 or more angular momenta. This can either be in angular momentum states, for example studying 2 particles with their spin and orbital angular momenta might require coupling four angular momenta together. We have already seen how irreducible spherical tensor operators can be coupled together and this also can require coupling more angular momenta. Combining the operators and states leads to even more possible couplings.

We already know how to couple three angular momenta. Given the uncoupled basis $|j_1 m_1 j_2 m_2 j_3 m_3\rangle$, we first couple two of the angular momenta together and then couple the third. There are obviously 3 ways to do this

- Couple $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ to give $\boldsymbol{J}_{12}$, and then couple these to $\boldsymbol{J}_3$, that is

$$|jm j_1 j_2(j_{12})j_3\rangle = \sum_{m_1 m_2 m_3 m_{12}} C^{jm}_{j_{12}m_{12}j_3m_3} C^{j_{12}m_{12}}_{j_1m_1j_2m_2} |j_1 m_1 j_2 m_2 j_3 m_3\rangle \qquad (35.85)$$

- Couple $\boldsymbol{J}_2$ and $\boldsymbol{J}_3$ to give $\boldsymbol{J}_{23}$, and then couple these to $\boldsymbol{J}_1$, that is

$$|jm j_1 j_2 j_3(j_{23})\rangle = \sum_{m_1 m_2 m_3 m_{23}} C^{jm}_{j_1m_1j_{23}m_{23}} C^{j_{23}m_{23}}_{j_2m_2j_3m_3} |j_1 m_1 j_2 m_2 j_3 m_3\rangle \qquad (35.86)$$

- Couple $\boldsymbol{J}_1$ and $\boldsymbol{J}_3$ to give $\boldsymbol{J}_{13}$, and then couple these to $\boldsymbol{J}_2$, that is

$$|jm j_1 j_3(j_{13})j_2\rangle = \sum_{m_1 m_2 m_3 m_{13}} C^{jm}_{j_{13}m_{13}j_2m_2} C^{j_{13}m_{13}}_{j_1m_1j_3m_3} |j_1 m_1 j_2 m_2 j_3 m_3\rangle \qquad (35.87)$$

Each of these choices gives a complete orthogonal set of states. We can transform between them as usual. Since all these sets are J^2 and J_z eigenstates, only the terms with the same j and m contribute so that we have, for example

$$|jm j_1 j_2(j_{12})j_3\rangle = \sum_{j_{23}} |jm j_1 j_2 j_3(j_{23})\rangle \langle jm j_1 j_2 j_3(j_{23})|jm j_1 j_2(j_{12})j_3\rangle \qquad (35.88)$$

---

[4]V.G.J. Stoks, R.A.M. Klomp, M.C.M. Rentmeester, and J.J. de Swart, *Partial-wave analytisi of all nucleon-nucleon scattering data below 350 MeV*, Phys. Rev. C **48** 792 (1993). R.B. Wiringa, V.G.J. Stoks, R. Schiavilla, *Accurate nucleon-nucleon potential with charge-independence breaking*, Phys. Rev. C **52**, 38 (1995).

Now apply $J_\pm A^{-1}(j, \pm m)$ to these states. Just as in our Wigner-Eckart theorem proof, the kets change m value, but the coefficients do not. Again, this shows that the coefficients are independent of m. The coefficents just depend on j, j_{12}, j_{23}, j_1, j_2, and j_3. Various ways have been used to write these down. Probably the most popular is the Wigner 6-j symbol

$$\langle jmj_1j_2j_3(j_{23})|j'm'j_1j_2(j_{12})j_3\rangle = \sum_{m_1m_2m_3m_{12}m_{23}} C^{jm}_{j_{12}m_{12}j_3m_3} C^{j_{12}m_{12}}_{j_1m_1j_2m_2} C^{j'm'}_{j_1m_1j_{23}m_{23}} C^{j_{23}m_{23}}_{j_2m_2j_3m_3}$$

$$= \delta_{j,j'}\delta_{m,m'}(-1)^{j_1+j_2+j_3+j}\sqrt{(2j_{12}+1)(2j_{23}+1)}\left\{\begin{matrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{matrix}\right\}$$

$$(35.89)$$

In terms of the 3-j symbols, the 6-j symbol is

$$\left\{\begin{matrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{matrix}\right\} = \sum_{m_1m_2m_{12}m_3mm_{23}} (-1)^{j_3+j+j_{23}-m_3-m-m_{23}} \left(\begin{matrix} j_1 & j_2 & j_{12} \\ m_1 & m_2 & -m_{12} \end{matrix}\right)\left(\begin{matrix} j_1 & j & j_{23} \\ m_1 & -m & m_{23} \end{matrix}\right).$$

$$\cdot\left(\begin{matrix} j_3 & j_2 & j_{23} \\ m_3 & m_2 & -m_{23} \end{matrix}\right)\left(\begin{matrix} j_3 & j & j_{12} \\ -m_3 & m & -m_{12} \end{matrix}\right)$$

$$(35.90)$$

where I have written the sum over all the m quantum numbers, but only three of them are independent, the others are set by the requirement that the sums of the values of the second row of the 3-j must be zero.

The 6-j symbols are therefore real. You can work out their symmetries from the corresponding ones for the Clebsch-Gordan or 3-j symbols.

While the 6-j can be used to transform between these different bases when needed, they are especially useful in simplifying the calculation of reduced matrix elements of products of tensor operators.

For example, we might want to calculate

$$\langle n'jm|\{U_{j_3}\otimes V_{j_2}\}_{j_{23}m_{23}}|nj_1m_1\rangle. \tag{35.91}$$

We can apply the Wigner-Eckart theorem directly, or we can write out the product of the operators, insert a complete set of states, and apply the Wigner-Eckart theorem to the resulting matrix elements. Equating these we get

$$\langle n'jm|\{U_{j_3}\otimes V_{j_2}\}_{j_{23}m_{23}}|nj_1m_1\rangle$$

$$= (-1)^{2j_{23}}\frac{1}{\sqrt{2j+1}}C^{jm}_{j_1m_1j_{23}m_{23}}\langle n'j||\{U_{j_3}\otimes V_{j_2}\}_{j_{23}}||nj_1\rangle$$

$$= \sum_{n''j_{12}m_{12},m_2m_3} C^{j_{23}m_{23}}_{j_3m_3j_2m_2}\langle n'jm|U_{j_3m_3}|n''j_{12}m_{12}\rangle\langle n''j_{12}m_{12}|V_{j_2m_2}|nj_1m_1\rangle$$

$$= \sum_{n''j_{12}m_{12},m_2m_3} C^{j_{23}m_{23}}_{j_3m_3j_2m_2}(-1)^{2j_3}\frac{1}{\sqrt{2j+1}}C^{jm}_{j_{12}m_{12}j_3m_3}\langle n'j||U_{j_3}||n''j_{12}\rangle.$$

$$\cdot(-1)^{2j_2}\frac{1}{\sqrt{2j_{12}+1}}C^{j_{12}m_{12}}_{j_1m_1j_2m_2}\langle n''j_{12}||V_{j_2}||nj_1\rangle \tag{35.92}$$

We can now use the orthogonality of the Clebsch-Gordan coefficients. We can multiply both sides by $C^{jm}_{j_1m_1j_{23}m_{23}}$ and sum over m_1, and m_{23}, giving

$$\frac{(-1)^{2j_{23}}}{\sqrt{2j+1}}\langle n'j||\{U_{j_3}\otimes V_{j_2}\}_{j_{23}}||nj_1\rangle = \sum_{n''j_{12}}\langle n'j||U_{j_3}||n''j_{12}\rangle\langle n''j_{12}||V_{j_2}||nj_1\rangle\,\cdot$$

$$\cdot\frac{(-1)^{2j_3+2j_2}}{\sqrt{(2j+1)(2j_{12}+1)}}\sum_{m_{12}m_1m_2m_3m_{23}}C^{jm}_{j_1m_1j_{23}m_{23}}C^{j_{23}m_{23}}_{j_3m_3j_2m_2}C^{jm}_{j_{12}m_{12}j_3m_3}C^{j_{12}m_{12}}_{j_1m_1j_2m_2}$$

$$(35.93)$$

The sum over Clebsch-Gordan coefficients is in the form of Eq. 35.89 that defined the 6-j coefficients, the second Clebsch-Gordan coefficient above has j_2m_2 and j_3m_3 reversed, so reversing these gives a factor $(-1)^{j_{23}-j_2-j_3}$, and

$$\sum_{m_{12}m_1m_2m_3m_{23}}C^{jm}_{j_1m_1j_{23}m_{23}}C^{j_{23}m_{23}}_{j_3m_3j_2m_2}C^{jm}_{j_{12}m_{12}j_3m_3}C^{j_{12}m_{12}}_{j_1m_1j_2m_2} =$$

$$(-1)^{j_1+j+j_{23}}\sqrt{(2j_{12}+1)(2j_{23}+1)}\left\{\begin{matrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{matrix}\right\}. \qquad (35.94)$$

Eq. 35.93 becomes

$$\langle n'j||\{U_{j_3}\otimes V_{j_2}\}_{j_{23}}||nj_1\rangle = (-1)^{j_1+j+j_{23}}\sqrt{2j_{23}+1}\,\cdot$$

$$\cdot\sum_{n''j_{12}}\left\{\begin{matrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{matrix}\right\}\langle n'j||U_{j_3}||n''j_{12}\rangle\langle n''j_{12}||V_{j_2}||nj_1\rangle$$

$$(35.95)$$

where we drop a factor of $(-1)^{2j_3+2j_2-2j_{23}}$ since $j_2 + j_3 - j_{23}$ is always an integer.

Substituting into the Wigner-Eckart theorem we get

$$\langle n'jm|\{U_{j_3}\otimes V_{j_2}\}_{j_{23}m_{23}}|nj_1m_1\rangle$$

$$= (-1)^{j_1+j-j_{23}}\sqrt{\frac{2j_{23}+1}{2j+1}}C^{jm}_{j_1m_1j_{23}m_{23}}\sum_{n''j_{12}}\left\{\begin{matrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{matrix}\right\}\langle n'j||U_{j_3}||n''j_{12}\rangle\langle n''j_{12}||V_{j_2}||nj_1\rangle$$

$$(35.96)$$

The 6-j symbol allows us to calculate the reduced matrix element of the irreducible spherical tensors formed from U and V using just the reduced matrix elements of U and V separately. The 6-j comes in because in our original matrix element we combined U and V to make a irreducible spherical tensor of rank j_{23}. This is analgous to adding j_2 and j_3 to get j_{23}. Operating $\{U_{j_3}\otimes V_{j_2}\}_{j_{23}}$ on the j_1 state is analogous to adding j_{23} to to j_1 and calculating the matrix element using the bra state $\langle n'jm|$ is analogous to forming the angular momentum j. The other break up operates on the j_1 state with the j_2 irreducible spherical tensor V and takes the matrix element with j_{12}. This is analogous to adding j_1 to j_2 to get

j_{12}. The further operation with the j_3 irreducible spherical tensor U is like adding j_3 to j_{12} to get j. We saw that these two different ways of adding angular momenta were related by a transformation with the 6-j symbol giving the transformation coefficients. Similarly, since the operator matrix elements obey the same algebra, they are also related by the 6-j symbol coefficients.

When adding 4 angular momenta, you can pair j_1 with any of the other 3 angular momenta, and add the remaining pair. The sum of these paired angular momenta give good J^2 states. For example, you can couple j_1 and j_2 to get j_{12}, couple j_3 and j_4 to get j_{34}, and then couple j_{12} and j_{34} to get j. You could also couple j_1 and j_3 to get j_{13}, couple j_2 and j_4 to get j_{24} and couple these together to get j. These two kinds of states are

$$|jmj_1j_2j_3j_4(j_{12})(j_{34})\rangle = \sum_{m_1m_2m_3m_4m_{12}m_{34}} C^{jm}_{j_{12}m_{12}j_{34}m_{34}}C^{j_{12}m_{12}}_{j_1m_1j_2m_2}C^{j_{34}m_{34}}_{j_3m_3j_4m_4}|j_1m_1j_2m_2j_3m_3j_4m_4\rangle$$

$$|jmj_1j_2j_3j_4(j_{13})(j_{24})\rangle = \sum_{m_1m_2m_3m_4m_{13}m_{24}} C^{jm}_{j_{13}m_{13}j_{24}m_{24}}C^{j_{13}m_{13}}_{j_1m_1j_3m_3}C^{j_{24}m_{24}}_{j_2m_2j_4m_4}|j_1m_1j_2m_2j_3m_3j_4m_4\rangle$$

$$(35.97)$$

As before we can transform between them

$$|jmj_1j_2j_3j_4(j_{12})(j_{34})\rangle = \sum_{j_{13},j_{24}}|jmj_1j_2j_3j_4(j_{13})(j_{24})\rangle\langle jmj_1j_2j_3j_4(j_{13})(j_{24})|jmj_1j_2j_3j_4(j_{12})(j_{34})\rangle$$

$$(35.98)$$

where again since both sets have good j and m, only those components contribute. Operating with $J_{\pm}A^{-1}(j, \pm m)$ changes the m values of the kets but not the coefficients, so the coefficients are independent of m. Calculating the coefficients explicitly shows that they are a sum over a product of 6 Clebsch-Gordan coefficients and they are given in terms of the 9-j symbols defined to be

$$\langle jmj_1j_2j_3j_4(j_{12})(j_{34})|j'm'j_1j_2j_3j_4(j_{13})(j_{24})\rangle$$

$$= \sum_{\text{all } m \text{ variables except } m,m'} C^{j_{12}m_{12}}_{j_1m_1j_2m_2}C^{j_{34}m_{34}}_{j_3m_3j_4m_4}C^{jm}_{j_{12}m_{12}j_{34}m_{34}}C^{j_{13}m_{13}}_{j_1m_1j_3m_3}C^{j_{24}m_{24}}_{j_2m_2j_4m_4}C^{j'm'}_{j_{13}m_{13}j_{24}m_{24}}$$

$$= \delta_{jj'}\delta_{mm'}\sqrt{(2j_{12}+1)(2j_{13}+1)(2j_{24}+1)(2j_{34}+1)}\left\{\begin{array}{ccc} j_1 & j_2 & j_{12} \\ j_3 & j_4 & j_{34} \\ j_{13} & j_{24} & j \end{array}\right\} \qquad (35.99)$$

As you should expect, you will find these when you simplify matrix elements formed with three irreducible spherical tensors.

Tensors which can be factored into parts that operate on different angular momenta of a composite system (like spin and orbital angular momenta or the angular momentum of particle 1 and the angular momentum of particle 2) can be evaluated as sums of products of the two parts separately with a 6-j or 9-j symbol factor.

35.10 Problems

Problem 35.1. Calculate the Clebsch-Gordan coefficients using a raising or lowering and orthogonalization algorithm along with the Condon and Shortley phase convention for $j_1 = \ell$ an integer and $j_2 = \frac{1}{2}$. Check your results for $\ell = 1$ against standard tables.

Problem 35.2. Verify by explicit calculation with the Clebsch-Gordan coefficients, the vector identities in the text for the coupling of two vectors, and the scalar couplings of three and four vectors.

Problem 35.3. Calculate the reduced matrix element corresponding to the angular momentum vector $\boldsymbol{J}$, $\langle j'||J_1||j\rangle$.

Problem 35.4. Verify that $\boldsymbol{p} = \hat{\boldsymbol{r}}(\hat{\boldsymbol{r}}\cdot\boldsymbol{p}) - r^{-1}\hat{\boldsymbol{r}}\times\boldsymbol{L}$. Use this and the Wigner-Eckart theorem to find an expression in terms of Clebsch-Gordan coefficients for $\langle n'\ell'm'|p_{m_1}|n\ell m\rangle$ where $p_0 = p_z$, $p_1 = -2^{-1/2}(p_x + ip_y)$, $p_{-1} = 2^{-1/2}(p_x - ip_y)$ and n and n' give the quantum numbers of the nonangular parts of the state. Use this to write an expression for the corresponding derivatives, $\partial_0 = \partial_z$, $\partial_{\pm 1} = \mp 2^{-1/2}(\partial_x \pm i\partial y)$ of $R_{nl}(r)Y_{\ell m}$ as a linear combination of terms consisting of $R_{nl}(r)$ and its derivatives multiplying spherical harmonics.

Problem 35.5. We worked out the transition rate for a hydrogenic n, ℓ m state to decay to an n' ℓ' m' state in the electric-dipole approximation, summing over polarizations to be

$$\Gamma = \frac{4e^2(E_n - E_{n'})^3}{3c^3\hbar^4}|\langle n'\ell'm'|\boldsymbol{r}|n\ell m\rangle|^2 \tag{35.100}$$

a. Sum over all m' for the final state to find the total rate. Your answer should be independent of m and equivalent to

$$\Gamma = \frac{4e^2(E_n - E_{n'})^3}{3c^3\hbar^4}\frac{1}{2\ell + 1}|\langle n\ell||r_1||n'\ell'\rangle|^2 . \tag{35.101}$$

b. If the hydrogenic bound states are written as

$$\langle r\theta\phi|n\ell m\rangle = R_{n\ell}(r)Y_{\ell m}(\theta, \phi) \tag{35.102}$$

calculate an expression for the reduced matrix element in terms of an integral over $R_{n'\ell'}$ and $R_{n\ell}$. Your answer should be

$$\langle n\ell||r_1||n'\ell'\rangle = \sqrt{2\ell' + 1}C^{\ell'0}_{\ell,0,1,0}\int_0^\infty drr^3 R^*_{n\ell}(r)R_{n'\ell'}(r) \tag{35.103}$$

Problem 35.6. An electronic state of a central potential makes electric dipole transitions from an $\ell = 2$, $m = 0$ state to one of three $\ell = 1$ states with different m values but the same radial quantum numbers. Calculate the branching ratios for the 3 final m values of $-1, 0, 1$ first using the Wigner-Eckart theorem and then by directly calculating the matrix elements. What is the angular distribution of the photons for each of the three transitions?

Problem 35.7.   a. Calculate the reduced matrix element for the Pauli matrices (i.e. the operator $\boldsymbol{\sigma} = 2\boldsymbol{S}/\hbar$ for spin 1/2.) Your result should be

$$\langle \tfrac{1}{2}||\boldsymbol{\sigma}||\tfrac{1}{2}\rangle = \sqrt{6} \tag{35.104}$$

b. In pion-nucleon scattering, the quarks in the nucleon can be excited so that it becomes a delta particle (or resonance). The delta has spin 3/2. In a nonrelativistic theory the transition between a nucleon and a delta is described by a transition spin operator $\boldsymbol{\Sigma}$ which has matrix elements between the nucleon spin 1/2 states and the delta spin 3/2 states, and the interaction hamiltonian contains the term

$$(\boldsymbol{\Sigma} + \boldsymbol{\Sigma}^+) \cdot \boldsymbol{\nabla}\phi_\pi \, . \tag{35.105}$$

ϕ_π is a pion field which has a form identical to our vector potential operator $\boldsymbol{A}$, except it does not have polarization vectors since the pion is a pseudoscalar, i.e. spin 0, particle, and $\hbar\omega_k = \sqrt{\hbar^2 k^2 c^2 + m_\pi^2 c^2}$. However, the only part of this that we care about here is that since the Hamiltonian must be rotationally invariant, and the gradient transforms like a vector, the transition spin operators must also be vectors. They can therefore be written as a spherical tensor of rank 1.

Show that the Wigner Eckart theorem requires that the matrix representation of $\langle \tfrac{3}{2}m|\boldsymbol{\Sigma}|\tfrac{1}{2}m'\rangle$ is proportional to

$$
\begin{pmatrix}
\langle \frac{3}{2},\frac{3}{2}|\boldsymbol{\Sigma}|\frac{1}{2},\frac{1}{2}\rangle & \langle \frac{3}{2},\frac{3}{2}|\boldsymbol{\Sigma}|\frac{1}{2},-\frac{1}{2}\rangle \\
\langle \frac{3}{2},\frac{1}{2}|\boldsymbol{\Sigma}|\frac{1}{2},\frac{1}{2}\rangle & \langle \frac{3}{2},\frac{1}{2}|\boldsymbol{\Sigma}|\frac{1}{2},-\frac{1}{2}\rangle \\
\langle \frac{3}{2},-\frac{1}{2}|\boldsymbol{\Sigma}|\frac{1}{2},\frac{1}{2}\rangle & \langle \frac{3}{2},-\frac{1}{2}|\boldsymbol{\Sigma}|\frac{1}{2},-\frac{1}{2}\rangle \\
\langle \frac{3}{2},-\frac{3}{2}|\boldsymbol{\Sigma}|\frac{1}{2},\frac{1}{2}\rangle & \langle \frac{3}{2},-\frac{3}{2}|\boldsymbol{\Sigma}|\frac{1}{2},-\frac{1}{2}\rangle
\end{pmatrix}
\propto
\begin{pmatrix}
-\frac{\hat{\boldsymbol{x}}-i\hat{\boldsymbol{y}}}{\sqrt{2}} & 0 \\
\sqrt{\frac{2}{3}}\hat{\boldsymbol{z}} & -\frac{1}{\sqrt{3}}\frac{\hat{\boldsymbol{x}}-i\hat{\boldsymbol{y}}}{\sqrt{2}} \\
\frac{1}{\sqrt{3}}\frac{\hat{\boldsymbol{x}}+i\hat{\boldsymbol{y}}}{\sqrt{2}} & \sqrt{\frac{2}{3}}\hat{\boldsymbol{z}} \\
0 & \frac{\hat{\boldsymbol{x}}+i\hat{\boldsymbol{y}}}{\sqrt{2}}
\end{pmatrix} .
$$
$$\tag{35.106}$$

Find the reduced matrix element if the matrix representation of $\boldsymbol{\Sigma}$ is given by the right hand side of Eq. 35.106.

Problem 35.8. The 6-j symbols are invariant under interchange of any two columns or interchange of the upper and lower components of any two columns. Show explicitly from the expansion of the 6-j symbols in terms of 3-j symbols, using the 3-j symbol symmetries, that these two examples of this symmetry hold

$$\left\{ \begin{matrix} a & b & c \\ d & e & f \end{matrix} \right\} = \left\{ \begin{matrix} b & a & c \\ e & d & f \end{matrix} \right\} = \left\{ \begin{matrix} d & e & c \\ a & b & f \end{matrix} \right\} .$$
$$\tag{35.107}$$

Problem 35.9.   a. Convert the Clebsch-Gordan summation for the 9-j symbols into 3-j symbols and verify that

$$
\begin{Bmatrix} a & b & c \\ d & e & f \\ g & h & j \end{Bmatrix} = \sum_{\text{all } m \text{ variables}} \begin{pmatrix} a & b & c \\ m_a & m_b & m_c \end{pmatrix} \begin{pmatrix} d & e & f \\ m_d & m_e & m_f \end{pmatrix} \begin{pmatrix} g & h & j \\ m_g & m_h & m \end{pmatrix}.
$$

$$
\cdot \begin{pmatrix} a & d & g \\ m_a & m_d & m_g \end{pmatrix} \begin{pmatrix} b & e & h \\ m_b & m_e & m_h \end{pmatrix} \begin{pmatrix} c & f & j \\ m_c & m_f & m \end{pmatrix} \tag{35.108}
$$

Since the first three 3-j symbols are constructed from the rows of the 9-j and the second three from the columns, the symmetries of the 3-j immediately show that the 9-j is invariant under even (cyclic) permutations of the rows or of the columns, and will give a factor of $(-1)^S$ where S is the sum of all 9 of the j values for an odd permutation of the rows or of the columns.

b. If one of the angular momenta in the 9-j symbol is 0, it collapses to a 6-j symbol. Show that

$$
\begin{Bmatrix} a & b & c \\ d & e & f \\ g & h & 0 \end{Bmatrix} = \delta_{cf}\delta_{gh}\frac{(-1)^{b+c+d+g}}{\sqrt{(2c+1)(2g+1)}} \begin{Bmatrix} a & b & c \\ e & d & g \end{Bmatrix}. \tag{35.109}
$$

Problem 35.10. The irreducible spherical tensor operator $U_{\ell_1 m_1}(1)$ operates on the subsystem described by n_1 and j_1 while the irreducible spherical tensor operator $V_{\ell_2 m_2}(2)$ operates on the subsystem described by n_2 and j_2.

a. Show that the matrix elements of the irreducible spherical tensor products in the coupled scheme become

$$
\langle n_1' n_2' j' m' j_1' j_2' | \{U_{\ell_1}(1) \otimes V_{\ell_2}(2)\}_{\ell m_\ell} | n_1 n_2 j m j_1 j_2 \rangle
$$

$$
= (-1)^{2\ell}\sqrt{(2\ell+1)(2j+1)}C_{jm\ell m_\ell}^{j'm'} \begin{Bmatrix} \ell_1 & \ell_2 & \ell \\ j_1' & j_2' & j' \\ j_1 & j_2 & j \end{Bmatrix} \langle n_1' j_1' || U_{\ell_1}(1) || n_1 j_1 \rangle \langle n_2' j_2' || V_{\ell_2}(2) || n_2 j_2 \rangle
$$

$$
\tag{35.110}
$$

b. Show that in the special case where $\ell_1 = \ell_2$ and $\ell = m_\ell = 0$, that this becomes

$$
\langle n_1' n_2' j' m' j_1' j_2' | \{U_{\ell_1}(1) \otimes V_{\ell_1}(2)\}_{00} | n_1 n_2 j m j_1 j_2 \rangle
$$

$$
= \delta_{jj'}\delta_{m'm}\frac{(-1)^{j+j_1+j_2'+\ell_1}}{\sqrt{2\ell_1+1}} \begin{Bmatrix} j_1' & j_1 & \ell_1 \\ j_2 & j_2' & j \end{Bmatrix} \langle n_1' j_1' || U_{\ell_1}(1) || n_1 j_1 \rangle \langle n_2' j_2' || V_{\ell_1}(2) || n_2 j_2 \rangle
$$

$$
\tag{35.111}
$$

411

Problem 35.11. You decide to calculate the bound state electronic energy levels of the helium atom. The first basis set you choose has basis states given by a product of orthogonal one-electron orbitals for each electron. These are of the form $R_{n\ell}(r)Y_{\ell m}(\theta, \phi)$

$$\langle \boldsymbol{r}_1 \boldsymbol{r}_2 | n_1 n_2 \ell_1 m_1 \ell_2 m_2 \rangle = R_{n_1\ell_1}(r_1)Y_{\ell_1 m_1}(\theta_1, \phi_1)R_{n_2\ell_2}(r_2)Y_{\ell_2 m_2}(\theta_2, \phi_2),$$

(35.112)

with

$$\int_0^\infty dr\, r^2 R_{n'\ell}^*(r)R_{n\ell}(r) = \delta_{n,n'}.$$

(35.113)

Since the eigenstates of the helium atom will have good total orbital angular momentum you combine these to get good j and m basis states,

$$|n_1 n_2 j m \ell_1 \ell_2 \rangle = \sum_{m_1 m_2} C_{\ell_1 m_1 \ell_2 m_2}^{jm} |n_1 n_2 \ell_1 m_1 \ell_2 m_2 \rangle.$$

(35.114)

Show that the nonzero matrix elements of the Hamiltonian

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \frac{2e^2}{r_1} - \frac{2e^2}{r_2} + \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}$$

(35.115)

are given by

$$\begin{aligned}
&\langle n_1' n_2' j m \ell_1' \ell_2' | H | n_1 n_2 j m \ell_1 \ell_2 \rangle \\
&= \delta_{n_2, n_2'} \delta_{\ell_1, \ell_1'} \delta_{\ell_2, \ell_2'} \int_0^\infty dr\, r^2 R_{n_1'\ell_1}^*(r) \left[-\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{\hbar^2 \ell_1(\ell_1+1)}{2mr^2} - \frac{2e^2}{r} \right] R_{n_1\ell_1}(r) \\
&+ \delta_{n_1, n_1'} \delta_{\ell_1, \ell_1'} \delta_{\ell_2, \ell_2'} \int_0^\infty dr\, r^2 R_{n_2'\ell_2}^*(r) \left[-\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{\hbar^2 \ell_2(\ell_2+1)}{2mr^2} - \frac{2e^2}{r} \right] R_{n_2\ell_2}(r) \\
&+ 4\pi e^2 \sum_\ell I_\ell(n_1, \ell_1, n_1', \ell_1', n_2, \ell_2, n_2', \ell_2') \frac{(-1)^{j+\ell_1+\ell_2'}}{2\ell+1} \begin{Bmatrix} \ell_1' & \ell_1 & \ell \\ \ell_2 & \ell_2' & j \end{Bmatrix} \langle \ell_1' || Y_\ell || \ell_1 \rangle \langle \ell_2' || Y_\ell || \ell_2 \rangle
\end{aligned}$$

(35.116)

where

$$\begin{aligned}
I_\ell(n_1, \ell_1, n_1', \ell_1', n_2, \ell_2, n_2', \ell_2') &= \int_0^\infty dr_1 r_1^2 \int_0^\infty dr_2 r_2^2 R_{n_1'\ell_1'}^*(r_1)R_{n_1\ell_1}(r_1)R_{n_2'\ell_2'}^*(r_2)R_{n_2\ell_2}(r_2)\frac{r_<^\ell}{r_>^{\ell+1}} \\
&= \int_0^\infty dr_1 R_{n_1'\ell_1'}^*(r_1)R_{n_1\ell_1}(r_1) \left[r_1^{1-\ell} \int_0^{r_1} dr_2 r_2^{2+\ell} R_{n_2'\ell_2'}^*(r_2)R_{n_2\ell_2}(r_2) \right. \\
&\quad \left. + r_1^{\ell+2} \int_{r_1}^\infty dr_2 r_2^{1-\ell} R_{n_2'\ell_2'}^*(r_2)R_{n_2\ell_2}(r_2) \right]
\end{aligned}$$

(35.117)

and the sum over ℓ is restricted by the triangle inequalities $|\ell_1 - \ell_1'| \le \ell \le \ell_1 + \ell_1'$, $|\ell_2 - \ell_2'| \le \ell \le \ell_2 + \ell_2'$.

Chapter 36

Spin dependent scattering

36.1 Spin 1/2 scattering from spin 0 particles

We follow Baym's example in his chapter 15 (Addition of angular momentum). He looks at scattering of a pion from a proton. Initially, the two particles are separated, so it is natural to use the separated angular momentum eigenstates which are not entangled, $|n\ell m_\ell s m_s\rangle$. That is, the wave function would depend on the center of mass and relative coordinates and the spin of the proton. In the center of mass frame, we would have the relative coordinate which can be separated into radial and angular components, and the proton spin. n describes the nonangular quantum numbers (the radial part of the relative coordinates), ℓ and m_ℓ the orbital angular momentum for the relative coordinate, and s, m_s give the proton spin (up or down).

The interaction conserves total angular momentum $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, but not $\boldsymbol{L}$ and $\boldsymbol{S}$ separately. The proton has spin $1/2$, so S^2 is conserved since we do not destroy the proton. The interaction, H_{int}, commutes with J_z, J^2, and S^2, and these all commute with each other. From the Wigner-Eckart theorem, since H_{int} is a scalar, i.e. an irreducible spherical tensor of rank 0, we get

$$
\begin{aligned}
\langle n'j'm'\ell's|H_{\text{int}}|njm\ell s\rangle &= \frac{C^{j'm'}_{jm00}}{\sqrt{2j'+1}}\langle n'j'\ell's||H_{\text{int}}||nj\ell s\rangle \\
&= \frac{\delta_{jj'}\delta_{mm'}}{\sqrt{2j+1}}\langle n'j\ell's||H_{\text{int}}||nj\ell s\rangle
\end{aligned}
$$

(36.1)

Experimentally, parity is conserved. The parity of a spherical harmonic is easily seen to be $(-1)^\ell$, so when we write $\langle r\theta\phi|n\ell m_\ell\rangle = R_n(r)Y_{\ell m}(r)$, we see that the parity of our states is also $(-1)^\ell$. Now since $s = 1/2$, the possible ℓ values are $\ell = j \pm \frac{1}{2}$. Flipping the spin can at most change ℓ' so that $\ell' = \ell \pm 1$ or $\ell' = 0$. However, with parity conserved, the two states must have the same parity and the matrix elements are nonzero only if $\ell' = \ell$. (This, of course, will give an experimental test of parity conservation in strong interactions.)

Therefore

$$\langle n'j'm'\ell's|H_{\text{int}}|njm\ell s\rangle \;\; = \;\; \frac{\delta_{jj'}\delta_{mm'}\delta_{\ell,\ell'}}{\sqrt{2j+1}}\langle n'j\ell s||H_{\text{int}}||nj\ell s\rangle$$

$$(36.2)$$

The angular part of the reduced matrix element is proportional to Baym's $B(\ell, j)$.

For the Born approximation we will need the matrix element (for either the lowest order T matrix to calculate the scattering amplitude, or for Fermi's golden rule)

$$
\begin{aligned}
\langle n'\ell'm'_\ell sm'_s|H_{\text{int}}|n\ell m_\ell sm_s\rangle \;\; &= \;\; \sum_{j,m,j'm'} \langle n'\ell'm'_\ell sm'_s|n'j'm'\ell's\rangle\langle n'j'm\ell's|H_{\text{int}}|njm\ell s\rangle\langle njm\ell s|n\ell m_\ell sm_s\rangle \\[2mm]
&= \;\; \sum_{j,m,j'm'} C^{j'm'}_{\ell'm'_\ell sm'_s}C^{jm}_{\ell m_\ell sm_s}\langle n'j'm\ell's|H_{\text{int}}|njm\ell s\rangle \\[2mm]
&= \;\; \sum_{j,m,j'm'} C^{j'm'}_{\ell'm'_\ell sm'_s}C^{jm}_{\ell m_\ell sm_s}\frac{\delta_{jj'}\delta_{mm'}\delta_{\ell,\ell'}}{\sqrt{2j+1}}\langle n'j\ell s||H_{\text{int}}||nj\ell s\rangle \\[2mm]
&= \;\; \sum_{j,m} C^{jm}_{\ell m'_\ell sm'_s}C^{jm}_{\ell m_\ell sm_s}\frac{1}{\sqrt{2j+1}}\langle n'j\ell s||H_{\text{int}}||nj\ell s\rangle \\[2mm]
&= \;\; C^{\ell+\frac{1}{2},m_\ell+m_s}_{\ell m'_\ell \frac{1}{2}m'_s}C^{\ell+\frac{1}{2},m_\ell+m_s}_{\ell m_\ell \frac{1}{2}m_s}\frac{\langle n'j=\ell+\frac{1}{2}\ell s||H_{\text{int}}||nj=\ell+\frac{1}{2},\ell s\rangle}{\sqrt{2j+1}} \\[2mm]
&\quad +C^{\ell-\frac{1}{2},m_\ell+m_s}_{\ell m'_\ell \frac{1}{2}m'_s}C^{\ell-\frac{1}{2},m_\ell+m_s}_{\ell m_\ell \frac{1}{2}m_s}\frac{\langle n'j=\ell-\frac{1}{2}\ell s||H_{\text{int}}||nj-\ell+\frac{1}{2},\ell s\rangle}{\sqrt{2j+1}} \quad (36.3)
\end{aligned}
$$

Chapter 37

Euler angles and Wigner rotations

37.1 Euler angles and the Rotation operator

As we have seen, rotations are generated by the angular momentum. Operating on a state

$$|\psi'\rangle = e^{-\frac{i}{\hbar}\boldsymbol{J}\cdot\boldsymbol{\alpha}}|\psi\rangle \tag{37.1}$$

gives $|\psi'\rangle$ which is rotated in J space (for example r space for orbital angular momenta, and spin space for spins) in the positive direction by an angle $|\boldsymbol{\alpha}|$ around the axis $\hat{\boldsymbol{\alpha}}$. Alternatively, since rotating the state in one direction is equivalent to rotating the coordinate system in the other, this same operator can be thought of as rotating the coordinate system in a negative direction around $\hat{\boldsymbol{\alpha}}$.

One way to check the sign to make sure you have it right is to look at the rotation of a good $|m\rangle$ state around z. In the usual coordinate system $\langle\phi|m\rangle = e^{im\phi}$, so the rotated state is

$$\langle\phi|\psi'\rangle = \langle\phi|e^{-\frac{i}{\hbar}L_z\alpha}|m\rangle = e^{-im\alpha}\langle\phi|m\rangle = e^{im(\phi-\alpha)} \tag{37.2}$$

The new value of the state at an angle ϕ is the old value at $\phi - \alpha$, so the state has been rotated in the positive direction around alpha. Alternatively, the old coordinate value ϕ has now been replaced by a new coordinate value $\phi' = \phi - \alpha$.

In order to specify a general orientation of a rigid body, Euler developed his angles. These then correspond to a general coordinate rotation and therefore, in quantum mechanics, a general state rotation. There are several conventions, but all correspond to rotations around coordinate axes. One common choice is the one given in Baym:

1. rotate α around the z axis.

2. rotate β around the *new y* axis (call it y').

3. rotate γ around the *new z* axis (call it z').

The quantum mechanical operator is

$$R = e^{-\frac{i}{\hbar}J'_z\gamma}e^{-\frac{i}{\hbar}J'_y\beta}e^{-\frac{i}{\hbar}J_z\alpha} \ . \tag{37.3}$$

The original reason to use the new axes was because these were developed in classical mechanics where the rotations are tied to the classical rigid body, and the rotations are done in the body frame. If the rigid body is initially specified in its principal axis frame, each rotation is about a principal axis. In quantum mechanics, it is usually easier to make all of the rotations around the original lab frame axes. There are two ways to transform to these axes. The fast way is to realize that you can think of rotating the coordinates starting with the final rotated state and the lab coordinates. If you rotate these coordinates exactly like you rotated the state, they would be in the same orientation with the state as the original lab coordinates and the original state. Therefore this transformation is the inverse of our state rotation,

$$R^{-1} = \left[e^{\frac{i}{\hbar}J_z\gamma} e^{\frac{i}{\hbar}J_y\beta} e^{\frac{i}{\hbar}J_z\alpha} \right] \tag{37.4}$$

where we use a positive sign in the exponents since these are coordinate rather than state rotations. Or,

$$R = \left[e^{\frac{i}{\hbar}J_z\gamma} e^{\frac{i}{\hbar}J_y\beta} e^{\frac{i}{\hbar}J_z\alpha} \right]^{-1} = e^{-\frac{i}{\hbar}J_z\alpha} e^{-\frac{i}{\hbar}J_y\beta} e^{-\frac{i}{\hbar}J_z\gamma} . \tag{37.5}$$

If this seems too tricky, we can follow Baym and calculate in a completely straightforward way. Starting with Eq. 37.3, we see that we need to first calculate J'_y in terms of the space fixed angular momenta. But J'_y is just the J_y operator rotated to the new coordinates, that is

$$
\begin{aligned}
J'_y &= e^{-\frac{i}{\hbar}J_z\alpha} J_y e^{\frac{i}{\hbar}J_z\alpha} \\
e^{-\frac{i}{\hbar}J'_y\beta} &= 1 - \frac{i}{\hbar}J'_y\beta + \left(-\frac{i}{\hbar}\right)^2 \frac{1}{2!} J'^2_y \beta^2 + ... \\
&= 1 - \frac{i}{\hbar}e^{-\frac{i}{\hbar}J_z\alpha} J_y\beta e^{\frac{i}{\hbar}J_z\alpha} + \left(-\frac{i}{\hbar}\right)^2 \frac{1}{2!} e^{-\frac{i}{\hbar}J_z\alpha} J_y\beta e^{\frac{i}{\hbar}J_z\alpha} e^{-\frac{i}{\hbar}J_z\alpha} J_y\beta e^{\frac{i}{\hbar}J_z\alpha} + ... \\
&= e^{-\frac{i}{\hbar}J_z\alpha} \left[1 - \frac{i}{\hbar}J_y\beta + \left(-\frac{i}{\hbar}\right)^2 \frac{1}{2!} J^2_y\beta^2 + ... \right] e^{\frac{i}{\hbar}J_z\alpha} \\
&= e^{-\frac{i}{\hbar}J_z\alpha} e^{-\frac{i}{\hbar}J_y\beta} e^{\frac{i}{\hbar}J_z\alpha} . \tag{37.6}
\end{aligned}
$$

Similarly

$$
\begin{aligned}
J'_z &= e^{-\frac{i}{\hbar}J'_y\beta} e^{-\frac{i}{\hbar}J_z\alpha} J_z e^{\frac{i}{\hbar}J_z\alpha} e^{\frac{i}{\hbar}J'_y\beta} \\
&= e^{-\frac{i}{\hbar}J'_y\beta} J_z e^{\frac{i}{\hbar}J'_y\beta} \\
&= e^{-\frac{i}{\hbar}J_z\alpha} e^{-\frac{i}{\hbar}J_y\beta} e^{\frac{i}{\hbar}J_z\alpha} J_z e^{-\frac{i}{\hbar}J_z\alpha} e^{\frac{i}{\hbar}J_y\beta} e^{\frac{i}{\hbar}J_z\alpha} \\
&= e^{-\frac{i}{\hbar}J_z\alpha} e^{-\frac{i}{\hbar}J_y\beta} J_z e^{\frac{i}{\hbar}J_y\beta} e^{\frac{i}{\hbar}J_z\alpha} \\
e^{-\frac{i}{\hbar}J'_z\gamma} &= e^{-\frac{i}{\hbar}J_z\alpha} e^{-\frac{i}{\hbar}J_y\beta} e^{-\frac{i}{\hbar}J_z\gamma} e^{\frac{i}{\hbar}J_y\beta} e^{\frac{i}{\hbar}J_z\alpha} \tag{37.7}
\end{aligned}
$$

416

Putting this together,

$$
\begin{aligned}
R &= e^{-\frac{i}{\hbar}J_z'\gamma}e^{-\frac{i}{\hbar}J_y'\beta}e^{-\frac{i}{\hbar}J_z\alpha} \\
&= \left[e^{-\frac{i}{\hbar}J_z\alpha}e^{-\frac{i}{\hbar}J_y\beta}e^{-\frac{i}{\hbar}J_z\gamma}e^{\frac{i}{\hbar}J_y\beta}e^{\frac{i}{\hbar}J_z\alpha} \right] \left[e^{-\frac{i}{\hbar}J_z\alpha}e^{-\frac{i}{\hbar}J_y\beta}e^{\frac{i}{\hbar}J_z\alpha} \right] e^{-\frac{i}{\hbar}J_z\alpha} \\
&= e^{-\frac{i}{\hbar}J_z\alpha}e^{-\frac{i}{\hbar}J_y\beta}e^{-\frac{i}{\hbar}J_z\gamma} ,
\end{aligned}
\tag{37.8}
$$

exactly as before.

37.2 Wigner rotation matrices

When operating on good angular momentum states $|jm\rangle$, since each of the components of $\boldsymbol{J}$ commute with J^2, the rotation cannot change the J^2 eigenvalue,

$$
J^2 R|jm\rangle = RJ^2|jm\rangle = \hbar^2 j(j+1)R|jm\rangle .
\tag{37.9}
$$

Therefore $R|jm\rangle$ must be a linear combination of $|jm'\rangle$, and we write the coefficients as $D^j_{m'm}(\alpha,\beta,\gamma)$,

$$
R|jm\rangle = \sum_{m'} |jm'\rangle\langle jm'|R|jm\rangle = \sum_{m'} |jm'\rangle D^j_{m'm}(\alpha,\beta,\gamma)
\tag{37.10}
$$

where

$$
\begin{aligned}
D^j_{m'm}(\alpha,\beta,\gamma) &= \langle jm'|e^{-\frac{i}{\hbar}J_z\alpha}e^{-\frac{i}{\hbar}J_y\beta}e^{-\frac{i}{\hbar}J_z\gamma}|jm\rangle \\
&= e^{-im'\alpha-im\gamma}\langle jm'|e^{-\frac{i}{\hbar}J_y\beta}|jm\rangle \\
&= e^{-im'\alpha-im\gamma}d^j_{m'm}(\beta) .
\end{aligned}
\tag{37.11}
$$

The small d function gives the rotation around y.

We can work out the values of the d function from the commutation relations for $\boldsymbol{J}$. For example, for spin $1/2$, we already know

$$
e^{-\frac{i}{\hbar}S_y\beta} = e^{-i\sigma_y\frac{\beta}{2}} = \cos\frac{\beta}{2} - i\sigma_y\sin\frac{\beta}{2}
\tag{37.12}
$$

The matrix representation is then

$$
\begin{pmatrix} d^{\frac{1}{2}}_{\frac{1}{2},\frac{1}{2}} & d^{\frac{1}{2}}_{\frac{1}{2},-\frac{1}{2}} \\ d^{\frac{1}{2}}_{-\frac{1}{2},\frac{1}{2}} & d^{\frac{1}{2}}_{-\frac{1}{2},-\frac{1}{2}} \end{pmatrix} = \begin{pmatrix} \cos\frac{\beta}{2} & -\sin\frac{\beta}{2} \\ \sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix} .
\tag{37.13}
$$

This agrees with the particle data group table.

There are general expressions for the d function.

For integer ℓ, the spherical harmonics are $Y_{\ell m}(\theta,\phi) = \langle\theta\phi|\ell m\rangle$. If we evaluate at $\theta=0$,

$$
Y_{\ell m}(0,0) = \langle 00|\ell m\rangle = \delta_{m,0}\sqrt{\frac{2\ell+1}{4\pi}} .
\tag{37.14}
$$

The rotation operator on the $\langle\theta\phi|$ rotates that state, while operating on the angular momentum states gives the Wigner matrix

$$
\begin{aligned}
\langle\theta = 0\phi = 0|R|\ell m\rangle &= \langle -\beta - \gamma|\ell m\rangle = \sum_{m'}\langle\theta = 0\phi = 0|\ell m'\rangle\langle\ell m'|R|\ell m\rangle = \sqrt{\frac{2\ell + 1}{4\pi}}d^\ell_{0m}(\beta)e^{-im\gamma} \\
(-1)^m Y^*_{\ell m}(\beta, \gamma) &= \sqrt{\frac{2\ell + 1}{4\pi}}D^\ell_{0m}(\alpha, \beta, \gamma) \tag{37.15}
\end{aligned}
$$

Similarly

$$
\begin{aligned}
D^\ell_{m,0}(\alpha, \beta, \gamma) &= (-1)^m\sqrt{\frac{4\pi}{2\ell + 1}}Y_{\ell,-m}(\beta, \alpha) = \sqrt{\frac{4\pi}{2\ell + 1}}Y^*_{\ell,m}(\beta, \alpha) \\
D^\ell_{0,m}(\alpha, \beta, \gamma) &= \sqrt{\frac{4\pi}{2\ell + 1}}Y_{\ell,-m}(\beta, \gamma) = (-1)^m\sqrt{\frac{4\pi}{2\ell + 1}}Y^*_{\ell,m}(\beta, \gamma) \tag{37.16}
\end{aligned}
$$

We can also look at rotations of other spherical harmonic expressions. For example if we rotate the z axis direction to θ', ϕ' with a rotation around the original z axis of α, we can write in the new coordinates

$$
\begin{aligned}
Y_{\ell m}(\gamma, \delta) &= \langle\gamma\delta|\ell m\rangle \\
&= \langle\gamma\delta|R^{-1}(\phi', \theta', \alpha)R(\phi', \theta', \alpha)|\ell m\rangle \tag{37.17}
\end{aligned}
$$

Calling the new directional angles for $\gamma\delta$ after rotation θ, ϕ, this shows that

$$
\begin{aligned}
Y_{\ell m}(\gamma, \delta) &= \langle\theta\phi|R(\phi', \theta', \alpha)|\ell m\rangle \\
&= \sum_{m'}\langle\theta\phi|\ell m'\rangle D^\ell_{m'm}(\phi', \theta', \alpha) \\
&= \sum_{m'}Y_{\ell m}(\theta, \phi)D^\ell_{m'm}(\phi', \theta', \alpha) \tag{37.18}
\end{aligned}
$$

For the special case where $m = 0$ this is

$$
Y_{\ell 0}(\gamma, \beta) = \sqrt{\frac{4\pi}{2\ell + 1}}\sum_{m'}Y_{\ell m'}(\theta\phi)Y^*_{\ell m'}(\theta', \phi') \tag{37.19}
$$

or

$$
P_\ell(\cos\gamma) = \frac{4\pi}{2\ell + 1}\sum_m Y_{\ell m}(\theta\phi)Y^*_{\ell m}(\theta', \phi') \tag{37.20}
$$

the spherical harmonic addition theorem, since γ was the angle between the z axis and the original direction and after rotation this is the angle between the unprimed and primed directions,

$$
\cos\gamma = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi - \phi'). \tag{37.21}
$$

418

37.3 Projections to J=0

If we have a state and average over all orientations, we must get a state that is invariant under rotations. That is, we must get a state that has $J = 0$. The Wigner rotation operator averaged over all possible rotations is therefore a projector onto $J = 0$. To see what the integration measure needs to be, we can first look at rotating a unit vector along z to θ, ϕ. Here we know that we first rotate around y by θ and then rotate around z by ϕ. To integrate over the surface of the sphere, we integrate over $d\cos\theta$ from -1 to 1 (or $\sin\theta d\theta$ from 0 to π) and ϕ from 0 to 2π or equivalent. To include the third Euler angle, we just have another integration over it from 0 to 2π. There is one caveat. Recall that when we rotated a spin 1/2 state by 2π, we got -1 instead of 1. So for spin 1/2 or any half integer spin, we need to integrate α from 0 to 4π instead of 0 to 2π (or we can integrate γ from 0 to 4π). Integrating integer spin from 0 to 4π just doubles the result, so we can average over all orientations for arbitrary angular momentum by integrating over

$$\frac{1}{16\pi^2} \int_0^{4\pi} d\alpha \int_{-1}^{1} d\cos\beta \int_0^{2\pi} d\gamma \tag{37.22}$$

where the factor out front makes the integration over 1 equal to 1. For integer j, you can change the 4π limits to 2π, and divide by $8\pi^2$ if you prefer. Note that for half integer j the result is always zero. Therefore many authors use the integer equation and require that j is integer for a nonzero result.

We can therefore write the projector as

$$\delta_{j,0}\delta_{m,0}|00\rangle = \frac{1}{16\pi^2} \int_0^{4\pi} d\alpha \int_{-1}^{1} d\cos\beta \int_0^{2\pi} d\gamma R(\alpha, \beta, \gamma)|jm\rangle \tag{37.23}$$

or operating from the left with $\langle jm'|$,

$$\frac{1}{16\pi^2} \int_0^{4\pi} d\alpha \int_{-1}^{1} d\cos\beta \int_0^{2\pi} d\gamma D^j_{m'm}(\alpha, \beta, \gamma) = \delta_{m'0}\delta_{m0}\delta_{j,0} \tag{37.24}$$

37.4 Rotations for angular momentum addition

For $\boldsymbol{J} = \boldsymbol{J}_1 + \boldsymbol{J}_2$, with $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ commuting, we get

$$\langle jmj_1j_2|R(\alpha, \beta, \gamma)|jm'j_1j_2\rangle = D^j_{mm'}(\alpha, \beta, \gamma)$$
$$= \sum_{m_1m_2m_1'm_2'} \langle jmj_1j_2|j_1m_1j_2m_2\rangle\langle j_1m_1j_2m_2|R(\alpha, \beta, \gamma)|j_1m_1'j_2m_2'\rangle\langle j_1m_1'j_2m_2'|jm'j_1j_2\rangle$$
$$= \sum_{m_1m_2m_1'm_2'} C^{jm}_{j_1m_1j_2m_2} C^{jm'}_{j_1m_1'j_2m_2'} D^{j_1}_{m_1m_1'}(\alpha, \beta, \gamma) D^{j_2}_{m_2m_2'}(\alpha, \beta, \gamma).$$

$$\tag{37.25}$$

In principle, this allows us to calculate any $D^j_{mm'}$ recursively from the spin 1/2 matrix and Clebsch-Gordan coefficients that we have already worked out. Writing the D matrices in terms of the d matrices, we find

$$d^j_{mm'}(\beta) = \sum_{m_1 m_2 m'_1 m'_2} C^{jm}_{j_1 m_1 j_2 m_2} C^{jm'}_{j_1 m'_1 j_2 m'_2} d^{j_1}_{m_1 m'_1}(\beta) d^{j_2}_{m_2 m'_2}(\beta) \tag{37.26}$$

since the Clebsch-Gordan coefficients are real, and the spin 1/2 matrices were real, we see that all of the $d^j_{mm'}(\beta)$ matrices are real. For spin 1/2, we can see from direct evaluation of Eq. 37.13 that changing the sign of all m values changes the sign by $(-1)^{2j-m-m'}$. We can then show this in general from induction. If it is true for all $j_1, j_2 < j$,

$$
\begin{aligned}
d^j_{-m,-m'}(\beta) &= \sum_{m_1 m_2 m'_1 m'_2} C^{j,-m}_{j_1-m_1 j_2-m_2} C^{j-m'}_{j_1-m'_1 j_2-m'_2} d^{j_1}_{-m_1,-m'_1}(\beta) d^{j_2}_{-m_2,-m'_2}(\beta) \\
&= \sum_{m_1 m_2 m'_1 m'_2} (-1)^{j-j_1-j_2} C^{j,m}_{j_1 m_1 j_2 m_2} (-1)^{j-j_1-j_2} C^{jm'}_{j_1-m_1 j_2-m_2} \cdot \\
&\quad \cdot (-1)^{2j_1-m_1-m'_1} d^{j_1}_{m_1,m'_1}(\beta)(-1)^{2j_2-m_2-m'_2} d^{j_2}_{m_2,m'_2}(\beta) \\
&= (-1)^{2j-m-m'} d^j_{m,m'}(\beta)
\end{aligned}
\tag{37.27}
$$

then it is true for all j.

Either using the orthogonality of the Clebsch-Gordan coefficients or writing out explicity:

$$
\begin{aligned}
\langle j_1 m_1 j_2 m_2 | R(\alpha,\beta,\gamma) | j_1 m'_1 j_2 m'_2 \rangle &= D^{j_1}_{m_1 m'_1}(\alpha,\beta,\gamma) D^{j_2}_{m_2 m'_2}(\alpha,\beta,\gamma) \\
&= \sum_{j,m,m'} \langle j_1 m_1 j_2 m_2 | j m j_1 j_2 \rangle \langle j m j_1 j_2 | R(\alpha,\beta,\gamma) | j m' j_1 j_2 \rangle \langle j m' j_1 j_2 | j_1 m'_1 j_2 m'_2 \rangle \\
&= \sum_{j,m,m'} C^{jm}_{j_1 m_1 j_2 m_2} C^{jm'}_{j_1 m'_1 j_2 m'_2} D^j_{mm'}(\alpha,\beta,\gamma)
\end{aligned}
\tag{37.28}
$$

we get the inverse relation.

Notice that this shows us how to write the product of two D matrices as a linear combination of D matrices of different j values.

If we integrate over the angles, we obtain

$$
\begin{aligned}
&\frac{1}{16\pi^2} \int_0^{4\pi} d\alpha \int_{-1}^1 d\cos\beta \int_0^{2\pi} D^{j_1}_{m_1 m'_1}(\alpha,\beta,\gamma) D^{j_2}_{m_2 m'_2}(\alpha,\beta,\gamma) \\
&= \sum_{j,m,m'} C^{jm}_{j_1 m_1 j_2 m_2} C^{jm'}_{j_1 m'_1 j_2 m'_2} \frac{1}{16\pi^2} \int_0^{4\pi} d\alpha \int_{-1}^1 d\cos\beta \int_0^{2\pi} D^j_{mm'}(\alpha,\beta,\gamma) \\
&= C^{00}_{j_1 m_1 j_2 m_2} C^{00}_{j_1 m'_1 j_2 m'_2} \\
&= \frac{(-1)^{j_2+m_2}(-1)^{j_2+m'_2}}{2j_2+1} \delta_{m_1,-m_2}\delta_{m'_1,-m'_2} = \frac{(-1)^{2j_2+m_2+m'_2}}{2j_2+1} \delta_{j_1,j_2}\delta_{m_1,-m_2}\delta_{m'_1,-m'_2}
\end{aligned}
\tag{37.29}
$$

Taking the complex conjugate of the rotation matrices, using the fact that d is real, we see that

$$D^{j*}_{mm'}(\alpha, \beta, \gamma) = e^{-im\alpha}e^{-im'\gamma}d^j_{mm'}(\beta) = D^j_{-m,-m'}(\alpha, \beta, \gamma)(-1)^{2j-m-m'}$$

(37.30)

so that

$$\frac{1}{16\pi^2} \int_0^{4\pi} d\alpha \int_{-1}^1 d\cos\beta \int_0^{2\pi} D^{j_2*}_{m_2m_2'}(\alpha, \beta, \gamma)D^{j_1}_{m_1m_1'}(\alpha, \beta, \gamma)$$

$$= \frac{1}{2j_2+1}\delta_{m_1,m_2}\delta_{m_1',m_2'}$$

(37.31)

which shows that the $D^j_{mm'}(\alpha, \beta, \gamma)$ functions are orthogonal. (In fact they are the eigenfunctions of the symmetric rotator as we will show later.)

Using the Clebsch-Gordan expansion, we can write the integral of three rotation matrices

$$\frac{1}{16\pi^2} \int_0^{4\pi} d\alpha \int_{-1}^1 d\cos\beta \int_0^{2\pi} D^{j_3*}_{m_3m_3'}(\alpha, \beta, \gamma)D^{j_2}_{m_2m_2'}(\alpha, \beta, \gamma)D^{j_1}_{m_1m_1'}(\alpha, \beta, \gamma)$$

$$= \sum_{jmm'} C^{jm}_{j_1m_1j_2m_2}C^{jm'}_{j_1m_1'j_2m_2'}\frac{1}{16\pi^2} \int_0^{4\pi} d\alpha \int_{-1}^1 d\cos\beta \int_0^{2\pi} D^{j_3*}_{m_3m_3'}(\alpha, \beta, \gamma)D^j_{mm'}(\alpha, \beta, \gamma)$$

$$= \frac{1}{2j_3+1}C^{j_3,m_1+m_2}_{j_1m_1j_2m_2}C^{j_3,m_1'+m_2'}_{j_1m_1'j_2m_2'}$$

(37.32)

For the special case where the all the m or all the m' variables are zero, the rotation matrices are proportional to spherical harmonics and this gives an alternative way of calculating the integral of three spherical harmonics. Since the spherical harmonics have integer angular momentum, you can restrict the integral now over 0 to 4π to 0 to 2π, and change the normalization factor from $16\pi^2$ to $8\pi^2$.

37.5 Rigid rotors

Molecules can often be approximated as rigid rotors. That is, first in the Born-Oppenheimer approximation, we solve for the electronic eigenstates with the nuclei fixed, and the eigenenergies as a function of the nuclear positions become the potential for the nuclei. Often this potential has a well defined deep minimum for the nuclear positions. The quantum states for the nucleus are localized around this minimum, and the molecule is approximately rigid. Since the fundamental Hamiltonian is rotationally invariant, so is this approximately rigid one. Therefore, the molecule is approximately a rigid rotor. You can go beyond the rigid rotor approximation by using the rigid rotor states as a basis to calculate rotational-vibrational coupling, etc.

It is convenient to take the center of mass position to be the origin for the nuclear coordinates. Let's look quickly at the classical problem. If the nuclei are at the positions $\boldsymbol{r}_i$ at some instant, and the molecule is rotating with angular velocity $\boldsymbol{\Omega}$, the kinetic energy and therefore the Lagrangian will be

$$
\begin{aligned}
\mathcal{L} &= \frac{1}{2}\sum_i m_i \left(\boldsymbol{\Omega}\times\boldsymbol{r}_i\right)^2 = \frac{1}{2}\sum_i m_i \left(\Omega^2 r_i^2 - \boldsymbol{\Omega}\cdot\boldsymbol{r}_i \boldsymbol{\Omega}\cdot\boldsymbol{r}_i\right) \\
&= \frac{1}{2}\sum_{\alpha,\beta}\Omega_\alpha \sum_i m_i \left(\delta_{\alpha\beta}r_i^2 - r_{i\alpha}r_{i\beta}\right)\Omega_\beta \equiv \frac{1}{2}\sum_{\alpha\beta}\Omega_\alpha I_{\alpha\beta}\Omega_\beta
\end{aligned}
\tag{37.33}
$$

where α and β indicate the usual x, y, z components, and $I_{\alpha\beta}$ is called the inertia tensor. The inertia tensor is obviously real and symmetric; we can diagonalize it and obtain real eigenvalues and real orthogonal eigenvectors. That is, we can write

$$
I_{\alpha\beta} = \sum_{n=1}^{3} p_{n\alpha}I_n p_{n\beta}
\tag{37.34}
$$

where I_1, I_2 and I_3 are the principal moments of inertia, and the $p_{n\alpha}$ give the three lab coordinates of each of the three principal or body axes. We will also sometimes write the principal or body axes as $\hat{\boldsymbol{p}}_n$, instead of giving their components. We can choose the principal axes signs and which we call 1, 2, or 3. For consistency with our usual notation, we choose them so that they define a right-handed coordinate system,

$$
\hat{\boldsymbol{p}}_i \times \hat{\boldsymbol{p}}_j = \epsilon_{ijk}\hat{\boldsymbol{p}}_k \, .
\tag{37.35}
$$

The Lagrangian becomes

$$
\mathcal{L} = \frac{1}{2}\sum_n I_n \left(\hat{\boldsymbol{p}}_n \cdot \boldsymbol{\Omega}\right)^2
\tag{37.36}
$$

If we define the body coordinate system from the principal axes, the Lagrangian in this instantaneous body frame is

$$
\mathcal{L} = \frac{1}{2}\sum_n I_n \Omega_{Bn}^2
\tag{37.37}
$$

where $\Omega_{Bn} = \hat{\boldsymbol{p}}_n \cdot \boldsymbol{\Omega}$ is a component of the angular velocity in the body frame. The corresponding angular momentum component is[1]

$$
J_{Bn} = \frac{\partial\mathcal{L}}{\partial\Omega_{Bn}} = I_n \Omega_{Bn}
\tag{37.38}
$$

---

[1]Normally we want to write the Lagrangian in terms of variables and their time derivatives, but since the only time derivatives are contained in $\boldsymbol{\Omega}$, we can still view the derivatives with respect to the Ω components as the momentum and not worry about defining the conjugate spatial components.

and the Hamiltonian is

$$H = \frac{J_{B1}^2}{2I_1} + \frac{J_{B2}^2}{2I_2} + \frac{J_{B3}^2}{2I_3} .$$ (37.39)

In the lab frame we can write this as

$$H = \sum_n \frac{(\hat{\boldsymbol{p}}_n \cdot \boldsymbol{J})^2}{2I_n}$$ (37.40)

where $\boldsymbol{J}$ is the lab frame angular momentum. This Hamiltonian is a scalar so it commutes with J^2 and J_z, and its eigenstates will have good total angular momentum and good J_z. That is they are $|njm\rangle$ states where n here corresponds to other quantum numbers. Physically you must have more quantum numbers since $|jm\rangle$ describe the angular momentum of a single particle, which has zero moment of inertia when rotated around its radial vector. A rigid body will have a nonzero moment around such a rotation and therefore have additional states.

The operators corresponding to the body axes angular momentum components are $J_{Bn} = \hat{\boldsymbol{p}}_n \cdot \boldsymbol{J}$. Notice that these are the dot products of two vectors and are therefore scalars. They commute with the angular momentum, $[\boldsymbol{J}, J_{Bn}] = 0$. The body axes are vectors. Therefore they have the standard vector commutation relations $[\boldsymbol{J}_\alpha, p_{n\beta}] = i\hbar\epsilon_{\alpha\beta\gamma}p_{n\gamma}$. (The summation convention is used here.)

We can therefore calculate the commutation relations between the body axes angular momentum components as

$$[J_{Bn}, J_{Bm}] = [p_{n\alpha}J_\alpha, p_{m\beta}J_\beta]$$ (37.41)

Applying the commutation relations in a straightforward but tedious way

$$
\begin{aligned}
p_{n\alpha}J_\alpha p_{m\beta}J_\beta &= p_{n\alpha}(p_{m\beta}J_\alpha + [J_\alpha, p_{m\beta}])J_\beta = p_{n\alpha}p_{m\beta}J_\alpha J_\beta + i\hbar\epsilon_{\alpha\beta\gamma}p_{n\alpha}p_{m\gamma}J_\beta \\
&= p_{n\alpha}p_{m\beta}(J_\beta J_\alpha + [J_\alpha, J_\beta]) + i\hbar\epsilon_{\alpha\beta\gamma}p_{n\alpha}p_{m\gamma}J_\beta \\
&= p_{n\alpha}p_{m\beta}J_\beta J_\alpha + i\hbar\epsilon_{\alpha\beta\gamma}p_{n\alpha}p_{m\beta}J_\gamma + i\hbar\epsilon_{\alpha\beta\gamma}p_{n\alpha}p_{m\gamma}J_\beta \\
&= p_{m\beta}(J_\beta p_{n\alpha} - [J_\beta, p_{n\alpha}])J_\alpha + i\hbar\epsilon_{\alpha\beta\gamma}p_{n\alpha}p_{m\beta}J_\gamma + i\hbar\epsilon_{\alpha\beta\gamma}p_{n\alpha}p_{m\gamma}J_\beta \\
&= p_{m\beta}J_\beta p_{n\alpha}J_\alpha + i\hbar\epsilon_{\alpha,\beta,\gamma}p_{m\beta}p_{n\gamma}J_\alpha + i\hbar\epsilon_{\alpha\beta\gamma}p_{n\alpha}p_{m\beta}J_\gamma + i\hbar\epsilon_{\alpha\beta\gamma}p_{n\alpha}p_{m\gamma}J_\beta \\
&= p_{m\beta}J_\beta p_{n\alpha}J_\alpha - i\hbar(\hat{\boldsymbol{p}}_n \times \hat{\boldsymbol{p}}_m) \cdot \boldsymbol{J}
\end{aligned}
$$ (37.42)

We get

$$[J_{Bn}, J_{Bm}] = -i\hbar(\hat{\boldsymbol{p}}_n \times \hat{\boldsymbol{p}}_m) \cdot \boldsymbol{J} = -i\hbar\epsilon_{nmo}J_{Bo}$$ (37.43)

which, except for the sign, looks like the angular momentum commutation relations. As an exercise, you can write $J_{B1} \pm iJ_{B2}$, and commuting with J_{B3}, you will find that these are lowering and raising operators for J_{B3}, with the plus sign the lowering operator. The usual angular momentum algebra goes through with some sign changes.[2] The restrictions

---

[2] One way to see this is to write the operators $j_1 = J_{B1}$, $j_2 = -J_{B2}$, and $j_3 = J_{B3}$. These have the standard commutation relations so $j_1 + ij_2$ is the raising operator for j_3 eigenvalues, and this is $j_1 + ij_2 = J_{B1} - iJ_{B2}$.

that $J^2 > 0$ requires, just as for normal angular momenta, that the eigenvalues of J_{B3} are $b\hbar$ where $-j \leq b \leq j$. Since J_{B3} commutes with J^2 and J_z, we can simultaneously diagonalize J^2, J_z and J_{B3} to obtain a basis $|jmb\rangle$, where both m and b change by integers from $-j \leq b, m \leq j$. Since we know the matrix elements of all the angular momentum operators in terms of their raising and lowering operators we can construct all of the matrix elements for the Hamiltonian in this basis.

One way to show this explicitly is to write the general Hamiltonian as

$$
\begin{aligned}
H &= \frac{J_{B1}^2}{2I_1} + \frac{J_{B2}^2}{2I_2} + \frac{J_{B3}^2}{2I_3} \\
&= \left[\frac{1}{4I_1} + \frac{1}{4I_2}\right] J^2 + \left[\frac{1}{2I_3} - \frac{1}{4I_1} - \frac{1}{4I_2}\right] J_{B3}^2 + \left[\frac{1}{8I_1} - \frac{1}{8I_2}\right] \left[(J_{B1} - iJ_{B2})^2 + (J_{B1} + iJ_{B2})^2\right] \\
&\equiv \alpha J^2 + \beta J_{B3}^2 + \gamma \left[J_{B+}^2 + J_{B-}^2\right]
\end{aligned}
\tag{37.44}
$$

Two symmetries are evident. First, the squared raising and lowering operators only couple states whose b values differ by 2. Therefore the Hamiltonian is block diagonal in even and odd b values. If we interchange $|jmb\rangle$ with $|jm - b\rangle$ the Hamiltonian is invariant, so that changing to these "parity" eigenstates $(|jmb\rangle \pm |jm - b\rangle)/\sqrt{2}$ will give a block diagonal Hamiltonian. These symmetries immediately give the energy eigenstates for $j = 1$, and simplify the matrices for larger j.

The general calculation still requires a numerical diagonalization which should not be too surprising since the corresponding calculation in classical mechanics is also rather messy. Lets look at the symmetric top which is simpler. In that case, two of the moments of inertia are the same and the third is different. We take the one that is different to be I_3, and the symmetric top Hamiltonian is

$$
\begin{aligned}
H_{ST} &= \frac{J_{B1}^2 + J_{B2}^2}{2I_1} + \frac{J_{B3}^2}{2I_3} \\
&= \frac{J_{B1}^2 + J_{B2}^2 + J_{B3}^2}{2I_1} + J_{B3}^2 \left[\frac{1}{2I_3} - \frac{1}{2I_1}\right] \\
&= \frac{J^2}{2I_1} + J_{B3}^2 \left[\frac{1}{2I_3} - \frac{1}{2I_1}\right]
\end{aligned}
\tag{37.45}
$$

This Hamiltonian is diagonal in our $|jmb\rangle$ basis with eigenvalues independent of m

$$
E(j, b) = \frac{\hbar^2[j(j+1) - b^2]}{2I_1} + \frac{\hbar^2 b^2}{2I_3}
\tag{37.46}
$$

with degeneracy $2j + 1$, from the m value for $b = 0$, and $2(2j + 1)$ for $|b| \neq 0$ since the values for b and $-b$ are degenerate. Notice that the spherical top where $I_1 = I_2 = I_3$ has energy levels proportional to $j(j+1)$, that are $(2j + 1)^2$ degenerate.

Sometimes it is useful to have the eigenfunctions. For example we might want

$$
\psi_{jmb}(\alpha, \beta, \gamma) = \langle \alpha\beta\gamma | jmb \rangle
\tag{37.47}
$$

where α, β and γ are now the Euler angles. We can produce the Euler angle basis by rotating the state with $\alpha = \beta = \gamma = 0$,

$$|\alpha\beta\gamma\rangle \;=\; e^{-\frac{i}{\hbar}J_z\alpha}e^{-\frac{i}{\hbar}J_y\beta}e^{-\frac{i}{\hbar}J_z\gamma}|000\rangle \tag{37.48}$$

which becomes

$$
\begin{aligned}
\langle\alpha\beta\gamma|jmb\rangle \;&=\; \langle 000|e^{\frac{i}{\hbar}J_z\gamma}e^{\frac{i}{\hbar}J_y\beta}e^{\frac{i}{\hbar}J_z\alpha}|jmb\rangle \\
&=\; \sum_{m'}\langle 000|e^{\frac{i}{\hbar}J_z\gamma}|jm'b\rangle\langle jm'b|e^{\frac{i}{\hbar}J_y\beta}e^{\frac{i}{\hbar}J_z\alpha}|jmb\rangle \\
&=\; \sum_{m'}\langle 000|jm'b\rangle e^{im'\gamma}e^{im\alpha}\langle jmb|e^{-\frac{i}{\hbar}J_y\beta}|jm'b\rangle^{*} \\
&=\; \sum_{m'} e^{im'\gamma}d^{j}_{mm'}(\beta)e^{im\alpha}\langle 000|jm'b\rangle \tag{37.49}
\end{aligned}
$$

where we note that $d^{j}_{mm'}(\beta)$ is real.

Let's look at

$$\hbar(b-m')\langle 000|jm'b\rangle \;=\; \langle 000|J_{B3} - J_z|jm'b\rangle = \langle 000|\boldsymbol{p}_3\cdot\boldsymbol{J} - J_z|jm'b\rangle \tag{37.50}$$

Since the unrotated molecule has its body axes along the coordinate axes, operating on the $\langle 000|$ bra with $\boldsymbol{p}_3$ gives $\hat{\boldsymbol{z}}$, and therefore $\langle 000|\boldsymbol{p}_3\cdot\boldsymbol{J} = \langle 000|J_z$:

$$\hbar(b-m')\langle 000|jm'b\rangle = 0 \tag{37.51}$$

We see that either $b = m'$ or $\langle 000|jm'b\rangle = 0$. We then have

$$\langle\alpha\beta\gamma|jmb\rangle = Ne^{ib\gamma}e^{im\alpha}d^{j}_{mb}(\beta) = ND^{j*}_{mb}(\alpha,\beta,\gamma) \tag{37.52}$$

where N is a normalization factor. Therefore the complex conjugate of the Wigner rotations give the eigenfunctions of the symmetric top. This is an alternative way of seeing that they must be orthogonal for different j, m and m' values.

37.6 Problems

Problem 37.1. Classically, a gyroscope in a gravitational field precesses with a precession axis given by the direction of the field. The usual gyroscope in this case is a symmetric top with one point fixed. Our quantum rigid rotors separate in the center of mass frame, so our rigid rotor wave functions would normally be multiplied by a center of mass wave function. To make contact with the classical problem, we can imagine that we add a very large point mass to our rigid rotor. In the limit of this mass going to infinity, its position becomes the center of mass and we can take this position as fixed. Taking its position as the origin of our coordinate system, the infinite mass does not change the moment of inertia. Therefore, taking one point fixed is equivalent to simply calculating our rigid rotor Hamiltonian with the fixed point as the origin.

For a symmetric top, the center of mass will be along the $\hat{\boldsymbol{p}}_3$ axis, so that in a gravitational field with acceleration g in the $-\hat{\boldsymbol{z}}$ direction, the potential energy will be $Mgh\cos\beta$, where M is the mass of the top, h the position along $\hat{\boldsymbol{p}}_3$ of the center of mass, and β the angle between $\hat{\boldsymbol{p}}_3$ and the z axis.

   a. Calculate, using perturbation theory, the lowest order change in energy of the energy eigenstates.

   b. Show that the for large angular momentum (i.e. the classic limit), with most of the angular momentum along the body symmetry axis, that the normally observed (slow) precession frequency agrees with the classical result $\omega_p \cong \pm\frac{Mgh}{I_3\omega_3}$ where ω_3 is the angular velocity around the symmetry axis with moment of inertia I_3, and the plus sign is for $\beta < \frac{\pi}{2}$ and the minus sign for $\beta > \frac{\pi}{2}$.

Problem 37.2. Solve for the energy levels and their degeneracies of an asymmetric rigid rotor $I_1 < I_2 < I_3$ for $j = 0, 1, 2$. Use the symmetries identified in the notes.

Chapter 38

Identical particles

38.1   Identical particles

Electrons are experimentally indistinguishable from each other. If we interchange the position and spin of two electrons in our wave function, the matrix elements do not change – all expectation values remain the same. In our Hamiltonians, the interactions cannot tell electrons apart. One way to formulate this in a consistent way is to use field theory, since all particles will be excitations of fields. The identity of particles comes about because they are excitations of the same underlying field. We will first look at this without field theory. Here we simply note that experimentally, the observables and interactions are all invariant under permutations of identical particle labels

$$O_{\text{obs}}(1, 2, 3, ..., N) = O_{\text{obs}}(\text{Permutation of } 1, 2, 3, ..., N) \tag{38.1}$$

where O_{obs} is some operator corresponding to an observable, and $1, 2, 3, ..., N$ are the particle labels. Some examples are the total momentum, the kinetic energy, and the coulomb interaction energy

$$
\begin{aligned}
K.E. &= \sum_i \frac{p_i^2}{2m} \\
\boldsymbol{P} &= \sum_i \boldsymbol{p}_i \\
V &= \sum_{i<j} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} = \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}
\end{aligned}
\tag{38.2}
$$

Notice that for the kinetic energy and V to be invariant under a permutation of the particle labels that the particles needed to have identical masses and charges.

The essentials are

1. We can write a wave function for N identical particles from the states $|\psi\rangle$ as

$$
\begin{aligned}
\psi(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2, ..., \boldsymbol{r}_N, s_N) &= \langle \boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2 \boldsymbol{r}_3 s_3, ..., \boldsymbol{r}_N s_N | \psi \rangle \\
&= \langle 1, 2, 3, ..., N | \psi \rangle
\end{aligned}
\tag{38.3}
$$

2. All $N!$ orderings of the particles must give the same physical results.

3. All $N!$ orderings can be produced by pairwise interchanges of $|1, 2, 3, ..., N\rangle$. We can show this by a demonstration. Take any ordering of $1, 2, 3, ..., N$. For 5 particles an example would be $|4, 3, 5, 1, 2\rangle$. To get a 4 in the first slot we operate with P_{14}, that interchanges the particle in slots 1 and 4 in the ket, $|4, 2, 3, 1, 5\rangle = P_{14}|1, 2, 3, 4, 5\rangle$. Now we look at the second slot we need it to be a 3, so we interchange with P_{23}, $|4, 3, 2, 1, 5\rangle = P_{23}P_{14}|1, 2, 3, 4, 5\rangle$. The next slot requires $|4, 3, 5, 1, 2\rangle = P_{35}P_{23}P_{14}|1, 2, 3, 4, 5\rangle$. At each step we check if the correct particle is in the desired slot and if not, we exchange with the slot having that particle.

4. The pairwise exchange operators commute with all totally symmetric operators.

5. Experimentally, all identical particles have wave functions that are either antisymmetric under all P_{ij} or symmetric under all P_{ij}.

   a. Particles that have wave functions that change sign under particle exchange are called fermions. They have spins that are odd multiples of $1/2$ (half integer spins).

   b. Particles that have wave functions that do not sign under particle exchange are called bosons. They have spins that are even multiples of $1/2$ (integer spins).

   c. Local relativistic field theories require this spin and statistics connection (essentially to have Hamiltonians made up of products of field operators and their derivatives that commute at space like intervals gives this requirement.) This makes the time-ordered product expansion manifestly covariant.

Examples:

S	particle	statistics
$\frac{1}{2}$	e^-	fermion
$\frac{1}{2}$	p	fermion
$\frac{1}{2}$	n	fermion
0	π	boson
1	γ	boson
$\frac{3}{2}$	Δ	fermion
$\frac{1}{2}$	^{108}Ag atom $(47\ p\ 47\ e^-\ 61\ n)$	fermion
$\frac{1}{2}$	^{6}Li $(3\ p\ 3e^-\ 3\ n)$	fermion
0	^{4}He $(2\ p\ 2e^-\ 2\ n)$	boson
$\frac{1}{2}$	^{3}He $(2\ p\ 2e^-\ 1\ n)$	fermion

Where composite particles can be viewed as an elementary particle if the energies of interest are much less than the excitation energies, so that the matrix elements divided by the energy denominators in perturbation theory can be neglected. In condensed matter, nuclei can almost always be considered elementary. In low energy nuclear physics protons, neutrons and pions can often be considered elementary particles. In cold atom studies the atoms can be considered elementary particles.

If you write down the Schrödinger equation,

$$H|\psi(t)\rangle = -\frac{\hbar}{i}\frac{\partial}{\partial t}|\psi(t)\rangle \tag{38.4}$$

since H commutes with all particle exchanges, an initially totally antisymmetric under interchange or totally symmetric under interchange state will remain such a state. If you solve this by expanding in eigenstates of the Hamiltonian

$$H|\Psi_n\rangle = E_n|\Psi_n\rangle \tag{38.5}$$

The eigenstates that are not totally antisymmetric for fermions and symmetric for bosons under interchange of identical particles will be orthogonal to the physical states. They can therefore be discarded, since they will never contribute in any completeness relation.

38.2 The very special case of 2 spin 1/2 particles

A big problem with this section is that it is easy to imagine that these results will generalize to more particles. They do not.

We can write down a general state for 2 electrons. $|\psi\rangle$, and the wave function as

$$\psi_{s_1,s_2}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \langle \boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2 | \psi \rangle \tag{38.6}$$

where $\boldsymbol{r}_i$ is the position and s_i is the $m_s = \pm\frac{1}{2}$ or $\uparrow$, $\downarrow$, for electron i. Therefore

$$
\begin{aligned}
|\psi\rangle &= \int d^3r_1 d^3r_2 \left[|\boldsymbol{r}_1 \uparrow \boldsymbol{r}_2 \uparrow\rangle\psi_{\uparrow\uparrow}(\boldsymbol{r}_1, \boldsymbol{r}_2) + |\boldsymbol{r}_1 \uparrow \boldsymbol{r}_2 \downarrow\rangle\psi_{\uparrow\downarrow}(\boldsymbol{r}_1, \boldsymbol{r}_2)\right. \\
&\quad \left. + |\boldsymbol{r}_1 \downarrow \boldsymbol{r}_2 \uparrow\rangle\psi_{\downarrow\uparrow}(\boldsymbol{r}_1, \boldsymbol{r}_2) + |\boldsymbol{r}_1 \downarrow \boldsymbol{r}_2 \downarrow\rangle\psi_{\downarrow\downarrow}(\boldsymbol{r}_1, \boldsymbol{r}_2)\right]
\end{aligned} \tag{38.7}
$$

Since electrons are fermions, we have $-P_{12}|\psi\rangle = |\psi\rangle$. This is

$$
\begin{aligned}
|\psi\rangle &= -\int d^3r_1 d^3r_2 \left[|\boldsymbol{r}_2 \uparrow \boldsymbol{r}_1 \uparrow\rangle\psi_{\uparrow\uparrow}(\boldsymbol{r}_1, \boldsymbol{r}_2) + |\boldsymbol{r}_2 \downarrow \boldsymbol{r}_1 \uparrow\rangle\psi_{\uparrow\downarrow}(\boldsymbol{r}_1, \boldsymbol{r}_2)\right. \\
&\quad \left. + |\boldsymbol{r}_2 \uparrow \boldsymbol{r}_1 \downarrow\rangle\psi_{\downarrow\uparrow}(\boldsymbol{r}_1, \boldsymbol{r}_2) + |\boldsymbol{r}_2 \downarrow \boldsymbol{r}_1 \downarrow\rangle\psi_{\downarrow\downarrow}(\boldsymbol{r}_1, \boldsymbol{r}_2)\right] \\
&= -\int d^3r_1 d^3r_2 \left[|\boldsymbol{r}_1 \uparrow \boldsymbol{r}_2 \uparrow\rangle\psi_{\uparrow\uparrow}(\boldsymbol{r}_2, \boldsymbol{r}_1) + |\boldsymbol{r}_1 \downarrow \boldsymbol{r}_2 \uparrow\rangle\psi_{\uparrow\downarrow}(\boldsymbol{r}_2, \boldsymbol{r}_1)\right. \\
&\quad \left. + |\boldsymbol{r}_1 \uparrow \boldsymbol{r}_2 \downarrow\rangle\psi_{\downarrow\uparrow}(\boldsymbol{r}_2, \boldsymbol{r}_1) + |\boldsymbol{r}_1 \downarrow \boldsymbol{r}_2 \downarrow\rangle\psi_{\downarrow\downarrow}(\boldsymbol{r}_2, \boldsymbol{r}_1)\right]
\end{aligned} \tag{38.8}
$$

where in the second expression the dummy integration arguments have been renamed so that $\boldsymbol{r}_1 \leftrightarrow \boldsymbol{r}_2$ (that is change variables to $\boldsymbol{r}_1 = \boldsymbol{r}_2'$, $\boldsymbol{r}_2 = \boldsymbol{r}_1'$, and then since $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are no longer used, you can drop the primes). Equating terms, the coefficients must satisfy

$$\psi_{s_1 s_2}(\boldsymbol{r}_1, \boldsymbol{r}_2) = -\psi_{s_2 s_1}(\boldsymbol{r}_2, \boldsymbol{r}_1) \tag{38.9}$$

Explicitly, we have

$$
\begin{aligned}
\psi_{\uparrow\uparrow}(\boldsymbol{r}1, \boldsymbol{r}_2) &= -\psi_{\uparrow\uparrow}(\boldsymbol{r}_2, \boldsymbol{r}_1) \\
\psi_{\uparrow\downarrow}(\boldsymbol{r}1, \boldsymbol{r}_2) &= -\psi_{\downarrow\uparrow}(\boldsymbol{r}_2, \boldsymbol{r}_1) \\
\psi_{\downarrow\uparrow}(\boldsymbol{r}1, \boldsymbol{r}_2) &= -\psi_{\uparrow\downarrow}(\boldsymbol{r}_2, \boldsymbol{r}_1) \\
\psi_{\downarrow\downarrow}(\boldsymbol{r}1, \boldsymbol{r}_2) &= -\psi_{\downarrow\downarrow}(\boldsymbol{r}_2, \boldsymbol{r}_1)
\end{aligned}
\tag{38.10}
$$

which shows that the wave functions have the same (anti)symmetry.

We can write the spin states in terms of the total spin angular momentum. That is, we can write

$$
\begin{aligned}
|\boldsymbol{r}_1 \uparrow \boldsymbol{r}_2 \uparrow\rangle &= |\boldsymbol{r}_1\boldsymbol{r}_2 S = 1, M_s = 1\rangle \\
|\boldsymbol{r}_1 \uparrow \boldsymbol{r}_2 \downarrow\rangle &= 2^{-1/2}\left(|\boldsymbol{r}_1\boldsymbol{r}_2 S = 1, M_s = 0\rangle + |\boldsymbol{r}_1\boldsymbol{r}_2 S = 0, M_s = 0\rangle\right) \\
|\boldsymbol{r}_1 \downarrow \boldsymbol{r}_2 \uparrow\rangle &= 2^{-1/2}\left(|\boldsymbol{r}_1\boldsymbol{r}_2 S = 1, M_s = 0\rangle - |\boldsymbol{r}_1\boldsymbol{r}_2 S = 0, M_s = 0\rangle\right) \\
|\boldsymbol{r}_1 \downarrow \boldsymbol{r}_2 \downarrow\rangle &= |\boldsymbol{r}_1\boldsymbol{r}_2 S = 1, M_s = -1\rangle
\end{aligned}
\tag{38.11}
$$

The state becomes

$$
\begin{aligned}
|\psi\rangle =& \int d^3 r_1 d^3 r_2 \left(\psi_{\uparrow\uparrow}(\boldsymbol{r}_1, \boldsymbol{r}_2)|\boldsymbol{r}_1\boldsymbol{r}_2 S = 1 M_s = 1\rangle\right. \\
&+ 2^{-1/2}\left[\psi_{\uparrow\downarrow}(\boldsymbol{r}_1, \boldsymbol{r}_2) + \psi_{\downarrow\uparrow}(\boldsymbol{r}_1, \boldsymbol{r}_2)\right]|\boldsymbol{r}_1\boldsymbol{r}_2 S = 1 M_s = 0\rangle \\
&+ \psi_{\downarrow\downarrow}(\boldsymbol{r}_1, \boldsymbol{r}_2)|\boldsymbol{r}_1\boldsymbol{r}_2 S = 1 M_s = -1\rangle \\
&\left.+ 2^{-1/2}\left[\psi_{\uparrow\downarrow}(\boldsymbol{r}_1, \boldsymbol{r}_2) - \psi_{\downarrow\uparrow}(\boldsymbol{r}_1, \boldsymbol{r}_2)\right]|\boldsymbol{r}_1\boldsymbol{r}_2 S = 0 M_s = 0\rangle\right)
\end{aligned}
\tag{38.12}
$$

Comparing with Eq. 38.10, we can define functions

$$
\begin{aligned}
\psi_{\uparrow\uparrow}(\boldsymbol{r}_1, \boldsymbol{r}_2) &\equiv \psi_{11}(\boldsymbol{r}_1, \boldsymbol{r}_2) = -\psi_{11}(\boldsymbol{r}_2, \boldsymbol{r}_1) \\
2^{-1/2}[\psi_{\uparrow\downarrow}(\boldsymbol{r}_1, \boldsymbol{r}_2) + \psi_{\downarrow\uparrow}(\boldsymbol{r}_1, \boldsymbol{r}_2)] &\equiv \psi_{10}(\boldsymbol{r}_1, \boldsymbol{r}_2) = -\psi_{10}(\boldsymbol{r}_2, \boldsymbol{r}_1) \\
\psi_{\downarrow\downarrow}(\boldsymbol{r}_1, \boldsymbol{r}_2) &\equiv \psi_{1,-1}(\boldsymbol{r}_1, \boldsymbol{r}_2) = -\psi_{1,-1}(\boldsymbol{r}_2, \boldsymbol{r}_1) \\
2^{-1/2}[\psi_{\uparrow\downarrow}(\boldsymbol{r}_1, \boldsymbol{r}_2) - \psi_{\downarrow\uparrow}(\boldsymbol{r}_1, \boldsymbol{r}_2)] &\equiv \psi_{00}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi_{00}(\boldsymbol{r}_2, \boldsymbol{r}_1)
\end{aligned}
\tag{38.13}
$$

If we write the states as a spin singlet and a spin triplet, since the spin triplets are symmetric under interchange of the spins, the spatial part is antisymmetric under interchange. The singlet state is antisymmetric under interchange of the spins and is therefore symmetric under spatial interchange. *This only happens for two particles.* You cannot write a general antisymmetric wave function for more then 2 spin 1/2 particles that has a antisymmetric spin wave function.

Let's look at the two electrons in a helium atom. Taking the helium nucleus to be infinitely massive and at the origin, we can write the electronic Hamiltonian as

$$
H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \frac{e^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} - \frac{2e^2}{|\boldsymbol{r}_1|} - \frac{2e^2}{|\boldsymbol{r}_2|}
\tag{38.14}
$$

The Hamiltonian is independent of spin. We can then write the general spin state as $S = 1$ with three m_s values, or $S = 0$. The spatial wave functions that multiply these are spatially symmetric or antisymmetric under interchange. We can therefore solve the spatial Hamiltonian for each of its states. Since the interchange operator commutes with H, $[P_{12}, H]$, we can simultaneously diagonalize them and the eigenfunctions will be symmetric or antisymmetric under spatial interchange. Multiplying the spatially symmetric functions times the spin singlet gives an antisymmetric under interchange state. Similarly, multiplying the spatially antisymmetric functions times These will give the complete set of states for the helium atom.

Spatially antisymmetric states must have nodes (zeroes). If we take the two electrons and move them along some path that interchanges them, the function changes sign. Therefore somewhere along the path it must go through zero. It is easy to see from the variational principle that a spatially symmetric wave function will always have a lower energy. If you assume the opposite, that the ground state is spatially antisymmetric, you can form a symmetric variational trial function by taking its absolute value. By integrating by parts, you can show that the energy of this symmetric state is the same as the antisymmetric state. However, this symmetric state has discontinuous first derivatives at its zeroes. Smoothing these will lower the kinetic energy without changing the potential energy, giving a lower energy, and the antisymmetric state cannot be the ground state.

The electronic ground state of the helium atom is therefore is not ferromagnetic, since it has net spin zero. Obviously these arguments break down for more electrons, since we can have ferromagnets for more electrons even though the interaction is not spin dependent.

For Helium, the lowest energy singlet state will have energy E_0, and there will be a lowest energy triplet state with energy E_1. Since the triplet state has $\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 = 1$, and the singlet state has $\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 = -3$, we can write an effective spin Hamiltonian for the ground states of the two electrons in helium as

$$H_{\text{spin}} = \frac{E_0 + 3E_1}{4} + \frac{E_1 - E_0}{4}\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 \qquad (38.15)$$

This is the usual origin of the Heisenberg spin interaction.

38.3   General solution

We can project out the totally antisymmetric or totally symmetric part of a state or wave function by summing over the permutations of the identical particles with the appropriate sign. For example, unnormalized wave functions can be constructed using

$$\langle \boldsymbol{r}_1, ..., r_N | \psi_A \rangle = \sum_P (-1)^P \langle P(\boldsymbol{r}_1, ..., r_N) | \psi \rangle$$

$$\langle \boldsymbol{r}_1, ..., r_N | \psi_S \rangle = \sum_P \langle P(\boldsymbol{r}_1, ..., r_N) | \psi \rangle . \qquad (38.16)$$

38.4   Exchange of distant particles

If a wave function has particles that are well separated, exchange of the particles is unimportant. This is of course necessary for cluster decomposition. If we needed to exchange all the electrons in the universe with each other to build our wave functions and calculate a result, calculations would be impossible.

Let's see how this works. Imagine we have a system consisting of two Hydrogen atoms. One is on earth and the other is on the moon. Normally when we have two well separated systems, separately prepared, we would write their wave function as the product of the wave functions of the two systems separately (equivalently we would write the Hilbert space as the outer product of the two separate Hilbert spaces.)

$$
\begin{aligned}
H &= H_e + H_m \\
H_e|\psi_e\rangle &= E_e|\psi_e\rangle \\
H_m|\psi_m\rangle &= E_m|\psi_m\rangle \\
H(|\psi_e\rangle \otimes |\psi_m\rangle) &= H|\psi_e\psi_m\rangle = (E_e + E_m)|\psi_e\psi_m\rangle
\end{aligned}
\tag{38.17}
$$

We write the wave functions as

$$
\begin{aligned}
\psi_e(\boldsymbol{r}_1, s_1) &= \langle r_1 s_1|\psi_e\rangle \\
\psi_m(\boldsymbol{r}_2, s_2) &= \langle r_2 s_2|\psi_m\rangle
\end{aligned}
\tag{38.18}
$$

and the totally antisymmetric wave function for both particles is

$$
\langle \boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2|\psi_A\rangle = \psi_A(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2) = \frac{1}{\sqrt{2}} \left[\psi_e(\boldsymbol{r}_1, s_1)\psi_m(\boldsymbol{r}_2, s_2) - \psi_e(\boldsymbol{r}_2, s_2)\psi_m(\boldsymbol{r}_1, s_1)\right] .
\tag{38.19}
$$

Let's calculate a matrix element, $\langle\psi_A|O|\psi_A\rangle$. Following Baym, we calculate the probability density of finding an electron at position $\boldsymbol{r}$. This is the expectation value of the density operator

$$
\rho_{\text{op}}(\boldsymbol{r}) = \sum_i \delta^3(\boldsymbol{r}_{i\text{op}} - \boldsymbol{r})
\tag{38.20}
$$

$$
\begin{aligned}
P(\boldsymbol{r}) &= \langle\psi_A|\rho_{\text{op}}(\boldsymbol{r})|\psi_A\rangle \\
&= \sum_{s_1,s_2} \int d^3r_1 \int d^3r_2 \langle\psi_A| \left[\delta^3(\boldsymbol{r}_1 - \boldsymbol{r}) + \delta^3(\boldsymbol{r}_2 - \boldsymbol{r})\right] |\boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2\rangle\langle\boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2|\psi_A\rangle \\
&= \sum_{s_1,s_2} \int d^3r_1 \int d^3r_2 \left[\delta^3(\boldsymbol{r}_1 - \boldsymbol{r}) + \delta^3(\boldsymbol{r}_2 - \boldsymbol{r})\right] |\psi_A(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2)|^2 \\
&= \sum_{s_1,s_2} \left[\int d^3r_2 \, |\psi_A(\boldsymbol{r}, s_1, \boldsymbol{r}_2, s_2)|^2 + \int d^3r_1 \, |\psi_A(\boldsymbol{r}_1, s_1, \boldsymbol{r}, s_2)|^2\right] .
\end{aligned}
\tag{38.21}
$$

To simplify, we can use the antisymmetry of the last term to interchange the arguments, since it is squared the sign change makes no difference. Then relabel the dummy summation and integration indices with $1 \leftrightarrow 2$, and the two terms are the same,

$$P(\boldsymbol{r}) \;=\; 2 \sum_{s_1, s_2} \int d^3 r_2 \, |\psi_A(\boldsymbol{r}, s_1, \boldsymbol{r}_2, s_2)|^2 \,. \tag{38.22}$$

We can now expand, using the product form

$$
\begin{aligned}
P(\boldsymbol{r}) \;=\;& \sum_{s_1, s_2} \int d^3 r_2 \, |\psi_e(\boldsymbol{r}, s_1)\psi_m(\boldsymbol{r}_2, s_2) - \psi_e(\boldsymbol{r}_2, s_2)\psi_m(\boldsymbol{r}, s_1)|^2 \\
=\;& \sum_{s_1, s_2} \int d^3 r_2 \, \big[|\psi_e(\boldsymbol{r}, s_1)|^2 |\psi_m(\boldsymbol{r}_2, s_2)|^2 + |\psi_e(\boldsymbol{r}_2, s_2)|^2 |\psi_m(\boldsymbol{r}, s_1)|^2 \\
& -\psi_e(\boldsymbol{r}, s_1)\psi_m(\boldsymbol{r}_2, s_2)\psi_e^*(\boldsymbol{r}_2, s_2)\psi_m^*(\boldsymbol{r}, s_1) - \psi_e(\boldsymbol{r}_2, s_2)\psi_m(\boldsymbol{r}, s_1)\psi_e^*(\boldsymbol{r}, s_1)\psi_m^*(\boldsymbol{r}_2, s_2) \big]
\end{aligned}
\tag{38.23}
$$

Notice that the cross terms are all zero since they contain products like $\psi_e(\boldsymbol{r}, s)\psi_m(\boldsymbol{r}, s)$, and for well separated systems, one of the two wave functions will be zero at every $\boldsymbol{r}$. We get

$$P(\boldsymbol{r}) \;=\; \sum_{s_1} |\psi_e(\boldsymbol{r}, s_1)|^2 + \sum_{s_1} |\psi_m(\boldsymbol{r}, s_1)|^2 \tag{38.24}$$

which is the sum of the densities we would have calculated separately. All of the cross terms that correspond to matrix elements between different permutations vanish, so we get the same answer for the probability density of finding an electron on earth whether we consider the identity with the electron on the moon or not. This is an example of cluster decomposition. Cluster decomposition is an important ingredient used to select fundamental theories (since, if it is violated, calculating anything will be hopeless).

38.5 Antisymmetrization operator

For fermions we want the totally antisymmetric part of this. We can project it out as before. It will be convenient to define an antisymmetrization operator $\mathcal{A}$ such that

$$\mathcal{A}|R\rangle = \mathcal{A}\Psi(R) = \frac{1}{N!} \sum_{\text{permutations } P} (-1)^P |PR\rangle \tag{38.25}$$

where the sum on permutations means all $N!$ orderings of the particle, R represents the set of coordinates $\boldsymbol{r}_1, s_1, ..., \boldsymbol{r}_N, s_N$, PR is the permuted set of coordinates, $(-1)^P$ is 1 for even permutations and (-1) for odd. For Bosons, the only change is that all of the terms have a positive sign.

Often we will just write $\sum_P$ to indicate the permutation sum.

The reason we divide by $N!$ in $\mathcal{A}$ is so that $\mathcal{A}^2 = \mathcal{A}$. You should verify that $\mathcal{A}$ is Hermitian, for example that $\langle R'|\mathcal{A}|R\rangle^*$.

As for the two-particle problem, we can write

$$|\Psi\rangle = \int dR|R\rangle\langle R|\psi\rangle = \int dR|R\rangle\Psi(R) \tag{38.26}$$

and

$$\begin{aligned}
\mathcal{A}|\Psi\rangle &= \sum_P (-1)^P \int dR|PR\rangle\langle R|\Psi\rangle = \sum_P (-1)^P \int d(PR)|PR\rangle\langle R|\Psi\rangle \\
&= \sum_P (-1)^P \int dR|R\rangle\langle P^{-1}R|\Psi\rangle = \sum_P (-1)^P \int dR|R\rangle\langle PR|\Psi\rangle \tag{38.27}
\end{aligned}$$

and the effect of $\mathcal{A}$ is to antisymmetrize the wave function.

It is sometimes convenient use the equivalent but sloppy notation

$$\mathcal{A}\Psi(R) = \frac{1}{N!}\sum_P (-1)^P \Psi(PR) \tag{38.28}$$

38.6 Matrix elements of spin-independent operators

Since the antisymmetrization operator is Hermition, and commutes with all physical observable operators (since they are symmetric under interchange) we have the general result that matrix elements between antisymmetric states constructed by antisymmetrizing nonantisymmetric states is

$$\begin{aligned}
\langle \Psi_A|\mathcal{A}^+ O\mathcal{A}|\Psi_B\rangle &= \langle \Psi_A|\mathcal{A}O\mathcal{A}|\Psi_B\rangle = \langle \Psi_A|O\mathcal{A}^2|\Psi_B\rangle \\
&= \langle \Psi_A|O\mathcal{A}|\Psi_B\rangle \tag{38.29}
\end{aligned}$$

which shows that we only have to antisymmetrize one of the states since this will project out the antisymmetric part of the other.

Often the the Hamiltonian and other operators of interest do not depend on the spin of the particles. For our purposes we can even be slightly more general and include any operator that does not flip spins. That is any operator O such that $[O, \sigma_{iz}] = 0$. An operator that contains σ_{1z} would therefore be allowed since it does not flip spins. Since the Hamiltonian commutes with S_z, we can simultaneously diagonalize S_z and H. Any such state will then have an S_z eigenvalue equal to $(N_\uparrow - N_\downarrow)\hbar/2$, the number of up particles minus the number of down particles times $\hbar/2$.

Let's imagine that we have such good S_z states so we know the number of up and down particles. If we were solving for the eigenstates of H, we could solve it by assuming particles particles 1 to $N_\uparrow$ have up spins, and particles $N_\uparrow + 1$ to $N_\uparrow + N_\downarrow$ have down spins. Since the Hamiltonian doesn't flip spins, these would be good eigenstates of H, but they would not be antisymmetric under interchange. We can produce the properly antisymmetric states by

application of our $\mathcal{A}$ operator. As we saw above, if we have two such states, we only need to antisymmetrize one of them in a matrix element of physical operators. Lets write two states each with particles $1 - N_\uparrow$ up and the rest down as $|A\rangle$ and $|B\rangle$. The matrix element for the antisymmetrized states is

$$\langle A|\mathcal{A}O\mathcal{A}|B\rangle = \langle A|O\mathcal{A}|B\rangle \qquad (38.30)$$

Since particle 1 is up in the state $|A\rangle$, if O does not flip the spin, only those parts of state $\mathcal{A}|B\rangle$ that have particle 1 up will contribute. That is, only those permutations of $\mathcal{A}$ that keep particle 1 up will contribute. Similarly for all the up particles in state $|A\rangle$. The down particles have the same property. The result is that the full matrix element with a spin independent operator calculating quantities with permutations just over the up particles and permutations over the down particles gives an identical answer to calculating with all permutations over all particles.

That is, if the Hamiltonian is spin independent and the operators of interest are spin independent, we get exactly the same result if we assume that up particles are different than down particles. Up particles are identical with each other and must be antisymmetrized, and down particles are identical with each other and must be antisymmetrized. However, since it is impossible with this set of operators to change an up to a down, we have no way of telling experimentally that the particles are identical, and similarly the theory gives the same result whether you consider them identical or not.

In much of condensed matter physics we ignore magnetic effects and the Hamiltonian does not flip spins. Electronic structure calculations can be done by assuming that up spin electrons are different kinds of particles than down spin electrons. For magnetic systems, spin-orbit or hyperfine interactions, we must antisymmetrize over spins. In nuclear and particle physics, the interactions nearly always flip spins and this simplification cannot be used.

38.7 Many-body states of noninteracting identical particles

Noninteracting identical particles are not very interesting. The main reason to study them is that we can solve for the energy eigenstates of the noninteracting Hamiltonian and these are a complete set of states and can provide a basis for other calculations. For example, if we solve the noninteracting problem we can then truncate this basis and calculate the Hamiltonian matrix for an interacting system. By finding the eigenvalues and eigenvectors of this matrix we have an approximate solution of the interacting system. If we can take a large enough set of the noninteracting eigenstates, we can get converged, accurate solutions. In atomic physics and chemistry, this method is called configuration interaction (the noninteracting states are also called configurations in chemistry, and the calculation of the Hamiltonian is where the word interaction comes from.) In nuclear physics these are called no-core shell model calculations.

Other schemes like many-body perturbation theory or coupled-cluster methods expand around the noninteracting state.

Methods based on density functional ideas almost always use a noninteracting calculation to define possible densities that can come from a properly symmetrized or antisymmetrized wave function.

The Hamiltonian for noninteracting identical particles can have the particles interacting with an external potential, but with no interaction between the particles,

$$H = \sum_i \left[\frac{p_i^2}{2m} + v(\boldsymbol{r}_i) \right] = \sum_i h_i \tag{38.31}$$

While we can formulate this for a spin dependent external potential, nuclear physics models usually include a spin orbit interaction, for simplicity, I will assume that the external potential is independent of spin.

Following Baym we can define one-body orbitals that factor into a spin part and a space part.

$$\begin{aligned} h_i |\psi_n\rangle &= \epsilon_n |\psi_n\rangle \\ \langle \boldsymbol{r}_i s_i | \psi_n \rangle &= \psi_n(\boldsymbol{r}_i, s_i) = \phi_n(\boldsymbol{r}_i)\chi_n(s_1) \end{aligned} \tag{38.32}$$

where $\chi_n(s_1) = \langle s_1| \uparrow \rangle$ or $\chi_n(s_1) = \langle s_1| \downarrow \rangle$.

Solutions with no particular symmetry are

$$\begin{aligned} H|\Psi\rangle &= E|\Psi\rangle \\ \Psi(R) &= \langle R|\Psi\rangle = \prod_i \phi_{n_i}(\boldsymbol{r}_i, s_i) \\ E &= \sum_i \epsilon_{n_i} \,. \end{aligned} \tag{38.33}$$

Applying the antisymmetrization operator $\mathcal{A}$, produces the Slater determinant

$$\mathcal{A} \prod_i \phi_{n_i}(\boldsymbol{r}_i, s_i) = \frac{1}{N!}\mathrm{det}S \tag{38.34}$$

where S is the matrix with elements $S_{ij} = \phi_{n_j}(i)$,

$$S = \begin{pmatrix} \phi_{n_1}(\boldsymbol{r}_1, s_1) & \phi_{n_2}(\boldsymbol{r}_1, s_1) & \phi_{n_3}(\boldsymbol{r}_1, s_1) & \ldots & \phi_{n_N}(\boldsymbol{r}_1, s_1) \\ \phi_{n_1}(\boldsymbol{r}_2, s_2) & \phi_{n_2}(\boldsymbol{r}_2, s_2) & \phi_{n_3}(\boldsymbol{r}_2, s_2) & \ldots & \phi_{n_N}(\boldsymbol{r}_2, s_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_{n_1}(\boldsymbol{r}_N, s_N) & \phi_{n_2}(\boldsymbol{r}_N, s_N) & \phi_{n_3}(\boldsymbol{r}_N, s_N) & \ldots & \phi_{n_N}(\boldsymbol{r}_N, s_N) \end{pmatrix} \tag{38.35}$$

The original product is the product of the diagonal terms, directly expanding the determinant out into all of its $N!$ terms shows that we reproduce the effect of the $N!$ permutations of the antisymmetrization operator.

Let's calculate the normalization of the Slater determinant. We can do this in a fast way or a more straightforward way. First the fast way. Note that $\mathcal{A}^2 = \mathcal{A}$. Any symmetric under interchange operator commutes with $\mathcal{A}$, $[O_S, \mathcal{A}] = 0$. Therefore a general matrix element of such an operator between totally antisymmetric under interchange states is

$$
\begin{aligned}
\langle \Phi | \mathcal{A}^+ O_S \mathcal{A} | \Psi \rangle &= \langle \Phi | \mathcal{A} O_S \mathcal{A} | \Psi \rangle = \langle \Phi | O_S \mathcal{A}^2 | \Psi \rangle \\
&= \langle \Phi | O_S \mathcal{A} | \Psi \rangle
\end{aligned}
\tag{38.36}
$$

and we only have to antisymmetrize one of our states. This will automatically project out the antisymmetric part of the other state.

For the normalization, we have

$$
C = \langle \Psi | \mathcal{A}^+ \mathcal{A} | \Psi \rangle = \langle \Psi | \mathcal{A} | \Psi \rangle
$$
$$
\int dR \prod_{i=1}^{N} \phi_{n_i}^*(\boldsymbol{r}_i, s_i) \mathcal{A} \prod_{j=1}^{N} \phi_{n_j}(\boldsymbol{r}_j, s_j) \, .
\tag{38.37}
$$

However, since our single particle orbitals are normalized, $\sum_s \int dr \phi_j^*(\boldsymbol{r}, s) \phi_k(\boldsymbol{r}, s) = \delta_{jk}$, and therefore only the diagonal term in the permutation sum will contribute, and we are left with just the $1/N!$ factor. We therefore get

$$
C = \frac{1}{N!}
\tag{38.38}
$$

and the normalized noninteracting many-body state is

$$
\Psi(R) = \frac{1}{\sqrt{N!}} \det S \equiv \frac{1}{\sqrt{N!}} \det[\phi_{n_j}(\boldsymbol{r}_i, s_i)] \, .
\tag{38.39}
$$

Notice that if two of the orbitals are the same, or if two of the positions and spins are the same, the wave function is zero. Noninteracting fermions cannot be in the same single-particle orbital.

If we sum over all the quantum numbers for all the ϕ_{n_j} that give a nonzero Slater determinant, taking each many-body state only once, we get a complete set of totally antisymmetric under interchange states.

One way to specify the many-body state uniquely is to give

1. The single particle quantum numbers that are filled (that is, the orbitals that will give the columns of the Slater matrix).

2. The ordering convention that is used. A common method is to first give an ordering for the single particle orbitals, and construct the Slater matrix with the orbitals in the columns arranged in the same order. For example, for the hydrogen atom you could write $n\ell m s$ orbitals in the order $100 \uparrow$, then $100 \downarrow$, then $200 \uparrow$, $200 \downarrow$, $21-1 \uparrow$, $21-1 \downarrow$, $210 \uparrow$, $210 \downarrow$, $211 \uparrow$, $211 \downarrow$,... The particular order is unimportant and can be chosen in any convenient way. Given a set of orbitals, you then write the one that appears first in the orbital ordering as the first column, the next as the next, etc.

The point of maintaining this ordering is to always use the same sign for each of your basis states, since interchanging the order of the orbitals can change the sign. That is, just like any other system, you have to maintain the same phase convention whenever you use the same state.

38.8   Matrix elements

To calculate matrix elements of these wave functions, will have an operator and one or two states.

To be specific, we will take the case of spin $1/2$ particles and calculate matrix elements of 1-body or two-body operators. A one-body operator is

$$O = \sum_i O_i \tag{38.40}$$

examples for electrons could be the charge density or the current density

$$
\begin{aligned}
\rho_{\text{op}}(\boldsymbol{r}) &= -e \sum_i \delta^3(\boldsymbol{r} - \boldsymbol{r}_{i \text{ op}}) \\
\boldsymbol{J}(\boldsymbol{r}) &= -\frac{e}{2} \sum_i \left[\boldsymbol{v}_{i \text{ op}} \delta^3(\boldsymbol{r} - \boldsymbol{r}_{i \text{ op}}) + \delta^3(\boldsymbol{r} - \boldsymbol{r}_{i \text{ op}}) v_{i \text{ op}} \right]
\end{aligned} \tag{38.41}
$$

where $\boldsymbol{v}_i = m^{-1}\left[\boldsymbol{p}_{i \text{ op}} + e\boldsymbol{A}(\boldsymbol{r}_{i \text{ op}})\right]$ and $\boldsymbol{A}$ could be a classical vector potential. A quantized vector potential could also be used, and while this would no longer be strictly a one-body operator since it contains the photon operators, it would still have a one-body structure in our electronic state basis.

The kinetic energy and an external potential have a one-body form

$$
\begin{aligned}
K.E. &= \sum_i \frac{p_{i \text{ op}}^2}{2m} \\
V_1 &= \sum_i v(\boldsymbol{r}_{i \text{ op}}) \,.
\end{aligned} \tag{38.42}
$$

An example of a two-body operator is the potential between two-particles

$$V_2 = \sum_{i<j} v(|\boldsymbol{r}_{i \text{ op}} - \boldsymbol{r}_{j \text{ op}}|) \tag{38.43}$$

To calculate $\langle \Psi_A | O_1 | \Psi_B \rangle$, we can insert a complete set of position states to form the Slater determinants and imagine writing out all $N!$ terms for both $|\Psi_A\rangle$ and $|\Psi_B\rangle$. We also expand out the N terms of O_1. Now look at one of the terms. If we have selected the particle i term form the operator sum, then all other particles must be in the same orbital in $|\Psi_A\rangle$ and $|\Psi_B\rangle$, or the matrix element will be zero. For the nonzero matrix elements, we can reorder the orbitals to make the identical ones in columns 2 through N. This will give a possible

sign change for the matrix element which we write as s. If we use our antisymmetrization operator, we can write for the nonzero case

$$
\begin{aligned}
\langle \Psi_A|O|\Psi_B\rangle &= sN! \int dR\phi_a^*(\boldsymbol{r}_1, s_1)\prod_{i=2}^{N}\phi_{n_i}^*(\boldsymbol{r}_i, s_i)\mathcal{A}^+\sum_j O_j\mathcal{A}\phi_b(\boldsymbol{r}_1, s_1)\prod_{i=2}^{N}\phi_{n_i}(\boldsymbol{r}_i, s_i) \\
&= sN! \int dR\phi_a^*(\boldsymbol{r}_1, s_1)\prod_{i=2}^{N}\phi_{n_i}^*(\boldsymbol{r}_i, s_i)\sum_j O_j\mathcal{A}\phi_b(\boldsymbol{r}_1, s_1)\prod_{i=2}^{N}\phi_{n_i}(\boldsymbol{r}_i, s_i)
\end{aligned}
$$

(38.44)

where we now only antisymmetrize on the right. We will first deal with the case where $\phi_a \neq \phi_b$. The only way to change the state of particle 1, is if O_1 operates, and since particles 2 through N are in states n_2 through n_N on the left, only those terms are picked out on the right, canceling the $N!$ denominator of $\mathcal{A}$ with the $N!$ prefactor, and doing the integrals, we have

$$
\langle \Psi_A|O|\Psi_B\rangle = s\sum_{s_1}\int d^3r_1\phi_a^*(\boldsymbol{r}_1, s_1)O_1\phi_b(\boldsymbol{r}_1, s_1)
$$

(38.45)

If $\psi_a = \psi_b$, then the argument is similar. In this case, for each term O_i, the other particles must be in the same state on the left and right. However, now, all of the O_i terms contribute. These give in turn the matrix element with ϕ_{n_i}, that is

$$
\langle \Psi_A|O|\Psi_A\rangle = \sum_i\sum_{s_1}\int d^3r_1\phi_{n_i}^*(\boldsymbol{r}_1, s_1)O_1\phi_{n_i}(\boldsymbol{r}_1, s_1)
$$

(38.46)

To summarize, for matrix elements of one-body operators

1. The two determinants can differ by at most one orbital. Otherwise the result is zero.

2. Reorder them so that all the orbitals that are the same are in the same columns on the left and right. Count the permutations to determine the sign s.

3. If the determinants differ by one orbital the result is given by Eq. 38.45.

4. If the determinants are equal, the result is given by Eq. 38.46

For example, we can now calculate the expectation value of the density using a Slater determinant wave function. Since $O_1 = \delta^3(\boldsymbol{r} - \boldsymbol{r}_i)$, we have immediately

$$
\langle \rho(\boldsymbol{r})\rangle = \sum_i\sum_s\langle|\phi_{n_i}(\boldsymbol{r}, s)\rangle|^2
$$

(38.47)

that is, you just add the single particle densities.

Similarly, the kinetic and 1-body potential energy expectation is

$$\langle \sum_i \frac{p_i^2}{2m} + v(\boldsymbol{r}_i)\rangle = \sum_i \sum_s \int d^3 r \phi^*_{n_i}(\boldsymbol{r}, s) \left[-\frac{\hbar^2}{2m}\nabla^2 + v(\boldsymbol{r}) \right] \phi_{n_i}(\boldsymbol{r}, s). \qquad (38.48)$$

Two-body operators are a little more interesting. Since the operator can change two orbitals, the determinants can differ at most by two orbitals. In general, we need to calculate the three cases where they differ by 0, 1, or 2 orbitals. Permuting the orbitals so that the ones that differ are in columns 1 and 2, and all the others are the same, we get a possible sign s, and antisymmetrizing only the $|\Psi_B\rangle$ we have

$$
\begin{aligned}
\langle \Psi_A | O | \Psi_B \rangle &= sN! \int dR \phi^*_{a_1}(\boldsymbol{r}_1, s_1)\phi^*_{a_2}(\boldsymbol{r}_2, s_2) \prod_{i=3}^{N} \phi^*_{n_i}(\boldsymbol{r}_i, s_i) \sum_{i<j} O_{ij} \cdot \\
&\quad \cdot \mathcal{A}\phi_{b_1}(\boldsymbol{r}_1, s_1)\phi_{b_2}(\boldsymbol{r}_2, s_2) \prod_{i=3}^{N} \phi_{n_i}(\boldsymbol{r}_i, s_i)
\end{aligned}
$$
$$(38.49)$$

If the states a_1 a_2 differ from b_1 and b_2, only the O_{12} term can give a nonzero result. Therefore particles 3 through N do not have any operators that can change them and we can integrate over their coordinates

$$
\begin{aligned}
\langle \Psi_A | O | \Psi_B \rangle &= s \sum_{s_1, s_2} \int d^3 r_1 \int d^3 r_2 \phi^*_{a_1}(\boldsymbol{r}_1, s_1)\phi^*_{a_2}(\boldsymbol{r}_2, s_2) O_{12} \\
&\quad [\phi_{b_1}(\boldsymbol{r}_1, s_1)\phi_{b_2}(\boldsymbol{r}_2, s_2) - \phi_{b_1}(\boldsymbol{r}_2, s_2)\phi_{b_2}(\boldsymbol{r}_1, s_1)]
\end{aligned}
$$
$$(38.50)$$

If $a_2 = b_2$, then any of the O_{1j} terms can change the particle 1 state. All of the other 2 through N except j must be in the same state. We get for that case

$$
\begin{aligned}
\langle \Psi_A | O | \Psi_B \rangle &= s \sum_{s_1, s_j} \sum_{j \neq 1} \int d^3 r_1 \int d^3 r_j \phi^*_{a_1}(\boldsymbol{r}_1, s_1)\phi^*_{n_j}(\boldsymbol{r}_j, s_j) O_{1j} \\
&\quad [\phi_{b_1}(\boldsymbol{r}_1, s_1)\phi_{n_j}(\boldsymbol{r}_j, s_j) - \phi_{b_1}(\boldsymbol{r}_j, s_j)\phi_{n_j}(\boldsymbol{r}_1, s_1)] \\
&= s \sum_{s_1, s_2} \sum_{j \neq 1} \int d^3 r_1 \int d^3 r_2 \phi^*_{a_1}(\boldsymbol{r}_1, s_1)\phi^*_{n_j}(\boldsymbol{r}_2, s_2) O_{12} \\
&\quad [\phi_{b_1}(\boldsymbol{r}_1, s_1)\phi_{n_j}(\boldsymbol{r}_2, s_2) - \phi_{b_1}(\boldsymbol{r}_2, s_2)\phi_{n_j}(\boldsymbol{r}_1, s_1)] \qquad (38.51)
\end{aligned}
$$

For the last case where all of the orbitals are the same all of the O_{ij} terms can contribute.

The orbitals associated with particles that are not i or j integrate to one leaving:

$$
\begin{aligned}
\langle \Psi_A | O | \Psi_A \rangle &= \sum_{i<j} \sum_{s_i s_j} \int d^3 r_i \int d^3 r_j \phi_{n_i}^*(\boldsymbol{r}_i, s_i) \phi_{n_j}^*(\boldsymbol{r}_j, s_j) O_{ij} \cdot \\
&\quad \cdot \left[\phi_{n_i}(\boldsymbol{r}_i, s_i) \phi_{n_j}(\boldsymbol{r}_j, s_j) - \phi_{n_i}(\boldsymbol{r}_j, s_j) \phi_{n_j}(\boldsymbol{r}_i, s_i) \right] \\
&= \sum_{i<j} \sum_{s_1 s_2} \int d^3 r_1 \int d^3 r_2 \phi_{n_i}^*(\boldsymbol{r}_1, s_1) \phi_{n_j}^*(\boldsymbol{r}_2, s_2) O_{12} \cdot \\
&\quad \cdot \left[\phi_{n_i}(\boldsymbol{r}_1, s_1) \phi_{n_j}(\boldsymbol{r}_2, s_2) - \phi_{n_i}(\boldsymbol{r}_2, s_2) \phi_{n_j}(\boldsymbol{r}_1, s_1) \right]
\end{aligned}
\tag{38.52}
$$

In each of the two-body operator matrix elements, there are terms with a minus sign that correspond to exchanging the two particles that interact. These are called the exchange term. The other term which, for the expectation value, corresponds to the interacting particles being in the same state in $\langle \Psi A |$ and $| \Psi_A \rangle$ is called the direct term.

38.9   Hartree-Fock Method

Since we now have a method for calculating the expectation value of Hamiltonians constructed with one- and two-body operators, we can apply the Rayleigh-Ritz variational principle to find the best variational wave function of the Slater determinant form. This is called the Hartree-Fock method. In chemistry, the terms Hartree-Fock and self-consistent field are often used interchangeably, and often mean instead of just one determinant, the smallest number of determinants to get the correct angular momentum (or other quantum numbers) are used. Here we will mean using only one Slater determinant.

Given a Hamiltonian of the form

$$
H = \sum_i \frac{p_i^2}{2m} + V_1 + V_2
\tag{38.53}
$$

we use our rules derived for one- and two-body operators to write the expectation value as

$$
\begin{aligned}
\langle H \rangle &= \sum_i \sum_s \int d^3 r \phi_{n_i}^*(\boldsymbol{r}, s) \left[-\frac{\hbar^2}{2m} \nabla^2 + v_1(\boldsymbol{r}) \right] \phi_{n_i}(\boldsymbol{r}, s) \\
&\quad + \sum_{i<j} \sum_{ss'} \int d^3 r \int d^3 r' \phi_{n_i}^*(\boldsymbol{r}, s) \phi_{n_j}^*(\boldsymbol{r}', s') v_2(\boldsymbol{r}, \boldsymbol{r}') \cdot \\
&\quad \cdot \left[\phi_{n_i}(\boldsymbol{r}, s) \phi_{n_j}(\boldsymbol{r}', s') - \phi_{n_i}(\boldsymbol{r}', s') \phi_{n_j}(\boldsymbol{r}, s) \right]
\end{aligned}
\tag{38.54}
$$

We worked out how to do a functional variation when looking for the minimum uncertainty wave packet. Taking the short cut there that we get the same result if we vary ϕ^* as an independent variable. We wish to vary the orbitals subject to the constraints that they

remain normalized, therefore we include Lagrange multipliers and vary

$$\langle H \rangle - \sum_i \lambda_i \sum_s \int d^3r \phi_{n_i}^*(\boldsymbol{r}, s)\phi_{n_i}(\boldsymbol{r}, s) \tag{38.55}$$

Varying with respect to $\phi_{n_k}^*(\boldsymbol{r}'', s'')$, we get

$$
\begin{aligned}
0 &= \left[-\frac{\hbar^2}{2m}\nabla''^2 + v_1(\boldsymbol{r}'') - \lambda_k \right] \phi_{n_k}(\boldsymbol{r}'', s'') \\
&+ \sum_{j \neq k} \sum_{s'} \int d^3r' v_2(\boldsymbol{r}'', \boldsymbol{r}')\phi_{n_j}^*(\boldsymbol{r}', s') \left[\phi_{n_k}(\boldsymbol{r}'', s'')\phi_{n_j}(\boldsymbol{r}', s') - \phi_{n_k}(\boldsymbol{r}', s')\phi_{n_j}(\boldsymbol{r}'', s'') \right]
\end{aligned}
\tag{38.56}
$$

At this point, we have no unprimed variables, so we can drop the double primes. Notice that the term $j = k$ in the summand is identically zero, so we can change the sum to include $j = k$. This change then makes the first term in the summand sum to the density expectation value, that is

$$
\begin{aligned}
0 &= \left[-\frac{\hbar^2}{2m}\nabla^2 + v_1(\boldsymbol{r}) + \int d^3r' v_2(\boldsymbol{r}, \boldsymbol{r}')\langle\rho(\boldsymbol{r}')\rangle - \lambda_k \right] \phi_{n_k}(\boldsymbol{r}, s) \\
&- \sum_j \phi_{n_j}(\boldsymbol{r}, s) \sum_{s'} \int d^3r' v_2(\boldsymbol{r}, \boldsymbol{r}')\phi_{n_j}^*(\boldsymbol{r}', s')\phi_{n_k}(\boldsymbol{r}', s') \\
\langle\rho(\boldsymbol{r})\rangle &= \sum_i \sum_s |\phi_{n_i}(\boldsymbol{r}, s)|^2
\end{aligned}
\tag{38.57}
$$

These are the Hartree-Fock equations.

If we apply these equations to atoms,

$$
\begin{aligned}
v_1(\boldsymbol{r}) &= -\frac{Ze^2}{r} \\
v_2(\boldsymbol{r}, \boldsymbol{r}') &= \frac{e^2}{|\boldsymbol{r} - \boldsymbol{r}'|} .
\end{aligned}
\tag{38.58}
$$

The first line looks like the Schrödinger equation for a hydrogenic atom with charge Z, and an additional potential

$$e^2 \int d^3r' \frac{\langle\rho(\boldsymbol{r}')\rangle}{|\boldsymbol{r} - \boldsymbol{r}'|} \tag{38.59}$$

This is simply the average coulomb interaction from the electrons. The additional "exchange term" includes the antisymmetry effects beyond the fact that each electron must be in a different orbital.

If this "exchange term" is dropped, (and we multiply the density by $(N-1)/N$ to account for the missing electron) we get the Hartree method. Hartree used physical ideas that each electron moved in the nuclear potential and the average field of the other electrons to motivate his method. The Hartree-Fock method is on firmer ground since it is based on the variational method.

38.10 Neon atom example

Let's develop the Hartree-Fock method a little farther for a nontrivial but still relatively simple system, the neon atom in its electronic ground state. We choose the neon atom since, if we think of filling hydrogenic orbitals, it has the $1s$, $2s$ and $2p$ orbitals completely filled. When electron-electron interactions are included the hydrogenic picture is no longer correct, but the qualitative features remain. The neon atom, having all of the $n = 2$ orbitals, that would be degenerate for hydrogen, filled, is difficult to perturb, so it doesn't react chemically very well – like the other fully closed shell atoms it is a noble gas.

This makes neon's electronic structure somewhat less interesting since we won't learn much chemistry by studying it, but the simplicity of not having to worry about the kinds of orbitals to include makes it ideal for our analysis.

Since neon is spherical, we will enforce this symmetry on our solutions. We can choose orbitals that have good L^2 and L_z. We expect the to have $1s$, $2s$ and $2p$ radial functions, these will be multiplied times spherical harmonics and up or down spin functions to get the 10 orbitals. The neon electronic ground state in the Hartree-Fock approximation is then described by the three unknown radial functions. Our 10 orbitals are

$$
\begin{aligned}
\phi_1(\boldsymbol{r}, s) &= R_{1s}(r)Y_{00}(\theta, \phi)\delta_{s,\uparrow} \\
\phi_2(\boldsymbol{r}, s) &= R_{2s}(r)Y_{00}(\theta, \phi)\delta_{s,\uparrow} \\
\phi_3(\boldsymbol{r}, s) &= R_{2p}(r)Y_{1-1}(\theta, \phi)\delta_{s,\uparrow} \\
\phi_4(\boldsymbol{r}, s) &= R_{2p}(r)Y_{10}(\theta, \phi)\delta_{s,\uparrow} \\
\phi_5(\boldsymbol{r}, s) &= R_{2p}(r)Y_{11}(\theta, \phi)\delta_{s,\uparrow} \\
\phi_6(\boldsymbol{r}, s) &= R_{1s}(r)Y_{00}(\theta, \phi)\delta_{s,\downarrow} \\
\phi_7(\boldsymbol{r}, s) &= R_{2s}(r)Y_{00}(\theta, \phi)\delta_{s,\downarrow} \\
\phi_8(\boldsymbol{r}, s) &= R_{2p}(r)Y_{1-1}(\theta, \phi)\delta_{s,\downarrow} \\
\phi_9(\boldsymbol{r}, s) &= R_{2p}(r)Y_{10}(\theta, \phi)\delta_{s,\downarrow} \\
\phi_{10}(\boldsymbol{r}, s) &= R_{2p}(r)Y_{11}(\theta, \phi)\delta_{s,\downarrow} \quad\quad (38.60)
\end{aligned}
$$

At this point, we can solve the Hartree-Fock differential equations with this ansatz, or we can calculate the energy expectation value and minimize it. In fact, one popular method simply parameterizes the radial functions, calculates the energy as a function of these parameters and minimizes it. Table 38.1 shows parameters calculated by Clementi and Roetti[4] using radial functions of the form

$$
R(r) = \sum_{k=1}^{5} C_k \sqrt{\frac{(2\zeta_k)^{2m_k+3}}{a_0^3(2m_k+2)!}} \left(\frac{r}{a_0}\right)^{m_k} \exp\left(-\zeta_k \frac{r}{a_0}\right) \quad\quad (38.61)
$$

where a_0 is the Bohr radius, and C_k, ζ_k and m_k are variational parameters, and the square root factor is a normalization.

The Hartree-Fock energy calculated using this direct energy minimization method is $E_{HF} = -128.5471 a.u.$ where one atomic unit is $e^2/a_0 \simeq 27.2\text{eV}$.

Orbital	R_{1s}	R_{2s}	R_{2p}
C_1	0.937170	-0.230930	0.217990
m_1	0	0	1
ζ_1	9.48486	9.48486	1.45208
C_2	0.048990	-0.006350	0.533380
m_2	0	0	1
ζ_2	15.56590	15.56590	2.38168
C_3	0.000580	0.186200	0.329330
m_3	1	1	1
ζ_3	1.96184	1.96184	4.48489
C_4	-0.000640	0.6689900	0.018720
m_4	1	1	1
ζ_4	2.86423	2.86423	9.13464
C_5	0.0005510	0.309100	-
m_5	1	1	-
ζ_5	4.82530	4.82530	-
C_6	0.019990	-0.138710	-
m_6	1	1	-
ζ_6	7.79242	7.79242	-

Table 38.1: The Clementi-Roetti calculated Hartree-Fock radial functions parameters for the electronic orbitals of the Ne atom.

The main advantage of solving the differential equation is that if you solve it with neglible numerical errors, you know that you have the optimal solution. For the direct variational method, you only know that you have an upper bound to the energy.

Since we derived the Hartree-Fock differential equation, let's look at how we would solve it. Since the equation is nonlinear, we will have to solve it iteratively. Typically, we make a linear approximation around our current approximate solution. We then solve that linear problem for a new approximate solution, and use this as the approximation for the next iteration[1]

That means we need an initial guess for the orbitals, and we use this guess to calculate the nonlinear terms. One possible initial guess would be hydrogenic orbitals with an effective nuclear charge given by the charge of the nucleus with a rough estimate of the canceling charge from electrons in the shells inside the orbital in question.

With our current approximation, we need to approximate the direct and exchange terms. A typical way is to use the current charge density to calculate the direct potential and the whole exchange term.

Recalling that coupling $\{Y_\ell \otimes Y_\ell\}_0$ is proportional to the sum over m of the squares of the spherical harmonics, we have

$$\sum_m |Y_{\ell m}(\theta, \phi)|^2 = \sum_m |Y_{\ell m}(0, 0)|^2 = \sum_m \left| \sqrt{\frac{2\ell + 1}{4\pi}} \delta_{m0} \right|^2 = \frac{2\ell + 1}{4\pi} \tag{38.62}$$

and the density is spherically symmetric as expected,

$$
\begin{aligned}
\rho(r) &= \frac{1}{4\pi} \left[2R_{1s}^2(r) + 2R_{2s}^2(r) + 6R_{2p}^2(r) \right] \\
&= 2 \sum_{n\ell} \frac{2\ell + 1}{4\pi} R_{n\ell}(r)^2
\end{aligned}
\tag{38.63}
$$

where the factor of 2 out front is from the spin summation.

The direct potential is then

$$V_{dir}(r) = e^2 \int d^3r' \frac{\rho(r')}{|\boldsymbol{r} - \boldsymbol{r}'|} = \frac{4\pi e^2}{r} \int_0^r dr' r'^2 \rho(r') + 4\pi e^2 \int_r^\infty dr' r' \rho(r') \tag{38.64}$$

The simplest way to evaluate the exchange term keeping just the spherically symmetric part is to multiply by the complex conjugate of the spherical harmonic of the desired term

---

[1]Sometimes the linear equation overestimates the changes to the solution and we need to employ methods such as relaxation where we take a weighted average of the old and new solutions for the next approximation. That is, we might need to take the sum of 0.90 times the old solution and 0.1 times the new as the next approimation in order to obtain a convergent result. There are a variety of such methods used to either obtain convergence or speed convergence to the solution. These are discussed in books on numerical methods.

and perform the angular integration. The exchange term becomes

$$
e^2 4\pi \sum_{n_j \ell_j m_j \ell m} \frac{1}{2\ell + 1} R_{n_j \ell_j}(r) \int dr'^2 dr' R_{n_j \ell_j}(r') R_{n_k, \ell_k}(r') \frac{r_<^\ell}{r_>^{\ell+1}} \cdot
$$

$$
\cdot \int d\cos\theta d\phi Y_{\ell_k m_k}^*(\theta, \phi) Y_{\ell_j m_j}(\theta, \phi) Y_{\ell m}(\theta, \phi) \int d\cos'\theta d\phi' Y_{\ell m}^*(\theta', \phi') Y_{\ell_j m_j}^*(\theta', \phi') Y_{\ell_k m_k}(\theta', \phi')
$$

$$(38.65)$$

where the j sum is over the states with the same spin as k. We calculated the spherical harmonic integrals before using the Wigner-Eckart theorem. The result is

$$
F_{n_k \ell_k}(r) = e^2 4\pi \sum_{n_j \ell_j m_j \ell m} R_{n_j \ell_j}(r) \int dr'^2 dr' R_{n_j \ell_j}(r') R_{n_k \ell_k}(r') \frac{r_<^\ell}{r_>^{\ell+1}} \cdot
$$

$$
\cdot \frac{(2\ell_j + 1)}{4\pi(2\ell_k + 1)} \left[C_{\ell_j 0 \ell 0}^{\ell_k 0} C_{\ell_j m_j \ell m}^{\ell_k m_k} \right]^2
$$

$$
= e^2 \sum_{n_j \ell_j \ell} R_{n_j \ell_j}(r) \int dr'^2 dr' R_{n_j \ell_j}(r') R_{n_k \ell_k}(r') \frac{r_<^\ell}{r_>^{\ell+1}} \frac{(2\ell_j + 1)}{2\ell_k + 1} \left[C_{\ell_j 0 \ell 0}^{\ell_k 0} \right]^2 \quad (38.66)
$$

where in the second expression, we use the completeness for the Clebsch-Gordan coefficients to do the m and m_j sums. We can use

$$
\frac{1}{2\ell_k + 1} \left[C_{\ell_j 0 \ell 0}^{\ell_k 0} \right]^2 = \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2 . \quad (38.67)
$$

Plugging these back, we get an ordinary differential equation for the R functions which we can integrate. We then use those functions to recalculate the direct potential and the exchange term and repeat until it converges:

$$
\left[-\frac{\hbar^2}{2m} \frac{d^2}{dr^2} + \frac{\hbar^2 \ell(\ell+1)}{2mr^2} - \frac{10e^2}{r} + V_{dir}(r) - \lambda_{n\ell} \right] r R_{n\ell}(r) = F_{n\ell}(r) \quad (38.68)
$$

$$
F_{n\ell}(r) = e^2 \sum_{n'\ell'} (2\ell' + 1) R_{n'\ell'}(r) \sum_{\ell''=|\ell-\ell'|}^{\ell+\ell'} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2 \cdot
$$

$$
\cdot \left[\frac{1}{r^{\ell''+1}} \int_0^r dr' r'^{\ell''+2} R_{n\ell}(r) R_{n'\ell'}(r) + r^{\ell''} \int_r^\infty dr' r'^{1-\ell''} R_{n\ell}(r) R_{n'\ell'}(r) \right] \quad (38.69)
$$

Recall the Clebsch-Gordan or 3-j coefficient in the exchange term requires $\ell + \ell' + \ell''$ to be even.

Hartree-Fock works pretty well for electronic systems. The exact energy calculated for the neon atom Hamiltonian is -128.9370 a.u. which differs from the Hartree-Fock value of -128.5471 by just 0.3899 a.u. However, this is nearly 11 eV. Hartree-Fock is actually most

accurate for the inner core electrons (the kinetic energy scales like electron spacing to the minus 2, and the electron-electron potential scales like the spacing to the minus 1, so Hartree-Fock becomes more accurate for electrons at high density), the low density region where all the chemistry takes place is less well done, and often, to get accurate predictions we need to go beyond Hartree-Fock. Diagonalizing the Hamiltonian in a truncated many-body basis made up of single particle orbitals is called configuration interaction or C.I. in chemistry and electronic structure physics. In nuclear physics it is called the nuclear shell model.

38.11   Hard core interactions

The direct use of a single particle basis fails for hard core interactions. For example, imagine that electrons interacted with each other with something like a Lennard-Jones interaction

$$v(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right] \tag{38.70}$$

Notice that all of the integrals for V_{dir} and $F_{n\ell}(r)$ would diverge. The method fails completely. To deal with the hard core, we must correlate the particles from the begining. We either need to develop an effective soft core Hamiltonian to use methods like configuration interaction and Hartree-Fock, or we need to use correlated bases or other many-body basis states. A set of methods that has been developed in the last 40 years based on Monte Carlo integration of the path integral is often effective in solving for some properties of these systems.

Chapter 39

Creation and destruction operators

39.1 Creation and destruction operators

We have already used creation and destruction operators for the harmonic oscillator and for photons where the harmonic oscillator modes of the electromagnetic field were quantized. The use of the creation and destruction operators is also known as second quantization. However, there is no additional quantization it is simply a different, but often convenient, noation. This notation does not change the physics it just makes the mathematical notation simpler. Simpler notation is of course extremely important since it often makes the fundamental ideas easier to understand without cluttering things up. Everything we do with creation and destruction operators can be done without them.

The photon case illustrates what happens when we "second quantize" bosons. The operators that created a photon with momentum $\boldsymbol{k}$ and polarization λ were $a^+_{\boldsymbol{k}\lambda}$, and these were multiplied by the vector $\hat{\boldsymbol{\lambda}}^*$ and $e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}$ when forming the vector potential. The $\hat{\boldsymbol{\lambda}}^*$ corresponds to the spin (or polarization) of the photon, and the plane wave is the spatial "orbital" when going over to the space and polarization representation. The commutation relations give a totally symmetric under interchange wave function for multiple photon states. That is

$$
\begin{aligned}
|\boldsymbol{k}_1\lambda_1\boldsymbol{k}_2\lambda_2\rangle &= a^+_{\boldsymbol{k}_1\lambda_1}a^+_{\boldsymbol{k}_2\lambda_2}|0\rangle \\
&= a^+_{\boldsymbol{k}_2\lambda_1}a^+_{\boldsymbol{k}_1\lambda_2}|0\rangle = |\boldsymbol{k}_2\lambda_2\boldsymbol{k}_1\lambda_1\rangle\,.
\end{aligned}
\tag{39.1}
$$

Photons do not interact with each other directly, so the noninteracting basis was a natural choice.

Similarly, to create a basis for noninteracting identical fermions, we would like to have

$$
\mathcal{N}\mathcal{A}|\boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2 ... \boldsymbol{r}_N s_N\rangle = \psi^+_{s_1}(\boldsymbol{r}_1)\psi^+_{s_2}(\boldsymbol{r}_2)...\psi^+_{s_N}(\boldsymbol{r}_N)|0\rangle
\tag{39.2}
$$

where $|0\rangle$ represents the "vacuum state" for fermions (i.e. no fermions), and $\psi^+_s(\boldsymbol{r})$ creates a fermion at the position $\boldsymbol{r}$ with spin state s (for spin $1/2$, s will usually be up or down but could be chosen to be any linear combination). $\mathcal{N}$ is a normalization which we will find is $\sqrt{N!}$. Destroying a fermion from the vacuum state gives the null state

$$
\psi_s(\boldsymbol{r})|0\rangle = 0\,.
\tag{39.3}
$$

In order for these operators to describe fermions and properly give the antisymmetrized state, we must have, for example,

$$P_{12}\mathcal{A}|\boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2 ... \boldsymbol{r}_N s_N\rangle = -\mathcal{A}|\boldsymbol{r}_2 s_2 \boldsymbol{r}_1 s_1 ... \boldsymbol{r}_N s_N\rangle = \psi_{s_2}^+(\boldsymbol{r}_2)\psi_{s_1}^+(\boldsymbol{r}_1)...\psi_{s_N}^+(\boldsymbol{r}_N)|0\rangle \quad (39.4)$$

so that the $\psi_s^+(\boldsymbol{r})$ must anticommute

$$\{\psi_s^+(\boldsymbol{r}), \psi_{s'}^+(\boldsymbol{r}')\} = 0. \quad (39.5)$$

You should convince yourself that this gives the correct sign change for any pair wise exchange. For example interchanging 1 and N above, you need to move $\psi_{s_1}^+(\boldsymbol{r}_1)$ to the right through all of the $N-1$ other ψ^+ terms, and then move the $\psi_{s_N}^+(\boldsymbol{r}_N)$ to the left back through $N-2$ other ψ^+ terms. You get $2N-3$ sign changes which gives a -1 as required. Since these give properly antisymmetric under interchange states

$$\psi_s^+(\boldsymbol{r})\psi_s^+(\boldsymbol{r}) = -\psi_s^+(\boldsymbol{r})\psi_s^+(\boldsymbol{r}) = 0. \quad (39.6)$$

We need to set the normalization for our creation and destruction operators. Our usual choice for one particle is

$$
\begin{aligned}
\delta_{s's}\delta^3(\boldsymbol{r}'-\boldsymbol{r}) &= \langle \boldsymbol{r}'s'|\boldsymbol{r}s\rangle = \langle 0|\psi_{s'}(\boldsymbol{r}')\psi_s^+(\boldsymbol{r})|\boldsymbol{r}s\rangle = \langle 0| \left[\{\psi_{s'}(\boldsymbol{r}'), \psi_s^+(\boldsymbol{r})\} - \psi_s^+(\boldsymbol{r})\psi_{s'}(\boldsymbol{r}')\right]|0\rangle \\
&= \langle 0|\{\psi_{s'}(\boldsymbol{r}'), \psi_s^+(\boldsymbol{r})\}|0\rangle \quad (39.7)
\end{aligned}
$$

so that we take

$$\{\psi_{s'}(\boldsymbol{r}'), \psi_s^+(\boldsymbol{r})\} = \delta_{s's}\delta^3(\boldsymbol{r}'-\boldsymbol{r}). \quad (39.8)$$

The fermion case looks like the boson case with all commutators replaced by anticommutators.

With this definition, we can calculate the normalization factors for the many-body states. In terms of the basis with no symmetry with their usual normalization[1]

$$\langle \boldsymbol{r}_1's_1'\boldsymbol{r}_2's_2'\boldsymbol{r}_3's_3'...\boldsymbol{r}_N's_N'|\boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2 \boldsymbol{r}_3 s_3 ...\boldsymbol{r}_N s_N\rangle = \delta_{NN'}\prod_i \delta^3(\boldsymbol{r}_i'-\boldsymbol{r}_i)\delta_{s_i',s_i}, \equiv \delta_{NN'}\delta(R'-R)$$

$$(39.9)$$

we have

$$
\begin{aligned}
&\langle \boldsymbol{r}_1's_1'\boldsymbol{r}_2's_2'\boldsymbol{r}_3's_3'...\boldsymbol{r}_{N'}'s_{N'}'|\mathcal{A}^+\mathcal{A}|\boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2 \boldsymbol{r}_3 s_3 ...\boldsymbol{r}_N s_N\rangle \\
&= \langle \boldsymbol{r}_1's_1'\boldsymbol{r}_2's_2'\boldsymbol{r}_3's_3'...\boldsymbol{r}_{N'}'s_{N'}'|\mathcal{A}|\boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2 \boldsymbol{r}_3 s_3 ...\boldsymbol{r}_N s_N\rangle \\
&= \frac{\delta_{NN'}}{N!}\sum_P (-1)^P \delta(R'-PR) \quad (39.10)
\end{aligned}
$$

---

[1] N.B., Baym's state $|\boldsymbol{r}_1 \boldsymbol{r}_2, ..., \boldsymbol{r}_N\rangle$ is our state $\mathcal{A}|\boldsymbol{r}_N \boldsymbol{r}_{N-1}, ..., \boldsymbol{r}_1\rangle$. That is he takes his position states for fermions to be already antisymmetrized, and he writes the quantum numbers $\boldsymbol{r}_1, ..., \boldsymbol{r}_N$ in the opposite order.

which agrees with Baym Eq. 19-42. The normalization for the second quantized version is calculated by applying the anticommutator to move each of the destruction operators to the right of all of the creation operators. The anticommutators give the minus signs from permutations, and all of the $N!$ combinations of delta functions,

$$\langle 0|\psi_{s'_{N'}}(\boldsymbol{r}'_{N'}),...,\psi_{s'_1}(\boldsymbol{r}'_1)\psi_{s_1}(\boldsymbol{r}_1),...,\psi_{s_N}(\boldsymbol{r}_N)|0\rangle = \delta_{NN'}\sum_P (-1)^P \delta(R' - PR)\,.$$

(39.11)

If this result is not obvious here, it will be an immediate consequence of Wick's theorem which we prove later.

We can transform to other single particle bases. For example, if we define the single particle kets $|n\rangle$ as $\langle \boldsymbol{r}s|n\rangle = \phi_n(\boldsymbol{r}, s)$. To see how the creation operators c_n^+ with $|n\rangle = c_n^+|0\rangle$ are related to $\psi_s^+(\boldsymbol{r})$, we write

$$\begin{aligned}|n\rangle &= c_n^+|0\rangle = \sum_s \int d^3r |\boldsymbol{r}s\rangle\langle \boldsymbol{r}s|n\rangle = \sum_s \int d^3r |\boldsymbol{r}s\rangle\phi_n(\boldsymbol{r}, s) \\ &= \sum_s \int d^3r \psi_s^+(\boldsymbol{r})|0\rangle\phi_n(\boldsymbol{r}, s)\end{aligned}$$

(39.12)

or

$$c_n^+ = \sum_s \int d^3r \psi_s^+(\boldsymbol{r})\phi_n(\boldsymbol{r}, s)\,.$$

(39.13)

Plugging in this expression and its Hermitian conjugate, and using the anticommutation relations for the ψ^+ and ψ, gives the anticommutation relations

$$\begin{aligned}\{c_n^+, c_m^+\} &= 0 \\ \{c_n, c_m\} &= 0 \\ \{c_n, c_m^+\} &= \sum_{ss'} \int d^3r \int d^3r' \phi_n^*(\boldsymbol{r}', s')\phi_m(\boldsymbol{r}, s)\{\psi_{s'}(\boldsymbol{r}'), \psi_s^+(\boldsymbol{r})\} \\ &= \sum_{ss'} \int d^3r \int d^3r' \phi_n^*(\boldsymbol{r}', s')\phi_m(\boldsymbol{r}, s)\delta^3(\boldsymbol{r} - \boldsymbol{r}')\delta_{ss'} = \sum_s \int d^3r \phi_n^*(\boldsymbol{r}, s)\phi_m(\boldsymbol{r}, s)\,.\end{aligned}$$

(39.14)

If we choose a set of orbitals normalized to 1, we get

$$\{c_n, c_m^+\} = \delta_{nm}\,.$$

(39.15)

If, for example, we use plane wave states and instead of normalizing them in the box, we use the plane waves without a factor, we would have

$$c_{\boldsymbol{k},\sigma}^+ = \sum_s \int d^3r e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\chi_\sigma(s)\psi_s^+(\boldsymbol{r})$$

(39.16)

451

and

$$\{c_{\boldsymbol{k}',\sigma'}, c_{\boldsymbol{k},\sigma}^+\} = (2\pi)^3 \delta^3(\boldsymbol{k} - \boldsymbol{k}') \delta_{\sigma,\sigma'} \tag{39.17}$$

The various choices we can use for the state normalization lead to corresponding anticommutation relations.

The totally-antisymmetric under interchange many-fermion state in the noninteracting basis is then

$$|n_1, n_2, n_3, ..., n_N\rangle = c_{n_1}^+ c_{n_2}^+ c_{n_3}^+ ... c_{n_N}^+ |0\rangle . \tag{39.18}$$

Taking the orbitals normalized to 1, we can multiply Eq. 39.13 by $\phi_n^*(\boldsymbol{r}', s')$ and sum over n,

$$
\begin{aligned}
\sum_n \phi_n^*(\boldsymbol{r}', s') c_n^+ &= \sum_s \int d^3r \, \psi_s^+(\boldsymbol{r}) \sum_n \langle \boldsymbol{r}'s'|n\rangle \langle n|\boldsymbol{r}'s'\rangle \\
&= \sum_s \int d^3r \, \psi_s^+(\boldsymbol{r}) \langle \boldsymbol{r}'s'|\boldsymbol{r}'s'\rangle = \sum_s \int d^3r \, \psi_s^+(\boldsymbol{r}) \delta_{ss'} \delta^3(\boldsymbol{r} - \boldsymbol{r}') \\
&= \psi_{s'}^+(\boldsymbol{r}') \tag{39.19}
\end{aligned}
$$

or

$$\psi_s^+(\boldsymbol{r}) = \sum_n \phi_n^*(\boldsymbol{r}, s) c_n^+ . \tag{39.20}$$

39.2 The completeness relations

Since the second quantized notation gives us bases for all numbers of particles, we can first write the completeness relation for position and spin states (with no particular symmetry) for any number of particles

$$
1 = |0\rangle\langle 0| + \sum_{s_1} \int d^3r_1 |\boldsymbol{r}_1 s_1\rangle\langle \boldsymbol{r}_1 s_1| + \sum_{s_1,s_2} \int d^3r_1 d^3r_2 |\boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2\rangle\langle \boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2| + ...
$$
$$\tag{39.21}$$

If we only operate on the subspace which is totally antisymmetric under interchange, then only the combinations that are totally antisymmetric will contribute, that is we can write

$$
\begin{aligned}
\mathcal{A} &= \mathcal{A}|0\rangle\langle 0|\mathcal{A} + + \sum_{s_1} \int d^3r_1 \mathcal{A}|\boldsymbol{r}_1 s_1\rangle\langle \boldsymbol{r}_1 s_1|\mathcal{A} + \sum_{s_1,s_2} \int d^3r_1 d^3r_2 \mathcal{A}|\boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2\rangle\langle \boldsymbol{r}_1 s_1 \boldsymbol{r}_2 s_2|\mathcal{A} + ... \\
&= 1_0 + 1_1 + 1_2 + ... = \sum_{N=0}^{\infty} 1_N \tag{39.22}
\end{aligned}
$$

and only use the fully antisymmetrized states. In the last line we define the properly antisymmetry projected unit operators 1_N, projected on particle number N.

39.3 Wick's theorem

Since we can build all of our states by operating with the creation operators on the vacuum state, we can write all matrix elements in our noninteracting basis as matrix elements between vacuum states. We can also write all of our 1-, 2-, etc. body operators in terms of the creation and destruction operators. Therefore, we can write the matrix elements of physical interest as vacuum expectation values of combinations of creation and destruction operators in a particular order. We can anticommute (for fermions) and commute (for bosons) the operators in order to bring them into normal ordering. Normal order means that the destruction operators are to the right of all creation operators, and we include the appropriate sign for the permutation of the operators. The vacuum expectation value of the normal ordered operators will be zero.[2] The anticommutators or commutators of the field operators are either zero, kronecker deltas or delta functions times constants, giving the value of the matrix element.

Wick's theorem is usually derived for time ordered operators, but it can be used for any ordering. For convenience we will call the desired order the time order. That is, assume that each operator has an associated time, and time order means they are ordered so that earlier operators are to the right. We define the contraction of two operators as the difference between their time order and their normal order,

$$\overset{\frown}{AB} = T(AB) - N(AB) \tag{39.23}$$

where T indicates the time order (i.e. the desired order), and N is the normal order. The contraction is denoted by connecting the operators by a line (usually above, but sometimes below). Wick's theorem (which was conjectured and proved to fourth order by Freeman Dyson) says that the time order is equal to the sum of all the normal orders with all possible contractions:

$$T(ABCD...YZ) = N(ABCD...YZ) + N(\overset{\frown}{A}BCD...YZ) + N(\overset{\frown}{AB}CD...YZ) + ...$$
$$+ ... + N(A\overset{\frown}{BC}D...YZ) + ... + N(\overset{\frown}{ABCD...YZ}) \tag{39.24}$$

The proof is straightforward. First notice that both sides are invariant under permutation of the operators inside the time and normal orders since the T and N will reorder them. We can therefore take the case where the time order is $ABCD...YZ$ without loss of generality. To normal order the operators, we need to anticommute for fermions (or commute for bosons)

---

[2]You will sometimes read about spontaneously broken symmetries with fields obtaining a vacuum expectation value. What this means is that we have changed the operators exactly as in the coherent state formalism that we used for describing classical fields within quantum electrodynamics. In that case we had $a_{k\lambda}|\alpha\rangle = \alpha_{k\lambda}|\alpha\rangle$, and transformed to new operators $b_{k\lambda} = a_{k\lambda} - \alpha_{k\lambda}$. The "vacuum" for the b operators is now $|\alpha\rangle$, and the a operators have an expectation value in the new vacuum $|\alpha\rangle$. Here we only deal with the case where our destruction operators destroy the vacuum, so we would need to rewrite everything in terms of the new b operators.

all creation operators with the destruction operators to their left. Each of these will give a contraction. For operators already in the correct order, the time order and normal order are the same, so the contraction is zero, and including these contractions does not change the result. Therefore we can sum over all possible contractions, which is Wick's theorem.

Wick's theorem is a general statement. For vacuum expectation values, the normal ordered operators give zero, so only the fully contracted term is nonzero.

We can use Wick's theorem to calculate the normalization of Eq. 39.11. Contractions between creation operators or between destruction operators give zero. Therefore, we must sum over all normal ordered products of the contractions between the destruction and the creation operators. The destruction operator for $\psi_{s'_1}(\mathbf{r}'_1)$ can be contracted with any of the N creation operators. The operator $\psi_{s'_2}(\mathbf{r}'_2)$ can be contracted with any of the remaining $N-1$ creation operators, and so on. We immediately get $N!$ terms corresponding to all permutations. The contractions give the delta functions of position and the Kronecker deltas for spin. The rule that the normal order contains the sign for all necessary permutations gives the correct sign.

39.4   Building operators

All physical operators for fermions can be written in terms of the creation and destruction operators. For example, the density operator

$$\rho_{\mathrm{op}}(\mathbf{r}) = \sum_{i=1}^{N} \delta^3(\mathbf{r} - \mathbf{r}_{i\mathrm{op}}) \tag{39.25}$$

can be written as

$$\rho_{\mathrm{op}}(\mathbf{r}) = \sum_{s} \psi_s^+(\mathbf{r})\psi_s(\mathbf{r}) \,. \tag{39.26}$$

This should look reasonable since the $\psi_s(\mathbf{r})$ will give zero except if a fermion is at the position $\mathbf{r}$, and the $\psi_s^+(\mathbf{r})$ recreates the fermion that the $\psi_s(\mathbf{r})$ destroyed.

We can verify this by calculating

$$
\begin{aligned}
\sum_{s}\langle\psi|\psi_s^+(\mathbf{r})\psi_s(\mathbf{r})|\psi\rangle &= \sum_{s}\langle\psi|\psi_s^+(\mathbf{r})\sum_{s_1,\dots s_{N-1}}\int d^3r_1\dots d^3r_N \mathcal{A}|\mathbf{r}_1,\dots\mathbf{r}_{N-1}\rangle\langle\mathbf{r}_1,\dots\mathbf{r}_{N-1}|\mathcal{A}\psi_s(\mathbf{r})|\psi\rangle \\
&= N\sum_{s}\langle\psi|\sum_{s_1,\dots s_{N-1}}\int d^3r_1\dots d^3r_N \mathcal{A}|\mathbf{r}s\mathbf{r}_1,\dots\mathbf{r}_{N-1}\rangle\langle\mathbf{r}s\mathbf{r}_1,\dots\mathbf{r}_{N-1}|\mathcal{A}|\psi\rangle \\
&= \sum_{s_1,\dots s_N}\int d^3r_1\dots d^3r_N \sum_{i}\delta(\mathbf{r}-\mathbf{r}_i)\langle\psi|\mathbf{r}_1,\dots\mathbf{r}_N\rangle\langle\mathbf{r}_1,\dots\mathbf{r}_N|\psi\rangle \tag{39.27}
\end{aligned}
$$

where in the last line we used the symmetry of the integral under interchange to convert the N into a sum over the particles.

Repeating the same analysis, we can write

$$
\begin{aligned}
\sum_i \frac{p_i^2}{2m} &= -\frac{\hbar^2}{2m} \sum_s \int d^3 r\, \psi^+(\boldsymbol{r}) \nabla^2 \psi(\boldsymbol{r}) \\
\sum_i v_1(\boldsymbol{r}_i) &= \sum_s \int d^3 r\, v_1(\boldsymbol{r}) \psi_s^+(\boldsymbol{r}) \psi_s(\boldsymbol{r}) \\
\sum_{i \neq j} v_2(\boldsymbol{r}_i - \boldsymbol{r}_j) &= \frac{1}{2} \sum_{ss'} \int d^3 r\, d^3 r'\, v_2(\boldsymbol{r}' - \boldsymbol{r}) \psi_{s'}^+(\boldsymbol{r}') \psi_s^+(\boldsymbol{r}) \psi_s(\boldsymbol{r}) \psi_{s'}(\boldsymbol{r}') \quad (39.28)
\end{aligned}
$$

If taking the derivative of a field operator seems odd, just remember that we can always view the derivative from its definition. That is writing $\psi_s(\boldsymbol{r}) = \psi_s(x, y, z)$ we can write

$$
\frac{\partial}{\partial x} \psi_s(x, y, z) = \lim_{h \to 0} \frac{\psi_s(x + h, y, z) - \psi_s(x, y, z)}{h} \tag{39.29}
$$

and we see that the derivative of the destruction operator is proportional to the difference between destroying a particle at $x + h$ and destroying it at x. The order of the operators for the two-body potential term is chosen so that having both destruction operators first puts the expression in normal order, and automatically excludes the self term. The overall order is so that it is equivalent to $\rho(\boldsymbol{r}) \rho(\boldsymbol{r}') v(\boldsymbol{r} - \boldsymbol{r}')$ when $\boldsymbol{r} \neq \boldsymbol{r}'$.

We can now calculate any matrix elements using Wick's theorem and the operator form.

Let's look at calculating the ground state expectation value of the energy of a Slater determinant state. Our state vector is

$$
|\Psi\rangle = \left[\prod_{i=1}^{N} \sum_{s_i} \int d^3 r_i\, \phi_{n_i}(\boldsymbol{r}_i, s_i) \psi_{s_i}^+(\boldsymbol{r}_i) \right] |0\rangle \tag{39.30}
$$

where the product means to keep the terms in left to right order from the limit $i = 1$ to $i = N$. We can write these in terms of the c_n operators of Eq. 39.13

$$
|\Psi\rangle = \left[\prod_{i=1}^{N} c_{n_i}^+ \right] |0\rangle \tag{39.31}
$$

First let's check the normalization. We find (notice that the order for the bra state is reversed from the hermitian conjugate operation,

$$
\langle \Psi | \Psi \rangle = \langle 0 | \left[\prod_{j=N}^{1} c_{n_j} \right] \left[\prod_{i=1}^{N} c_{n_i}^+ \right] |0\rangle \tag{39.32}
$$

Wick's theorem tells us to form all of the contractions, but since

$$
\contraction{}{c}{_{n_j}}{c} c_{n_j} c_{n_i}^+ = c_{n_j} c_{n_i}^+ - \left[-c_{n_i}^+ c_{n_j} \right] = \{ c_{n_j} c_{n_i}^+ \} = \delta_{n_j, n_i}, \tag{39.33}
$$

if $n_j \neq n_i$, the contraction is zero. Only one of the $N!$ orderings is nonzero and that term is 1 (the sign is positive because we can first contract the adjacent n_1 terms, then the now adjacent n_2 terms, etc. Therefore $\langle \Psi | \Psi \rangle = 1$.

To be specific, let's think of the ground-state energy expectation value for a Slater determinant state for the neon atom with 10 electrons and the nucleus fixed at the origin. The three types of terms are given in Eq. 39.28. The kinetic energy is

$$KE = -\frac{\hbar^2}{2m} \sum_s \int d^3r \langle 0| \left[\prod_{j=N}^1 c_{n_j} \right] \psi_s^+(\boldsymbol{r}) \nabla^2 \psi_s(\boldsymbol{r}) \left[\prod_{i=1}^N c_{n_i}^+ \right] |0\rangle \qquad (39.34)$$

From Wick's theorem, the $\psi_s(\boldsymbol{r})$ must be contracted with a c_n^+ (it is already normal ordered with ψ^+ so that term would give zero). Similarly, the $\psi_s^+(\boldsymbol{r})$ must be contracted with a c_n. The remaining c_n terms must be contracted with a c_n^+. The additional contractions we need are

$$\overbrace{\psi_s(\boldsymbol{r}) c_n^+} = \psi_s(\boldsymbol{r}) \sum_{s'} \int d^3r' \phi_n(\boldsymbol{r}', s') \psi_{s'}^+(\boldsymbol{r}') + \sum_{s'} \int d^3r' \phi_n(\boldsymbol{r}', s') \psi_{s'}^+(\boldsymbol{r}') \psi_s(\boldsymbol{r})$$

$$= \sum_{s'} \int d^3r' \delta_{s,s'} \delta^3(\boldsymbol{r} - \boldsymbol{r}') \phi_n(\boldsymbol{r}', s') = \phi_n(\boldsymbol{r}, s) \qquad (39.35)$$

and

$$\overbrace{c_n \psi_s^+(\boldsymbol{r})} = \psi_s^+(\boldsymbol{r}) \sum_{s'} \int d^3r' \phi_n^*(\boldsymbol{r}', s') \psi_{s'}(\boldsymbol{r}') + \sum_{s'} \int d^3r' \phi_n^*(\boldsymbol{r}', s') \psi_{s'}(\boldsymbol{r}') \psi_s^+(\boldsymbol{r})$$

$$= \sum_{s'} \int d^3r' \delta_{s,s'} \delta^3(\boldsymbol{r} - \boldsymbol{r}') \phi_n^*(\boldsymbol{r}', s') = \phi_n^*(\boldsymbol{r}, s) \qquad (39.36)$$

We pick one of the c_{n_j} to be contracted with the $\psi_s^+(\boldsymbol{r})$, the other c_n operators must be contracted with their matching c_n^+ operator or the result will be zero, therefore the $\psi_s(\boldsymbol{r})$ operator must be contracted with the corresponding $c_{n_j}^+$ whose c_{n_j} was contracted with $\psi_s^+(\boldsymbol{r})$. Since the ∇^2 takes the derivatives of this operator, it will take the derivatives with respect to the components of $\boldsymbol{r}$ of the corresponding contraction. To determine the sign, we start by moving the c_{n_1} term (assuming it is not contracted with $\psi_s^+(\boldsymbol{r})$) through the two $\psi^+\psi$ operators giving a net plus sign. Its contraction with $c_{n_1}^+$ then gives 1. We repeat until we reach the operator that is contracted with $\psi_s^+(\boldsymbol{r})$, since they are now adjacent, its sign is plus 1, the $\psi_s(\boldsymbol{r})$ is also now adjacent to the corresponding $c_{n_j}^+$, so its contraction is 1. Each of the remaining c_n terms are adjacent to their corresponding c_n^+ terms and give 1. The result, when summing over the c_{n_j} that contract with $\psi_s^+(\boldsymbol{r})$ is

$$KE = -\frac{\hbar^2}{2m} \sum_j \sum_s \int d^3r \phi_{n_j}^*(\boldsymbol{r}, s) \nabla^2 \phi_{n_j}(\boldsymbol{r}, s) \qquad (39.37)$$

exactly as when we used the Slater determinant formalism.

The one-body potential term goes through nearly identically. In place of the $-\frac{\hbar^2}{2m}\nabla^2$ we have the potential $v_1(\boldsymbol{r})$

$$V_1 = \langle\Psi|\sum_s \int d^3r\psi_s^+(\boldsymbol{r})v_1(\boldsymbol{r})\psi_s(\boldsymbol{r})|\Psi\rangle = \sum_j\sum_s \int d^3r\phi_{n_j}^*(\boldsymbol{r},s)v_1(\boldsymbol{r})\phi_{n_j}(\boldsymbol{r},s)\,. \quad (39.38)$$

The two-body potential term is slightly more complicated but handled in much the same way. We have

$$V_2 = \frac{1}{2}\sum_{ss'} \int d^3rd^3r'v_2(\boldsymbol{r}'-\boldsymbol{r})\langle 0|\left[\prod_{j=N}^{1} c_{n_j}\right]\psi_{s'}^+(\boldsymbol{r}')\psi_s^+(\boldsymbol{r})\psi_s(\boldsymbol{r})\psi_{s'}(\boldsymbol{r}')\left[\prod_{i=1}^{N} c_{n_i}^+\right]|0\rangle\,.$$

$$(39.39)$$

As before, the two ψ^+ operators have to be contracted with two of the c_n operators, The other c_n operators must be contracted with their corresponding c_n^+ operators or the result will be zero, the ψ operators must then be contracted with the two c_n^+ operators corresponding to the two c_n operators contracted with the ψ^+. That is, after eliminating the contractions between the c_n and c_n^+ operators, the only surviving terms have the form

$$V_2 = \frac{1}{2}\sum_{ss'} \int d^3rd^3r'v_2(\boldsymbol{r}'-\boldsymbol{r})\sum_{i<j}[\langle 0|\overbracket{c_{n_j}\overbracket{c_{n_i}\psi_{s'}^+(\boldsymbol{r}')}\psi_s^+(\boldsymbol{r})}\overbracket{\psi_s(\boldsymbol{r})\overbracket{\psi_{s'}(\boldsymbol{r}')c_{n_i}^+}c_{n_j}^+}|0\rangle$$

$$+\langle 0|\overbracket{c_{n_j}\overbracket{c_{n_i}\psi_{s'}^+(\boldsymbol{r}')}\psi_s^+(\boldsymbol{r})}\overbracket{\psi_s(\boldsymbol{r})\overbracket{\psi_{s'}(\boldsymbol{r}')c_{n_i}^+}c_{n_j}^+}|0\rangle$$

$$+\langle 0|\overbracket{c_{n_j}\overbracket{c_{n_i}\psi_{s'}^+(\boldsymbol{r}')}\psi_s^+(\boldsymbol{r})}\overbracket{\psi_s(\boldsymbol{r})\overbracket{\psi_{s'}(\boldsymbol{r}')c_{n_i}^+}c_{n_j}^+}|0\rangle$$

$$+\langle 0|\overbracket{c_{n_j}\overbracket{c_{n_i}\psi_{s'}^+(\boldsymbol{r}')}\psi_s^+(\boldsymbol{r})}\overbracket{\psi_s(\boldsymbol{r})\overbracket{\psi_{s'}(\boldsymbol{r}')c_{n_i}^+}c_{n_j}^+}|0\rangle]\,.$$

$$(39.40)$$

We can now plug in for the contractions, noting that the first term has all positive signs, the second has a negative sign from the needed interchange of two operators to bring them adjacent for contraction, the third also has a negative sign for the same reason. The last term has two negative signs since two exchanges are needed to bring the contracted operators

adjacent, so the result is

$$
\begin{aligned}
V_2 &= \frac{1}{2} \sum_{ss'} \int d^3r d^3r' v_2(\boldsymbol{r}' - \boldsymbol{r}) \sum_{i<j} \Big[\phi_{n_i}^*(\boldsymbol{r}', s') \phi_{n_j}^*(\boldsymbol{r}, s) \phi_{n_i}(\boldsymbol{r}', s') \phi_{n_j}(\boldsymbol{r}, s) \\
&\quad - \phi_{n_j}^*(\boldsymbol{r}', s') \phi_{n_i}^*(\boldsymbol{r}, s) \phi_{n_i}(\boldsymbol{r}', s') \phi_{n_j}(\boldsymbol{r}, s) \\
&\quad - \phi_{n_i}^*(\boldsymbol{r}', s') \phi_{n_j}^*(\boldsymbol{r}, s) \phi_{n_j}(\boldsymbol{r}', s') \phi_{n_i}(\boldsymbol{r}, s) \\
&\quad + \phi_{n_j}^*(\boldsymbol{r}', s') \phi_{n_i}^*(\boldsymbol{r}, s) \phi_{n_j}(\boldsymbol{r}', s') \phi_{n_i}(\boldsymbol{r}, s) \Big] \\
&= \sum_{ss'} \int d^3r d^3r' v_2(\boldsymbol{r}' - \boldsymbol{r}) \sum_{i<j} \Big[|\phi_{n_i}(\boldsymbol{r}', s')|^2 |\phi_{n_j}(\boldsymbol{r}, s)|^2 \\
&\quad - \phi_{n_i}^*(\boldsymbol{r}', s') \phi_{n_j}^*(\boldsymbol{r}, s) \phi_{n_j}(\boldsymbol{r}', s') \phi_{n_i}(\boldsymbol{r}, s) \Big]
\end{aligned}
$$

$$(39.41)$$

Not surprisingly, the result is the same that we obtained using the Slater determinants in first quantized form.

39.5 Particle-hole transformation

Often we want to calculate the ground state and low excited states of a system. For example, we might want to calculate the ground-state energy of the electrons to use as a Born-Oppenheimer potential for the nuclei. If the interaction is not too strong, the ground state can be expanded around the noninteracting ground state. It is convenient to make this state the new "vacuum."

Fermion operators have the following structure.

$$
\begin{aligned}
\{c, c^+\} &= 1 \\
c^+|0\rangle &= |1\rangle \\
c^+|1\rangle &= 0 \\
c|1\rangle &= |0\rangle \\
c|0\rangle &= 0
\end{aligned}
$$

$$(39.42)$$

The key observation, is that if we write new operators $b = c^+$, $b^+ = c$, $|0_b\rangle = |1\rangle$, $|1_b\rangle = |0\rangle$, then we get the same relationships, and the b operators are perfectly good fermion operators. Their vaccum (or unoccupied) states are the occupied states for the original c operators.

We can interpret c_n^+ as creating a particle or destroying a hole[3]

---

[3]Since creating a particle with spin angular momentum $\frac{1}{2}\hbar$, must be the same as destroying a particle with spin angular momentum $-\frac{1}{2}\hbar$, we choose the spin of the hole to be opposite that of the corresponding particle. Similarly, for the orbitals, we have the complex conjugate state. Both of these are generally the time reversed orbitals.

In our case, we want the new vacuum to be

$$|0'\rangle = \prod_{i=1}^{N} c_{n_i}^{+}|0\rangle \tag{39.43}$$

To have a specific example, lets look at the case where $v_1 = 0$, and the orbitals are plane waves times spinors. To keep the total energy finite, we will put our system in a large box of size L, and let $L \to \infty$. We often want the noninteracting ground state to be the vacuum,

$$|0'\rangle = \prod_{s} \prod_{|\boldsymbol{k}|<k_F} c_{ks}^{+}|0\rangle . \tag{39.44}$$

We then write

$$b_{ks} = c_{-\boldsymbol{k},-s}^{+} \qquad k < k_F \tag{39.45}$$

where $-s$ means that the spin is flipped. The field operator becomes

$$\psi_s^{+}(\boldsymbol{r}) = \sum_{k} \left[\Theta(k - k_F) \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{L^{3/2}} c_{ks}^{+} + \Theta(k_F - k) \frac{e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}}{L^{3/2}} b_{\boldsymbol{k}-s} \right] . \tag{39.46}$$

We can calculate the contractions we will need

$$
\begin{aligned}
\overbracket{\psi_s^{+}(\boldsymbol{r})\psi_{s'}^{+}(\boldsymbol{r}')} &= \overbracket{\psi_s(\boldsymbol{r})\psi_{s'}(\boldsymbol{r}')} = 0 \\
\overbracket{\psi_s^{+}(\boldsymbol{r})\psi_{s'}(\boldsymbol{r}')} &= L^{-3} \sum_{k} \Theta(k_F - k) e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} \delta_{ss'} = \int \frac{d^3k}{(2\pi)^3} \Theta(k_F - k) e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} \delta_{ss'} \\
&= \frac{k_F^3}{6\pi^2} \ell(k_F|\boldsymbol{r} - \boldsymbol{r}'|) \delta_{ss'} \\
\overbracket{\psi_s(\boldsymbol{r})\psi_{s'}^{+}(\boldsymbol{r}')} &= \delta^3(\boldsymbol{r} - \boldsymbol{r}') \delta_{ss'} - \frac{k_F^3}{6\pi^2} \ell(k_F|\boldsymbol{r} - \boldsymbol{r}'|) \delta_{ss'} \tag{39.47}
\end{aligned}
$$

and

$$\ell(x) = \frac{3}{x^3} \left[\sin(x) - x\cos(x) \right] \tag{39.48}$$

is called the Slater function.

Let's calculate the Hartree-Fock energy for such a system interacting via a two-body potential that depends on the pair distance

$$V_2 = \frac{1}{2} \sum_{ss'} \int d^3r \int d^3r' v_2(|\boldsymbol{r} - \boldsymbol{r}'|) \psi_s^{+}(\boldsymbol{r}) \psi_{s'}^{+}(\boldsymbol{r}') \psi_{s'}(\boldsymbol{r}') \psi_s(\boldsymbol{r}) . \tag{39.49}$$

The kinetic energy is

$$
\begin{aligned}
\langle KE \rangle &= -\frac{\hbar^2}{2m} \langle 0' | \sum_s \int d^3 r \, \psi_s^+(\boldsymbol{r}) \nabla^2 \psi_s(\boldsymbol{r}) | 0' \rangle \\
&= -\frac{\hbar^2}{2m} \sum_s \int d^3 r' \int d^3 r \, \delta^3(\boldsymbol{r}' - \boldsymbol{r}) \nabla^2 \langle 0' | \overline{\psi_s^+(\boldsymbol{r}')} \psi_s(\boldsymbol{r}) | 0' \rangle \\
&= -\frac{k_F^3 \hbar^2}{12\pi^2 m} \sum_s \int d^3 r' \int d^3 r \, \delta^3(\boldsymbol{r}' - \boldsymbol{r}) \nabla^2 \ell(k_F | \boldsymbol{r} - \boldsymbol{r}' |) \\
&= -\frac{k_F^3 \hbar^2}{6\pi^2 m} \int d^3 r \left[\nabla'^2 \ell(k_F r') \big|_{r'=0} \right] \\
&= -\frac{k_F^3 \hbar^2 L^3}{6\pi^2 m} \left[\nabla'^2 \ell(k_F r') \big|_{r'=0} \right] \tag{39.50}
\end{aligned}
$$

where in the second line, we insert a delta function to make sure that we take the derivative of just the $\psi_s(\boldsymbol{r})$. Once you have done this a few times, you will see how it works out and you can skip putting in the delta function. Expanding

$$
\begin{aligned}
\ell(x) &= \frac{3}{x^3} \left(x - \frac{x^3}{6} + \frac{x^5}{120} + \ldots - x + \frac{x^3}{2} - \frac{x^5}{24} + \ldots \right) \\
&= \left(1 - \frac{x^2}{10} + \ldots \right) \tag{39.51}
\end{aligned}
$$

and

$$
\langle KE \rangle = \frac{3}{5} \frac{\hbar^2 k_F^2}{2m} \frac{k_F^3 L^3}{3\pi^2} . \tag{39.52}
$$

Often we want the energy per particle. The number of particles is

$$
\begin{aligned}
N &= \sum_s \int d^3 r \langle 0' | \psi_s^+(\boldsymbol{r}) \psi_s(\boldsymbol{r}) | 0' \rangle \\
&= \sum_s \int d^3 r \langle 0' | \overline{\psi_s^+(\boldsymbol{r})} \psi_s(\boldsymbol{r}) | 0' \rangle \\
&= \frac{k_F^3}{6\pi^2} \sum_s \int d^3 r \, \ell(0) \\
&= \frac{k_F^3 L^3}{3\pi^2} . \tag{39.53}
\end{aligned}
$$

We get the usual spin $1/2$ free gas result

$$
n = \frac{N}{L^3} = \frac{k_F^3}{3\pi^2} \tag{39.54}
$$

which defines the fermi wave vector as

$$
k_F = (3\pi^2 n)^{1/3} . \tag{39.55}
$$

The kinetic energy per particle is then

$$\left\langle \frac{KE}{N} \right\rangle = \frac{3}{5} \frac{\hbar^2 k_F^2}{2m}. \tag{39.56}$$

The 2-body potential energy is

$$\langle PE \rangle = \frac{1}{2} \sum_{ss'} \int d^3r d^3r' v_2(|\boldsymbol{r} - \boldsymbol{r}'|) \langle 0'|\psi_s^+(\boldsymbol{r})\psi_{s'}^+(\boldsymbol{r}')\psi_{s'}(\boldsymbol{r}')\psi_s(\boldsymbol{r})|0'\rangle \tag{39.57}$$

The matrix element we need is

$$
\begin{aligned}
\langle 0'|\psi_s^+(\boldsymbol{r})\psi_{s'}^+(\boldsymbol{r}')\psi_{s'}(\boldsymbol{r}')\psi_s(\boldsymbol{r})|0'\rangle &= \langle 0'|\overbrace{\psi_s^+(\boldsymbol{r})\overbrace{\psi_{s'}^+(\boldsymbol{r}')\psi_{s'}(\boldsymbol{r}')\psi_s(\boldsymbol{r})}}|0'\rangle \\
&\quad + \langle 0'|\overbrace{\psi_s^+(\boldsymbol{r})\overbrace{\psi_{s'}^+(\boldsymbol{r}')\psi_{s'}(\boldsymbol{r}')}\psi_s(\boldsymbol{r})}|0'\rangle \\
&= \frac{k_F^6}{36\pi^4} \left[\ell^2(0) - \delta_{ss'}\ell^2(k_F|\boldsymbol{r} - \boldsymbol{r}'|) \right] \\
&= \frac{k_F^6}{36\pi^4} \left[1 - \delta_{ss'}\ell^2(k_F|\boldsymbol{r} - \boldsymbol{r}'|) \right] \tag{39.58}
\end{aligned}
$$

The integral is most easily done by converting to relative and center of mass coordinates for the pair. The center of mass integral gives a factor of L^3, leaving

$$\langle PE \rangle = \frac{n^2 L^3}{2} \int d^3r v_2(r) - \frac{n^2 L^3}{4} \int d^3r v_2(r)\ell^2(k_F r) = N \left[\frac{n}{2} \int d^3r v_2(r) - \frac{n}{4} \int d^3r v_2(r)\ell^2(k_F r) \right]. \tag{39.59}$$

Since the Hamiltonian is translationally invariant, unless there is a spontaneously broken symmetry, we expect the orbitals to be eigenstates of the translation operator, therefore they must be plane waves. This is therefore the self-consistent Hartree-Fock result for a uniform system.

The jellium model is a crude model of a metal. Instead of the positive nuclei (or ions) we imagine that the canceling positive charge is smeared out into a uniform positive jelly. This cancels the diverging $k = 0$ term in the Fourier transform of the electron-electron coulomb potential. We can see this explicitly in our calcuation since the direct term in the potential is

$$V_d = \frac{1}{2} \int d^3r' \int d^3r \frac{(-en)^2}{|\boldsymbol{r} - \boldsymbol{r}'|} \tag{39.60}$$

The positive charged background would give exactly the same energy, while the electron-positive charge interaction would give two times this with a negative sign. Therefore, the jellium model Hartree-Fock energy per particle is

$$\left\langle \frac{E}{N} \right\rangle = \frac{3}{5} \frac{\hbar^2 k_F^2}{2m} - \frac{ne^2}{4} \int d^3r \frac{\ell^2(k_F r)}{r}. \tag{39.61}$$

The integral is straightforward

$$\left\langle \frac{E}{N} \right\rangle = \frac{3}{5}\frac{\hbar^2 k_F^2}{2m} - \frac{9\pi n e^2}{4k_F^2}$$

$$= \frac{3}{5}\frac{\hbar^2 k_F^2}{2m} - \frac{3k_F e^2}{4\pi} \tag{39.62}$$

Traditionally, the energy is written in Rydbergs $(e^2/2a_0)$, and instead of k_F, the quantity r_s is used where

$$n = \frac{3}{4\pi a_0^3 r_s^3} \tag{39.63}$$

Plugging these in we get in Rydbergs,

$$\left\langle \frac{E}{N} \right\rangle = \frac{3}{5}\left(\frac{9\pi}{4}\right)^{2/3} r_s^{-2} - \frac{3}{2\pi}\left(\frac{9\pi}{4}\right)^{1/3} r_s^{-1}$$

$$= \frac{2.2099}{r_s^2} - \frac{0.9163}{r_s} \tag{39.64}$$

Already our result shows one of the oddities of the electron gas. At high densities, r_s becomes small, and the kinetic energy dominates. That is, the free gas approximation becomes better at high densities for the electron gas. In fact, the jellium model crystallizes at low densities as Wigner showed.

Chapter 40

Green's functions in quantum mechanics

40.1 Motivation

Quantum Green's functions were originally motivated by perturbation theory arguments. If we write a time dependent perturbation theory treatment where we slowly turn on our interactions, we can write all our matrix elements as the expectation value of time-ordered operators made up of creation and anihilation operators, $\psi_s^+(\boldsymbol{r})$ and $\psi_s(\boldsymbol{r})$. Wick's theorem will then lead to contractions of these operators. If we arrange for the normal-ordered operators to give zero, we can express all of our contractions as expectations of the time-ordered product of pairs of creation and destruction operators. These are the one-body Green's functions.

Alternatively we can realize that these one-body Green's functions allow us to calculate many of the quantities that we most want.

The expectation value of all one-body operators can be obtained from matrix elements where we destroy a particle at some position $\boldsymbol{r}$ and time t and create a particle at a position $\boldsymbol{r}'$ and time t'. It is convenient to go into the Heisenberg picture (and for perturbation theory, the interaction picture). We therefore write

$$iG_{s,s'}(\boldsymbol{r}, t, \boldsymbol{r}', t') = \frac{\langle \Psi | T[\psi_{Hs}(\boldsymbol{r}, t)\psi_{Hs'}^+(\boldsymbol{r}', t')] | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{40.1}$$

The capital H subscript is to indicate that these are Heisenberg operators. The factor of i is a conventional choice. We have written the Green's function with the time ordering operator so that we can have simultaneously the result when we first create a particle and then destroy one and the result when we first destroy and then create. All of our physical matrix elements will have the operators in time order. It is also possible to write retarded and advanced Green's functions which are zero if t is less than t' or vice versa.

40.2 Ground-state energy

We could go on and write two-body Green's functions which would allow us to calculate the expectation values of all two-body operators. Often we want the ground-state energy for a Hamiltonian that has a two-body potential. It turns out that this can be extracted from the one-body Green's function.

$$\frac{\partial}{\partial t}\psi_{Hs}(\boldsymbol{r},t) \;=\; \frac{i}{\hbar}[H,\psi_{Hs}(\boldsymbol{r},t)] \tag{40.2}$$

Chapter 41

Relativistic hydrogen atom

41.1 Introduction

Let's look at the bound states of the coulomb problem using the Klein-Gordon and Dirac equations. Since we are most interested in the nonrelativistic limit, we take Coulomb gauge, and drop the vector potential terms. To go further, we would need to use renormalized quantum electrodynamics and include all the relativistic effects – the quantized electromagnetic, the quantized Klein-Gordon field or the Dirac sea of electrons, or the equivalent field theories.

41.2 The Klein-Gordon bound-state energy

Coupling the scalar potential to the energy gives the Klein-Gordon energy eigenstate equation as

$$(p^2c^2 + m^2c^4)|\psi\rangle = \left(E + \frac{Ze^2}{r}\right)^2 |\psi\rangle. \tag{41.1}$$

Sometimes it is convenient to write the energy as $E = mc^2 + \epsilon$, so that the Klein-Gordon equation becomes

$$\left[\frac{p^2}{2m} - \frac{Ze^2}{r} - \epsilon - \frac{1}{2mc^2}\left(\epsilon^2 + \frac{2\epsilon Ze^2}{r} + \frac{Z^2e^4}{r^2}\right)\right]|\psi\rangle = 0 \tag{41.2}$$

Notice that dropping the c^{-2} terms will give the nonrelativistic Schrödinger equation as expected.

The Klein-Gordon hydrogen atom equation is invariant under spatial rotations, so we can find simultaneous eigenvectors of L^2, L_z, and energy. The eigenfunctions of L^2 and L_z are $R_\ell(r)Y_{\ell m}(\theta, \phi)$, giving the radial equation

$$\left[-\frac{1}{r}\frac{d^2}{dr^2}r + \frac{\ell(\ell+1)}{r^2} - \frac{1}{\hbar^2c^2}\left(E^2 - m^2c^4 + \frac{2EZe^2}{r} + \frac{Z^2e^4}{r^2}\right)\right]R_\ell(r) = 0. \tag{41.3}$$

We can now follow the nonrelativistic solution method. For the bound states, if we ignore the inverse powers of r at large r, the normalizable solution decays exponentially. We therefore try the power series solution

$$R_\ell(r) = \sum_{j=0} a_j r^{\gamma+j} e^{-\beta r}, \tag{41.4}$$

with

$$\beta^2 \equiv \frac{m^2 c^4 - E^2}{\hbar^2 c^2}. \tag{41.5}$$

Substituting we obtain the recursion relation (writing the fine-structure constant $e^2/\hbar c = \alpha$)

$$a_{j+2}\left[(\gamma+j+1)(\gamma+j) - \ell(\ell+1) + Z^2\alpha^2\right] + a_{j+1}\left[\frac{2EZ\alpha}{\hbar c} - 2(\gamma+j+1)\beta\right] = 0. \tag{41.6}$$

Just as in the nonrelativistic case, in order to be sensible at the origin, $rR \to 0$ at the origin, so

$$(\gamma+1)\gamma - \ell(\ell+1) + Z^2\alpha^2 = 0, \tag{41.7}$$

with solution normalizable at the origin for small enough $Z\alpha$,

$$\gamma = \sqrt{\left(\ell + \frac{1}{2}\right)^2 - Z^2\alpha^2} - \frac{1}{2}. \tag{41.8}$$

Again, as in the nonrelativistic case, if the series does not truncate, the function is not normalizable, therefore we require the coefficient of a_{j+1} above to be zero for some value of j. Solving we find

$$E_{n\ell} = \left[\frac{1}{1 + \frac{Z^2\alpha^2}{(\gamma+n)^2}}\right]^{1/2} mc^2 = \left[\frac{1}{1 + \frac{Z^2\alpha^2}{\left(\sqrt{(\ell+\frac{1}{2})^2 - Z^2\alpha^2} - \frac{1}{2} + n\right)^2}}\right]^{1/2} mc^2 \tag{41.9}$$

where n takes the values $1, 2, ...$

Expanding in powers of α,

$$E_{n\ell} = mc^2 - \frac{1}{2(n+\ell)^2}Z^2\alpha^2 mc^2 - \frac{8n - 3 + 2\ell}{8(\ell+n)^4(2\ell+1)}Z^4\alpha^4 mc^2 + ... \tag{41.10}$$

The first term is the rest energy, the next is the nonrelativistic Bohr energy, and the next term gives the nonradiative relativistic fine structure. As Schrödinger realized immediately, this does not agree with the experimental hydrogen atom fine structure. It does, along with radiative corrections, give the fine structure of pionic atoms[1].

---

[1] see, for example, L. Delker, G. Dugan, C.S. Wu, D.C. Lu, A.J Caffrey, Y.T. Cheng, and Y.K Lee, *Experimental verification of the relativistic fine-structure term of the Klein-Gordon equation in pionic titanium atoms*, Phys. Rev. Lett. **42**, 89 (1979).

41.3 The Dirac bound-state energy

The corresponding Dirac equation is

$$\boldsymbol{\sigma} \cdot \boldsymbol{p} c |\psi_l\rangle = \left(E - mc^2 + \frac{Ze^2}{r} \right) |\psi_u\rangle$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{p} c |\psi_u\rangle = \left(E + mc^2 + \frac{Ze^2}{r} \right) |\psi_l\rangle . \tag{41.11}$$

It is convenient to write the Dirac equation in terms of the upper component,

$$|\psi_l\rangle = \frac{1}{E + mc^2 + \frac{Ze^2}{r}} \boldsymbol{\sigma} \cdot \boldsymbol{p} c |\psi_u\rangle$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{p} c \frac{1}{E + mc^2 + \frac{Ze^2}{r}} \boldsymbol{\sigma} \cdot \boldsymbol{p} c |\psi_u\rangle = \left(E - mc^2 + \frac{Ze^2}{r} \right) |\psi_u\rangle \tag{41.12}$$

We can write this as

$$\left\{ p^2 c^2 - i\hbar c^2 \frac{1}{E + mc^2 + \frac{Ze^2}{r}} \frac{Ze^2}{r^3} \left[\boldsymbol{r} \cdot \boldsymbol{p} + i\boldsymbol{\sigma} \cdot (\boldsymbol{r} \times \boldsymbol{p}) \right] \right\} |\psi_u\rangle = \left(E^2 - m^2 c^4 + \frac{2EZe^2}{r} + \frac{Z^2 e^4}{r^2} \right) |\psi_u\rangle . \tag{41.13}$$

This is not invariant under spatial rotations alone. The spin-orbit term couples space and spin, but it is invariant under rotation of both space and spin. The generator of these rotations is $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$. The spin orbit term will be diagonal for eigenstates of J^2, L^2 and S^2, since

$$2\boldsymbol{S} \cdot \boldsymbol{L} = J^2 - L^2 - S^2 = \hbar^2 \left[j(j+1) - \ell(\ell+1) - \frac{3}{4} \right] . \tag{41.14}$$

We therefore write eigenstates of J^2, J_z, L^2 and S^2. Just like the orbital states, these are angular momentum eigenstates for any radial function, so a general energy eigenstate will have the form

$$R_{jm}^\ell(r) |jm\ell s\rangle . \tag{41.15}$$

Since s is always $1/2$, we don't display it in the radial indices. $j = \ell \pm \frac{1}{2}$, except of course for $\ell = 0$ where $j = \frac{1}{2}$.

Substituting, we get the radial equation

$$\left\{ -\frac{1}{r} \frac{d^2}{dr^2} r + \frac{\ell(\ell+1)}{r^2} + \frac{Z\alpha r^{-3}}{\frac{E+mc^2}{\hbar c} + Z\alpha r^{-1}} \left(-r \frac{d}{dr} + \left[j(j+1) - \ell(\ell+1) - \frac{3}{4} \right] \right) \right\} R_{jm}^\ell(r)$$

$$= \left(\frac{E^2 - m^2 c^4}{\hbar^2 c^2} + \frac{2EZ\alpha}{\hbar c r} + \frac{Z^2 \alpha^2}{r^2} \right) R_{jm}^\ell(r) \tag{41.16}$$

$$\left\{\left(\frac{E+mc^2}{\hbar c}+Z\alpha r^{-1}\right)\left(-\frac{1}{r}\frac{d^2}{dr^2}r+\frac{\ell(\ell+1)}{r^2}-\frac{E^2-m^2c^4}{\hbar^2 c^2}-\frac{2EZ\alpha}{\hbar cr}-\frac{Z^2\alpha^2}{r^2}\right)\right.$$

$$\left.+Z\alpha r^{-3}\left(-r\frac{d}{dr}+\left[j(j+1)-\ell(\ell+1)-\frac{3}{4}\right]\right)\right\}R_{jm}^\ell(r)=0$$

$$(41.17)$$

At large r we require $R \to e^{-\beta r}$ with $\beta^2 = \frac{m^2c^4-E^2}{\hbar^2 c^2}$, so we try the power series

$$R_{jm}^\ell(r)=\sum_{k=0}a_k r^{\gamma+k}e^{-\beta r}\,. \qquad (41.18)$$

Plugging this in, the coefficient of $a_0 r^{\gamma-3}e^{-\beta r}e^{-\beta r}$ must be zero which gives the equation

$$\gamma^2+2\gamma-j(j+1)+\frac{3}{4}+Z^2\alpha^2=0 \qquad (41.19)$$

with solution

$$\gamma_\pm+1=\pm\sqrt{\left(j+\frac{1}{2}\right)^2-Z^2\alpha^2} \qquad (41.20)$$

The negative sign gives an unnormalizable function at the origin except for the case where $j=\frac{1}{2}$ and $Z\alpha=0$, but in that case it gives a r^{-1} behavior, which would give a delta function in the ∇^2, and no such delta function exists in the Hamiltonian. Therefore we must take the positive sign. This is analogous to taking the r^ℓ and rejecting the $r^{-\ell-1}$ solutions for the nonrelativistic problem. Notice that for $j=\frac{1}{2}$, that the radial function diverges at the origin – divergence by itself is not a problem as long as it is a solution to the differential equation and does not give rise to pathological terms that do not occur in the Hamiltonian. However, if $Z\alpha>1$, the function is not normalizable and $Z\alpha=1$ has a pathological delta function in its ∇^2. The Dirac equation solution therefore requires $Z\alpha<1$, or $Z<137$. Real nuclei are not point charges, and in any case the nuclei with the largest charge in the standard periodic table has $Z=111$.

We have

$$\gamma=\sqrt{\left(j+\frac{1}{2}\right)^2-Z^2\alpha^2}-1 \qquad (41.21)$$

Chapter 42

Semiclassical method

42.1 Introduction

Obviously quantum mechanics must reproduce all the results of classical mechanics for systems which are well described by classical mechanics. Roughly this will be when the action or at least changes in the action are much larger than $\hbar$.

The idea of semiclassical methods is expand our equations and results in powers of $\hbar$. The results that are independent of $\hbar$ are (with a few pathological exceptions[1]) are given by classical mechanics. The terms proportional to higher powers of $\hbar$ then give quantum corrections. If these are small, then the quantum corrected classical result is called a semiclassical result.

We will begin by looking at the one-dimensional Schrödinger equation

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) + \frac{\hbar}{i} \frac{\partial}{\partial t} \right] \psi(x, t) = 0 \tag{42.1}$$

We can, without loss of generality, write

$$\psi(x, t) = e^{\frac{i}{\hbar} W(x,t)} \tag{42.2}$$

since any complex value can be written as the exponential of a complex number. We choose the prefactor of the exponent so that the result of substituting into the Schrödinger equation gives terms with no $\hbar$ factors. Substituting we get

$$\frac{1}{2m} \left[\left(\frac{\partial W(x,t)}{\partial x} \right)^2 + i\hbar \frac{\partial^2 W(x,t)}{\partial x^2} \right] + V(x) + \frac{\partial W(x,t)}{\partial t} = 0 \,. \tag{42.3}$$

If we drop the term with the $\hbar$ factor, the remaining equation is the Hamilton-Jacobi equation of classical mechanics[?]. It's solution gives a complete solution of the classical mechanics problem. The term with $\hbar$ therefore generates the quantum corrections.

---

[1]Like the one-dimensional scattering from a potential step where the reflection and transmission probabilities are independent of $\hbar$

Typically, we will use the semiclassical approximation to calculate approximate energy eigenfunctions and occasionally energy eigenvalues. In that case, we write

$$W(x,t) = S(x) - Et\,. \tag{42.4}$$

The Solution of the time independent Schrödinger equation for the energy eigenfunctions becomes

$$\frac{1}{2m}\left[\left(\frac{dS(x,t)}{dx}\right)^2 + i\hbar\frac{d^2S(x)}{dx^2}\right] + V(x) - E = 0 \tag{42.5}$$

The method of solution of this equation is usually called the WKB method, for the initials of Wenzel, Kramers, and Brillouin, who worked on this problem. Occasionally Jeffries' name is added, and you will also see reference to the JWKB or WKBJ method. These all refer to the same method.

42.2   The WKB method

The WKB method starts with the obvious expansion in powers of $\hbar$. We write

$$S(x) = S_0(x) + \hbar S_1(x) + \hbar^2 S_2(x) + ... \tag{42.6}$$

Substituting this into Eq. 42.5 and equating the coefficients of powers of $\hbar$, we get

$$\begin{aligned}
[S_0'(x)]^2 &= 2m[E - V(x)] \\
2S_0'(x)S_1'(x) + iS_0''(x) &= 0 \\
2S_0'(x)S_2'(x) + [S_1'(x)]^2 + iS_1''(x) &= 0 \\
&\quad ..., \text{etc.}.
\end{aligned} \tag{42.7}$$

The first equation has solutions

$$S_0(x) = \begin{cases} \pm\int dx\sqrt{2m[E - V(x)]} + \text{Constant} & E > V(x) \\ \pm i\int dx\sqrt{2m[V(x) - E]} + \text{Constant} & E < V(x) \end{cases} \quad cc\,. \tag{42.8}$$

Notice that in the first expression the integrand is the magnitude of the classical momentum. If we had a constant potential, the result would be a plane wave with momentum eigenvalue of either $\pm\sqrt{2m(E - V)}$. For a potential that changes with position, the result says the phase change will be given by the local value of the classical momentum. The form for $E < V(x)$ corresponds to the classically forbidden region, so it does not have a classical analog, but again, for a constant potential, it says the local exponential decay is set by the local value of the potential.

If we continue, the equation corresponding to the terms of order $\hbar$, can be written as

$$\frac{dS_1(x)}{dx} = -\frac{i}{2}\frac{S_0''(x)}{S_0'(x)} = -\frac{i}{2}\frac{d}{dx}\ln[S_0'(x)] \tag{42.9}$$

This can be immediately integrated to give

$$S_1(x) = -\frac{i}{2}\ln[S_0'(x)] + \text{Constant} = -\frac{i}{2}\ln|S_0'(x)| + \text{Constant}. \qquad (42.10)$$

where the phase of $S_0'(x)$ in the last expression has been included in the arbitrary additive constant.

Truncating our expansion at this level, the corresponding approximate solutions to the original Schrödinger equation are

$$\psi(x) = \begin{cases} \dfrac{A}{\sqrt{k(x)}}e^{i\int dx k(x)x}\dfrac{B}{\sqrt{k(x)}}e^{-i\int dx k(x)x} & E > V(x) \\[3mm] \dfrac{A}{\sqrt{\kappa(x)}}e^{i\int dx \kappa(x)x}\dfrac{B}{\sqrt{\kappa(x)}}e^{-i\int dx \kappa(x)x} & E < V(x) \end{cases}.$$

with

$$k(x) = \sqrt{\frac{2m(E - V(x))}{\hbar^2}}$$

$$\kappa(x) = \sqrt{\frac{2m(V(x) - E)}{\hbar^2}}. \qquad (42.11)$$

The factor from $S_1(x)$ also has a simple classical interpretation in the classically allowed region. A particle moving in a potential $V(x)$ with $E > V(x)$, will spend a time $dt = \frac{dx}{v(x)}$ in the region between x and $x+dx$, where $v(x)$ is its velocity. Therefore, the probability density of finding the particle at a position x will be proportional to its inverse velocity. When the terms of the wave function $\psi(x)$ above are squared, the terms from $S_1(x)$ give exactly this relative probability. Again, the expression in the classically forbidden region does not have a simple analog, but the mathematics is the same.

Next, we need to look at where this approximation is valid. A simple estimate is when $\hbar|S_1'(x)| \ll |S_0'(x)|$. Evaluating, dropping overall constants, we require

$$\frac{\hbar V'(x)}{\sqrt{m}|E - V(x)|^{3/2}} \ll 1. \qquad (42.12)$$

We see that for $E - V(x)$ positive, this corresponds to the particle De Broglie wavelength $\lambda_D = \frac{2\pi\hbar}{p} = \frac{2\pi\hbar}{\sqrt{2m[E-V(x)]}}$ being large compared to the distance over which the potential changes substantially $\frac{V(x)}{V'(x)}$.

Eq. 42.12 also shows us that we can never satisfy this requirement near the classical turning point when $E \sim V(x)$. This is a major problem, since we might want to use the semiclassical method to calculate bound state energy eigenfunctions and eigenvalues. Without a solution at the turning points, there is no way to match the semiclassical solution for $E > V(x)$ to the semiclassical solution for $E < V(x)$.

The WKB method circumvents this problem by breaking the solution up into regions where Eq. 42.12 is satisfied, and regions around any classical turning points. The potential

is Taylor series expanded around the classical turning point. The result is a linear potential. The solutions to the linear potential problem are Airy functions. This, in itself, isn't particularly helpful, but the Airy functions have an easily derived asymptotic expansion which applies far from the turning point. These asymptotic expansions are exactly the semiclassical results above. What this analysis then gives is a way to match the semiclassical solutions across the turning points. The Airy function solutions give equations relating the coefficients of the $E > V(x)$ form to the coefficients of the $E < V(x)$ form. These are called the WKB connection relations.

42.3 Linear potential

Appendix A

Analytic functions and contour integration

A.1 Analytic Functions

A complex function $f(z) = f(x, y)$, of a complex variable, $z = x + iy$, is analytic at the point z if the derivative

$$\frac{df(z)}{dz} = \lim_{h \to 0} \frac{f(z + h) - f(z)}{h} \tag{A.1}$$

is independent of how the limit is taken. In particular, if h is purely real or purely imaginary, the result must be the same for the function to be analytic

$$\frac{\partial f(x, y)}{\partial x} = \frac{\partial f(x, y)}{\partial iy} . \tag{A.2}$$

Taking the real and imaginary parts of f, we have the Cauchy-Riemann equations,

$$
\begin{aligned}
\frac{\partial \mathrm{Re} f(x, y)}{\partial x} &= \frac{\partial \mathrm{Im} f(x, y)}{\partial y} \\
\frac{\partial \mathrm{Im} f(x, y)}{\partial x} &= -\frac{\partial \mathrm{Re} f(x, y)}{\partial y} .
\end{aligned} \tag{A.3}
$$

Taylor series expanding $f(z + h)$ for a general $h = h_r + ih_i$, we have

$$\lim_{h \to 0} \frac{f(z + h) - f(z)}{h} = \left[\frac{\partial f(x, y)}{\partial x} h_r + \frac{\partial f(x, y)}{\partial y} h_i \right] \frac{1}{h_r + ih_i} . \tag{A.4}$$

This is independent of the ratio of h_r and h_i if the Cauchy-Riemann equations are satisfied showing they are also sufficient conditions for the function f to be analytic at z.

Taking the x or y derivative of the first Cauchy-Riemann condition and substituting the

y or x derivative of the second shows that

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right] \mathrm{Re} f(x, y) = 0$$

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right] \mathrm{Im} f(x, y) = 0 \tag{A.5}$$

so that both the real and imaginary parts of an analytic function satisfy the two-dimensional Laplace's equation.

A.2 Contour integration

We now need to show that the integral of a complex function around a closed path in the complex plane is zero if the function is analytic in the interior of the path.

The integral along the path is given by dividing the path into infinitesimal segments, multiplying the complex Δz of the segment times the value of function on the segment, and summing to give the integral.

We can define a path by a parametric equation $z(\alpha) = f(\alpha)$ where α is real with a given range, and f is a complex function. For example the path around a circle, of radius a, centered on the origin, could by given by

$$z(\alpha) = a\cos(\alpha) + ia\sin(\alpha) = ae^{i\alpha}, \quad 0 \le \alpha \le 2\pi \tag{A.6}$$

The integral is then

$$\int_{\text{path}} dz f(z) = \int d\alpha \frac{dz(\alpha)}{d\alpha} f(z(\alpha)) \tag{A.7}$$

For example the integral around a circle of radius a centered on the origin of the function $1/z$, is

$$\int_{\text{circle}} dz \frac{1}{z} = \int_0^{2\pi} d\alpha \frac{d(ae^{i\alpha})}{d\alpha} \frac{1}{ae^{i\alpha}} = i \int_0^{2\pi} d\alpha = 2\pi i \tag{A.8}$$

independent of the radius a.

A.3 Contour integration of analytic functions

Now notice that the Cauchy-Rieman conditions look like the $\hat{z}$ component of a curl set to zero.[1] To take advantage of this, we can define a two-dimensional vector field in terms of $f(z)$ as

$$\boldsymbol{F}(x, y) = f(z)\left[\hat{\boldsymbol{x}} + i\hat{\boldsymbol{y}}\right]. \tag{A.9}$$

---

[1]The notation here is a little awkward. $\hat{z}$ will correspond to the direction perpenducular to our complex $x - y$ plane. Without the hat z will by $z = x + iy$. Luckily, we only need to use this notation in one equation.

You can verify that the Cauchy-Rieman conditions ensure that its curl is zero if $f(z)$ is analytic. Let's integrate its curl over a surface in the x, y-plane. For convenience take the direction of the normal to the surface to be along positive $\hat{z}$, so that

$$0 = \int_S d\boldsymbol{S} \cdot \boldsymbol{\nabla} \times \boldsymbol{F}(x, y) = \int_S dx dy \hat{z} \cdot [\boldsymbol{\nabla} \times \boldsymbol{F}(x, y)] \tag{A.10}$$

if $f(z)$ is analytic everywhere on the integrated surface. We can use Stokes theorem (or prove it by integrating by parts or integrating over a small region and Taylor expanding $\boldsymbol{F}(x, y)$) to write

$$0 = \oint d\boldsymbol{\ell} \cdot \boldsymbol{F}(x, y) \tag{A.11}$$

where the integral is around the boundary of the surface. This line integral would be parameterized by giving $\boldsymbol{\ell}(\alpha)$ as

$$\boldsymbol{\ell}(\alpha) = \hat{\boldsymbol{x}} \ell_x(\alpha) + \hat{\boldsymbol{y}} \ell_y(\alpha) \tag{A.12}$$

over some range of α^2

If we plug in the form of Eq. A.9, Eq. A.11 becomes

$$\begin{aligned}
0 &= \int d\alpha \left[\frac{d\ell_x(\alpha)}{d\alpha} + i \frac{d\ell_y(\alpha)}{d\alpha} \right] f(\ell_x(\alpha) + i\ell_y(\alpha)) \\
&= \int d\alpha \frac{z(\alpha)}{d\alpha} f(z(\alpha)) \\
&= \oint_{\text{closed contour}} dz f(z) \tag{A.15}
\end{aligned}$$

where $z(\alpha) = \ell_x(\alpha) + i\ell_y(\alpha)$.

We have shown that the integration around a closed contour is zero if the integrand is analytic in the interior of the contour[3].

---

[2]For example a circle of radius a centered on the origin, could be given by

$$\boldsymbol{\ell}(\alpha) = a \cos(\alpha) \hat{\boldsymbol{x}} + a \sin(\alpha) \hat{\boldsymbol{y}}, \ 0 \le \alpha \le 2\pi \tag{A.13}$$

and the integral done

$$\int_0^{2\pi} d\alpha \frac{d\boldsymbol{\ell}(\alpha)}{d\alpha} \cdot \boldsymbol{F}(\ell_x(\alpha), \ell_y(\alpha)). \tag{A.14}$$

[3]This has many applications in two-dimensional electrostatics because the real and imaginary parts of an analytic function separately satisfy Laplace's equation so they can represent electrostatic potentials in charge free regions. In addition, if an electric field is constructed from the real and imaginary parts, it can be derived from another electrostatic potential given by the contour integration. The result is independent of contour so the integration around a closed path is zero.
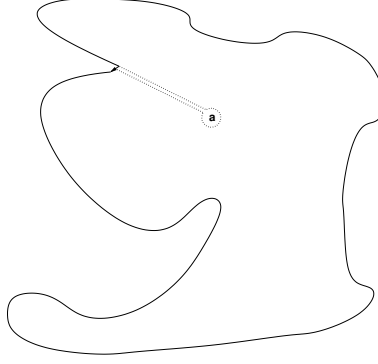
Figure A.1: The solid line shows the integration contour around a pole at $z = a$. The dotted line shows how we change to a new contour that has no poles inside. The two straight lines leading to the pole are shown apart, but are taken infinitesimally close together so their contributions cancel. Since the integrand is analytic inside the new contour the circular integration around the pole must be equal and opposite to the integration around the original contour. The circular integration is done as in Eq. A.8.

A.4 Singularities

Often a function that is analytic in some regions has singularities. For example z^{-n} for positive integer n has a singularity at the origin. You can easily verify it is analytic at all other points. A function with a singularity that goes like $(z - a)^{-1}$ has a simple pole at a. If it goes like $(z - a)^{-n}$ for a positive integer n, it has a higher order pole. The contour integration around any of these singularities is given by Cauchy's theorem.

Cauchy's theorem says that if $f(z)$ is analytic, then

$$f(a) = \frac{1}{2\pi i} \oint dz \frac{f(z)}{z - a} \tag{A.16}$$

where the contour encloses the point a and is in the positive (counter clockwise) direction. We can show this by modifying the original contour by including a small circle around the point a, and connecting both contours by a pair of lines infinitesimally close together as shown in fig. A.1. The new contour does not enclose the singularity at $z = a$, so its integral is zero. The integrations along the lines joining the original contour and the small circle cancel since they are done in opposite directions. Therefore the integration around the small circle must cancel the original contour. Around the small circle, we can Taylor series expand $f(z)$, and only the $f(a)$ term will contribute. The integration around the circle is exactly like that of Eq. A.8, giving the Cauchy's theorem result.

We can either repeat this analysis for the denominator $(z - a)^n$ or simply take the derivative with respect to a of Cauchy's theorem to get

$$\frac{d^n f(a)}{da^n} = \frac{n!}{2\pi i} \oint dz \frac{f(z)}{(z - a)^{(n+1)}} . \tag{A.17}$$

476

These arguments go through exactly the same way for multiple poles. To integrate a function around a closed contour which is analytic except at a set of poles, we change to a contour like that of fig. A.1, where the function is analytic everywhere in the interior. The original contour must then be canceled by the integrations around the infinitesimal circles around the poles. Notice that the direction (clockwise or counterclockwise) of the integration around the poles will be opposite the direction of the original contour. That means that we can replace the original integration by the integrations around the little circles with the integrations around the circles in the same direction as the original contour.

To repeat, the circular integrations are all done the same way. For a pole at $z = a$, we write the circular contour as $z(\alpha) = a + Re^{i\alpha}$, with R real, the radius of the circle. Taking α from 0 to 2π gives a counterclockwise integration. Since we can take $R \to 0$, we can expand the rest of the integrand in a Taylor expansion. That is writing the mth derivative of $f(z)$ as $f^{(m)}(z)$ we can write the integrand for an nth order pole as

$$
\begin{aligned}
\oint dz \frac{f(z)}{(z-a)^n} &= \oint dz \frac{\sum_{m=0}^{\infty} \frac{1}{m!} f^{(m)}(a)(z-a)^m}{(z-a)^n} \\
&= \int_0^{2\pi} d\alpha i R e^{i\alpha} \frac{\sum_{m=0}^{\infty} \frac{1}{m!} f^{(m)}(a) R^m e^{im\alpha}}{R^n e^{i\alpha n}} \\
&= \int_0^{2\pi} d\alpha \sum_{m=0}^{\infty} \frac{1}{m!} f^{(m)}(a) R^{m-(n-1)} e^{i(m-(n-1))\alpha} \\
&= 2\pi i \frac{f^{(n-1)}(a)}{(n-1)!}
\end{aligned}
\tag{A.18}
$$

where, as usual, $f^{(0)}(a) = f(a)$. The expression multiplying the $2\pi i$ is called the residue.

A.4.1 Examples

Let's calculate a few integrals. The integral

1.

$$
\int_0^\infty \frac{dx}{1+x^4}
\tag{A.19}
$$

can be done without contour integration. To employ contour integration, we first note that the integrand is even under $x \to -x$, so we can write

$$
\int_0^\infty \frac{dx}{1+x^4} = \frac{1}{2} \int_{-\infty}^\infty \frac{dx}{1+x^4}
\tag{A.20}
$$

Since the integrand dies off faster than x^{-1} for large x, we can change x to z and close the contour with a large semicircle in either the upper or lower half plane. The poles of the integrand are at $z^4 = -1$ or $z^4 = e^{i\pi+2\pi n}$ so that there are poles at $e^{i\pi/4} = (1+i)/\sqrt{2}$, $e^{i3\pi/4} = (-1+i)/\sqrt{2}$, $e^{-i\pi/4} = (1-i)/\sqrt{2}$, $e^{-i3\pi/4} = (-1-i)/\sqrt{2}$, Taking the contour in the upper half plane, only the poles with positive imaginary

parts contribute, and have residues of $-i2^{-5/2}(1-i)$ and $-i2^{-5/2}(1+i)$ so that integral is

$$\int_0^\infty \frac{dx}{1+x^4} = \frac{\pi}{2\sqrt{2}}. \tag{A.21}$$

2. The integral

$$\int_0^\infty dx \frac{\sin x}{x} \tag{A.22}$$

is symmetric so we can extend the integral to the whole real axis

$$\int_0^\infty dx \frac{\sin x}{x} = \frac{1}{2} \int_{-\infty}^\infty dx \frac{\sin x}{x}. \tag{A.23}$$

If we try to calculate the integral on the large semicircles, we find that the sine diverges for either semicircle. To get around this we can write $\sin(x) = (e^{ix} - e^{-ix})/2i$. The exponential e^{ix} can be closed with a large semicircle in the upper half plane since the exponential will give zero contribution there. Similarly e^{-ix} can be closed in the lower half plane. We have one other problem, while $\sin(x)/x$ is well behaved at $x \to 0$, $e^{\pm ix}/x$ has a pole there. A simple way to deal with this is to realize that since there is no real singularity, $\sin(x)/x$, is analytic near the origin we can change the contour to be a small semicircle around the origin. In either case, only one of the two terms will now contain the pole at the origin. Specifically, distorting the contour at the origin to be a small semicircle in the lower half plane, only the e^{ix} term will contribute. The residue for e^{ix}/x will be 1, and with the overall factor of 1/4i, the result for the integral

$$\int_0^\infty dx \frac{\sin x}{x} = \frac{\pi}{2}. \tag{A.24}$$

3.

$$\int_{-\infty}^\infty dx \frac{e^{ax}}{1+e^x}. \tag{A.25}$$

We now take a to have small positive imaginary part, so that we can close the integral with a large semicircle in the upper half plane. Notice that $1+e^x$ will have zeroes for $x = in\pi$ for odd integer n, so we encircle an infinite number of poles on the imaginary axis. The integration is then (with η a positive infinitesimal and a real),

$$\int_{-\infty}^\infty dx \frac{e^{(a+i\eta)x}}{1+e^x} = -2\pi i \sum_{n=1,3,5}^\infty e^{ina\pi} \tag{A.26}$$

This is a geometric series and the sum is

$$\int_{-\infty}^\infty dx \frac{e^{(a+i\eta)x}}{1+e^x} = -2\pi i \frac{e^{ia\pi}}{1-e^{i2a\pi}} = \frac{\pi}{\sin(\pi a)} \tag{A.27}$$

478

4. The integral

$$\int_0^\infty dx \frac{\cos(x)}{1 + x^2} \tag{A.28}$$

is symmetric so we can extend the integration to the whole real axis and divide by two. We could write $2\cos(x) = e^{ix} + e^{-ix}$, and do the two integrals separately. It is slightly easier to write

$$\int_0^\infty dx \frac{\cos(x)}{1 + x^2} = \mathrm{Re} \frac{1}{2} \int_{-\infty}^\infty dx \frac{e^{ix}}{1 + x^2} \tag{A.29}$$

The contour can now be closed in the upper half plane. The pole at $x = i$ is encircled, with residue $e^{-1}/2i$, so that the result is

$$\int_0^\infty dx \frac{\cos(x)}{1 + x^2} = \frac{\pi}{2e} \tag{A.30}$$

5.

$$\int_0^{2\pi} \frac{d\phi}{1 - 2a\cos\phi + a^2} \tag{A.31}$$

with real a. Here we can take $z = e^{i\phi}$, so that $\cos\phi = (z + z^{-1})/2$, and this integral becomes the integral around the unit circle of

$$\oint dz \frac{1}{i(1 - az)(z - a)} . \tag{A.32}$$

For $|a| < 1$, the only pole encircled will be at $z = a$, with the result $2\pi(1 - a^2)^{-1}$. For $|a| > 1$, the only pole encircled is at $z = 1/a$, with residue $1/i(a^2 - 1)$. The result is

$$\int_0^{2\pi} \frac{d\phi}{1 - 2a\cos\phi + a^2} = \begin{cases} \frac{2\pi}{1-a^2} & |a| < 1 \\ \frac{2\pi}{a^2-1} & |a| > 1 \end{cases} . \tag{A.33}$$

6. Often we have integrals like

$$i \int_{-\infty}^\infty \frac{d\omega}{2\pi} \frac{e^{-i\omega t}}{\omega + i\eta} \tag{A.34}$$

where η is a positive infinitesimal, so there is pole with residue $i/2\pi$ just below the real axis. For $t < 0$, we close the contour with a large semicircle in the upper half plane which gives zero contribution, and the function is analytic inside, and the full integral is zero. For $t > 0$, we close with a large semicircle in the lower half plane which gives zero contribution. We now enclose the pole and the result is 1. Our integral is therefore

$$i \int_{-\infty}^\infty \frac{d\omega}{2\pi} \frac{e^{-i\omega t}}{\omega + i\eta} = \Theta(t) = \begin{cases} 0 & t < 0 \\ 1 & t > 0 \end{cases} . \tag{A.35}$$

Linear response functions are causal, so for example the dielectric response must occur after the application of the driving field. If you Fourier transform the frequency response to get the response as a function of time, the requirements of causality that there can be no response before the applied field force the response functions to be analytic in the upper half plane.

A.5 Multiple valued functions and branch cuts

Some functions can give multiple values for the same argument. Examples are $\sqrt{z}$ and $\ln(z)$. If you randomly picked the sign for $\sqrt{z}$, the derivative, Eq. A.1 would not be well defined. To analyze $\sqrt{z}$ and find out where it is an analytic function you need to choose a clear rule for how to select one of the multiple values for each value of z in the complex plane (or at least the part that you are interested in). For the square root, you might choose the following. For a value of z, you write $\rho e^{i\phi} = z$ with ρ and ϕ real, and $-\pi < \phi \leq \pi$. The square root is then taken to be $\sqrt{z} = \sqrt{\rho}e^{i\phi/2}$ where the square root of ρ is taken to be a positive real number. With this rule, you can calculate that $\sqrt{z}$ is analytic everywhere except on the negative real axis. Points z infinitesimally below the negative real axis will have values $-i\sqrt{|z|}$, while those infinitesimally above will have values $i\sqrt{|z|}$. This means that the square root is not analytic anywhere on the negative real axis. We say there is a branch cut there. When we choose contours so that the function is analytic on the interior they cannot cross the branch cut. However, we get to choose how to assign the values to the function, so that we choose the location of the branch cut. Sometimes it is possible to run the branch cut from one singularity to another[4].

A.5.1 Examples with a branch cut

1.

$$\int_0^\infty dx \frac{\sqrt{x}}{(1+x)^2} \, . \tag{A.36}$$

The most convenient place to put the branch cut is on the positive real axis. We then write the original integral as the integral infinitesimally above the real axis, with $z = \rho e^{i\phi}$, $0 \leq \phi < 2\pi$. The square root is then given by $z^{1/2} = \sqrt{\rho}e^{i\phi/2}$. We take a contour running infinitesimally above the real axis, then a big circle all the way around to just below the positive real axis, and then running back to the origin and circling around to connect to the original integral. Notice that since the $\sqrt{z}$ changes sign on the two sides of the branch cut, that the integration around this contour will give twice the value of the original integral. The circular parts of the contour at infinity and around the origin give zero contribution. There is a second order pole at $z = e^{i\pi}$, with residue $e^{-i\pi/2}/2 = -i/2$, so the result is

$$\int_0^\infty dx \frac{\sqrt{x}}{(1+x)^2} = \frac{\pi}{2} \tag{A.37}$$

---

[4]For example, if you have the function $\sqrt{z-1}\sqrt{z+1}$, you can define the square root so that the branch cut for both square roots runs along the real axis to $-\infty$. The two cuts together running along the negative real axis from -1 to $-\infty$ each introduce a sign change, so taken together, there is no discontinuity, and only the cut running between $z = 1$ and $z = -1$ is required. However, you must always check whether this can work. For example the effects of the cuts from $\ln(z)$ and $\sqrt{z}$ do not cancel and running these cuts along the same contour will just give a different branch cut.

2. The integral

$$I = \int_0^\infty dx \frac{\ln^2(x)}{1 + x^2} \tag{A.38}$$

can be done by taking $z = \rho e^{i\phi}$ and with $-\pi/2 < \phi < 3\pi/2$ we have put the branch cut on the negative imaginary axis and $\ln(z) = \ln \rho + i\phi$. Anywhere else in the lower half plane will do. We integrate along the real axis with a large semicircle in the upper half plane. The integral around this contour encloses the pole at $z = i = e^{i\pi/2}$, with residue $i\pi^2/8$.

$$\begin{aligned}
-\frac{\pi^3}{4} &= \oint dz \frac{\ln^2(z)}{1 + z^2} \\
&= \int_\infty^0 d\rho e^{i\pi} \frac{[\ln(\rho) + i\pi]^2}{1 + \rho^2} + \int_0^\infty d\rho \frac{\ln^2(\rho)}{1 + \rho^2} \\
&= 2 \int_0^\infty dx \frac{\ln^2(x)}{1 + x^2} + \int_0^\infty dx \frac{2i\pi \ln(x)}{1 + x^2} - \int_0^\infty dx \frac{\pi^2}{1 + x^2}
\end{aligned} \tag{A.39}$$

The last integral can be done by contour integration exactly like the integral of Eq. A.19 above. The integrand is symmetric and closing in the upper half plane there is a pole with residue $1/2i$. The result is

$$\int_0^\infty dx \frac{1}{1 + x^2} = \frac{\pi}{2}. \tag{A.40}$$

We can integrate the term with a single $\ln(x)$ in the same way that we are using for the $\ln^2(x)$ term. The contour encloses only the pole at $x = i$, with residue $\pi/4$, so that

$$\begin{aligned}
\frac{i\pi^2}{2} &= \oint dz \frac{\ln(z)}{1 + z^2} \\
&= \int_\infty^0 d\rho e^{i\pi} \frac{\ln(\rho) + i\pi}{1 + \rho^2} + \int_0^\infty d\rho \frac{\ln(\rho)}{1 + \rho^2}
\end{aligned} \tag{A.41}$$

or

$$\int_0^\infty dx \frac{\ln(x)}{1 + x^2} = \frac{i\pi^2}{4} - \frac{1}{2}i\pi \int_0^\infty dx \frac{1}{1 + x^2} = 0. \tag{A.42}$$

So that

$$\int_0^\infty dx \frac{\ln^2(x)}{1 + x^2} = \frac{\pi^3}{8} \tag{A.43}$$

Appendix B

Classical driven harmonic oscillator

B.1 Introduction

Much of the complex analysis used in the solution of quantum scattering problems as well as classical electromagnetic radiation and scattering can be developed for the related problem of the driven classical harmonic oscillator.

B.2 Fourier Transforms

The Fourier transform is most easily developed from the limit of a Fourier series. The Fourier series basis functions are sines, cosines, or complex exponentials. They are the solution of the Sturm-Liouville eigenvalue differential equation

$$-\frac{d^2}{dx^2}\psi_n(x) = k_n^2\psi_n(x) \tag{B.1}$$

along with boundary conditions. Particularly for quantum problems, but also for others, the equations simplify if the basis functions are eigenfunctions of the derivative d/dx. That is we would like to use the complex exponentials for the eigenfunctions. Periodic boundary conditions, where $-L/2 \leq x \leq L/2$, with the functions and their derivatives matched at $\pm L/2$ leads to the solutions

$$\psi_n(x) = e^{ik_n x} \tag{B.2}$$

with $k_n = 2\pi n/L$, and n an integer. The different $\psi_n(x)$ are orthogonal. Since these functions are complete we can expand any function in the range $-L/2 < x < L/2$ as

$$f(x) = \sum_{n=-\infty}^{\infty} e^{ik_n x}\tilde{f}_n \tag{B.3}$$

where $\tilde{f}_n$ are the expansion coefficients. Multiplying by $e^{-ik_m x}$ and integrating, we have

$$\int_{-L/2}^{L/2} dx e^{-ik_m x} f(x) = \int_{-L/2}^{L/2} dx e^{-ik_m x} \sum_{n=-\infty}^{\infty} e^{ik_n x}\tilde{f}_n = L\tilde{f}_m \equiv C\tilde{f}(k_n), \tag{B.4}$$

with C an arbitrary constant.

We now take the limit that $L \to \infty$. Since physical functions are not periodic with a period $L \to \infty$, this only makes sense for functions that go to a constant at large $|x|$. Typically, the Fourier transform is only useful if the functions are localized and go to zero for large $|x|$. For $L \to \infty$, the spacing between the k_n goes to zero, and the sum over n can be replaced by an integral over n,

$$
\begin{aligned}
f(x) &= \sum_{n=-\infty}^{\infty} e^{ik_n x} \tilde{f}_n = \int_{-\infty}^{\infty} dn\, e^{ik_n x} \frac{C}{L} \tilde{f}(k_n) \\
&= C \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ikx} \tilde{f}(k)\,.
\end{aligned}
\tag{B.5}
$$

The transform pair is then

$$
\begin{aligned}
\tilde{f}(k) &= \frac{1}{C} \int_{-\infty}^{\infty} dx\, e^{-ikx} f(x) \\
f(x) &= C \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ikx} \tilde{f}(k)\,.
\end{aligned}
\tag{B.6}
$$

Since mathematically, the symbols x and k are on equal footing, it is your choice whether to define your forward transform with e^{-ikx} and the reverse transform with e^{ikx} or vice versa. For example since a forward propagating 1-dimensional plane wave is usually written in the physics literature as $e^{ikx-i\omega t}$, opposite conventions are typically used for space and time so that a plane wave propagating along positive x has both positive k and ω. Engineering literature generally assumes a $e^{j\omega t}$ time dependence, but since the symbol j is used conversion to the usual physics convention simply requires taking $j = -i$.

The most commonly used conventions for the constant for the transform pair are $C = 1$, which puts the $(2\pi)^{-1}$ factor on the k integral, $C = \sqrt{2\pi}$, which puts $(2\pi)^{-1/2}$ symmetrically on both integrals, and $C = 2\pi$ which puts the $(2\pi)^{-1}$ factor on the x integral. I nearly always put the 2π factors on the k integral, for a k - x pair, or the ω integral for an ω - t pair. This agrees with the normalizations for the position and momentum eigenstates of $\langle x|x'\rangle = \delta(x-x')$ with $\langle x|p\rangle = e^{\frac{i}{\hbar}px}$, and $\langle p|p'\rangle = 2\pi\hbar\delta(p-p')$.

B.3 Solution of the driven harmonic oscillator by Fourier transforms

Newton's equation for the position $x(t)$ of a driven undamped harmonic oscillator of mass m and spring constant $m\omega_0^2$ is

$$
m\frac{d^2 x(t)}{dt^2} = -m\omega_0^2 x(t) + F(t)
\tag{B.7}
$$

where $F(t)$ is the driving force at time t.

For the Fourier transform to be well defined, we require that the function to be transformed go to zero at $t \to \pm\infty$. We therefore imagine the force is applied for a finite amount of time. However, we know that if we excite the harmonic oscillator and thereafter have the driving force zero, the undamped oscillator will continue to oscillate forever, and the Fourier transform of $x(t)$ will not be well defined. The standard method of handling this is to modify the equation of motion, usually to include a small amount of damping, so that $x(t)$ goes to zero. Here we will examine the analytic properties of the Fourier transformed solution to obtain the required result.

Fourier transforming the equation of motion and integrating the acceleration term by parts twice gives

$$
\begin{aligned}
0 &= \int_{-\infty}^{\infty} dt e^{i\omega t} \left[m\frac{d^2 x(t)}{dt^2} + m\omega_0^2 x(t) - F(t) \right] \\
&= \int_{-\infty}^{\infty} dt e^{i\omega t} m \left\{ x(t)[\omega_0^2 - \omega^2] - F(t) \right\} \\
&= \tilde{x}(\omega) m[\omega_0^2 - \omega^2] - \tilde{F}(\omega)
\end{aligned}
\tag{B.8}
$$

where we assume (incorrectly) that $x(t)$ and $dx(t)/dt$ are zero at $t \to \pm\infty$ so that the surface terms can be dropped. Solving this algebraic equation gives

$$
\tilde{x}(\omega) = \frac{\tilde{F}(\omega)}{m(\omega_0^2 - \omega^2)} .
\tag{B.9}
$$

Notice this diverges at $\omega = \pm\omega_0$ reflecting the problem with defining the Fourier transform. To see how to modify the equations around $\omega = \pm\omega_0$, we Fourier transform back to $x(t)$ to give

$$
x(t) \stackrel{?}{=} \frac{1}{m} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} \frac{\tilde{F}(\omega)}{\omega_0^2 - \omega^2} .
\tag{B.10}
$$

where the question mark indicates we have yet to deal with the divergent integrand. Notice that integration around the poles at $\omega = \pm\omega_0$ will give contributions proportional to $e^{\pm i\omega_0 t}$ which are the homogeneous solutions to the differential equations. Modifying how we handle the poles corresponds to selecting the amounts of each of these solutions to add, and sets the boundary conditions.

If the oscillator is at rest at the origin before the driving force is applied, the boundary condition is $x(t \to -\infty) = 0$. We look at our integral solution at large negative t and note that we can add to the integral along the real ω axis, an integral around the large semicircle in the upper half plane, since the factor $e^{-i\omega t}$ will make that integral zero. To enforce our boundary condition the contour integration around this semicircle and the real ω axis must go to zero for all large negative times. Notice that any poles in the upper half plane with a finite imaginary part will give residues that are damped exponentially at large negative t. However, poles infinitesimally close to the real axis will not be damped. To satisfy the boundary condition we move the singularities at $\omega = \pm\omega_0$ slightly into the lower half plane

so their residues cannot contribute at large negative times. We get the solution

$$x(t) = \frac{1}{m} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} \frac{\tilde{F}(\omega)}{\omega_0^2 - (\omega + i\eta)^2} , \qquad (B.11)$$

where η is a positive infinitesimal. The Fourier transform is

$$\tilde{x}(\omega) = \frac{1}{m} \frac{\tilde{F}(\omega)}{\omega_0^2 - (\omega + i\eta)^2} . \qquad (B.12)$$

An alternative way to get to this equation is to change the equation of motion so that the function $x(t)$ can be Fourier transformed without difficulty. If we had instead solved the problem of a harmonic oscillator initially at rest at the origin with a frictional damping force $-m\gamma v(t)$, the boundary condition enforces $x(t \to -\infty) = 0$, and the damping force ensures that $x(t \to \infty) = 0$. Therefore this $x(t)$ can be transformed immediately. Setting $\gamma = \eta$, a positive infinitesimal, at the end of the calculation, gives Eq. B.11.

B.4   Energy transfered to the harmonic oscillator

Let's look at the solution at $t \to \infty$. From the argument above, only the poles infinitesimally close to the real axis will survive. We close in the lower half plane, and the residues give

$$x(t \to \infty) = \frac{i}{2m\omega_0} \left[\tilde{F}(\omega_0)e^{-i\omega_0 t} - \tilde{F}(-\omega_0)e^{i\omega_0 t} \right] . \qquad (B.13)$$

Since $F(t)$ is real, $F(-\omega) = F^*(\omega)$, and we have

$$x(t \to \infty) = -\frac{1}{m\omega_0}\text{Im}\left[\tilde{F}(\omega_0)e^{-i\omega_0 t} \right] = \frac{1}{m\omega_0} \left[\text{Re}F(\omega_0)\sin(\omega_0 t) - \text{Im}F(\omega_0)\cos(\omega_0 t) \right] .$$
$$(B.14)$$

The energy of the harmonic oscillator evaluated from $E = \frac{1}{2}mv(t)^2 + \frac{1}{2}m\omega_0^2 x(t)^2$ at long times is then

$$E = \frac{1}{2m}|\tilde{F}(\omega_0)|^2 . \qquad (B.15)$$

We see that only the ω_0 component of $\tilde{F}(\omega)$ contributes to the energy of the oscillator after the force is zero.

Looking directly at the residual energy in the oscillator is one method of finding the energy transfered to the oscillator. An alternative is to calculate work done by the force which is

$$\begin{aligned} W &= \int_{-\infty}^{\infty} dt v(t)F(t) \\ &= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \tilde{v}(\omega)e^{-i\omega t} \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} \tilde{F}^*(\omega')e^{i\omega' t} \\ &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \tilde{v}(\omega)\tilde{F}^*(\omega) = \frac{1}{\pi}\text{Re} \int_0^{\infty} d\omega \tilde{v}(\omega)\tilde{F}^*(\omega) \end{aligned} \qquad (B.16)$$

486

where both $v(t)$ and $F(t)$ are real so we have used $v(-\omega) = v^*(\omega)$ and $F^*(-\omega) = F(\omega)$ to reduce this to the real part of the integral over positive frequencies. The Fourier transform of the velocity is

$$\tilde{v}(\omega) = -i\omega\tilde{x}(\omega) = \frac{-i\omega}{m}\frac{\tilde{F}(\omega)}{\omega_0^2 - (\omega + i\eta)^2}\,. \tag{B.17}$$

Plugging this in, we have

$$W = \operatorname{Im} \int_0^\infty d\omega\, \frac{\omega}{m\pi}\frac{|\tilde{F}(\omega)|^2}{\omega_0^2 - (\omega + i\eta)^2}\,. \tag{B.18}$$

We can make the formal replacement under an integral

$$\frac{1}{x \pm i\eta} = \mathrm{P}\frac{1}{x} \mp i\pi\delta(x)\,, \tag{B.19}$$

where P indicates the principal parts integration.[1]

Our result is then

$$W = \frac{1}{2m}|\tilde{F}(\omega_0)|^2\,. \tag{B.22}$$

This agrees, as it must from energy conservation, with our previous result.

Be sure to be able to calculate quantities with these too methods since, for other problems, one of them may be easier, or they may complement each other in helping your understanding.

---

[1]To show this we can require that the function above is multiplied by a function analytic in the region around $x = 0$. Since the integrand is analytic, we can then distort the contour into a small semicircle at $x = 0$, where the semicircle is in the lower half plane for the $-i\eta$ term, and the semicircle in the upper half plane for the $i\eta$ term, since these distortions will not cause the contour to cross the pole at $\pm i\eta$. Now take the limit that η goes to zero, and then the limit that the semicircle radius shrinks to zero. The portion of the contour along the real axis gives the principal parts integral, while the semicircle gives $\mp i\pi$ times the residue, and the residue is picked up by the delta function. Alternatively, without using contour integration, divide the integral into a principal parts integration and the integration on the segment between the principal parts limits. The integral becomes

$$\int_{-\epsilon}^{\epsilon} dx f(x)\frac{1}{x \pm i\eta} = \int_{-\epsilon}^{\epsilon} dx f(x)\frac{x \mp i\eta}{x^2 + \eta^2}\,. \tag{B.20}$$

For small enough ϵ, a well behaved $f(x)$, can be replaced by $f(0)$, the x term in the numerator gives zero. Changing variables to $u = x/\eta$ and taking the limit $\eta \to 0$ this becomes

$$\int_{-\epsilon}^{\epsilon} dx f(x)\frac{1}{x \pm i\eta} = f(0)\int_{-\infty}^{\infty} du\frac{\mp i}{u^2 + 1} = \mp i\pi f(0)\,, \tag{B.21}$$

and the formal replacement is again shown.

B.5 Green's function Equation

We can evaluate Eq. B.11 as

$$
\begin{aligned}
x(t) &= \frac{1}{m} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} \frac{\tilde{F}(\omega)}{\omega_0^2 - (\omega + i\eta)^2} \\
&= \frac{1}{m} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} \frac{1}{\omega_0^2 - (\omega + i\eta)^2} \int_{-\infty}^{\infty} dt' e^{i\omega t'} F(t') \\
&= \int_{-\infty}^{\infty} dt' F(t') \underbrace{\frac{1}{m} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{e^{-i\omega(t-t')}}{\omega_0^2 - (\omega + i\eta)^2}}_{\equiv G(t-t')} .
\end{aligned}
\tag{B.23}
$$

For $t < t'$, we can close the contour for the Green's function integral with a large semicircle in the upper half plane. The poles are in the lower half plane, so the integral gives zero. For $t > t'$, we close with a large semicircle in the lower half plane. The poles are at $\pm\omega - i\eta$, and the residues after taking $\eta \to 0$ are $e^{-i(\pm\omega_0)(t-t')}/(\mp 4\pi m\omega_0)$. The integral is $-2\pi i$ times the sum of the residues giving

$$
G(t - t') = \begin{cases} 0 & t < t' \\ \frac{1}{m\omega_0} \sin[\omega_0(t - t')] & t > t' \end{cases} .
\tag{B.24}
$$

Notice that we could have written this retarded Green's function almost immediately since it is the solution for $x(t)$ initially at rest with a unit impulsive force acting at $t = t'$. Before the impulse the solution is 0. After the impulse, the solution must be a free oscillator solution and be continuous at $t = t'$, i.e. $A \sin[\omega_0(t - t')]$. The unit impulse gives a unit momentum just after it is applied. Taking the derivative at $t = t'^+$, gives a velocity $A\omega_0$, so $mA\omega_0 = 1$, and $A = 1/m\omega_0$.

B.6 Application to Classical Radiation

The power of these methods can be shown by looking at the problem of classical radiation from a charge and current distribution. We imagine that the charge and current densities are localized in space and time, so they can be Fourier transformed. Maxwell's equations are usually written as

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot \boldsymbol{E}(\boldsymbol{r}, t) &= 4\pi\rho(\boldsymbol{r}, t) \\
\boldsymbol{\nabla} \cdot \boldsymbol{B}(\boldsymbol{r}, t) &= 0 \\
\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}, t) &= -\frac{1}{c} \frac{\partial \boldsymbol{B}(\boldsymbol{r}, t)}{\partial t} \\
\boldsymbol{\nabla} \times \boldsymbol{B}(\boldsymbol{r}, t) &= \frac{4\pi}{c} \boldsymbol{J}(\boldsymbol{r}, t) + \frac{1}{c} \frac{\partial \boldsymbol{E}(\boldsymbol{r}, t)}{\partial t} .
\end{aligned}
\tag{B.25}
$$

Since the divergence of $\boldsymbol{B}$ is zero, it can be written as the curl of a vector potential,

$$
\boldsymbol{B}(\boldsymbol{r}, t) = \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r}, t) .
\tag{B.26}
$$

Plugging into the curl equation for $\boldsymbol{E}$, we have

$$\boldsymbol{\nabla} \times \left[\boldsymbol{E}(\boldsymbol{r},t) + \frac{1}{c}\frac{\partial}{\partial t}\boldsymbol{A}(\boldsymbol{r},t)\right] = 0 \tag{B.27}$$

A function whose curl is zero can be written as the gradient of a scalar potential, so that

$$\boldsymbol{E}(\boldsymbol{r},t) = -\boldsymbol{\nabla}\Phi(\boldsymbol{r},t) - \frac{1}{c}\frac{\partial}{\partial t}\boldsymbol{A}(\boldsymbol{r},t). \tag{B.28}$$

In Lorentz gauge, Maxwell's equations are

$$\begin{aligned}
\boldsymbol{\nabla}\cdot\boldsymbol{A}(\boldsymbol{r},t) + \frac{1}{c}\frac{\partial}{\partial t}\Phi(\boldsymbol{r},t) &= 0 \\
\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\boldsymbol{A}(\boldsymbol{r},t) &= -\frac{4\pi}{c}\boldsymbol{J}(\boldsymbol{r},t) \\
\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\Phi(\boldsymbol{r},t) &= -4\pi\rho(\boldsymbol{r},t).
\end{aligned} \tag{B.29}$$

Fourier transforming by multiplying by $e^{-i\boldsymbol{k}\cdot\boldsymbol{r}+i\omega t}$ and integrating over space and time gives the result

$$\begin{aligned}
(\omega^2 - k^2 c^2)\tilde{\boldsymbol{A}}(\boldsymbol{k},\omega) &= -4\pi c\tilde{\boldsymbol{J}}(\boldsymbol{k},\omega) \\
\omega\tilde{\Phi}(\boldsymbol{k},\omega) &= \boldsymbol{k}\cdot\tilde{\boldsymbol{A}}(\boldsymbol{k},\omega) \\
\tilde{\boldsymbol{E}}(\boldsymbol{k},\omega) &= -i\boldsymbol{k}\tilde{\Phi}(\boldsymbol{k},\omega) + \frac{i\omega}{c}\tilde{\boldsymbol{A}}(\boldsymbol{k},\omega) \\
\tilde{\boldsymbol{B}}(\boldsymbol{k},\omega) &= i\boldsymbol{k}\times\tilde{\boldsymbol{A}}(\boldsymbol{k},\omega).
\end{aligned} \tag{B.30}$$

The first of these equations is identical in form to the driven harmonic oscillator. Therefore, to match the boundary condition that the fields are zero at $t \to -\infty$, we handle the poles exactly as for the harmonic oscillator, with solution

$$\tilde{\boldsymbol{A}}(\boldsymbol{k},\omega) = \frac{4\pi c}{k^2 c^2 - (\omega + i\eta)^2}\tilde{\boldsymbol{J}}(\boldsymbol{k},\omega) \tag{B.31}$$

so that

$$\begin{aligned}
\tilde{\boldsymbol{B}}(\boldsymbol{k},\omega) &= -\frac{i4\pi c}{k^2 c^2 - (\omega + i\eta)^2}\boldsymbol{k}\times\tilde{\boldsymbol{J}}(\boldsymbol{k},\omega) \\
\tilde{\boldsymbol{E}}(\boldsymbol{k},\omega) &= \frac{4\pi i}{\omega}\frac{\omega^2\tilde{\boldsymbol{J}}(\boldsymbol{k},\omega) - c^2\boldsymbol{k}\boldsymbol{k}\cdot\tilde{\boldsymbol{J}}(\boldsymbol{k},\omega)}{k^2 c^2 - (\omega + i\eta)^2}.
\end{aligned} \tag{B.32}$$

The work done by the charges on the electromagnetic field must be the negative of the work done on the charges by the field or

$$I = -\int_{-\infty}^{\infty} dt \sum_i \boldsymbol{v}_i(t)\cdot\boldsymbol{F}_i(t). \tag{B.33}$$

489

Plugging in the result of the Lorentz force, $F_i(t) = q_i \left[\boldsymbol{E}(\boldsymbol{r}_i(t), t) + \frac{\boldsymbol{v}_i(t)}{c} \times \boldsymbol{B}(\boldsymbol{r}_i(t), t) \right]$, this becomes

$$I = - \int_{-\infty}^{\infty} dt \sum_i q_i \boldsymbol{v}_i(t) \cdot \boldsymbol{E}(\boldsymbol{r}_i(t), t) \,. \tag{B.34}$$

Using the current density for a set of charges $\boldsymbol{J}(\boldsymbol{r}, t) = \sum_i q_i \boldsymbol{v}_i(t) \delta[\boldsymbol{r} - \boldsymbol{r}_i(t)]$, this becomes just as in the energy expression for the harmonic oscillator

$$
\begin{aligned}
I &= - \int d^3r \int dt \boldsymbol{J}(\boldsymbol{r}, t) \cdot \boldsymbol{E}(\boldsymbol{r}, t) = -\mathrm{Re} \int_0^{\infty} \frac{d\omega}{\pi} \int \frac{d^3k}{(2\pi)^3} \tilde{\boldsymbol{J}}^*(\boldsymbol{k}, \omega) \cdot \tilde{\boldsymbol{E}}(\boldsymbol{k}, \omega) \\
&= -\mathrm{Re} \int_0^{\infty} \frac{d\omega}{\pi} \int \frac{d^3k}{(2\pi)^3} \frac{4\pi i}{\omega} \frac{\omega^2 \left| \tilde{\boldsymbol{J}}(\boldsymbol{k}, \omega) \right|^2 - c^2 \left| \boldsymbol{k} \cdot \tilde{\boldsymbol{J}}(\boldsymbol{k}, \omega) \right|^2}{k^2 c^2 - (\omega + i\eta)^2} \,.
\end{aligned} \tag{B.35}
$$

As in the harmonic oscillator case, only the delta function from the $i\eta$ gives a real result. The delta function makes the length $kc = \omega$, and only the transverse part of the current contributes. We get

$$
\begin{aligned}
I &= \int_0^{\infty} \frac{d\omega}{\pi} \int \frac{d^3k}{(2\pi)^3} \frac{4\pi}{\omega} \left[\omega^2 \left| \tilde{\boldsymbol{J}}(\boldsymbol{k}, \omega) \right|^2 - c^2 \left| \boldsymbol{k} \cdot \tilde{\boldsymbol{J}}(\boldsymbol{k}, \omega) \right|^2 \right] \pi \delta(k^2 c^2 - \omega^2) \\
&= \int_0^{\infty} d\omega \int d\Omega \frac{\omega^2}{4\pi^2 c^3} \left| \hat{\boldsymbol{k}} \times \hat{\boldsymbol{k}} \times \tilde{\boldsymbol{J}} \left(\hat{\boldsymbol{k}} \frac{\omega}{c}, \omega \right) \right|^2
\end{aligned} \tag{B.36}
$$

where $d\Omega = d\cos\theta d\phi$, and $\hat{\boldsymbol{k}} = \hat{\boldsymbol{x}} \sin\theta \cos\phi + \hat{\boldsymbol{y}} \sin\theta \sin\phi + \hat{\boldsymbol{z}} \cos\theta$. The intensity distribution in frequency and angle is

$$\frac{d^2 I}{d\Omega d\omega} = \frac{\omega^2}{4\pi^2 c^3} \left| \hat{\boldsymbol{k}} \times \hat{\boldsymbol{k}} \times \tilde{\boldsymbol{J}} \left(\hat{\boldsymbol{k}} \frac{\omega}{c}, \omega \right) \right|^2 \,. \tag{B.37}$$

Notice this is equal to the result in Jackson's book, Eq. 14.70.

B.7   Green's function for the wave equation

The Green's function for the wave equation can be calculated as for the harmonic oscillator,

$$A(\boldsymbol{r}, t) = \int d^3r dt G(\boldsymbol{r}, t, \boldsymbol{r}', t') \boldsymbol{J}(\boldsymbol{r}', t') \tag{B.38}$$

with the simplifying definitions $\boldsymbol{R} = \boldsymbol{r} - \boldsymbol{r}'$ and $T = t - t'$,

$$G(\boldsymbol{r}, t, \boldsymbol{r}', t') = \int \frac{d^3k}{(2\pi)^3} \frac{d\omega}{2\pi} e^{i\boldsymbol{k} \cdot \boldsymbol{R}} e^{-i\omega T} \frac{4\pi c}{k^2 c^2 - (\omega + i\eta)^2} \tag{B.39}$$

The ω integral is identical to the one done for the harmonic oscillator with the result

$$
\begin{aligned}
G(\boldsymbol{r}, t, \boldsymbol{r}', t') &= \int \frac{d^3 k}{(2\pi)^3} e^{i\boldsymbol{k}\cdot\boldsymbol{R}} \frac{4\pi}{k} \sin(kcT)\Theta(T) \\
&= \frac{2}{\pi R}\Theta(T) \int_0^\infty dk \, \sin(kR)\sin(kcT) \\
&= \frac{1}{\pi R}\Theta(T) \int_{-\infty}^\infty dk \, \sin(kR)\sin(kcT) \\
&= \frac{1}{2\pi R}\Theta(T) \int_{-\infty}^\infty dk \, \{\cos[k(R-cT)] - \cos[k(R+cT)]\} \\
&= \frac{1}{R}\Theta(T) \left[\delta(R-cT) - \delta(R+cT)\right] .
\end{aligned}
\tag{B.40}
$$

Since $R > 0$ and $\Theta(T)$ is zero for $T < 0$, the second delta function never contributes. Dropping the second delta function, the first delta function will be nonzero only for $T > 0$, so the Theta function can be dropped as well. The result is

$$
G(\boldsymbol{r}, t, \boldsymbol{r}', t') = \frac{1}{R}\delta(R-cT)
\tag{B.41}
$$

which is retarded Coulomb's law.

Appendix C

Plane wave expansion

C.1 Method 1

We want a Bessel function expansion of the plane wave $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$. The plane wave is a solution of Helmholtz equation

$$\left[\nabla^2 + k^2\right] e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = 0 \,. \tag{C.1}$$

If we separate variables, we can derive the general solution of the Helmholtz equation as a linear combination of spherical Bessel functions $j_\ell(kr)$, and $n_\ell(kr)$ multiplied by spherical harmonics. That is, a general solution is

$$\Phi(\boldsymbol{r}) = \sum_{\ell m} \left[A_{\ell m} j_\ell(kr) + B_{\ell m} n_\ell(kr)\right] Y_{\ell m}(\theta, \phi) \,. \tag{C.2}$$

Taking the z axis along $\boldsymbol{k}$, the plane wave does not depend on ϕ, so that we can write

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = e^{ikr\cos\theta} = \sum_{\ell m} \left[A_{\ell m} j_\ell(kr) + B_{\ell m} n_\ell(kr)\right] Y_{\ell m}(\theta, \phi) \,. \tag{C.3}$$

The plane wave is well behaved at the origin, so the $B_{\ell m}$ terms must all be zero. Integrating over ϕ, only the $Y_{\ell 0}$ terms survive, and since these are proportional to $P_\ell(\cos\theta)$ we can write

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \sum_{\ell m} C_\ell j_\ell(kr) P_\ell(\cos\theta) \,. \tag{C.4}$$

Therefore, using the orthogonality of the P_ℓ, and writing $\cos\theta = \mu$, we have

$$\int_{-1}^{1} e^{ikr\mu} P_\ell(\mu) d\mu = \frac{2}{2\ell + 1} C_\ell j_\ell(kr) \,. \tag{C.5}$$

Since this is true for all kr, we can expand in powers of kr and equate the lowest surviving power. Using

$$j_\ell(kr) = \frac{(kr)^\ell}{(2\ell + 1)!!} + \dots \tag{C.6}$$

we have

$$\frac{i^\ell}{\ell!} \int_{-1}^{1} d\mu \, \mu^\ell P_\ell(\mu) = \frac{2}{(2\ell+1)(2\ell+1)!!} C_\ell \,. \tag{C.7}$$

The Rodrigues formula for the Legendre Polynomials is

$$P_\ell(\mu) = \frac{1}{2^\ell \ell!} \frac{d^\ell}{d\mu^\ell} (\mu^2 - 1)^\ell \,. \tag{C.8}$$

This shows that the μ^ℓ term in $P_\ell(\mu)$ comes from taking all ℓ derivatives of the $\mu^{2\ell}$ term, so that

$$
\begin{aligned}
\mu^\ell &= \frac{2^\ell \ell!^2}{(2\ell)!} P_\ell(\mu) + \text{Lower order } P_n(\mu) \text{ terms} \\
&= \frac{\ell!}{(2\ell-1)!!} P_\ell(\mu) + \text{Lower order } P_n(\mu) \text{ terms} \,.
\end{aligned}
\tag{C.9}
$$

The lower order terms are orthogonal to $P_\ell(\mu)$, so the integral is

$$\frac{i^\ell}{\ell!} \int_{-1}^{1} d\mu \, P_\ell(\mu) \mu^\ell = \frac{2}{2\ell+1} \frac{i^\ell}{(2\ell-1)!!} = \frac{2}{(2\ell+1)(2\ell+1)!!} C_\ell \tag{C.10}$$

and

$$C_\ell = i^\ell (2\ell+1) \,. \tag{C.11}$$

The plane wave is

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \sum_\ell i^\ell (2\ell+1) j_\ell(kr) P_\ell(\cos\theta) \tag{C.12}$$

where $\cos\theta$ is $\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{k}}$. The spherical harmonic addition theorem can be used to write this as

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = 4\pi \sum_\ell i^\ell j_\ell(kr) Y_{\ell m}(\theta_r, \phi_r) Y_{\ell m}^*(\theta_k, \phi_k) \,. \tag{C.13}$$

C.2   Method 2

We can solve for the retarded Green's function for the Helmholtz equation

$$\left[\nabla^2 + k^2\right] G(\boldsymbol{r}, \boldsymbol{r}') = \delta^3(\boldsymbol{r} - \boldsymbol{r}') \tag{C.14}$$

in two ways. Taking the origin to be at $\boldsymbol{r}'$, we have the equation

$$\left[\nabla^2 + k^2\right] G(\boldsymbol{r} - \boldsymbol{r}', 0) = \delta^3(\boldsymbol{r} - \boldsymbol{r}') \tag{C.15}$$

which shows that $G(\boldsymbol{r}, \boldsymbol{r}') = G(\boldsymbol{r} - \boldsymbol{r}', 0)$. Solving with $\boldsymbol{r}' = 0$ we have

$$\left[\nabla^2 + k^2\right] G(\boldsymbol{r}, 0) = \delta^3(\boldsymbol{r}) \tag{C.16}$$

and since neither the equation nor the boundary conditions depend on direction, we have $G(\boldsymbol{r}, 0) = g(r)$, with

$$\frac{1}{r}\frac{d^2}{dr^2}rg(r) = \delta^3(\boldsymbol{r}). \tag{C.17}$$

Away from $r = 0$, the solution is

$$rg(r) = Ae^{ikr} + Be^{-ikr} \tag{C.18}$$

with the boundary condition that the retarded Green's function has only outgoing waves, we have $rg(r) = Ae^{ikr}$, so that

$$G(\boldsymbol{r}, 0) = A\frac{e^{ikr}}{r} \tag{C.19}$$

Integrating over a small spherical volume V of radius R around the origin, and taking the limit the volume goes to zero, the delta function gives 1, and using the divergence theorem,

$$
\begin{aligned}
1 &= \lim_{R \to 0} \int_V d^3r \left[\nabla^2 + k^2 \right] A\frac{e^{ikr}}{r} \\
&= \lim_{R \to 0} \int_V d^3r \, \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} A\frac{e^{ikr}}{r} \\
&= \lim_{R \to 0} 4\pi A R^2 \frac{d}{dR}\frac{e^{ikR}}{R} = -4\pi A
\end{aligned}
\tag{C.20}
$$

and the Green's function is

$$G(\boldsymbol{r}, \boldsymbol{r}') = -\frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{4\pi|\boldsymbol{r} - \boldsymbol{r}'|}. \tag{C.21}$$

We now repeat the solution using spherical coordinates. Using the general solution and enforcing the boundary conditions that the function is well behaved at the origin, i.e. only j_ℓ terms there, and contains only outgoing waves at large r, i.e. only $h_\ell^{(1)} = j_\ell + in_\ell$ terms there, we can write

$$G(\boldsymbol{r}, \boldsymbol{r}') = \sum_{\ell m} Y_{\ell m}(\theta, \phi) \begin{cases} A_{\ell m} j_\ell(kr) & r < r' \\ B_{\ell m} h_\ell^{(1)}(kr) & r > r' \end{cases}. \tag{C.22}$$

Substituting this into the Green's function equation, we have

$$\sum_{\ell m} \left[\frac{1}{r}\frac{d^2}{dr^2}r - \frac{\ell(\ell+1)}{r^2} + k^2 \right] Y_{\ell m}(\theta, \phi) \begin{cases} A_{\ell m} j_\ell(kr) & r < r' \\ B_{\ell m} h_\ell^{(1)}(kr) & r > r' \end{cases} = \frac{\delta(r-r')\delta(\phi-\phi')\delta(\cos\theta - \cos\theta')}{r'^2} \tag{C.23}$$

Multiplying both sides by $Y_{\ell' m'}^*(\theta, \phi)$, integrating over the solid angle, and then dropping the primes on ℓ and m, we get

$$\left[\frac{1}{r}\frac{d^2}{dr^2}r - \frac{\ell(\ell+1)}{r^2} + k^2 \right] \begin{cases} A_{\ell m} j_\ell(kr) & r < r' \\ B_{\ell m} h_\ell^{(1)}(kr) & r > r' \end{cases} = Y_{\ell m}^*(\theta', \phi')\frac{\delta(r-r')}{r'^2}. \tag{C.24}$$

Since the second derivative gives a delta function, the first derivative must be discontinuous, and the function continous at $r = r'$. Continuity requires

$$A_{\ell m} j_\ell(kr') = B_{\ell m} h_\ell^{(1)}(kr') \tag{C.25}$$

and integrating over r from $r' - \epsilon$ to $r' + \epsilon$ and taking the limit that $\epsilon \to 0$ gives

$$B_{\ell m} \frac{d'}{dr'} h_\ell^{(1)}(kr') - A_{\ell m} \frac{d'}{dr'} j_\ell(kr') = r'^{-2} Y_{\ell m}^*(\theta', \phi') \tag{C.26}$$

with solution

$$A_{\ell m} = \frac{r'^{-2} h_\ell(kr')}{D_\ell} Y_{\ell m}^*(\theta', \phi')$$

$$B_{\ell m} = \frac{r'^{-2} j_\ell(kr')}{D_\ell} Y_{\ell m}^*(\theta', \phi') \tag{C.27}$$

with the Wronskian

$$D_\ell = j_\ell(kr') \frac{d'}{dr'} h_\ell^{(1)}(kr') - h_\ell^{(1)}(kr') \frac{d'}{dr'} j_\ell(kr') . \tag{C.28}$$

The Wronskian is evaluated by writing

$$0 = \int_a^b dr\, r j_\ell(kr) \left[\frac{d^2}{dr^2} - \frac{\ell(\ell+1)}{r^2} + k^2 \right] r h_\ell^{(1)}(kr) \tag{C.29}$$

Integrating by parts twice gives

$$0 = \left[r^2 h_\ell^{(1)}(kr) \frac{d}{dr} j_\ell(kr) - r^2 j_\ell(kr) \frac{d}{dr} h_\ell^{(1)}(kr) \right] \Big|_a^b + \int_a^b dr\, r h_\ell^{(1)}(kr) \left[\frac{d^2}{dr^2} - \frac{\ell(\ell+1)}{r^2} + k^2 \right] r j_\ell(kr) \tag{C.30}$$

and since the integral is zero from the differential equation, the surface term must cancel for any a and b. Therefore it is independent of the argument so that

$$j_\ell(kr) \frac{d}{dr} h_\ell^{(1)}(kr) - h_\ell^{(1)}(kr) \frac{d}{dr} j_\ell(kr) = \frac{C_\ell}{r^2} \tag{C.31}$$

where C_ℓ is a constant.

Since it is independent of argument, we can evaluate the constant in the Wronskian at any r. One choice is to use $r \to \infty$ where

$$h_\ell(kr) \to (-i)^{\ell+1} \frac{e^{ikr}}{kr}$$

$$\frac{d}{dr} h_\ell(kr) \to (-1)^{\ell+1} ik \frac{e^{ikr}}{kr}$$

$$j_\ell(kr) \to \frac{\sin(kr - \ell\pi/2)}{kr}$$

$$\frac{d}{dr} j_\ell(kr) \to \frac{k \cos(kr - \ell\pi/2)}{kr} \tag{C.32}$$

and
$$D_\ell = \frac{i}{kr^2} \tag{C.33}$$

so that the Green's function is

$$G(r, \theta, \phi, r'\theta', \phi') = ik \sum_{\ell m} j_\ell(kr_<)h_\ell^{(1)}(kr_>)Y_{\ell m}(\theta, \phi)Y_{\ell m}^*(\theta', \phi') , \tag{C.34}$$

where $r_<$ is the smaller of r and r' and $r_>$ is the larger.

Equating the two forms of the Green's function and taking the limit that $r' \to \infty$, we have

$$-\frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{4\pi|\boldsymbol{r}-\boldsymbol{r}'|} \quad \to \quad -\frac{e^{ikr'}}{4\pi r'}e^{-ik\hat{\boldsymbol{r}}'\cdot\boldsymbol{r}}$$

$$= ik \sum_{\ell m} j_\ell(kr)(-i)^{\ell+1}\frac{e^{ikr'}}{kr'}Y_{\ell m}(\theta, \phi)Y_{\ell m}^*(\theta', \phi') \tag{C.35}$$

or

$$e^{-ik\hat{\boldsymbol{r}}'\cdot\boldsymbol{r}} = 4\pi \sum_{\ell m}(-i)^\ell j_\ell(kr)Y_{\ell m}(\theta, \phi)Y_{\ell m}^*(\theta', \phi') \tag{C.36}$$

Taking the complex conjugate of both sides and defining $\boldsymbol{k} = k\hat{\boldsymbol{r}}'$, gives

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = 4\pi \sum_{\ell} i^\ell j_\ell(kr)Y_{\ell m}(\theta_r, \phi_r)Y_{\ell m}^*(\theta_k, \phi_k) . \tag{C.37}$$

Appendix D

Electromagnetic analogies

D.1 Introduction

Since in linear problems classical electromagnetic fields satisfy wave equations, the mathematics of scattering theory has many simularities. Julian Schwinger worked at the MIT radiation laboratory on radar problems and the theory of wave guides and the effect of obstacles in wave guides during world war two. His work helped integrate some of the same concepts of using Green's functions and the scattering matrix to describe both quantum scattering and scattering in these classical systems.

We begin by looking at Maxwell's equations for an ideal cylindrical wave guide made up of perfect conductors and filled with a material with isotropic material with dielectric constant ϵ and permeability μ. Since the wave guide does not change with time, Maxwell's equations are invariant under time translation, the time translation eigenstates correspond to $e^{-i\omega t}$ time dependence. Alternatively, you can Fourier transform the equations in time, and either case solve for a single frequency at a time. We will not deal explicitly with the static case where $\omega = 0$, since, for any physical system, we can always imagine that the fields were turned on slowly from zero in the distant past and turned off again in the distant future. In this case, Maxwell's divergence equations are contained in the curl equations

$$
\begin{aligned}
\boldsymbol{\nabla} \times \boldsymbol{E}_\omega &= \frac{i\omega\mu}{c}\boldsymbol{H}_\omega \\
\boldsymbol{\nabla} \times \boldsymbol{H}_\omega &= -\frac{i\omega\epsilon}{c}\boldsymbol{E}_\omega \,.
\end{aligned}
\tag{D.1}
$$

All fields are zero in the perfect conductors. The surface currents and charges give discontinuities in the normal electric and tangential magnetic fields. The surface currents and charges will be such that the boundary conditions that the tangential components of $\boldsymbol{E}_\omega$ and normal components of $\boldsymbol{H}_\omega$ are zero at the conductor boundaries.

We take the z axis to be along the cylindrical guide axis. The guide is translationally invariant, and the translation operator eigenstates e^{ikz}, so we can again expand (or Fourier

transform) so that our general solution is given by

$$
\begin{aligned}
\boldsymbol{E}(\boldsymbol{r}, t) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \int_{-\infty}^{\infty} \frac{dk}{2\pi} \boldsymbol{E}_{k\omega}(x, y) \\
\boldsymbol{H}(\boldsymbol{r}, t) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \int_{-\infty}^{\infty} \frac{dk}{2\pi} \boldsymbol{H}_{k\omega}(x, y)
\end{aligned} \tag{D.2}
$$

Dropping the $k\omega$ subscripts that occur everywhere

$$
\begin{aligned}
\partial_x E_y - \partial_y E_x &= \frac{i\omega\mu}{c} H_z \\
\partial_x H_y - \partial_y H_x &= -\frac{i\omega\epsilon}{c} E_z \\
\partial_y E_z - ik E_y &= \frac{i\omega\mu}{c} H_x \\
\partial_y H_z - ik H_y &= -\frac{i\omega\epsilon}{c} E_x \\
ik E_x - \partial_x E_z &= \frac{i\omega\mu}{c} H_y \\
ik H_x - \partial_x H_z &= -\frac{i\omega\epsilon}{c} E_y
\end{aligned} \tag{D.3}
$$

Notice that the last four equations give four equations for the four components H_x, H_y, E_x, E_y, in terms of the z components and their derivatives if $\omega^2 \epsilon \mu \neq k^2$. Solving, we have

$$
\begin{aligned}
E_x &= \frac{ic(\omega\mu\partial_y H_z + kc\partial_x E_z)}{\omega^2\epsilon\mu - k^2 c^2)} \\
E_y &= \frac{ic(-\omega\mu\partial_x H_z + kc\partial_y E_z)}{\omega^2\epsilon\mu - k^2 c^2} \\
H_x &= \frac{ic(-\omega\epsilon\partial_y E_z + kc\partial_x H_z)}{\omega^2\epsilon\mu - k^2 c^2} \\
H_y &= \frac{ic(\omega\epsilon\partial_x E_z + kc\partial_y H_z)}{\omega^2\epsilon\mu - k^2 c^2}
\end{aligned} \tag{D.4}
$$

and substituting into the first two equations of Eqs. D.3 gives Helmholtz equations

$$
\begin{aligned}
\left(\partial_x^2 + \partial_y^2 + \frac{\omega^2\epsilon\mu}{c^2} - k^2 \right) E_z &= 0 \\
\left(\partial_x^2 + \partial_y^2 + \frac{\omega^2\epsilon\mu}{c^2} - k^2 \right) H_z &= 0 \,.
\end{aligned} \tag{D.5}
$$

We will come back to the case where $\omega^2 \epsilon \mu = k^2$ in a moment.

We can write these equations in any type of cylindrical coordinates by writing the transverse part of the gradient as $\boldsymbol{\nabla}_t = \hat{\boldsymbol{x}}\partial_x + \hat{\boldsymbol{y}}\partial_y$, so that the transverse fields are

$$
\begin{aligned}
\boldsymbol{E}_t &= \frac{ic(-\omega\mu\boldsymbol{\nabla}_t \times H_z + kc\boldsymbol{\nabla}_t E_z)}{\omega^2\epsilon\mu - k^2c^2} \\
\boldsymbol{H}_t &= \frac{ic(\omega\mu\boldsymbol{\nabla}_t \times E_z + kc\boldsymbol{\nabla}_t H_z)}{\omega^2\epsilon\mu - k^2c^2}
\end{aligned}
\tag{D.6}
$$

and the z components satisfy

$$
\begin{aligned}
0 &= \left(\nabla_t^2 + \frac{\omega^2\epsilon\mu}{c^2} - k^2\right)E_z \\
0 &= \left(\nabla_t^2 + \frac{\omega^2\epsilon\mu}{c^2} - k^2\right)H_z\,.
\end{aligned}
\tag{D.7}
$$

For ideal wave guides, with perfectly conducting boundaries, the transverse electric fields are continuous, $E_z = 0$ on the boundaries. For a boundary with normal $\hat{\boldsymbol{n}}$ pointing out of the boundary into the wave guide region, the $\hat{\boldsymbol{n}} \times \hat{\boldsymbol{z}}$ component of the electric field is zero, which means that $(\hat{\boldsymbol{n}} \times \hat{\boldsymbol{z}}) \cdot \boldsymbol{E}_t = 0$ on the boundary. The $\boldsymbol{\nabla}_t E_z$ component along the boundary must be zero since E_z is zero, therefore we require $(\hat{\boldsymbol{n}} \times \hat{\boldsymbol{z}}) \cdot (\boldsymbol{\nabla}_t \times H_z) = 0$, or

$$
\hat{\boldsymbol{n}} \cdot \nabla_t H_z = 0\,.
\tag{D.8}
$$

We therefore have two two-dimensional Helmholtz equations for E_z and H_z with the E_z value specified and the normal derivative of H_z specified. We can write them as

$$
\begin{aligned}
0 &= (\nabla_t^2 + \gamma^2)E_z \\
0 &= (\nabla_t^2 + \gamma^2)H_z
\end{aligned}
\tag{D.9}
$$

In general, since the boundary conditions are different, we would expect the γ values to be different for these two equations, therefore, for any particular γ value, usually only E_z or H_z cann be nonzero. Even for the case where E_z and H_z solutions share the same γ value, the possible solutions can be any linear combination of the E_z and H_z solutions alone. We therefore can divide the solutions into modes where $E_z = 0$, called transverse electric, or TE, modes, and $H_z = 0$, called transverse magnetic, or TM, modes.

We can now come back to the case where $\omega^2\epsilon\mu = k^2c^2$, which corresponds to $\gamma = 0$. In that case E_z and H_z must satisfy Laplace's equation. Since E_z must be zero on any conductors, and must go to zero far from any conductors (if we have propagation on the outside of the guide), only $E_z = 0$ will be a solution. Similarly, the only solution consistent with the normal derivative of H_z being zero on the conductors and $H_z = 0$ far from the conductors is constant H_z inside regions surrounded by a conductor. However, integrating Eq. D.3 around the boundary where tangential $\boldsymbol{E}$ is zero and using Stokes theorem shows that ωH_z integrated over the area must be zero. Therefore the constant H_z solution can only occur for $\omega = 0$. It is the usual static field from a static solenoidal current. The solutions

501

with $\omega^2 \mu \epsilon = k^2 c^2 \neq 0$ must then have both $E_z = H_z = 0$. Since there are no z components of the fields for this case, any modes like these are called the transverse electromagnetic or TEM modes. We then have $\partial_x E_y^{(TEM)} - \partial_y E_x^{(TEM)} = 0$, so that the transverse electric field can be derived from a potential. Substituting into Eq. D.3 gives

$$
\begin{aligned}
\boldsymbol{E}_t^{(TEM)} &= -\boldsymbol{\nabla}_t \Phi^{(TEM)} \\
\boldsymbol{H}_t^{(TEM)} &= -\frac{kc}{\omega\mu} \hat{\boldsymbol{z}} \times \boldsymbol{\nabla} \Phi^{(TEM)} \\
\nabla^2 \Phi^{(TEM)} &= 0 \, .
\end{aligned}
\tag{D.10}
$$

For the electric field to be zero in a conductor, $\Phi^{(TEM)}$ must be constant there. Only if we have two or more conductors will $\Phi^{(TEM)}$ be nonconstant outside the conductors. The number of independent TEM modes is therefore one less than the number of conductors.

Since the solutions of the Helmholtz equation with boundary conditions form a complete set, the TE modes give a complete set of solutions that have any nonzero H_z, the TM modes give a complete set of solutions that have any nonzero E_z, and the TEM modes give the solutions with $E_z = H_z = 0$. Substituting a linear combination of these, back into Maxwell's equations then gives the general solution

$$
\boldsymbol{E}(x, y, z, t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \sum_n \boldsymbol{E}_n(x, y) \left[F_n e^{ik_n z - i\omega t} + R_n e^{-ik_n z - i\omega t} \right]
\tag{D.11}
$$

where $\boldsymbol{E}_n(x, y)$ are the fields for all of the possible TE, TM, and TEM modes. Each separate mode is labeled by a different n. The k_n values are

$$
k_n = \sqrt{\epsilon\mu} \frac{\omega}{c} \qquad \text{for TEM modes}
$$

$$
k_n = \sqrt{\frac{\omega^2 \epsilon \mu}{c^2} - \gamma_n^2} \quad \text{for TE and TM modes} \quad \frac{\omega^2 \epsilon \mu}{c^2} > \gamma_n^2
$$

$$
k_n = i\sqrt{\frac{-\omega^2 \epsilon \mu}{c^2} + \gamma_n^2} \quad \text{for TE and TM modes} \quad \frac{\omega^2 \epsilon \mu}{c^2} < \gamma_n^2 \, .
\tag{D.12}
$$

Appendix E

Small angle expression for Legendre Polynomials

E.1   Small angle expansion of $P_\ell(\cos\theta)$

For small θ, we can approximate

$$P_\ell(\cos\theta) = J_0([\ell + \tfrac{1}{2}]\theta) + ... \tag{E.1}$$

where the ... terms are of order θ^2.

One way to derive this result is to use the recursion relation for the Legendre Polynomials,

$$(\ell + 1)P_{\ell+1}(\cos\theta) + \ell P_{\ell-1}(\cos\theta) - (2\ell + 1)\cos\theta P_\ell(\cos\theta) = 0 \tag{E.2}$$

and view ℓ as a continuous variable. We can then write

$$P_{\ell\pm1}(\cos\theta) = P_\ell(\cos\theta) \pm \partial_\ell P_\ell(\cos\theta) + \frac{1}{2}\partial_\ell^2 P_\ell(\cos\theta) \pm \frac{1}{6}\partial_\ell^3 P_\ell(\cos\theta) + ... \tag{E.3}$$

We can also expand $\cos\theta = 1 - \frac{\theta^2}{2} + \frac{\theta^4}{24} - ...$ in the last term of the recursion relation for small θ. Substituting these expressions the recursion relation becomes

$$\left[\partial_\ell + (\ell + \tfrac{1}{2})\partial_\ell^2 + (\ell + \tfrac{1}{2})\theta^2\right] P_\ell(\cos\theta) - \left[\tfrac{1}{6}\partial_\ell^3 + \frac{\ell + \frac{1}{2}}{12}\theta^4 + ...\right] P_\ell(\cos\theta) = 0 \tag{E.4}$$

We now change variables to $z = (\ell + \frac{1}{2})\theta$ so that $\partial_\ell = \theta\partial_z$, with the result that can be written two ways

$$
\begin{aligned}
0 &= \left[\partial_z + z\partial_z^2 + z - \left(\theta^2\partial_z^3 + \tfrac{1}{12}\theta^2 z + ...\right)\right] P_{\frac{z}{\theta} - \frac{1}{2}}(\cos\theta) \\
&= \left[\partial_z + z\partial_z^2 + z - \left(\frac{z^2}{(\ell + \frac{1}{2})^2}\partial_z^3 + \frac{z^3}{12(\ell + \frac{1}{2})^2} + ...\right)\right] P_{\frac{z}{\theta} - \frac{1}{2}}(\cos\theta). \tag{E.5}
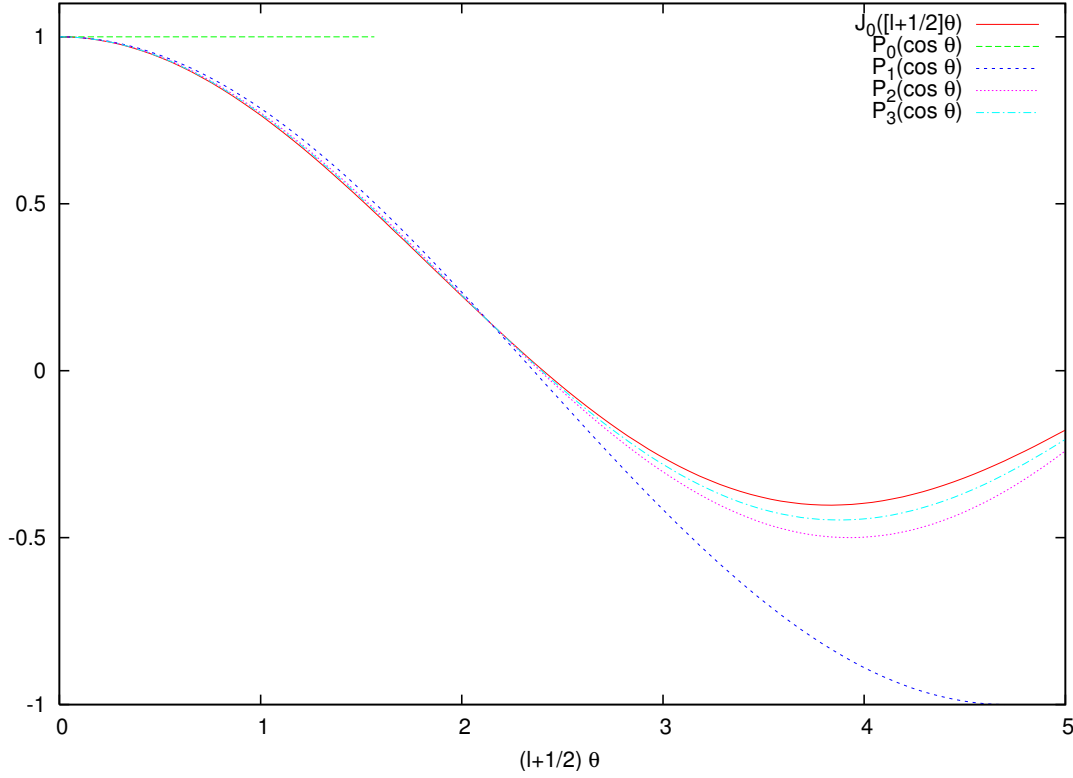\end{aligned}
$$

Figure E.1: A comparison of $J_0[(\ell + \frac{1}{2})\theta]$ with the Legendre polynomials $P_\ell(\cos\theta)$ for $\ell \le 3$.

The ... terms are higher order in θ, so for small θ or large ℓ we can drop the term in parentheses, and we find that for $\theta \ll 1$,

$$F(z) = P_{\frac{z}{\theta} - \frac{1}{2}}(\cos\theta)$$
$$\left[\partial_z + z\partial_z^2 + z\right] F(z) = 0 \tag{E.6}$$

The differential equation is Bessel's differential equation for order zero. The general solution is $F(z) = A J_0(z) + B N_0(z)$ for any constants A and B. For $\theta \to 0$, $N_0 \to -\infty$, so $B = 0$, and to match $P_\ell(\cos 0) = 1$, we must choose $A = 1$. We then find the small θ or large ℓ approximation $J_0(z) = P_{\frac{z}{\theta} - \frac{1}{2}}(\cos\theta)$ or

$$P_\ell(\cos\theta) = J_0([\ell + \tfrac{1}{2}]\theta) + ... \tag{E.7}$$

To give an idea of the quality of these approximations, we can plot the two expressions in fig. E.1.

Appendix F

Converting sums to integrals

F.1 Introduction

In quantum mechanics, it is often easier to start with a discrete set of states and take a limit to make a continuum. For example, to obtain the appropriate completeness relations for free particle solutions, it is often simplest to confine the particles in a box (usually periodic so that we can use momentum eigenstates) find the discrete set of states in the box, write the completeness relation as a sum over states, and finally, by taking the size of the box to infinity, we can convert the discrete sum to an integral and have a continuum of states.

Physically, since the visible universe is finite, truncation of our basis to a large and finite box cannot change the results of an experiment performed in a laboratory with all the constituents confined to the laboratory. Often, however, we can perform integrations easier than sums.

F.2 Conversion of a sum to an integral when the spacing goes to zero

A typical case is our k summation where for periodic boundary conditions in a box of side L, the wave functions are $e^{ik_n x}$, to be periodice, $k_n L = 2\pi n$, with n any integer from $-\infty < n < \infty$. As the box size L goes to infinity, the summation over n of any function becomes $\sum_n f(n)$. We can rewrite $f(n)$ as $f(\frac{kL}{2\pi}) = g(k)$. Next we need to look at physical quantities. If L is the size of the galaxy, then to describe a physical function in our laboratory, we will need to have many oscillations of $e^{ik_n x}$ over the laboratory, so that $L/2\pi n$ is of order the size of the laboratory. This means that n will need to be extremely large. As L gets larger, so does the required n values. The difference between the function $g(k)$ with $n = 10^{10}$, and $n = 10^{10} + 1$ will likely be in the tenth decimal place or less. As L gets larger, these differences will vanish, and we can replace the summation with the an integral. If n values

change by 1, we can simply write

$$\sum_n \to \int dn \, . \tag{F.1}$$

Usually we then change to more physical variables, so if $k = \frac{2\pi n}{L}$, $dn = \frac{L}{2\pi} dk$, and we have

$$\sum_k \to \frac{L}{2\pi} \int dk \, . \tag{F.2}$$

An equivalent way to write this that displays clearly what we are doing in this latter case is to write the a sum over $k_n = \frac{2\pi n}{L}$ values as

$$\sum_{k_n} f(k_n) = \frac{1}{\Delta k} \sum_{k_n} \Delta k f(k_n) \tag{F.3}$$

where Δk is the spacing between k_n values, $\Delta k = k_{n+1} - k_n = \frac{2\pi}{L}$. Taking the limit that $L \to \infty$ is the same as taking $\Delta k \to 0$. Therefore, the sum is

$$\lim_{L \to \infty} \sum_{k_n} \Delta k f(k_n) = \lim_{\Delta k \to 0} \sum_{k_n} \Delta k f(k_n) = \int dk f(k) \tag{F.4}$$

where we see that the last summation is the definition of an integral. We then have for this case

$$\sum_{k_n} f(k_n) \to \frac{L}{2\pi} \int dk f(k) \tag{F.5}$$

as before. For physical results there will always be all of the L factors will cancel in the limit of $L \to \infty$.

F.3   Euler-Maclaurin summation

Sometimes we will not be able to take the limit that the relative spacing goes to zero. Sometimes we have, instead, a small but fixed spacing. The Euler-Maclaurin summation formula can help us estimate the error of replacing a sum by an integral. The usual equation is

$$\int_a^b f(x)dx = h \left[\tfrac{1}{2}f(a) + \sum_{n=1}^{N-1} f(a+nh) + \tfrac{1}{2}f(b) \right] - \frac{B_2}{2!} h^2 f'(x) \Big|_a^b - \frac{B_4}{4!} h^4 f'''(x) \Big|_a^b - \dots \tag{F.6}$$

where $h = (b-a)/N$. The right hand side is the trapezoidal rule for the integral with the error terms given by the Bernoulli number expressions. This expression typically leads to an asymptotic series, therefore taking more terms in the Bernoulli number expression will usually not converge, however, the last term neglected gives a good estimate of the error, and this goes to zero with the step size.

One way to derive this is to use our quantum mechanical expressions $f(x) = \langle x|f\rangle$, with the translation operator by h, $T(h) = e^{\frac{i}{\hbar}ph}$, where p is the momentum conjugate to x so that it generates the translation. We then have $e^{-\frac{i}{\hbar}ph}|x\rangle = |x+h\rangle$, and $\langle x|e^{\frac{i}{\hbar}ph}|f\rangle = f(x+h)$. All of the translations and p commute, so we do not have to worry about their order. The trapezoidal rule can be written as

$$h\left[\tfrac{1}{2}f(a) + \sum_{n=1}^{N-1} f(a+nh) + \tfrac{1}{2}f(b)\right] = h\langle a|\left[\tfrac{1}{2} + T(h) + T^2(h) + T^3(h) + ... + \tfrac{1}{2}T^N(h)\right]|f\rangle$$

$$= h\langle a|\left[T^N(h) - 1\right]\left(\frac{1}{2} + \frac{1}{T(h) - 1}\right)|f\rangle. \qquad (F.7)$$

The integral is defined to be the limit of this expression when $N \to \infty$, so that $h \to 0$. In that case, the $T(h) - 1 \to \frac{i}{\hbar}ph + ...$, so that

$$\int_a^b dx f(x) = \langle a|\frac{T^N(h) - 1}{\frac{i}{\hbar}p}|f\rangle + O(h). \qquad (F.8)$$

since $h \to 0$, we can drop the $O(h)$ term. Subtracting these expressions, writing $T^N(h) = T(Nh)$, and factoring out $\frac{\hbar}{ip}$, we have

$$\int_a^b dx f(x) - h\left[\tfrac{1}{2}f(a) + \sum_{n=1}^{N-1} f(a+nh) + \tfrac{1}{2}f(b)\right] = \langle a|[T(Nh) - 1]\frac{\hbar}{ip}\left[1 - h\frac{ip}{\hbar}\left(\frac{1}{2} + \frac{1}{e^{\frac{i}{\hbar}ph} - 1}\right)\right]|f\rangle$$

$$(F.9)$$

The Bernoulli numbers can be defined by the coefficients of the power series

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!}x^n \qquad (F.10)$$

which leads to $B_3 = B_5 = B_7 = ... = 0$, $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$, $B_6 = \frac{1}{42}$, etc.[1] Therefore the operator corresponding to the difference between our integral and sum is

$$[T(Nh) - 1]\frac{\hbar}{ip}\left[1 - h\frac{ip}{\hbar}\left(\frac{1}{2} + \frac{1}{e^{\frac{i}{\hbar}ph} - 1}\right)\right] = [T(Nh) - 1]\left[-\frac{B_2}{2!}\left(\frac{i}{\hbar}ph\right) - \frac{B_4}{2!}\left(\frac{i}{\hbar}ph\right)^3 - ...\right]$$

$$(F.12)$$

---

[1] Other conventions use the closely related power series $\frac{x}{2}\cot\frac{x}{2} = 1 - B_1'\frac{x^2}{2!} - B_2'\frac{x^4}{4!} - B_3'\frac{x^6}{6!} - ...$ to define the Bernoulli numbers (indicated here with a prime to differentiate from our convention) in which case they are all positive and $B_n' = |B_{2n}|$.

The first few Bernoulli numbers can then be calculated by multiplying through by $\sin\frac{x}{2}$, expanding the sine and cosine, and equating the coefficients of the terms,

$$\frac{x}{2}\left[1 - \frac{x^2}{2^2 \cdot 2!} + \frac{x^4}{2^4 \cdot 4!} - ...\right] = \left[1 - B_1\frac{x^2}{2!} - B_2\frac{x^4}{4!} - B_3\frac{x^6}{6!} - ...\right]\left[\frac{x}{2} - \frac{x^3}{2^3 \cdot 3!} + \frac{x^5}{2^5 \cdot 5!} - ...\right] \qquad (F.11)$$

Since $\langle x|p|f\rangle = -i\hbar f'(x)$, substituting back gives the Euler-Maclaurin summation formula. Since a quantum mechanical wave function can be any well behaved (i.e. continuosly differentiable) function, our proof is good for any such well behaved function.

Various other forms can be derived. The midpoint rule can be derived as the difference between twice the trapezoidal rule with half the step size and the trapezoidal rule. This leads to

$$\int_a^b f(x)dx = h\sum_{n=1}^N f[a + (n - \tfrac{1}{2})h] + \left(1 - \frac{1}{2^1}\right)\frac{B_2}{2!}h^2 f'(x)\Big|_a^b + \left(1 - \frac{1}{2^3}\right)\frac{B_4}{4!}h^4 f'''(x)\Big|_a^b + ...$$

(F.13)

As an aside, these show why the trapezoidal and midpoint rules work so well for periodic functions where a and b refer to the same point where the difference in the a and b derivatives are zero, and for functions like gaussians when integrated to $\pm\infty$, beyond four or five standard deviations, all of the derivatives are well approximated by zero, and the trapezoidal rule converges exponentially fast. It also shows why, for these sorts of integrations, using Simpson's rule or similar is a complete waste of effort.

Appendix G

Determinants

G.1 Determinant

The determinant of an $N \times N$ matrix can be defined as

$$\text{Det} A = \sum_{\{i\}} A_{i_1,1} A_{i_2,2} A_{i_3,3} ... A_{i_N,N} \epsilon_{i_1,i_2,i_3,...,i_N} \tag{G.1}$$

where the sum over $\{i\}$ indicates summing over all possibilities for $i_1, ..., i_N$, and the Levi-Civita symbol $\epsilon_{i_1,i_2,i_3,...,i_N}$ is 1 for an even permutation of the integers $1, 2, 3, ..., N$, -1 for an odd permutation, and zero otherwise.

We can determine if a permutation of the integers is even or odd, by counting the number of pair exchanges needed to bring it into the form $1, 2, 3, ..., N$. For example, for $N = 5$, given the permutation $5, 3, 2, 4, 1$, we make the pair interchange to put the number 1 in the first position, $1, 3, 2, 4, 5$, next we make the interchange to put 2 in the second position, $1, 2, 3, 4, 5$, which shows that two pair interchanges gives the standard order, so $5, 3, 2, 4, 1$ is an even permutation of $1, 2, 3, 4, 5$.

Notice that interchanging two of the second subscripts of Eq. G.1 gives

$$\sum_{\{i\}} A_{i_1,2} A_{i_2,1} A_{i_3,3} ... A_{i_N,N} \epsilon_{i_1,i_2,i_3,...,i_N} = \sum_{\{i\}} A_{i_1,1} A_{i_2,2} A_{i_3,3} ... A_{i_N,N} \epsilon_{i_2,i_1,i_3,...,i_N} = -\text{Det} A \, .$$

$$\tag{G.2}$$

Similarly if two of the second subscripts are identical we have

$$\sum_{\{i\}} A_{i_1,1} A_{i_2,1} A_{i_3,3} ... A_{i_N,N} \epsilon_{i_1,i_2,i_3,...,i_N} = \sum_{\{i\}} A_{i_1,1} A_{i_2,1} A_{i_3,3} ... A_{i_N,N} \epsilon_{i_2,i_1,i_3,...,i_N} = 0 \tag{G.3}$$

Repeating these arguments for any set of indices shows that

$$\text{Det} A \epsilon_{j_1,j_2,j_3,...,j_N} = \sum_{\{i\}} A_{i_1,j_1} A_{i_2,j_2} A_{i_3,j_3} ... A_{i_N,j_N} \epsilon_{i_1,i_2,i_3,...,i_N} \, . \tag{G.4}$$

Multiplying both sides by $\epsilon_{j_1,j_2,j_3,...,j_N}$ and summing over all possible $\{j\}$ values gives us the desired result

$$\text{Det}\, A = \frac{1}{N!} \sum_{\{i\}\{j\}} A_{i_1,j_1} A_{i_2,j_2} A_{i_3,j_3}...A_{i_N,j_N} \epsilon_{i_1,i_2,i_3,...,i_N} \epsilon_{j_1,j_2,j_3,...,j_N} \,. \tag{G.5}$$

If we had defined the determinant as

$$\text{Det}\, A = \sum_{\{j\}} A_{1,j_1} A_{2,j_2} A_{3,j_3}...A_{N,j_N} \epsilon_{j_1,j_2,j_3,...,j_N} \,, \tag{G.6}$$

we could have gone through the same argument above, this time with the row indices. This would again lead to expression Eq. G.5, and the expression

$$\text{Det}\, A \epsilon_{i_1,i_2,i_3,...,i_N} = \sum_{\{j\}} A_{i_1,j_1} A_{i_2,j_2} A_{i_3,j_3}...A_{i_N,j_N} \epsilon_{j_1,j_2,j_3,...,j_N} \tag{G.7}$$

and we see that Eqs. G.1, G.4, G.5, G.6, and G.7 all give valid expressions for the determinant.

G.2   Product rule and invariance

The determinant of the product of two $N \times N$ matrices is the product of their determinants. That is if

$$C_{i,j} = \sum_k A_{i,k} B_{k,j} \tag{G.8}$$

then

$$\text{Det}\, C = \text{Det}\, A \, \text{Det}\, B \,. \tag{G.9}$$

To show this we write

$$
\begin{aligned}
\text{Det}\, C \;=\;& \frac{1}{N!} \sum_{\{i\},\{j\}} C_{i_1,j_1} C_{i_2,j_2}...C_{i_N,j_N} \epsilon_{i_1,i_2,i_3,...,i_N} \epsilon_{j_1,j_2,j_3,...,j_N} \\
=\;& \frac{1}{N!} \sum_{\{i\},\{j\},\{k\}} A_{i_1,k_1} B_{k_1,j_1} A_{i_2,k_2} B_{k_2,j_2} A_{i_3,k_3} B_{k_3,j_3}...A_{i_N,k_N} B_{k_N,j_N} \epsilon_{i_1,i_2,i_3,...,i_N} \epsilon_{j_1,j_2,j_3,...,j_N} \\
=\;& \frac{1}{N!} \sum_{\{k\}} \left[\sum_{\{i\}} A_{i_1,k_1} A_{i_2,k_2} A_{i_3,k_3}...A_{i_N,k_N} \epsilon_{i_1,i_2,i_3,...,i_N} \right] \\
& \cdot \left[\sum_{\{j\}} B_{k_1,j_1} B_{k_2,j_2} B_{k_3,j_3}...B_{k_N,j_N} \epsilon_{j_1,j_2,j_3,...,j_N} \right] \\
=\;& \frac{1}{N!} \sum_{\{k\}} \left[\det A \epsilon_{k_1,k_2,k_3,...,k_N} \right] \left[\det B \epsilon_{k_1,k_2,k_3,...,k_N} \right] \tag{G.10}
\end{aligned}
$$

The k values are summed over all possibilities, but we see in the last line that only the permutations of $1, 2, 3, ..., N$ give nonzero results. For any set of k values that is an even permutation of $1, 2, 3..., N$, both Levi-Civita symbols give $+1$, and for odd permutations both give -1, so the sum over $\{k\}$ gives a factor of $N!$, and the result is $\mathrm{Det}C = \mathrm{Det}A\mathrm{Det}B$.

We can apply this result immediately to a matrix and its inverse. We have

$$\sum_k A_{i,k} A_{k,j}^{-1} = \delta_{ij} , \tag{G.11}$$

and since the determinant of the identity matrix from Eq. G.1 is $\epsilon_{1,2,3,...,N} = 1$,

$$\mathrm{Det}A^{-1} = \frac{1}{\mathrm{Det}A} . \tag{G.12}$$

This then shows that the determinant is invariant under a similarity transform

$$\mathrm{Det}A^{-1}BA = \mathrm{Det}A\mathrm{Det}B\mathrm{Det}A^{-1} = \mathrm{Det}B . \tag{G.13}$$

G.3   Expansion in minors

From its definition, a determinant has $N!$ terms. A convenient way to enumerate these is to use what is called the expansion in minors. To see how this works, we can use Eq. G.6, and write out the j_1 sum explicitly

$$\mathrm{Det}A = A_{1,1} \sum_{\{j\}} A_{2,j_2} A_{3,j_3}...A_{N,j_N} \epsilon_{1,j_2,j_3,...,j_N} + A_{1,2} \sum_{\{j\}} A_{2,j_2} A_{3,j_3}...A_{N,j_N} \epsilon_{2,j_2,j_3,...,j_N}$$
$$+ A_{1,3} \sum_{\{j\}} A_{2,j_2} A_{3,j_3}...A_{N,j_N} \epsilon_{3,j_2,j_3,...,j_N} + ... + A_{1,N} \sum_{\{j\}} A_{2,j_2} A_{3,j_3}...A_{N,j_N} \epsilon_{N,j_2,j_3,...,j_N}$$
$$\tag{G.14}$$

In the first term, the $\epsilon_{1,j_2,j_3,...,j_N}$ will be one (minus one) for $j_2, j_3, ..., j_N$ even (odd) permutations of $2, 3, ..., N$. So this is the determinant of the matrix obtained by striking out the first row and first column of the original matrix. The determinant of this submatrix is called the minor (or first minor) for the element $A_{1,1}$. Similarly, in the second term, if we look at the term with $j_2 = 1, j_3 = 3, j_4 = 4, ..., j_N = N$, we see that $\epsilon_{2,j_2,j_3,...,j_N} = -1$ since interchanging 2 and j_2 brings us back to $1, 2, 3, ..., N$. Even permutations of this $j_2, ..., j_N$ will then give a negative sign, and odd permutations will give a positive sign, so $A_{1,2}$ is multiplied by the negative of its minor. Repeating this argument for the other $A_{1,k}$ terms, we find that the determinant of the matrix can be found by multiplying each element of the first row by a sign which alternates between positive and negative and elements minor. There is nothing special about the first row. We could have used any row with Eq. G.1 or any column with Eq. G.6. The result is that the determinant can be calculated by multiplying the elements of a row or the elements of a column by their minors and a sign which is given by $(-1)^{i+j}$ (i.e. the signs alternate like the colors of a chess board with the diagonals positive). Since the minors are determinants they can be expanded in turn. The minors times the sign $(-1)^{i+j}$ are called the cofactors $C_{i,j}$ of the elements $A_{i,j}$.

G.4 Simplifying determinants and numerical calculations

Since a determinant has $N!$ terms, analytic calculations of a general determinant quickly become impossible when N gets large. Sometimes it is possible to simplify the determinants of certain special forms. For numerical calculations, a general determinant can be calculated easily in order N^3 operations.

We can show the following rules

1. Interchanging two rows or two columns of a matrix changes the sign of the determinant.

2. Multiplying a row or a column by a constant multiplies the determinant by that constant.

3. Adding a constant times a row to another row, or adding a constant times a column to another column, does not change the determinant.

While we can prove these directly from the definitions Eqs. G.1 and G.6, an alternative is to use our determinant product rule. For example, we can interchange rows i and j by multiplying from the left by the matrix with $M_{i,j} = M_{j,i} = 1$, and $M_{k,k} = 1$ for $k \neq i$ or j. That is to interchange rows 1 and 2 we can multiply from the left by

$$
\begin{pmatrix}
0 & 1 & 0 & 0 & \dots & 0 \\
1 & 0 & 0 & 0 & \dots & 0 \\
0 & 0 & 1 & 0 & \dots & 0 \\
0 & 0 & 0 & 1 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \dots & 1
\end{pmatrix}
\tag{G.15}
$$

Calculating the determinant of this matrix we get -1, and our matrix product rule tells us that interchanging two rows changes the sign of the determinant. Multiplying AM instead of MA interchanges the corresponding colummns, so interchanging two columns changes the sign. Multiplying a row or column by a constant α corresponds to left or right multiplying by the matrix which is the identity except with that row or column's diagonal element changed to α. The determinant of this matrix is α, so this multiplies the determinant by α. For the last rule, multiplying a matrix by

$$
\begin{pmatrix}
1 & \alpha & 0 & 0 & \dots & 0 \\
0 & 1 & 0 & 0 & \dots & 0 \\
0 & 0 & 1 & 0 & \dots & 0 \\
0 & 0 & 0 & 1 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \dots & 1
\end{pmatrix}
\tag{G.16}
$$

from the left will add a α times the second row to the first row. Similar matrices can add a constant times any row or column to any other row or column. The determinant of these matrices is readily calculated using minors to be 1. Therefore this does not change the determinant.

The gaussian elimination method using row (or column) reduction is one standard way to calculate the determinant numerically. A typical method would begin with the first row $i = 1$, and

1. Find the largest magnitude element in the current row i and interchange its column with the ith column so this largest magnitude value is on the diagonal. If the largest magnitude is zero, the determinant is zero. The interchange (unless the largest magnitude was found on the diagonal) will multiply the determinant by -1.

2. For each row $j > i$, subtract A_{ji}/A_{ii} times the ith row. That is the new row will be

$$A'_{jk} = A_{jk} - \frac{A_{ji}A_{ik}}{A_{ii}} \tag{G.17}$$

   where the prime indicates the new, and the unprimed the old matrix. This will zero out the A_{ji} elements for $j > i$, so for subsequent rows only the $k \geq j$ elements will be affected.

3. Repeat steps 1 and 2 incrementing the row i until $i = N$. The resulting matrix is upper triangular with all elements $A_{ij} = 0$ for $i > j$. Using the expansion in minors, the determinant of the final matrix is the product of the diagonal elements. It is then multiplied by (-1) for each interchange required in step 1 to give the determinant of the original matrix.

The number of operations for an $N \times N$ matrix it is order N^3. This is the same order as the number of operations needed to multiply two dense $N \times N$ matrices together.

G.5 Inverse and Cofactors

If we multiply the transpose of the cofactor matrix times the matrix, the diagonal elements will be the sum of the elements of a column of the matrix times their cofactors, so all of the diagonal elements will be the determinant. The off diagonal elements will be the sum of the elements of a column of the matrix times the cofactors of a different column. This is equivalent to the determinant of a matrix which has that column written twice – but this is zero. So all of the off diagonal elements will be zero. The transpose of the cofactor matrix is therefore the determinant of the matrix times its inverse. For example, if we have $\sum_j A_{ij}x_j = y_i$ with A_{ij} and y_i given, we could solve by calculating the inverse of A, $x_i = \sum_j A_{ij}^{-1}y_j$. Substituting the cofactors divided by the determinant for A_{ij}^{-1}, we get that x_k is equal to the determinant of the matrix with y_j substituded for the A_{kj} elements divided by the determinant of the matrix A. This result is known as Kramer's rule.

Appendix H

Coordinate Systems

H.1 Introduction

It is useful to tabulate the various standard coordinate systems and their differential vector operations. While all of these results could be obtained by repeatedly applying the chain rule for all of the derivatives, a more organized method uses the metric. Nearly all of our work is for nonrelativistic systems which are in or embedded in three spatial dimensions with a Euclidean metric so that the distance ds between two infinitesimally separated points x, y, z, and $x + dx, y + dy, z + dz$ is

$$ds^2 = dx^2 + dy^2 + dz^2 \,. \tag{H.1}$$

It is helpful to use the notation $x = x_1$, $y = x_2$, $z = x_3$. We can describe points in other coordinate systems where the coordinates would be u_1, u_2, u_3, where each physical point cooresponds to a unique set of u_i values. Given u_1, u_2, and u_3, we have a physical point which has corresponding x_1, x_2, x_3 values. We can then write x_1, x_2, and x_3 as functions of u_1, u_2, and u_3. For example we can take $u_1 = r$, $u_2 = \theta$, $u_3 = \phi$, to be the standard spherical coordinates where

$$
\begin{aligned}
x &= r \sin\theta \cos\phi \\
y &= r \sin\theta \sin\phi \\
z &= r \cos\theta \,.
\end{aligned}
\tag{H.2}
$$

In general, we can write

$$dx_\alpha = \sum_{\beta=1}^{3} \frac{\partial x_\alpha}{\partial u_\beta} du_\beta \tag{H.3}$$

and substituting into our expression for ds^2,

$$ds^2 = \sum_{\alpha,\beta=1}^{3} g_{\alpha\beta} du_\alpha du_\beta \tag{H.4}$$

where the metric tensor $g_{\alpha\beta}$ is

$$g_{\alpha\beta} = \sum_{\gamma=1}^{3} \frac{\partial x_\gamma}{\partial u_\alpha} \frac{\partial x_\gamma}{\partial u_\beta} \,. \tag{H.5}$$

Notice that once we have calculated the metric tensor in terms of the u_α, we no longer have x, y, or z appearing. It is therefore unnecessary to use them at all. They are, however, a helpful crutch.

We can define unit vectors in the directions $\hat{u}_1$, $\hat{u}_2$, and $\hat{u}_3$. The vector that runs from u_1, u_2, u_3 to $u_1 + du_1, u_2, u_3$ for an infinitesimal du_1 points in the $\hat{u}_1$ direction. This is the Cartesian vector $\sum_{\alpha=1}^{3} \frac{\partial x_\alpha}{\partial u_1} \hat{x}_\alpha$. The unit vectors are therefore

$$\begin{aligned}
\hat{u}_\beta &= \left[\sum_{\gamma=1}^{3} \frac{\partial x_\gamma}{\partial u_\beta} \frac{\partial x_\gamma}{\partial u_\beta} \right]^{-1/2} \sum_\alpha \frac{\partial x_\alpha}{\partial u_\beta} \hat{x}_\alpha \\
&= \frac{1}{\sqrt{g_{\beta\beta}}} \sum_\alpha \frac{\partial x_\alpha}{\partial u_\beta} \hat{x}_\alpha \,.
\end{aligned} \tag{H.6}$$

H.2   Orthogonal Coordinates

The most useful coordinates for quantum mechanics are those where the Schrödinger equation separates. Most systems don't separate. Such systems are usually formulated in cartesian coordinates. For special cases where there is a separation, the coordinates are nearly always orthogonal. These are the usual coordinates that we study, which include circularly cylindrical, spherical, ellipsoidal, and parabolic. Orthogonal coordinates means that the unit vectors at a point are orthogonal, $\hat{u}_\alpha \cdot \hat{u}_\beta = \delta_{\alpha\beta}$. Substituting Eq. H.6, the requirement is that $g_{\alpha\beta} = 0$ if $\alpha \neq \beta$. That is, orthogonal coordinates have a diagonal metric tensor.

We can check that this is true for our spherical coordinates. Directly calculating the derivatives, we get

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2 \tag{H.7}$$

so that only the diagonal terms of the metric tensor are nonzero. They are

$$\begin{aligned}
g_{rr} &= 1 \\
g_{\theta\theta} &= r^2 \\
g_{\phi\phi} &= r^2 \sin^2\theta \,.
\end{aligned} \tag{H.8}$$

Notice that these are all positive. This is necessarily true since we started with the Euclidean metric which has all three terms positive (the Lorentz metric is an example where all the terms of the metric are not positive), and linearly transformed the differentials. It is then convenient to define $h_\alpha = \sqrt{g_{\alpha\alpha}}$. For spherical coordinates we have

$$\begin{aligned}
h_r &= 1 \\
h_\theta &= r \\
h_\phi &= r \sin\theta \,.
\end{aligned} \tag{H.9}$$

H.2.1 The gradient

We can define a componenent of the gradient by the change in a function with a change in distance when we infinitesimally change its coordinates, that is

$$F(u_1 + du_1, u_2, u_3) = F(u_1, u_2, u_3) + \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{\nabla} F ds \tag{H.10}$$

where ds is the distance between the coordinates $(u_1 + du_1, u_2, u_3)$ and the coordinates (u_1, u_2, u_3). From our metric expression for orthogonal coordinates, $ds = h_1 du_1$. Therefore

$$\hat{\boldsymbol{u}}_1 \cdot \boldsymbol{\nabla} F = \frac{1}{h_1} \frac{\partial F}{\partial u_1} . \tag{H.11}$$

Generalizing to all three coordinates,

$$\nabla F = \sum_{\alpha=1}^{3} \frac{1}{h_\alpha} \frac{\partial F}{\partial \alpha} \hat{\boldsymbol{u}}_\alpha . \tag{H.12}$$

Applying this to spherical coordinates, we have

$$\boldsymbol{\nabla} F = \left[\hat{\boldsymbol{r}} \frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\boldsymbol{\phi}} \frac{1}{r \sin\theta} \frac{\partial}{\partial \phi} \right] F(r, \theta, \phi) , \tag{H.13}$$

which is our usual spherical coordinate gradient.

H.2.2 The divergence

To calculate the divergence, we use the divergence theorem

$$\int dV \boldsymbol{\nabla} \cdot \boldsymbol{F} = \oint d\boldsymbol{S} \cdot \boldsymbol{F} \tag{H.14}$$

and look at an infinitesimal volume (and corresponding areas) given by the eight points where we either add zero or du_α to u_α. The volume integral becomes

$$\int dV \boldsymbol{\nabla} \cdot \boldsymbol{F} = du_1 h_1 du_2 h_2 du_3 h_3 \boldsymbol{\nabla} \cdot \boldsymbol{F} \tag{H.15}$$

while the surface integrals are

$$\begin{aligned}
\oint d\boldsymbol{S} \cdot \boldsymbol{F} = {}& \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{F} h_2 du_2 h_3 du_3 |_{(u_1+du_1, u_2, u_3)} - \hat{\boldsymbol{u}}_1 \cdot \boldsymbol{F} h_2 du_2 h_3 du_3 |_{(u_1, u_2, u_3)} \\
& + \hat{\boldsymbol{u}}_2 \cdot \boldsymbol{F} h_1 du_1 h_3 du_3 |_{(u_1, u_2+du_2, u_3)} - \hat{\boldsymbol{u}}_2 \cdot \boldsymbol{F} h_1 du_1 h_3 du_3 |_{(u_1, u_2, u_3)} \\
& + \hat{\boldsymbol{u}}_3 \cdot \boldsymbol{F} h_1 du_1 h_2 du_2 |_{(u_1, u_2, u_3+du_3)} - \hat{\boldsymbol{u}}_3 \cdot \boldsymbol{F} h_1 du_1 h_2 du_2 |_{(u_1, u_2, u_3)}
\end{aligned} \tag{H.16}$$

Dividing through by $du_1 h_1 du_2 h_2 du_3 h_3$, and writing $\hat{\boldsymbol{u}}_\alpha \cdot \boldsymbol{F} = F_\alpha$, we have

$$\boldsymbol{\nabla} \cdot \boldsymbol{F} = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial u_1} h_2 h_3 F_1 + \frac{\partial}{\partial u_2} h_1 h_3 F_2 + \frac{\partial}{\partial u_3} h_1 h_2 F_3 \right] . \tag{H.17}$$

Applying this to spherical coordinates

$$\boldsymbol{\nabla} \cdot \boldsymbol{F} = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 F_r + \frac{1}{r \sin\theta} \frac{\partial}{\partial \theta} \sin\theta F_\theta + \frac{1}{r \sin\theta} \frac{\partial}{\partial \phi} F_\phi \tag{H.18}$$

which agrees with the standard result.

H.2.3   The Laplacian

The Laplacian ∇^2 can be calculated as the divergence of the gradient. For orthogonal coordinates it is

$$\nabla^2 = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial u_1} \frac{h_2 h_3}{h_1} \frac{\partial}{\partial u_1} + \frac{\partial}{\partial u_2} \frac{h_1 h_3}{h_2} \frac{\partial}{\partial u_2} + \frac{\partial}{\partial u_3} \frac{h_1 h_2}{h_3} \frac{\partial}{\partial u_3} \right] . \tag{H.19}$$

For spherical coordinates we get

$$\begin{aligned}
\nabla^2 &= \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} . \\
&= \frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} .
\end{aligned} \tag{H.20}$$

H.2.4   The curl

The curl can be calculated by using Stokes' theorem,

$$\int dS \cdot (\nabla \times F) = \oint d\ell \cdot F \tag{H.21}$$

for an infinitesimal area in a similar manner to what we used for the divergence. Looking at the $\hat{u}_1$ component, the surface along $\hat{u}_1$ is $du_2 h_2 du_3 h_3$,

$$\int dS_1 (\nabla \times F)_1 = du_2 h_2 du_3 h_3 \hat{u}_1 \cdot (\nabla \times F) \tag{H.22}$$

and the corresponding integral around the boundary is

$$\oint d\ell \cdot F = F_3 h_3 du_3 |_{u_1, u_2+du_2, u_3} - F_3 h_3 du_3 |_{u_1, u_2, u_3} + F_2 h_2 du_2 |_{u_1, u_2, u_3+du_3} - F_2 h_2 du_2 |_{u_1, u_2, u_3} \tag{H.23}$$

This becomes

$$\hat{u}_1 \cdot (\nabla \times F) = \frac{1}{h_2 h_3} \frac{\partial}{\partial u_2} h_3 F_3 - \frac{1}{h_2 h_3} \frac{\partial}{\partial u_3} h_2 F_2 \tag{H.24}$$

or

$$\nabla \times F = \hat{u}_1 \frac{1}{h_2 h_3} \left[\frac{\partial}{\partial u_2} h_3 F_3 - \frac{\partial}{\partial u_3} h_2 F_2 \right] + \hat{u}_2 \frac{1}{h_1 h_3} \left[\frac{\partial}{\partial u_3} h_1 F_1 - \frac{\partial}{\partial u_1} h_3 F_3 \right] + \hat{u}_3 \frac{1}{h_1 h_2} \left[\frac{\partial}{\partial u_1} h_2 F_2 - \frac{\partial}{\partial u_2} h_1 F_1 \right] . \tag{H.25}$$

Applying this to spherical coordinates

$$\nabla \times F = \hat{r} \frac{1}{r \sin \theta} \left[\frac{\partial}{\partial \theta} \sin \theta F_\phi - \frac{\partial}{\partial \phi} F_\theta \right] + \hat{\theta} \frac{1}{r} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \phi} F_r - \frac{\partial}{\partial r} r F_\phi \right] + \hat{\phi} \frac{1}{r} \left[\frac{\partial}{\partial r} r F_\theta - \frac{\partial}{\partial \theta} F_r \right] . \tag{H.26}$$

H.3 Tabulation for standard coordinates

H.3.1 Cartesian

$$
\begin{aligned}
h_x &= h_y = h_z = 1 \\
\boldsymbol{\nabla} &= \hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z} \\
\boldsymbol{\nabla} \cdot \boldsymbol{F} &= \frac{\partial}{\partial x}F_x + \frac{\partial}{\partial y}F_y + \frac{\partial}{\partial z}F_z \\
\nabla^2 &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \\
\boldsymbol{\nabla} \times \boldsymbol{F} &= \hat{\boldsymbol{x}}\left[\frac{\partial}{\partial y}F_z - \frac{\partial}{\partial z}F_y\right] + \hat{\boldsymbol{y}}\left[\frac{\partial}{\partial z}F_x - \frac{\partial}{\partial x}F_z\right] + \hat{\boldsymbol{z}}\left[\frac{\partial}{\partial x}F_y - \frac{\partial}{\partial y}F_x\right] \quad (\text{H}.27)
\end{aligned}
$$

H.3.2 Spherical

$$
\begin{aligned}
x &= r\sin\theta\cos\phi \\
y &= r\sin\theta\sin\phi \\
z &= r\cos\theta \\
h_r &= 1 \\
h_\theta &= r \\
h_\phi &= r\sin\theta \\
\boldsymbol{\nabla} &= \left[\hat{\boldsymbol{r}}\frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial}{\partial\theta} + \hat{\boldsymbol{\phi}}\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right] \\
\boldsymbol{\nabla} \cdot \boldsymbol{F} &= \frac{1}{r^2}\frac{\partial}{\partial r}r^2 F_r + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta F_\theta + \frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}F_\phi \\
\nabla^2 &= \frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}. \\
\boldsymbol{\nabla} \times \boldsymbol{F} &= \hat{\boldsymbol{r}}\frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}\sin\theta F_\phi - \frac{\partial}{\partial\phi}F_\theta\right] + \hat{\boldsymbol{\theta}}\frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}F_r - \frac{\partial}{\partial r}rF_\phi\right] + \hat{\boldsymbol{\phi}}\frac{1}{r}\left[\frac{\partial}{\partial r}rF_\theta - \frac{\partial}{\partial\theta}F_r\right].
\end{aligned}
$$
$$(\text{H}.28)$$

H.3.3 Circular cylindrical

$$
\begin{aligned}
x &= \rho \cos\phi \\
y &= \rho \sin\phi \\
z &= z \\
h_\rho &= 1 \\
h_\phi &= \rho \\
h_z &= 1 \\
\boldsymbol{\nabla} &= \hat{\boldsymbol{\rho}}\frac{\partial}{\partial\rho} + \hat{\boldsymbol{\phi}}\frac{1}{\rho}\frac{\partial}{\partial\phi} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z} \\
\boldsymbol{\nabla}\cdot\boldsymbol{F} &= \frac{1}{\rho}\frac{\partial}{\partial\rho}\rho F_\rho + \frac{1}{\rho}\frac{\partial}{\partial\phi}F_\phi + \frac{\partial}{\partial z}F_z \\
\nabla^2 &= \frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial z^2} \\
\boldsymbol{\nabla}\times\boldsymbol{F} &= \hat{\boldsymbol{\rho}}\left[\frac{1}{\rho}\frac{\partial}{\partial\phi}F_z - \frac{\partial}{\partial z}F_\phi\right] + \hat{\boldsymbol{\phi}}\left[\frac{\partial}{\partial z}F_\rho - \frac{\partial}{\partial\rho}F_z\right] + \hat{\boldsymbol{z}}\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\rho F_\phi - \frac{1}{\rho}\frac{\partial}{\partial\phi}F_\rho\right]
\end{aligned}
\tag{H.29}
$$

H.3.4 Oblate spheroidal coordinates

Here c is a positive constant scaling factor. If $c \to 0$, with $r = c\xi$, these become spherical coordinates with $\gamma = \theta$.

$$
\begin{aligned}
x &= c\sqrt{1+\xi^2}\sin\gamma\cos\phi \\
y &= c\sqrt{1+\xi^2}\sin\gamma\sin\phi \\
z &= c\xi\cos\gamma \\
h_\xi &= c\sqrt{\frac{\cos^2\gamma + \xi^2}{1+\xi^2}} \\
h_\gamma &= c\sqrt{\cos^2\gamma + \xi^2} \\
h_\phi &= c\sin\gamma\sqrt{1+\xi^2} \\
\boldsymbol{\nabla} &= \hat{\boldsymbol{\xi}}\frac{1}{c}\sqrt{\frac{1+\xi^2}{\cos^2\gamma+\xi^2}}\frac{\partial}{\partial\xi} + \hat{\boldsymbol{\gamma}}\frac{1}{c\sqrt{\cos^2\gamma+\xi^2}}\frac{\partial}{\partial\gamma} + \hat{\boldsymbol{\phi}}\frac{1}{c\sin\gamma\sqrt{1+\xi^2}}\frac{\partial}{\partial\phi} \\
\boldsymbol{\nabla}\cdot\boldsymbol{F} &= \frac{1}{c(\cos^2\gamma+\xi^2)}\frac{\partial}{\partial\xi}\sqrt{1+\xi^2}\sqrt{\cos^2\gamma+\xi^2}F_\xi + \frac{1}{c\sin\gamma(\cos^2\gamma+\xi^2)}\frac{\partial}{\partial\gamma}\sin\gamma\sqrt{\cos^2\gamma+\xi^2}F_\gamma \\
&\quad + \frac{1}{c\sin\gamma\sqrt{1+\xi^2}}\frac{\partial}{\partial\phi}F_\phi \\
\nabla^2 &= \frac{1}{c^2(\xi^2+\cos^2\gamma)}\left\{\frac{\partial}{\partial\xi}\left[(\xi^2+1)\frac{\partial}{\partial\xi}\right] + \frac{1}{\sin\gamma}\frac{\partial}{\partial\gamma}\left[\sin\gamma\frac{\partial}{\partial\gamma}\right]\right\} + \frac{1}{c^2(\xi^2+1)\sin^2\gamma}\frac{\partial^2}{\partial\phi^2} \\
\boldsymbol{\nabla}\times\boldsymbol{F} &=
\end{aligned}
\tag{H.30}
$$

H.4 More general coordinates

Usually, nonorthogonal coordinates are not as useful. sometimes are useful to describe many-particle variational calculations. Occasionally we want to have particles that move in a curved geometry like on the surface of a sphere. If you know the mathematics of general relativity, you already know how to deal with general coordinates. We will use some of that notation here to simplify the description. Curved space can always be thought of as motion on a hypersurface embedded in a higher dimensional flat cartesian space, so it can be convenient (although perhaps an unnecessary crutch) to assume that there is always a flat cartesian space available to describe our system.

We define continuous coordinates x^α where $\alpha = 1, 2, 3$ for normal three-dimensional space. For the rest of this section, we will assume the number of dimensions is 3, but it is easy to generalize the results to more or fewer dimensions. All points in the region of interest are specified by a unique set of coordinates. At each point, we define 3 basis vectors $\boldsymbol{e}_\alpha$ which point in the direction of increasing x^α. We further define the differential vector from the point given by coordinates x^α to the point given by coordinates $x^\alpha + dx^\alpha$ to be

$$d\boldsymbol{r} = dx^\alpha \boldsymbol{e}_\alpha \tag{H.31}$$

where we employ Einstein's summation convention such that repeated indices with one upper and one lower indicates summation. The form of the vector where the coefficients have upper indices and the basis lower indices we call the contravariant form of the vector as opposed to the covariant form that we will write later. The distance squared between these two points is given by the dot product of the vector $d\boldsymbol{r}$ with itself

$$ds^2 = d\boldsymbol{r} \cdot d\boldsymbol{r} = \boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta dx^\alpha dx^\beta \tag{H.32}$$

and comparing with the definition of the metric tensor, Eq. H.4, we see that

$$g_{\alpha\beta} = \boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta \,. \tag{H.33}$$

Thinking geometrically, as long as the basis vectors $\boldsymbol{e}_\alpha$ are linearly independent, we can find a vectors that are orthogonal to all but one of them and choose any nonzero length. We write this new set of basis vectors as $\boldsymbol{e}^\alpha$ with the normalization requirement

$$\boldsymbol{e}^\alpha \cdot \boldsymbol{e}_\beta = \delta^\alpha_\beta = \left\{ \begin{array}{ll} 1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{array} \right. . \tag{H.34}$$

For three dimensions, we almost always take righthanded coordinates, and write the $\boldsymbol{e}^\alpha$ as

$$\boldsymbol{e}^1 = \frac{\boldsymbol{e}_2 \times \boldsymbol{e}_3}{\boldsymbol{e}_1 \cdot (\boldsymbol{e}_2 \times \boldsymbol{e}_3)}$$
$$\boldsymbol{e}^2 = \frac{\boldsymbol{e}_3 \times \boldsymbol{e}_1}{\boldsymbol{e}_1 \cdot (\boldsymbol{e}_2 \times \boldsymbol{e}_3)}$$
$$\boldsymbol{e}^3 = \frac{\boldsymbol{e}_1 \times \boldsymbol{e}_2}{\boldsymbol{e}_1 \cdot (\boldsymbol{e}_2 \times \boldsymbol{e}_3)} \tag{H.35}$$

which you may recognize as the definition of the reciprocal lattice vectors in terms of the physical lattice vectors using the crystallographer's convention from solid-state physics.

Notice that we have a completeness relation

$$\boldsymbol{e}_\beta (\boldsymbol{e}^\beta \cdot \boldsymbol{V}) = \boldsymbol{e}_\beta (\boldsymbol{e}^\beta \cdot \boldsymbol{e}_\alpha) V^\alpha = V^\beta \boldsymbol{e}_\beta = \boldsymbol{V} \tag{H.36}$$

so that $\boldsymbol{e}_\beta \otimes \boldsymbol{e}^\beta = \overset{\leftrightarrow}{I}$ is the identity tensor. The outer product symbol $\otimes$ is just there to remind us that we need to dot one or both of the vectors with some other vector or tensor – we did not use it in Eq. H.36 where we showed completeness. We can define $g^{\alpha\beta}$ as

$$g^{\alpha\beta} = \boldsymbol{e}^\alpha \cdot \boldsymbol{e}^\beta \tag{H.37}$$

and multiplying and using the completeness relation

$$g^{\alpha\beta} g_{\beta\gamma} = \boldsymbol{e}^\alpha \cdot \boldsymbol{e}^\beta \boldsymbol{e}_\beta \cdot \boldsymbol{e}_\gamma = \boldsymbol{e}^\alpha \cdot \boldsymbol{e}_\gamma = \delta^\alpha_\gamma \tag{H.38}$$

shows that the matrix $g^{\alpha\beta}$ is the inverse of the matrix $g_{\alpha,\beta}$.

Since the $\boldsymbol{e}^\alpha$ form an independent basis that spans our space, we can equally well write our vectors in terms of them

$$\boldsymbol{V} = V_\alpha \boldsymbol{e}^\alpha \tag{H.39}$$

which is the covariant form. We can convert from one form to the other by equating $V_\alpha \boldsymbol{e}^\alpha$ with $V^\gamma \boldsymbol{e}_\gamma$, and dotting with $\boldsymbol{e}_\beta$,

$$\boldsymbol{e}_\beta \cdot \boldsymbol{e}^\alpha V_\alpha = V^\gamma \boldsymbol{e}_\beta \cdot \boldsymbol{e}_\gamma \tag{H.40}$$

or

$$V_\beta = V^\gamma g_{\gamma\beta} \tag{H.41}$$

and multiplying through by the inverse

$$V^\beta = V_\gamma g^{\gamma\beta} \tag{H.42}$$

which shows that the metric tensor and its inverse can be used to lower and raise the indices.

H.4.1   Volume element

Since the metric tensor is symmetric, we can always find its real eigenvalues λ_n with real orthonormal eigenvectors $\psi^{(n)}_\alpha$ at any point in space. We can write the locally transformed coordinates as

$$y^n = \psi^{(n)}_\alpha x^\alpha \tag{H.43}$$

Moving in the y^n direction by dy will correspond to a change of length by $\sqrt{|\lambda_n|} dy^n$. Since infinitesimal motion along different y_n coordinates is orthogonal, the physical volume element for d dimensions is

$$d^d r = \prod_{n=1}^d \sqrt{|\lambda_n|} dy^n = \sqrt{|\det g|} \prod_n dy^n = \sqrt{|\det g|} \prod_\alpha dx^\alpha \tag{H.44}$$

where in the last line we transform back to x^α, and note that the Jacobian of the transformation is the absolute value of the determinant of the eigenvector matrix. Since the eigenvectors are orthonormal, this matrix is orthogonal, and the Jacobian is one.

H.4.2   The gradient

If we move by $d\boldsymbol{x} = dx^\alpha \boldsymbol{e}_\alpha$, a scalar function S of these coordinates will change by

$$\frac{\partial S}{\partial x^\alpha} = \boldsymbol{e}_\alpha \cdot \boldsymbol{\nabla} S \,. \tag{H.45}$$

It is useful when using just one set of coordinates to define

$$\partial_\alpha \equiv \frac{\partial}{\partial x^\alpha} \,. \tag{H.46}$$

Multiplying Eq. H.45 by $\boldsymbol{e}^\alpha$ and summing gives the expression for the gradient

$$\boldsymbol{\nabla} S = \boldsymbol{e}^\alpha \partial_\alpha S \,. \tag{H.47}$$

H.4.3   Reduction to orthogonal coordinates

For orthogonal coordinates, $g_{\alpha,\beta}$ is diagonal. Therefore

$$\begin{aligned}
\boldsymbol{e}_\alpha &= \sqrt{g_{\alpha\alpha}}\,\hat{\boldsymbol{u}}_\alpha \,. \\
g^{\alpha\alpha} &= \frac{1}{g_{\alpha\alpha}} \\
\boldsymbol{e}^\alpha &= \sqrt{g^{\alpha\alpha}}\,\hat{\boldsymbol{u}}_\alpha \,.
\end{aligned} \tag{H.48}$$

The gradient is

$$\boldsymbol{\nabla} = \sum_\alpha \frac{\hat{\boldsymbol{u}}_\alpha}{\sqrt{g_{\alpha\alpha}}} \frac{\partial}{\partial x^\alpha} \,. \tag{H.49}$$

Appendix I

Linear one-dimensional differential equations with constant coefficients

I.1 Solutions of linear ordinary differential equations with constant coefficients

The equations are of the form

$$\sum_{n=0}^{N} \alpha_n \frac{d^n}{dx^n} y(x) = 0 \tag{I.1}$$

where α_n are, in general, complex constants. Without loss of generality, we can assume that $\alpha_N = 1$, since if $\alpha_N = 0$, we can change the maximum index to $N - 1$. If $\alpha_N \neq 0$, we can divide through by it without changing the result.

Let's look at the first interesting equation where $N = 1$,

$$\frac{d}{dx} y(x) + \alpha_0 y(x) = 0 \tag{I.2}$$

We can integrate this immediately

$$\int \frac{dy}{y} = -\alpha_0 \int dx \tag{I.3}$$

to get the general solution

$$y(x) = A e^{-\alpha_0 x} \,. \tag{I.4}$$

Next let's look at the equation with $N = 2$.

$$\left[\frac{d^2}{dx^2} + \alpha_1 \frac{d}{dx} + \alpha_0 \right] y(x) = 0 \tag{I.5}$$

The key insight is to notice that the quantity inside the brackets looks like a quadratic polynomial in the "variable" $\frac{d}{dx}$. That is we can write

$$P(u) = u^2 + \alpha_1 u + \alpha_0 \,. \tag{I.6}$$

We know that we can write this polynomial in terms of its two roots u_1 and u_2

$$P(u) = (u - u_1)(u - u_2) \tag{I.7}$$

where $P(u_1) = P(u_2) = 0$. You can check this explicitly if you need reassurance: using the quadratic equation we have

$$u_1 = \frac{-\alpha_1 + \sqrt{\alpha_1^2 - 4\alpha_0}}{2}$$
$$u_2 = \frac{-\alpha_1 - \sqrt{\alpha_1^2 - 4\alpha_0}}{2} \tag{I.8}$$

and substituting above gives the original polynomial. We can then write our differential equation as

$$\left(\frac{d}{dx} - u_2\right)\left(\frac{d}{dx} - u_1\right) y(x) = \left(\frac{d}{dx} - u_1\right)\left(\frac{d}{dx} - u_2\right) y(x) = 0 \tag{I.9}$$

where either order of the factors works.

Now if we have a solution to

$$\left(\frac{d}{dx} - u_1\right) y(x) = 0 \tag{I.10}$$

we have a solution to the first ordering, while if we have a solution to

$$\left(\frac{d}{dx} - u_2\right) y(x) = 0 \tag{I.11}$$

we have a solution to the second ordering. Both of these can be given exactly as for the case where $N = 1$, that is

$$\begin{aligned} y_1(x) &= A_1 e^{u_1 x} \\ y_2(x) &= A_2 e^{u_2 x} \, . \end{aligned} \tag{I.12}$$

Further since our equations are linear, $y(x) = A_1 e^{u_1 x} + A_2 e^{u_2 x}$ is also a solution.

Let's now assume that $u_1 \neq u_2$ and come back to that case later. If $u_1 \neq u_2$, then the two solutions are different, and our solution above, which has two integration constants, for a second order differential equation must be the general solution.

For arbitrary N, we have to find the roots of the the corresponding $P(u)$ which is called the characteristic polynomial. Analytically, that can be done in general to order $N = 4$. For special cases where there is some symmetry that allows us to figure out the roots, we can sometimes get analytic results for higher N. We can always calculate the roots numerically for numerical values of α_n. Once we have the roots, assuming they are distinct, we then write down the general solution as above $y(x) = \sum_{m=1}^{N} A_m e^{u_m x}$.

Let's deal with the case where we have two or more roots that are exactly equal. We could go back and try to solve the second order differential equation. However, we are interested in solving physics problems, and in physics, this cannot be different than the case where the roots are very very close. Therefore, lets look at the $N = 2$ case where the roots are u_1 and $u_1 + \epsilon$ with ϵ very small, i.e. we will take the limit that it goes to zero – but it is *never* exactly zero. In that case our general solution will be

$$A_1 e^{u_1 x} + A_2 e^{(u_1 + \epsilon)x} \tag{I.13}$$

For any x range, we can imagine ϵ small enough that ϵx is small, and we can expand

$$e^{\epsilon x} = 1 + \epsilon x + ... \tag{I.14}$$

The one term multiplying $A_2 e^{u_1 x}$ has the same form as the first solution, so we can combine it with the first solution. The second term gives an independent solution, therefore we have the general solution in the case where $\epsilon \to 0$,

$$y(x) = A_1 e^{u_1 x} + A_2 x e^{u_1 x} \tag{I.15}$$

Continuing on, we find the additional solutions are given by the u_1 derivatives of $e^{u_1 x}$, i.e. $x^m e^{u_1 x}$. So for each distinct root we have a solution e^{ux}, and for multiple roots with multiplicity M, we multiply these by by 1, x, x^2, x^3 etc. up to x^{M-1}. A linear combination of these solutions is the general solution.[1]

I.2 Boundary conditions

In quantum mechanics we typically have second order equations for the wave function. These will have boundary conditions where they typically are either zero, have slope zero, or have their value and slope matched to other solutions. Enforcing the boundary conditions will then pick out allowed values for the coefficients (the A_m above) and often will constrain the allowed values for the coefficients in the equation.

For example, the eigenvalue equation for one particle in a 1-dimensional infinite square well running from $-L/2 < x < L/2$ is

$$\left[\frac{d^2}{dx^2} + \frac{2mE_n}{\hbar^2} \right] = 0 \tag{I.16}$$

with boundary conditions that $\psi_n(-L/2) = \psi_n(L/2) = 0$. The general solution is

$$\psi(x) = A_1 e^{i\sqrt{\frac{2mE_n}{\hbar^2}}x} + A_2 e^{-i\sqrt{\frac{2mE_n}{\hbar^2}}x} \tag{I.17}$$

---

[1]Now that we see what the solutions are, it's easy to motivate trying the solution $f(x)e^{ax}$ for $\left(\frac{d}{dx} - a\right)^n y(x) = 0$. Taking an unknown function multiplying a known solution will nearly always eliminate some terms in the differential equation making finding a solution for $f(x)$ easier. In this case, we find that $\left(\frac{d}{dx} - a\right) f(x)e^{ax}$ becomes $\frac{df(x)}{dx}$, so that $\left(\frac{d}{dx} - a\right)^n f(x)e^{ax} = 0$ becomes $\frac{d^n f(x)}{dx^n} = 0$, which has solutions $A + Bx + Cx^2 + ...Dx^{n-1}$ giving our previous solution.

for $E_n > 0$, and

$$\psi(x) = A_1 e^{\sqrt{\frac{-2mE_n}{\hbar^2}}x} + A_2 e^{-\sqrt{\frac{-2mE_n}{\hbar^2}}x} \tag{I.18}$$

for $E_n < 0$. In the second case, we cannot match the boundary conditions; since one of the exponentials is monotonically increasing and the other montonically decreasing a linear combination can only be zero at one point. Therefore, there is no negative E_n value allowed.

For positive energies, the requirement at $x = -L/2$ leads to

$$\psi(x) = A_1 \left(e^{i\sqrt{\frac{2mE_n}{\hbar^2}}(x+L/2)} - e^{-i\sqrt{\frac{2mE_n}{\hbar^2}}(x+L/2)} \right) \tag{I.19}$$

A_1 is a normalization, and matching at $x = L/2$ gives

$$0 = \left(e^{i\sqrt{\frac{2mE_n}{\hbar^2}}L} - e^{-i\sqrt{\frac{2mE_n}{\hbar^2}}L} \right) = 2i \sin\left(\sqrt{\frac{2mE_n}{\hbar^2}}L \right) \tag{I.20}$$

The sine will be zero when its argument is an integer multiple of π.

It is important to examine all possible solutions. We have

$$\sqrt{\frac{2mE_n L^2}{\hbar^2}} = n\pi \tag{I.21}$$

with n an integer. The normalized eigenfunctions are

$$\psi_n(x) = \begin{cases} \sqrt{\frac{2}{L}} \cos\left(\frac{n\pi}{L}x\right) & n \text{ odd} \\ \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi}{L}x\right) & n \text{ even} \end{cases} \tag{I.22}$$

Notice that $n = 0$ gives $\psi(x) = 0$, so it is not normalizable and not included. Similarly, negative n values do not give independent functions. That is $n = -1$ eigenfunction is the identical to the $n = 1$ eigenfunction, while the $n = -2$ eigenfunction is the negative of the $n = 2$ eigenfunction. Therefore we take only $n > 0$. The corresponding energies are from Eq. I.21

$$E_n = \frac{\hbar^2 \pi^2 n^2}{2mL^2} . \tag{I.23}$$

Appendix J

Baker-Campbell-Hausdorff and related formulas

J.1   Introduction

The Baker-Campbell-Hausdorff formula

$$e^{At}e^{Bt} = e^{(A+B)t+\frac{t^2}{2}[A,B]+\frac{t^3}{12}([A,[A,B]]+[B,[B,A]])+\dots} \tag{J.1}$$

relates the product of the exponentials of two noncommuting operators to a single exponential. The derivation of the general expression is not trivial, but, in physics, an analytic expression is typically useful only if the expansion truncates after a few terms or if the higher-order terms are small enough to be neglected. Therefore, we will only derive the first few terms of the expansion.

J.2   Taylor Expansion method

We first rewrite the equation as

$$e^{At}e^{Bt} = e^{S(t)} \tag{J.2}$$

and our task is to find the expansion of S in powers of t. We can write

$$S(t) = \ln\left(e^{At}e^{Bt}\right) \tag{J.3}$$

and use the Taylor series expansions

$$
\begin{aligned}
\exp(x) &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \\
\ln(1+x) &= \int_0^x dx' \frac{1}{1+x'} = \int_0^x dx' \sum_{n=0}^{\infty}(-1)^n x'^n = -\sum_{n=1}^{\infty} \frac{(-1)^n}{n} x^n
\end{aligned}
\tag{J.4}
$$

Substituting, the result is

$$S(t) = -\sum_{n=1}^{\infty} \frac{(-1)^n}{n} \left[\sum_{j=0,k=0,j+k>0}^{\infty} (At)^j (Bt)^k \right]^n . \tag{J.5}$$

To obtain all of the terms up to order t^3, we first keep terms only to that order in the exponentials

$$
\begin{aligned}
\exp(At)\exp(Bt) - 1 &= \left[1 + At + \tfrac{1}{2}A^2 t^2 + \tfrac{1}{6}A^3 t^3 + ...\right] \\
&\quad \left[1 + Bt + \tfrac{1}{2}B^2 t^2 + \tfrac{1}{6}B^3 t^3 + ...\right] - 1 \\
&= (A+B)t + \tfrac{1}{2}(A^2 + 2AB + B^2)t^2 \\
&\quad + \tfrac{1}{6}(A^3 + 3A^2 B + 3AB^2 + B^3)t^3 + ...
\end{aligned}
\tag{J.6}
$$

A faster way to this result is to realize that if the operators commuted, we would just have the exponential expansion of $\exp[(A+B)t]$. In the product of the exponentials, we have all of those terms. The only difference is that the A operators are to the left of all the B operators. The logartithm expansion to this order is

$$\ln(1+x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 + ... \tag{J.7}$$

so that

$$
\begin{aligned}
S &= \ln\left[\exp(At)\exp(Bt)\right] \\
&= (A+B)t + \tfrac{1}{2}(A^2 + 2AB + B^2)t^2 + \tfrac{1}{6}(A^3 + 3A^2 B + 3AB^2 + B^3)t^3 \\
&\quad - \tfrac{1}{2}\left[(A+B)t + \tfrac{1}{2}(A^2 + 2AB + B^2)t^2\right]^2 + \tfrac{1}{3}\left[(A+B)t\right]^3 + ... \\
&= (A+B)t + \tfrac{1}{2}\left[A^2 + 2AB + B^2 - (A+B)^2\right]t^2 \\
&\quad + \tfrac{1}{12}\left[2A^3 + 6A^2 B + 6AB^2 + 2B^3 - 3(A^2 + 2AB + B^2)(A+B)\right. \\
&\quad \left. -3(A+B)(A^2 + 2AB + B^2) + 4(A+B)^3\right]t^3 + ...
\end{aligned}
\tag{J.8}
$$

Expanding and cancelling terms this becomes

$$
\begin{aligned}
S &= (A+B)t + \tfrac{1}{2}\left(AB - BA\right)t^2 \\
&\quad + \tfrac{1}{12}\left(A^2 B - 2ABA + BA^2 + AB^2 - 2BAB + B^2 A\right)t^3 + ... \\
&= (A+B)t + \tfrac{1}{2}[A,B]t^2 + \tfrac{1}{12}\left([A,[A,B]] + [B,[B,A]]\right)t^3 + ...
\end{aligned}
\tag{J.9}
$$

Notice that in Eq. J.9, the odd order terms in t are invariant under interchange of A and B, while the even order term in t changes sign. This behavior is true to all orders as can be understood by writing

$$T(A, B, t) = e^{At} e^{Bt} = e^{S(t)} \tag{J.10}$$

If we interchange A and B and change the sign of t, we get

$$T(B, A, -t) = e^{-Bt} e^{-At} . \tag{J.11}$$

530

We immediately see that $T(B, A, -t)T(A, B, t) = 1$, so $T(B, A, -t)$ is the inverse of $T(A, B, t)$, or

$$T(B, A, -t) = e^{-S(t)} \tag{J.12}$$

so interchanging A and B while changing the sign of t must change the sign of every term in $S(t)$. For the odd order terms in t, the t sign change gives a minus one, so the term is invariant under interchange of A and B. For the even order terms in t, the t sign change does nothing, so the term must change sign under interchange of A and B.

J.3 Nested commutator expansion

From our result above, and thinking a bit about the structure, it should seem plausible that all of the terms in S beyond the first $(A + B)t$ term will be sums of nested commutators of A and B. We will follow the general idea of D. Ž. Djoković[1].

It is helpful to define a commutation operator $[O$, with multiple such operators terminated with a single $]$, which represents as many right brackets as necessary to terminate the nested commutators. That is

$$
\begin{aligned}
([A,)\ \ B] &= [A, B] \\
([A,)([B,)\ \ C] &= [A, [B, C]]
\end{aligned}
\tag{J.13}
$$

and we define functions of this operator by their power series expansions

$$
\begin{aligned}
e^{[A,}\ B] &= 1 + [A, B] + \tfrac{1}{2!}[A, [A, B]] + \frac{1}{3!}[A.[A, [A, B]]] + ... \\
\frac{1}{1 - [A,}\ B] &= 1 + [A, B] + [A, [A, B]] + [A, [A, [A, B]]] + ...
\end{aligned}
\tag{J.14}
$$

In the mathematics literature this operator is often written ad[2]

So for example, Taylor series expanding $e^{At}Be^{-At}$, the derivative with respect to t is

$$\frac{\partial}{\partial t} e^{At}Be^{-At} = e^{At}ABe^{-At} - e^{At}BAe^{-At} = e^{At}[A, B]e^{-At} \tag{J.15}$$

repeating, we get a set of nested commutators, so expanding about $t = 0$ gives

$$e^{At}Be^{-At} = B + [A, B]t + \frac{1}{2!}[A, [A, B]]t^2 + \frac{1}{3!}[A, [A, [A, B]]]t^3 + ... = e^{t[A,}\ B]. \tag{J.16}$$

Let's write $e^{S(t)}\frac{\partial}{\partial t}e^{-S(t)}$ in two ways. First, we can write

$$
\begin{aligned}
e^{S(t)}\frac{\partial}{\partial t}e^{-S(t)} &= e^{S(t)}\frac{\partial}{\partial t}e^{-Bt}e^{-At} = -e^{S(t)}Be^{-S(t)} - A = -e^{At}Be^{-At} - A \\
&= -A - e^{[S(t),}\ B] = -A - e^{t[A,}\ B].
\end{aligned}
\tag{J.17}
$$

[1]D. Ž. Djoković, Math. Z. **143** 209(1975).
[2]ad comes from the adjoint representation of a Lie group, but that is not important for our purposes.

Next we write

$$e^{-S(t)} = \lim_{N\to\infty}\left(1 - \frac{S(t)}{N}\right)^N \tag{J.18}$$

so that

$$\frac{\partial}{\partial t}e^{-S(t)} = \frac{\partial}{\partial t}\lim_{N\to\infty}\left(1 - \frac{S(t)}{N}\right)^N = -\frac{1}{N}\lim_{N\to\infty}\sum_{j=1}^{N}\left(1 - \frac{S(t)}{N}\right)^{N-j}\frac{\partial S(t)}{\partial t}\left(1 - \frac{S(t)}{N}\right)^{j-1}$$

$$= -\int_0^1 d\lambda e^{-(1-\lambda)S(t)}\frac{\partial S(t)}{\partial t}e^{-\lambda S(t)}. \tag{J.19}$$

Therefore

$$e^{S(t)}\frac{\partial}{\partial t}e^{-S(t)} = -\int_0^1 d\lambda e^{\lambda S(t)}\frac{\partial S(t)}{\partial t}e^{-\lambda S(t)} = -\int_0^1 d\lambda e^{\lambda[S(t),}\ \frac{\partial S(t)}{\partial t}]$$

$$= \frac{1 - e^{[S(t),}}{[S(t),}\ \frac{\partial S(t)}{\partial t}] \tag{J.20}$$

where you should check that power series expanding, then integrating gives the same result as power series expanding the result. The power series for $(e^x - 1)/x$ is invertible, so equating Eqs. J.17 and J.20,

$$\frac{\partial S(t)}{\partial t} = \frac{[S(t),}{\exp([S(t),) - 1}\ \left(A + e^{t[A,}B)\right].$$

$$= \sum_{n=0}^{\infty}\frac{B_n}{n!}\left([S(t),)^n\ \left(A + e^{t[A,}B)\right], \tag{J.21}$$

where B_n are Bernoulli numbers, $B_3 = B_5 = B_7 = ... = 0$, $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$, $B_6 = \frac{1}{42}$, etc.

Writing

$$S(t) = \sum_{n=1}^{\infty}S_n t^n \tag{J.22}$$

we can solve recursively. That is, the S_n term on the right only contributes to terms of order t^m where $m \geq n$. Therefore knowing S_m for $m < n$, we can substitute the series for $S(t)$ truncated at $m = n - 1$, on the right, and solve for S_n on the left. For example, we keep terms to first order in t everywhere on both sides,

$$S_1 + 2S_2 t + ... = \left(1 - \frac{1}{2}t[S_1,\right)\ \left(A + (1 + t[A,)B)\right] + ... \tag{J.23}$$

The terms with no t on the right are $A+B$, so $S_1 = A+B$. We can now substitute $S_1 = A+B$ on the right to find that

$$2S_2 = [A, B] \tag{J.24}$$

as before.

Here we just want to note that t only occurs in the commutation operators on the right. Therefore, as we iterate, we will only have terms that are made up of commutators of A and B.

J.4 Operators that commute with their commutators

Sometimes A and B commute with $[A, B]$. A typical example is if A is proportional to the momentum and B proportional to the conjugate position. The momentum and conjugate position do not commute, but, since their commutator is proportional to the identity, they both commute with their commutator. In that case, from our previous result, we know all of the higher order terms must be made up of terms that have as their last commutator in the nested commutators either $[A, B]$ or $[B, A]$. The next operator in the nest must also be A or B, but if they commute with their commutator, all of these terms must be zero. Therefore, if A and B commute with $[A, B]$, we have

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]} = e^{A+B} e^{\frac{1}{2}[A,B]} \tag{J.25}$$

or

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]} . \tag{J.26}$$

J.5 Dynkin's replacement

Dynkin developed a method that gives explicit expressions for the commutators at each order in the Baker-Campbell-Hausdorff formula[3]. He showed that whenever we have a sum over combinations of products of operators that can be written as a sum over nested commutators, then you can find the sum over nested commutators by replacing each of the operator products with the corresponding nested commutator and dividing by the total number of operators.

For example, you can make the replacements everywhere

$$
\begin{aligned}
DCBA &\rightarrow \tfrac{1}{4}[D, [C, [B, A]]] \\
ABCDEFGHIJ &\rightarrow \tfrac{1}{10}[A, [B, [C, [D, [E, [F, [G, [H, [I, [J]]]]]]]]]]
\end{aligned}
\tag{J.27}
$$

etc.

To show this, imagine that you started with a nested commutator expression, but have expanded it out into operator products. If you make Dynkin's replacement changing each of the operator products into a nested commutator and dividing by the number of operators you should get back the original nested commutator.

This result is trivially true for two operators. The commutator expression is $[B, A] = BA - AB$, replacing each of these by one half of the corresponding commutator

$$[B, A] = BA - AB \overset{?}{=} \tfrac{1}{2}[B, A] - \tfrac{1}{2}[A, B] = [B, A] \tag{J.28}$$

---

[3]E.B. Dynkin, Dokl. Akad. Nauk. SSSR **57**, 323 (1947); English translation in *Selected papers of E.B. Dynkin with commentary*, Ed. E.B. Dynkin, A.A. Yushkevich, G.M. Seitz, and A.L. Onishchik (American Mathematical Society, Providence, RI, International Press, Cambridge MA, 2000).

immediately shows that the replacement is correct. Let's check the replacement for three and four operator nested commutators

$$
\begin{aligned}
[C,[B,A]] &= CBA - CAB - BAC + ABC \\
&\stackrel{?}{=} \tfrac{1}{3}\left([C,[B,A]] - [C,[A,B]] - [B,[A,C]] + [A,[B,C]]\right) \\
&= \tfrac{1}{3}\left(3[C,[B,A]] - [C,[B,A]] - [B,[A,C]] - [A,[C,B]]\right) \\
&= [C,[B,A]]
\end{aligned}
\tag{J.29}
$$

Where we have used the Jacobi identity

$$
[C,[B,A]] + [B,[A,C]] + [A,[C,B]] = 0\,,
\tag{J.30}
$$

which can easily be proved by writing out all of the operator terms.

For four operators, we have

$$
\begin{aligned}
[D,[C,[B,A]]] &= DCBA - DCAB - DBAC + DABC - CBAD + CABD + BACD - ABCD \\
&\stackrel{?}{=} \tfrac{1}{4}\big([D,[C,[B,A]]] - [D,[C,[A,B]]] - [D,[B,[A,C]]] + [D,[A,[B,C]]] \\
&\qquad - [C,[B,[A,D]]] + [C,[A,[B,D]]] + [B,[A,[C,D]]] - [A,[B,[C,D]]]\big)
\end{aligned}
\tag{J.31}
$$

We can use our result for 3 operators to replace the first 4 nested commutators by $3[D,[C,[B,A]]]$, and the Jacobi identity changes the following two into one,

$$
[D,[C,[B,A]]] \stackrel{?}{=} \tfrac{1}{4}\big(3[D,[C,[B,A]]] + [C,[D,[B,A]]] + [B,[A,[C,D]]] - [A,[B,[C,D]]]\big)
\tag{J.32}
$$

The last two nested commutators can be simplified by again using the Jacobi identity for the three operators, A, B, and $[C,D]$, which says

$$
[B,[A,[C,D]]] - [A,[B,[C,D]]] = [[C,D],[A,B]]
\tag{J.33}
$$

we then have

$$
\begin{aligned}
[C,[D,[B,A]]] + [B,[A,[C,D]]] - [A,[B,[C,D]]] &= [C,[D,[B,A]]] - [[C,D],[B,A]] \\
&= CD[B,A] - C[B,A]D - D[B,A]C + [B,A]DC \\
&\quad - CD[B,A] + DC[B,A] + [B,A]CD - [B,A]DC \\
&= -C[B,A]D - D[B,A]C + DC[B,A] + [B,A]CD = [D,[C,[B,A]]]
\end{aligned}
\tag{J.34}
$$

so that again, we verify Dynkin's replacement. The full proof will be sketched below, but since fourth order is the farthest we will actually calculate, the above explicit results are sufficient for that case.

J.6   Oteo Replacement

Oteo conjectured[4] and Wyrauch and Scholz proved[5] that for linear combinations of operator products of two operators A and B that form nested commutators, each product of operators can be replaced by the right-nested commutator with a coefficient. The coefficient is nonzero only if the product ends in BA, and in that case the coefficient is one over the number of A operators.

$$
\begin{aligned}
BAAB &\rightarrow 0 \\
ABBA &\rightarrow \tfrac{1}{2}[A,[B,[B,A]]] \\
ABABAAABA &\rightarrow \tfrac{1}{6}[A,[B,[A,[B,[A,[A,[A,[B,A]]]]]]]] \\
ABABAAAAB &\rightarrow 0
\end{aligned}
\tag{J.35}
$$

J.7   Sketch of the general proofs

We follow the discussion in Weyrauch and Scholz. They define the Dynkin replacement as the map Θ and the Oteo replacement as the map Φ,

$$
\begin{aligned}
\Theta\left(\sum a(X_{i_n}, X_{i_{n-1}}, ..., X_{i_1})X_{i_n}...X_{i_1}\right) &= \sum \frac{a(X_{i_n}, X_{i_{n-1}}, ..., X_{i_1})}{n}[X_{i_{n-1}}, [X_{i_{n-1}}, ...[X_{i_2}, X_{i_1}]...] \\
\Phi\left(\sum a(X_{i_n}, X_{i_{n-1}}, ..., X_{i_1})X_{i_n}...X_{i_1}\right) &= \sum \frac{a(X_{i_n}, X_{i_{n-1}}, ..., X_{i_1})}{N_A(X_{i_n}, X_{i_{n-1}}, ..., X_{i_1})}[X_{i_{n-1}}, [X_{i_{n-1}}, ...[B, A]...]
\end{aligned}
\tag{J.36}
$$

where the X_i for Θ can be any one of any number of different operators, but the X_i for Φ must be either A or B. Of course, the Baker-Campbell-Hausdorff expansion only contains A and B, so both replacements apply to it. The operator N_A counts the number of A operators, and as noted above, and shown explicitly in the definition of Φ, only the terms that end in BA contribute.

(a) The right-nested commutator can be written out in operator products as

$$
[X_n, [X_{n-1}, ...[X_2, X_1]...] = X_n X_{n-1}...X_2 X_1 + ...
\tag{J.37}
$$

where the other terms do not end in X_1.

(b) There is an expansion of any right-nested commutator $[X_n, [X_{n-1}, ...[X_2, X_1]...]$ in terms of a linear combination of right-nested commutators where any element X_k is the right most element in the expansion, i.e.

$$
[X_n, [X_{n-1}, ...[X_2, X_1]...] = \sum c(X_{j_n}, ..., X_{j_2}, X_k)[X_{j_n}, [X_{j_{n-1}}, ...[X_{j_2}, X_k]...]
\tag{J.38}
$$

and the sum is over all $j_2, ...j_n$ permutations of the integers $1, ..., k-1, k+1, ...n$. This can be shown by induction and repeated use of the Jacobi identity.

---

[4]J. Oteo, J. Math. Phys. **32** 419 (1991).

[5]M. Weyrauch and D. Scholz, Comp. Phys. Comm. **180**, 1558 (2009)

(c) we write out a general right-nested commutator as

$$[X_n, [X_{n-1}, ...[X_2, X_1]...] = \sum a(X_{i_n}, ..., X_{i_2}, X_{i_1})X_{i_n}...X_{i_2}X_{i_1} \qquad \text{(J.39)}$$

where the sum is over all $i_1, .., i_n$ permutations of the integers $1, ..., n$ Using (a) and (b) above,

$$\sum a(X_{i_n}, ..., X_{i_2}, X_{i_1})X_{i_n}...X_{i_2}X_{i_1} = \sum c(X_{j_n}, ..., X_{j_2}X_k,)[X_{j_n}, [X_{j_{n-1}}, ...[X_{j_2}, X_k]...]$$
$$= \sum c(X_{j_n}, ..., X_{j_2}, X_k)\left(X_{j_n}X_{j_{n-1}}...X_{j_2}X_k + ...\right)$$
$$\qquad \text{(J.40)}$$

This shows that $a(X_{j_n}, ..., X_{j_2}, X_k) = c(X_{j_n}, ..., X_{j_2}, X_k)$ and therefore

$$[X_n, [..., [X_2, X_1]...] = \sum a(X_{j_n}, X_{j_{n-1}}, ..., X_{j_2}, X_k)[X_{j_n}, [..., [X_{j_2}, X_k]...] \qquad \text{(J.41)}$$

These were the first steps used by Dynkin in his proof.

To prove that the mapping Θ operating on a right-nested commutator gives back the right-nested commutator we sum Eq. J.41 over all k values and divide by the total number of operators

$$\sum_k \frac{1}{n}[X_n, [..., [X_2, X_1]...] = \sum_k \sum \frac{a(X_{j_n}, X_{j_{n-1}}, ..., X_{j_2}, X_k)}{n}[X_{j_n}, [..., [X_{j_2}, X_k]...] \qquad \text{(J.42)}$$

the left side has no k, so the sum cancels the $\frac{1}{n}$ and gives back the nested commutator. The sum over k on the right completes the sum over all permutations of the operators so the right hand side is exactly Dynkin's replacement.

For the special case where the X_j are only A or only B, we can use (b) above to rewrite any nested commutator into a right-nested commutator where the rightmost element is A. We then sum Eq. J.41 over all k values that correspond to an A operator, and divide by the number of A operators

$$\sum_k \frac{1}{N_A(X_n...X_2A)}[X_n, [..., [X_2, A]...] = \sum_k \sum \frac{a(X_{j_n}, X_{j_{n-1}}, ..., X_{j_2}, A)}{N_A(X_{j_n}...X_{j_2}A)}[X_{j_n}, [..., [X_{j_2}, A]...] \qquad \text{(J.43)}$$

The left side is the original right-nested anticommutator with right-most operator A, and the right-side is exactly Oteo's replacement.

J.8 Application to the Baker-Campbell-Hausdorff expansion

Our straightforward Taylor series expansion Eq. J.1 gave

$$\begin{aligned} S &= (A+B)t + \tfrac{1}{2}(AB - BA)t^2 \\ &+ \tfrac{1}{12}\left(A^2B - 2ABA + BA^2 + AB^2 - 2BAB + B^2A\right)t^3 + ... \end{aligned} \qquad \text{(J.44)}$$

Making Dynkin's replacement everywhere except for the $(A+B)t$ term

$$
\begin{aligned}
S &= (A+B)t + \tfrac{1}{4}\left([A,B]-[B,A]\right)t^2 \\
&\quad + \tfrac{1}{36}\left([A,[A,B]]-2[A,[B,A]]+[B,[A,A]]+[A,[B,B]]-2[B,[A,B]]+[B,[B,A]]\right)t^3 + ... \\
&= (A+B)t + \tfrac{1}{2}[A,B]t^2 + \tfrac{1}{36}\left([A,[A,B]]+2[A,[A,B]]+2[B,[B,A]]+[B,[B,A]]\right)t^3 + ... \\
&= (A+B)t + \tfrac{1}{2}[A,B]t^2 + \tfrac{1}{12}\left([A,[A,B]]+[B,[B,A]]\right)t^3 + ... \qquad\qquad\text{(J.45)}
\end{aligned}
$$

The Oteo replacement gives many fewer terms

$$
S = (A+B)t - \tfrac{1}{2}[B,A]t^2 + \tfrac{1}{12}\left(-[A,[B,A]]+[B,[B,A]]\right)t^3 + ... \qquad\text{(J.46)}
$$

Changing two of the $[B,A]$ to $-[A,B]$ gives the previous result.

We can use Dynkin's replacement in the full Taylor series expansion, here written explicitly and dropping terms which have commutators like $[A,A]$ or $[B,B]$,

$$
S_n = \frac{1}{n}\sum_{p=0}^{n-1} P_{p,n-p}(A,B) \qquad\qquad\text{(J.47)}
$$

where

$$
\begin{aligned}
P_{p,q}(A,B) &= -\sum_{\substack{j>0,p_j>0 \\ \{p_k+q_k>0\} \\ p_1+...+p_j=p \\ q_1+...+q_{j-1}=q-1}} \frac{(-1)^j}{j}\,\frac{([A,\,)^{p_1}}{p_1!}\frac{([B,\,)^{q_1}}{q_1!}\cdots\frac{([A,\,)^{p_{j-1}}}{p_{j-1}!}\frac{([B,\,)^{q_{j-1}}}{q_{j-1}!}\frac{([A,\,)^{p_j}}{p_j!}\,B\Bigg] \\
&\quad -\sum_{\substack{j>0 \\ \{p_k+q_k>0\} \\ p_1+...+p_{j-1}=p-1 \\ q_1+...+q_{j-1}=q}} \frac{(-1)^j}{j}\,\frac{([A,\,)^{p_1}}{p_1!}\frac{([B,\,)^{q_1}}{q_1!}\cdots\frac{([A,\,)^{p_{j-1}}}{p_{j-1}!}\frac{([B,\,)^{q_{j-1}}}{q_{j-1}!}\,A\Bigg] \qquad\text{(J.48)}
\end{aligned}
$$

For Oteo's replacement,

$$
S_n = \sum_{p=0}^{n-1} P^O_{p,n-p}(A,B) \qquad\qquad\text{(J.49)}
$$

where

$$
P^O_{p,q}(A,B) = -\sum_{\substack{j>0 \\ \{p_k+q_k>0\} \\ p_1+...+p_{j-1}=p-1 \\ q_1+...+q_{j-1}=q}} \frac{1}{p}\frac{(-1)^j}{j}\,\frac{([A,\,)^{p_1}}{p_1!}\frac{([B,\,)^{q_1}}{q_1!}\cdots\frac{([A,\,)^{p_{j-1}}}{p_{j-1}!}\frac{([B,\,)^{q_{j-1}}}{q_{j-1}!}\,A\Bigg] \quad\text{(J.50)}
$$

Let's use Oteo's replacement to go to fourth order in the Baker-Campbell-Hausdorff expansion. It is somewhat easier to just expand $x = e^{At}e^{Bt}-1$ and keep only the terms that

can contribute after the Oteo replacement, i.e. besides the $(A+B)t$ term, terms that have their rightmost element A and an adjacent B, we have

$$
\begin{aligned}
x &\rightarrow (A+B)t \\
-\tfrac{1}{2}x^2 &\rightarrow -\tfrac{1}{2}\left[Bt + \tfrac{1}{2}(2AB+B^2)t^2 + \tfrac{1}{6}(3A^2B + 3AB^2 + B^3)t^3\right] At \\
&\rightarrow -\tfrac{t^2}{2}BA - \tfrac{t^3}{4}(2ABA + B^2A) - \tfrac{t^4}{12}(3A^2BA + 3AB^2A + B^3A) \\
\tfrac{1}{3}x^3 &\rightarrow \tfrac{1}{3}\left[(A+B)t + \tfrac{1}{2}(A^2 + 2AB + B^2)t^2\right]\left[Bt + \tfrac{1}{2}(2AB+B^2)t^2\right]At \\
&\rightarrow \tfrac{t^3}{3}(ABA + B^2A) + \tfrac{t^4}{6}(3A^2BA + 3AB^2A + 2B^3A + 2BABA) \\
-\tfrac{1}{4}x^4 &\rightarrow -\tfrac{1}{4}(A+B)t(A+B)tBtAt \\
&\rightarrow -\tfrac{t^4}{4}(A^2BA + AB^2A + BABA + B^3A) \tag{J.51}
\end{aligned}
$$

adding these we get

$$
x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \tfrac{1}{4}x^4 \rightarrow (A+B)t - \tfrac{1}{2}BAt^2 - \tfrac{1}{6}ABAt^3 + \tfrac{1}{12}B^2At^3 + \tfrac{1}{12}BABAt^4 \tag{J.52}
$$

and making Oteo's replacement

$$
\begin{aligned}
S(t) &= (A+B)t - \tfrac{t^2}{2}[B,A] + \tfrac{t^3}{12}\left(-[A,[B,A]] + [B,[B,A]]\right) + \tfrac{t^4}{24}[B,[A,[B,A]]] + ... \\
&= (A+B)t + \tfrac{t^2}{2}[A,B] + \tfrac{t^3}{12}\left([A,[A,B]] + [B,[B,A]]\right) + \tfrac{t^4}{24}[B,[A,[B,A]]] + ...
\end{aligned} \tag{J.53}
$$

Appendix K

Additional Problems

K.1   Delta Function Problems

1.   a. Show that $\delta(ax) = |a^{-1}|\delta(x)$.

   b. Show that $\delta[f(x)] = \sum_j \left| \frac{df(x_j)}{dx_j} \right| \delta(x - x_j)$ where x_j are the zeroes of $f(x)$, i.e. $f(x_j) = 0$.

2. The three dimensional delta function $\delta^3(\boldsymbol{r} - \boldsymbol{r}')$ is zero except at $\boldsymbol{r} = \boldsymbol{r}'$ and integrates to unity,

$$\int d^3 r \delta^3(\boldsymbol{r} - \boldsymbol{r}') = 1 . \tag{K.1}$$

   Write the three dimensional delta function as a product of three one dimensional delta functions, one for each coordinate, multiplied by a factor, and find the factor, for

   a. Cartesian coordinates x, y, z and the corresponding primed coordinates.

   b. Spherical coordinates r, θ, ϕ and the corresponding primed coordinates, where $x = r \sin\theta \cos\phi$, $y = r \sin\theta \sin\phi$, $z = r \cos\theta$.

   c. In spherical coordinates using r, $\cos\theta$ and the corresponding primed coordinates.

   d. In cylindrical coordinates using ρ, z, and ϕ and the corresponding primed coordinates, where $x = \rho \cos\phi$, $y = \rho \sin\phi$.

3.   a. Calculate the integral

$$I(T) = \int_{-T/2}^{T/2} dt e^{ixt} \tag{K.2}$$

   and verify that for $x \neq 0$ its limit for $T \to \infty$ is ero, while for $x = 0$ its limit for $T \to \infty$ is infinite.

b. The result of part a suggests that $I(T)$ is proportional to a delta function when $T \to \infty$. Integrate it over x and find the proportionality constant.

c. Occasionally you may find that a calculation gives a result that is a square of a delta function which is, of course, nonsense. This happens when you improperly interchange limits. For example you might have a calculation where you need the limit of $T \to \infty$ of the combination $T^{-1}|I(T)|^2$ Naively taking the limits above, you would get a delta function squared from the $|I(t)|^2$. Show that by taking the limit correctly you get a result that is equivalent to $\delta^2(x) \to \frac{T}{2\pi}\delta(x)$.

4. Any function that has a limit that fulfills the definition of a delta function can be used to deduce the properties of a delta function. The derivative of a delta function can be defined as the limit of the derivative of such a function. One such function is the pulse of width w around $x = 0$,

$$
\begin{aligned}
p(x, w) &= w^{-1}\left[\theta(x + w/2) - \theta(x - w/2)\right] \\
&= \begin{cases} 0 & |x| > \frac{w}{2} \\ w^{-1} & |x| < \frac{w}{2} \end{cases} ,
\end{aligned} \tag{K.3}
$$

where the function $\theta(x)$ is 1 for positive x and 0 for negative x.

a. Show that

$$
\lim_{w\to 0} p(x, w) = \delta(x) . \tag{K.4}
$$

b. Verify that

$$
\frac{d}{dx}p(x, w) = w^{-1}\left[\delta(x + w/2) - \delta(x - w/2)\right] \tag{K.5}
$$

and use this to show that

$$
\int_{-\infty}^{\infty} dx f(x)\frac{d\delta(x)}{dx} = -\frac{df(x)}{dx}\bigg|_{x=0} \tag{K.6}
$$

c. Show that you get the same result by integrating $\int_{-\infty}^{\infty} dx f(x)\frac{d\delta(x)}{dx}$ by parts.

Appendix L

Assessment

L.1 Quantum Mechanics Preparation

1. If $\psi(\boldsymbol{r}, t)$ is the wave function for a particle at $\boldsymbol{r}$ at time t, What is the physical meaning, if any, of

   a. $|\psi(\boldsymbol{r}, t)|^2$

   b. $\frac{\hbar}{2mi} \left[\psi^*(\boldsymbol{r}, t) \boldsymbol{\nabla} \psi(\boldsymbol{r}, t) - \psi(\boldsymbol{r}, t) \boldsymbol{\nabla} \psi^*(\boldsymbol{r}, t) \right]$

2. To what physical system, if any, would a model of one particle moving in one dimension apply? An example is the solution to the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + v(x) \right] \psi(x, t) = -\frac{\hbar}{i} \frac{\partial}{\partial t} \psi(x, t) . \tag{L.1}$$

3. Given half an hour each, which of the following, if any, could you calculate analytically without looking at *any* reference material?

   a. The energy eigenfunctions and the energy eigenvalues of the one-dimensional infinite square well.

   b. The energy eigenfunctions and the energy eigenvalues of the one-dimensional harmonic oscillator.

   c. The reflection and transmission probabilities for a particle of mass m moving in one-dimension with energy E incident from negative x on a potential step

$$v(x) = \begin{cases} 0 & x < 0 \\ V_0 & x > 0 \end{cases} \tag{L.2}$$

4. If two states have degenerate eigenvalues, what does that mean?

541

5. An oxygen atom consists of 8 electrons and the oxygen nucleus all interacting via electromagnetic forces. Is a system consisting of just one oxygen atom described by just one wave function for the entire oxygen atom or are there different wave functions for each particle?

6. If angular momentum is conserved, the Hamiltonian is invariant under what operations?

7. If you measure the value of a property (like position, momentum, energy, angular momentum, etc.) corresponding to an operator O, what are the possible values of the result?

8. How would you measure the expectation value of an operator corresponding to an observable?

9. What can hermitian operators represent in quantum mechanics?

10. What can unitary operators represent in quantum mechanics?

L.2   Mathematical Preparation

11. What is the sum of the series $1 + x + x^2 + x^3 + ...$?

12. What are the first three terms in the power series expansion in x of $(1 + x)^y$?

13. Can you numerically integrate ordinary differential equations using, for example, Euler's method? For example, given $\frac{d^2}{dx^2} y(x) = xy(x)$, and boundary conditions like $y(0) = 1$, $y'(0) = 0$, can you describe a method for approximate numerical calculation of $y(x)$ for other values of x?

14. Do you know what Fourier transforms and series are?

15. Can you derive the Fourier transform pairs

$$\tilde{f}(k) = \int_{-\infty}^{\infty} dx f(x) e^{-ikx}, \quad f(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \tilde{f}(k) e^{ikx}, \tag{L.3}$$

without looking at any reference material.

16. Do you know what spherical harmonics are?

17. Do you know how to find the eigenvalues and eigenvectors of a matrix?

18. What is the largest size general matrix for which you could find the eigenvalues and eigenvectors analytically?

19.   a. Explain how you could use matrix methods to solve numerically for the energy eigenvalues of the one-dimensional Hamiltonian, $H = \frac{p_x^2}{2m} + \lambda|x|$.

      b. How could you use these matrix methods to plot the eigenfunctions?

20. Which of the following are hermitian matrices?

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & i & 1 \\ i & 0 & 2i \\ 1 & 2i & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & i & 1 \\ -i & 0 & 2i \\ 1 & -2i & 3 \end{pmatrix}.$$

21. Which of the following are unitary matrices?

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & i & 1 \\ -i & 0 & 2i \\ 1 & -2i & 3 \end{pmatrix}. \qquad \text{(L.4)}$$

22. Which of the following are analytic functions of $z = x + iy$ inside the unit circle $|z| < 1$,

    a. $f(z) = x = \text{Re } z$,

    b. $f(z) = \frac{1}{\pi^2 - z^2}$,

    c. $f(z) = z^* = x - iy$.

23. Do you know how to calculate $\int_0^\infty dx \frac{\cos(\alpha x)}{\beta^2 + x^2}$ for real α and β using contour integration?

24. What is the general solution of the differential equation $\frac{d^2}{dx^2} y(x) + k^2 y(x) = 0$.

L.3   Classical Mechanics Preparation

25. Given a Lagrangian $\mathcal{L}(x, \dot{x})$, what is the canonical momentum conjugate to x. For a nonrelativistic particle is it always $m\dot{x}$?

26. Do you know how to Lorentz transform physical quantities?

L.4   Classical Electromagnetism Preparation

27. If you use SI units, what are the units of μ_0, ϵ_0, the electric field $\boldsymbol{E}$, the displacement $\boldsymbol{D}$, the magnetic induction $\boldsymbol{B}$, and the magnetic field $\boldsymbol{H}$ in terms of kilograms, meters, seconds and amperes? What are a Volt and an Ohm in terms of kg, m, s, A?

If you use Gaussian or Heaviside Lorentz units, what are the units of the speed of light c, the electric charge e, the electric field $\boldsymbol{E}$, the displacement $\boldsymbol{D}$, the magnetic induction $\boldsymbol{B}$, and the magnetic field $\boldsymbol{H}$ in terms of grams, centimeters and seconds? What are the units of resistance?

28. If you have a Hamiltonian for a particle of charge q, whose momentum and position are $\boldsymbol{p}_1$ and $\boldsymbol{r}_1$, $H(\boldsymbol{p}_1, \boldsymbol{r}_1)$, do you know how to calculate the Hamiltonian when an external electric field $\boldsymbol{E}(\boldsymbol{r}, t)$ and magnetic induction $\boldsymbol{B}(\boldsymbol{r}, t)$ are turned on?

L.5 Physical Magnitudes

29. Within an order of magnitude or two, what is the size (or radius) of

   a. A hydrogen atom.

   b. A uranium atom.

   c. An atomic nucleus.

   d. The wavelength of visible light.

30. Within an order of magnitude or two, what is the ratio of the maximum electric field of a 1 milliwatt laser pointer focused to a spot of area 1 millimeter^2, compared to the average magnitude electric field felt by an electron in a hydrogen atom?

Bibliography

[1] G. Baym, *Lectures on Quantum Mechanics*, 1st ed. (Addison-Wesley, New York, 1990).

[2] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (John Wiley and Sons, New York, 1999).

[3] E. U. Condon and Q. W. Shortley, *Theory of atomic spectra* (Cambridge University Press, Cambridge, 1935).

[4] E. Clementi and C. Roetti, At. Data Nucl. Data Tables **14**, 177 (1974).

545

Index