

Quantum Mechanics

Thomas DeGrand

April 30, 2025

Contents

1	Introduction	5
2	Quantum mechanics in the language of Hilbert space	11
3	Time dependence in quantum mechanics	47
4	Propagators and path integrals	73
5	Density matrices	89
6	Wave mechanics	103
7	Angular momentum	153
8	Identical particles	191
9	Time independent perturbation theory	201
10	Variational methods	227
11	Time dependent perturbation theory	237
12	Electromagnetic interactions in semiclassical approximation	255

13 Scattering	287
14 Classical waves – quantum mechanical particles	341
15 Atoms and molecules	371

Chapter 1

Introduction

The subject of most of this book is the quantum mechanics of systems with a small number of degrees of freedom. The book is a mix of descriptions of quantum mechanics itself, of the general properties of systems described by quantum mechanics, and of techniques for describing their behavior. The examples illustrating quantum mechanical properties are often presented as specific physical systems. They will be drawn from many areas of physics, just to illustrate the fact that many apparently different physical systems naturally behave quite similarly.

Ehrenfest famously said that “Physics is simple but subtle,” but he did not mean that it was easy. You cannot learn quantum mechanics out of one book, nor out of these notes. You will have to compare the explanations from many sources and integrate them together into your own whole. You should not be surprised to find that any two books are likely to present completely contradictory explanations for the same topic, and that both explanations are at least partially true.

The quantum mechanical description of nature is fundamentally different from a classical description, in that it involves probabilistic statements. The usual causal story of classical mechanics is that in specifying a set of initial conditions, one completely specifies the evolution of the system for all time. That is not possible in quantum mechanics, simply because it is not possible to completely specify all the initial conditions. For example, the uncertainty principle $\Delta p \Delta x \geq \hbar$ forbids us from simultaneously knowing the coordinates and momentum of a particle at any time. However, the evolution of the probability itself is causal, and is encoded in the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \psi(t)}{\partial t} = \hat{H} \psi(t). \quad (1.1)$$

Once we specify the wave function (probability amplitude) $\psi(t)$ at some time t_0 , we know it for all later times. This is of course difficult for us, macroscopic beings that we are, to deal with.

Once we have gotten over our classical discomfort with the quantum world, we notice several striking features which recur again and again:

- The wide applicability of the same ideas, or the same physical systems, to many areas of physics.
- The presence of symmetries and their consequences for dynamics.

I like to take advantage of the first of these features and think about applications of quantum mechanics in terms of a few paradigmatic systems, which are approximations to nearly all physical systems you might encounter. It is worthwhile to understand these systems completely and in many different ways. The two most important such systems are

- The simple harmonic oscillator. Most weakly coupled systems behave like a set of coupled harmonic oscillators. This is easiest to visualize classically, for a system of particles interacting with a potential. If the system has some equilibrium structure, and if the energy of the system is small, then the particles will be found close to their equilibrium locations. This means that the potential energy is close to its minimum. We can do a Taylor expansion about the minimum, and then

$$V(x) \simeq V(x_0) + \frac{1}{2}V''(x_0)(x - x_0)^2 + \dots \quad (1.2)$$

which is the potential of an oscillator. The same situation will happen for a quantum mechanical system. An obvious example would be the motion of nuclear degrees of freedom in a molecule, the origin of the “vibrational spectrum” of molecular excitations, typically seen in the infrared. But there are many other less-obvious examples.

Classical systems described by a set of equations of motion

$$\frac{d^2 y_i}{dt^2} = -\omega_i^2 y_i \quad (1.3)$$

can be thought of as collections of harmonic oscillators. Classical wave systems have such equations of motion. Their quantum analogues are also oscillators, and so their quantum descriptions will involve oscillator-like energies and degrees of freedom. Do you recall that the quantum mechanical energy spectrum for an oscillator is a set of equally spaced levels, $E = n\epsilon$ (up to an overall constant) where ϵ is an energy scale, $\epsilon = \hbar\omega_i$ in Eq. 1.3, for example, and n is an integer $0, 1, 2, \dots$? Do you also recall the story of the photon, that the quantum electromagnetic field is labeled by an integer, the number of photons in a particular allowed state? This integer is the n of the harmonic oscillator. We can take the analogy between “counting” and oscillator states still further. Imagine that we had a system which seemed to have nothing to do with an oscillator, like a hydrogen atom. It has an energy spectrum $E = \epsilon_i$, where i labels the quantum number of the state. Now imagine that we have a whole collection of hydrogen atoms, and imagine also that these atoms do not interact. The energy of the collection is $E = \sum n_i \epsilon_i$, where again n_i , the number of particles in the collection with

energy ϵ_i , is an integer. If we forget about where the numbers ϵ_i came from, we will be making an oscillator-like description of the collection.

- The two state system. In an undergraduate textbook, the paradigm for this system is the spin-1/2 particle in an external magnetic field. The electron is either spin-up or spin-down, and energy eigenstates are states where the spin aligns along the magnetic field. But this description is often used in situations not involving spin, typically as a way of approximating some complicated dynamics. For example, it often happens that a probe can excite a system from one state to a single different state, or that the probe can excite the state to many different ones, but that there is only one strong transition, and all other ones are small. Then it makes sense to replace the more complicated system by one which contains only two states, and assume that the system is confined only to those states. These states could be different atomic levels (perhaps interacting with a radiation field), different kinds of elementary particles which can exchange roles (neutrino oscillations are a recent example) or could even be systems which are naively very different (cold gases, which can condense as either molecules or as individual atoms, for another recent example). And of course the “qubit” of quantum computing jargon is a two state system.

There are other paradigmatic systems. Often, we study collections of particles which are weakly interacting. A good approximate starting point for these systems is to ignore the interactions, giving a set of free particles as zeroth order states. This is often how we describe scattering.

Hydrogen, despite its appearance in every undergraduate quantum mechanics course, is not so paradigmatic. Hydrogen is not even a very typical atom and its “ $1/n^2$ ” Rydberg spectrum is unique, a consequence of a particular special symmetry. It is useful to know about it, though, because it is simple enough that we can solve for its properties completely.

The second feature – physics arising from symmetry – appears again and again in these notes. If a physical system has an underlying symmetry, seemingly different processes can be related. An example is the relative intensity of spectral lines of families of states with the same total angular momentum and different z-components: the pattern of spectral lines for an atom decaying from an excited state, placed in an external magnetic field. Often, without knowing anything about a dynamical system other than the symmetries it obeys, one can determine what processes are forbidden. (People speak of “selection rules” as a shorthand for these consequences; “forbidden” means that the rate is zero.) Such predictions are more

reliable than ones of the absolute rates for processes actually occur. Knowing how to use symmetry to make predictions is at the bottom of most of the technical parts of this course.

There is a deep connection between symmetry and conservation laws. We usually describe this in quantum mechanics by saying that the Hamiltonian possess a symmetry and the presence of a symmetry of the Hamiltonian directly leads to a conservation law. And most of the time we say “Here is a Hamiltonian, what are the consequences?” This is a perfectly good question, except that it may be putting the cart before the horse. Where do Hamiltonians come from, anyway? I think most people would say that it is the symmetries which are at the bottom; a Hamiltonian is a construct which builds the symmetries into the dynamics from the beginning. And one should use dynamical variables that naturally encode the symmetry, or transform in some simple way under the symmetry. Figuring out an appropriate set of variables for dealing with perhaps the most important symmetry transformation – rotational invariance – will be a big part of the course. It would not be an exaggeration to say that the job of the physicist is to recognize some symmetry, write down a mathematical system which encodes it, and then look for the consequence of the symmetry in new places.

I’ve wandered away from remarks about the content of the course, to try to disturb you. Let me say one more disturbing thing, before we get to work: Typically, physics at different energy scales decouples. To study atoms, one generally needs to know little about nuclei other than their charge and mass – even though nuclei themselves are complicated bound states of protons and neutrons, which are themselves complicated bound states of quarks and gluons. Atoms themselves are complicated, and the interactions of atoms – atom - atom scattering – is even more complicated. However, to describe the properties of condensed atom gases, all that is needed from the complicated scattering of atoms is a single number, the scattering length. Perhaps this is not so strange. After all, atomic physics happens over length scales of Angstroms and typical energies are electron volts. Nuclei have sizes of a few fermis, 10^{-5} times smaller, and typical excitation energies of KeV or MeV. So nuclear levels can’t be excited in ordinary “atomic - scale” processes. There is a similar hierarchy of energy scales in condensed atom gases, where the energy scale associated with their collective behavior is much much smaller than a fraction of an electron volt.

But perhaps this explanation is a little too simple: when we learn about perturbation theory, we will discover that arbitrarily high energy states can exist as so-called “virtual states,” and that physical systems can fluctuate in and out of these virtual states during their evolution. The states are present in the formulas. Nevertheless, most of what is happening at high energies and short distances usually has little effect on low energy, long

distance phenomena. When we are doing a typical calculation of the spectrum of an atom, it is still the case all that we need to know about the nucleus is its charge. There is an interesting story here, apparently! And a practical one – you have a system whose dynamics occurs at some far away scale from where you are working. What do you need to know about it? How could you calculate what you need to know (and nothing else)?

These notes are based on my graduate quantum mechanics course at the University of Colorado. They started as a set of real class notes: what I taught, no more. They have grown as I found more applications of quantum mechanics that I thought could potentially be interesting to a reader. So there is too much material to lecture about.

The first semester of this course used the text by Sakurai as background. Some attempt was made to make my notation coincide with its conventions. But I am not trying to provide an outline of any specific text. Like anyone who teaches a quantum mechanics course, I have my own prejudices about what is important and how it should be presented. These notes had their genesis from many sources. When I was learning quantum mechanics, I was particularly influenced by the books by Schiff and Baym, for practical applications, and by the text of Dirac for its poetry. I also want to specifically acknowledge two of my teachers: at the University of Tennessee, Edward Harris, and at MIT, John Negele. Finally, I owe a considerable debt of gratitude to Joseph Seele, a student in my class in 2003-2004, who made the first electronic version of these notes. If he hadn't typed in all the equations, this book would not exist.

Chapter 2

Quantum mechanics in the language of Hilbert space

Wave mechanics

Let's begin by recalling what we were taught about wave mechanics in our introductory class. All the properties of a system of N particles are contained in a wave function $\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N, t)$. The probability that the system is found between some differential N -volume element \vec{x} and $\vec{x} + d\vec{x}$ is proportional to the modulus squared of the amplitude of the wave function.

$$Prob(\vec{x}) \propto |\Psi(\vec{x}, t)|^2 d^3x \quad (2.1)$$

In practice, one usually defines $\Psi(\vec{x}, t)$ such that its integrand over the entire space in question is unity,

$$\int d^3x |\Psi(\vec{x}, t)|^2 = 1. \quad (2.2)$$

With this normalization convention, the definition that the square modulus is proportional to the probability can be replaced by the statement that it is equal to the probability. To have a probabilistic interpretation, the un-normalized $|\Psi(x, t)|^2$ should be a bounded function, so that the integral is well defined. This usually means that the wave function must die away to zero at spatial infinity.

In quantum mechanics, all information about the state is contained in $\Psi(x, t)$. Dynamical variables are replaced by operators, quantities which act on the state. For a quantum analog of a non-relativistic particle in a potential, the dynamical variables at the heart of any physical description are the coordinate x and the momentum p . The operators corresponding to x and p obey a fundamental commutation relation

$$\hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar. \quad (2.3)$$

One can pick a basis where either p or x is “diagonal,” meaning that the operator is just the variable itself. In the coordinate-diagonal basis implicitly taken at the start of this section, we satisfy the commutation relation with the identifications

$$\hat{x} = x \quad \hat{p} = \frac{\hbar}{i} \frac{\partial}{\partial x}. \quad (2.4)$$

Quantum mechanics is probabilistic, so we can only talk about statistical averages of observables. The expectation value or average value of an observable $O(p, x)$ represented by an operator \hat{O} is

$$\langle \hat{O} \rangle = \int d^3x \psi^*(x, t) \hat{O} \psi(x, t). \quad (2.5)$$

$\langle \hat{O} \rangle$ represents the average of a set of measurements on independent systems of the observable $O(p, x)$.

While the motion of particles can only be described probabilistically, the evolution of the probability itself evolves causally, through the Schrödinger equation (or ‘time-dependent Schrödinger equation’)

$$\hat{H}\Psi(x, t) = i\hbar \frac{\partial \Psi(x, t)}{\partial t}. \quad (2.6)$$

The quantity \hat{H} is the Hamiltonian operator.

If \hat{H} is not an explicit function of time, the solution for time evolution of the wave function can be written as a superposition of energy eigenfunctions,

$$\psi(x, t) = \sum_n c_n \psi_n(x) e^{-\frac{iE_n t}{\hbar}} \quad (2.7)$$

where the energies E_n are given by solving the eigenvalue equation (called the ‘time-independent Schrödinger equation,’)

$$\hat{H}\psi_n(x) = E_n \psi_n(x). \quad (2.8)$$

In the case of a system described by a classical Hamiltonian $H(x, p)$ the quantum Hamiltonian operator is constructed by replacing the classical variables by their the operator expressions, so that the Hamiltonian for a single non-relativistic particle in an external potential is

$$\begin{aligned} H &= \frac{p^2}{2m} + V(x) \\ &= -\frac{\hbar^2}{2m} \nabla^2 + V(x). \end{aligned} \quad (2.9)$$

In this case the time-independent Schrödinger equation is a partial differential eigenvalue equation. Hence the title of Schrödinger’s papers – ‘Quantization as an eigenvalue problem.’

However, this is far from the whole story. Where did the transcription of observable to operator come from? Many of the interesting problems in quantum mechanics do not have classical analogues. We need a more general formalism.

As an example of a physical system which is difficult to describe with wave mechanics, consider the Stern-Gerlach experiment. (I am borrowing this story from Volume III of the

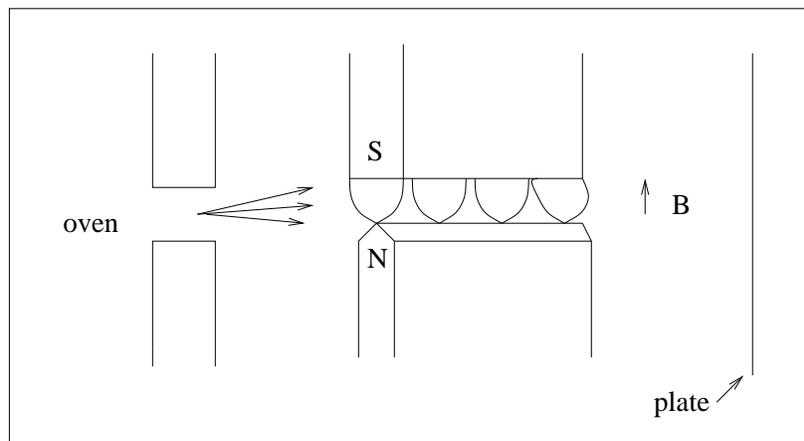


Figure 2.1: A Stern-Gerlach apparatus.

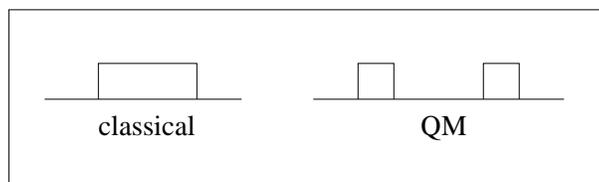


Figure 2.2: A cartoon of the spots in a detector, for classical or quantum spins in a Stern-gerlach apparatus.

Feynman lectures, so you should really read it there.) The experimental apparatus is shown schematically in the figure. Atoms are heated in the oven and emerge collimated in a beam; they pass along the axis of a magnet and are detected by striking a plate. Imagine that we use an atom with 1 valence electron so that to a good approximation the magnetic moment of the atom is that of the electron. Then the potential energy is $U = -\vec{\mu} \cdot \vec{B}$ and if the magnetic field is inhomogeneous, there is a force acting on the atom along the direction of the field derivative $F_z = -\partial U / \partial z = \mu_z \frac{\partial B_z}{\partial z}$. We know that the magnetic moment is connected to the spin, $\mu_z = \frac{e\hbar}{mc} S_z$, and quantum mechanically, the spin is quantized, $S_z = \pm \frac{1}{2} \hbar$. We can see what we expect both classically and quantum mechanically in Fig. 2.2, labeling the intensity of the spots of atoms deposited on the plate. Classically we would expect a continuous distribution, but in real life we find that we get only two different values (this is a signature of the quantization of spin, and was regarded as such as soon as the experiments were first done). Apparently the magnetic field has separated the beam into two states, one of which has its spin pointed along the direction of the magnetic field, and the other state with an

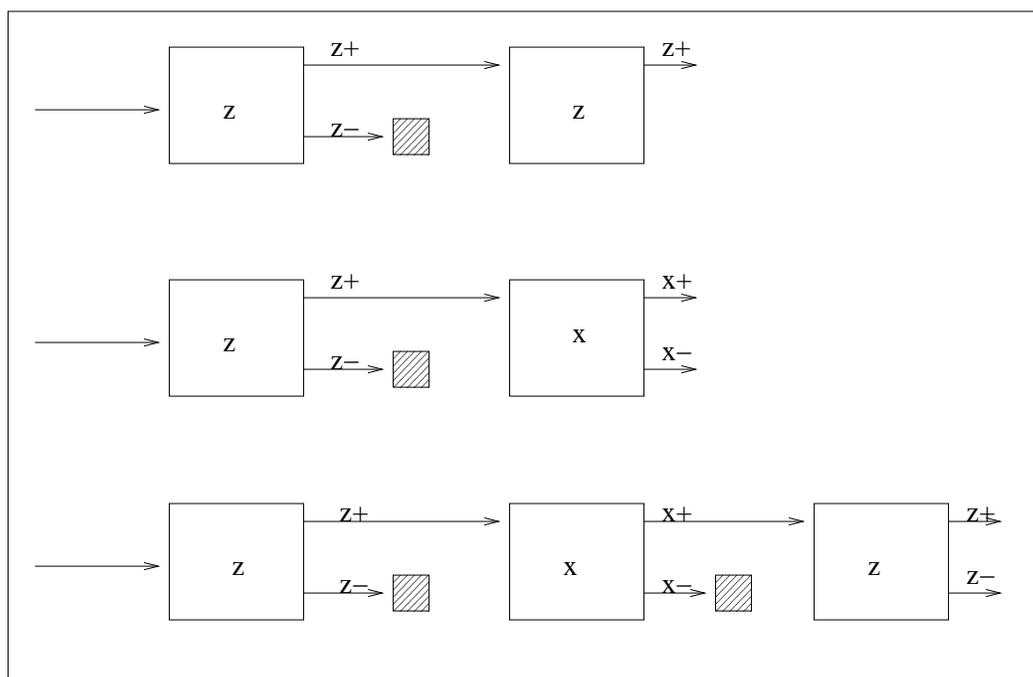


Figure 2.3: A series of Stern-Gerlach apparatuses, with and without beam stops.

opposite orientation.

Now imagine sequential Stern-Gerlach apparatuses. We label the orientation of the magnetic fields with the letter in the box, and we either stop one of the emerging beams (the black box) or let them continue on into the next magnet. If we allow the $+1/2$ spin component of a beam emerging from a magnetic field pointing in the z direction pass through a second z magnet, only a $+1/2$ spin component reappears. However, if we rotate the second magnet, two spin states emerge, aligned and anti-aligned (presumably, since the spots are aligned with its direction) with the magnetic field in the second apparatus. See Fig. 2.3. If we now consider the third picture, we produce a $+z$ beam, split it into a $+x$ and $-x$ beams, and then convert the $+x$ beam into a mixture of $+z$ and $-z$ beams. Apparently from the experiment the S_z^+ state is “like” a superposition of the S_x^+ and S_x^- states and the S_x^+ state is “like” a superposition of the S_z^+ and S_z^- states.

Next consider in Fig. 2.4 an “improved” Stern-Gerlach device, which actually appears to do nothing: In this device the beams are recombined at the end of the apparatus. In cases (a) and (b), we again insert beam stops to remove spin components of the beam. However, in case (c) we let both beams propagate through the middle apparatus. See Fig. 2.5. We

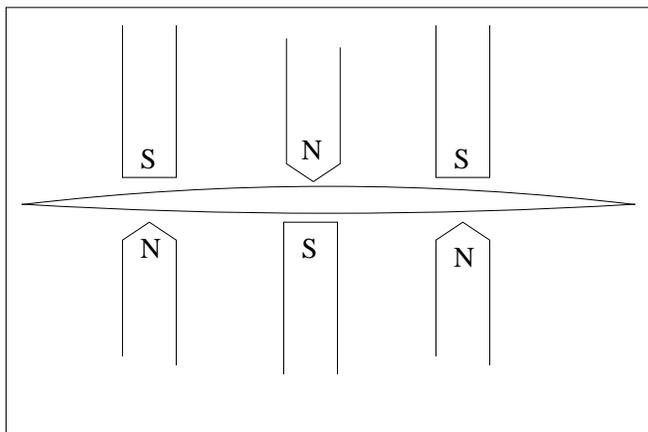


Figure 2.4: An “improved” Stern-Gerlach apparatus.

discover that only the $+z$ state is present in the rightmost magnetic field!

Quantum mechanics and Hilbert space

We now begin a more formal discussion of quantum mechanics., which will encode the physics we saw in our example. We start by defining (rather too casually) a Hilbert space: a complex linear vector space whose dimensionality could be finite or infinite. We postulate that all quantum mechanical states are represented by a ray in a Hilbert space. Rays are presented in print as “ket vectors” (labeled $|a\rangle$). Corresponding to, but not identical to, the ket vectors are the “bra vectors” $\langle a|$ which lie in a dual Hilbert space.

Vectors in a Hilbert space obey the following relations: (We quote them without proof, but the proofs can easily be obtained from the axioms of linear algebra and vector spaces.)

Rays obey commutativity

$$|a\rangle + |b\rangle = |b\rangle + |a\rangle, \quad (2.10)$$

and associativity properties:

$$(|a\rangle + |b\rangle) + |c\rangle = |a\rangle + (|b\rangle + |c\rangle). \quad (2.11)$$

There is a null vector $|0\rangle$ such that

$$|0\rangle + |a\rangle = |a\rangle. \quad (2.12)$$

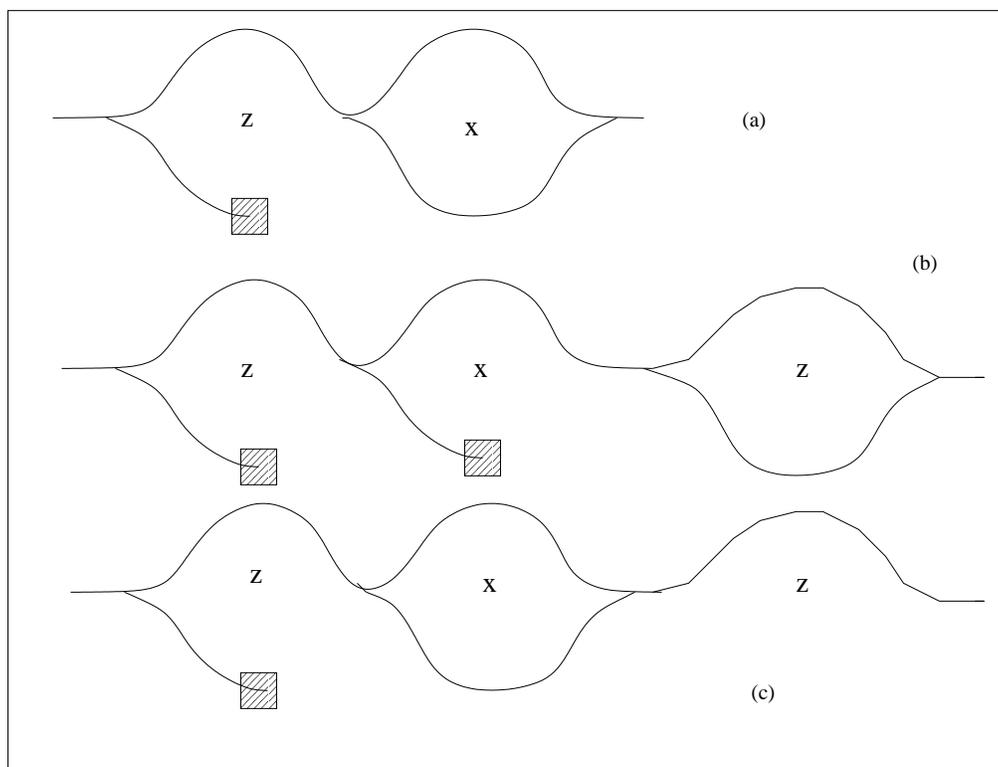


Figure 2.5: Particle trajectories through the “improved” Stern-Gerlach apparatus.

Each element has an inverse

$$|a\rangle + |-a\rangle = |0\rangle. \quad (2.13)$$

Finally, lengths of rays may be rescaled by multiplication by complex constants:

$$\begin{aligned} \lambda(|a\rangle + |b\rangle) &= \lambda|a\rangle + \lambda|b\rangle \\ (\lambda + \mu)|a\rangle &= \lambda|a\rangle + \mu|a\rangle \\ \lambda\mu|a\rangle &= \lambda(\mu|a\rangle) = (\lambda\mu)|a\rangle \end{aligned} \quad (2.14)$$

and so on. One defines the inner product of two kets $|a\rangle, |b\rangle$ using the bra vector for one of them,

$$(|a\rangle, |b\rangle) \equiv \langle a|b\rangle \quad (2.15)$$

The quantity $\langle a|b\rangle$ is postulated to be a complex number \mathbf{C} .

We next postulate that $\langle a|b\rangle = \langle b|a\rangle^*$. (Notice that $\langle a|a\rangle$, essentially the squared length of the vector, is thus always real.) From a practical point of view we can do all calculations involving a bra vector $\langle b|$ using the ket vector $|b\rangle$, by using inner products. Begin with

$$|c\rangle = A|a\rangle + B|b\rangle \quad A, B \in \mathbf{C} \quad (2.16)$$

then

$$\langle c| = A^*\langle a| + B^*\langle b|. \quad (2.17)$$

It is convenient (but not necessary) to require that kets labeling physical states be normalized. We define a normalized vector by

$$|\bar{a}\rangle = \frac{1}{\sqrt{\langle a|a\rangle}} |a\rangle. \quad (2.18)$$

Two vectors are said to be orthogonal if $\langle a|b\rangle = 0$ and similarly a set of vectors are said to be orthonormal if they satisfy $\langle a|a\rangle = 1$ as well as obeying the orthogonality condition. If a state can be written as a superposition of a set of (normalized, orthogonal) basis states

$$|\psi\rangle = \sum_i c_i |\psi_i\rangle \quad c_i \in \mathbf{C}, \quad (2.19)$$

and

$$\sum_i |c_i|^2 = 1, \quad (2.20)$$

then we interpret the squared modulus of the coefficient c_i as the probability that the system (represented by ket $|\psi\rangle$) is to be found (measured to be in) the state represented by $|\psi_i\rangle$. As an example, suppose we had a Hilbert space consisting of two states, $|+\rangle$ and $|-\rangle$. The electron is a spin 1/2 particle, so the states could correspond to the two states of an electron's spin, up or down. Imagine that we could prepare a large number of electrons as an identical superposition of these two states,

$$|\psi\rangle = \sum_{i=\pm} c_i |\psi_i\rangle. \quad (2.21)$$

We then perform a set of measurements of the spin of the electron. The probability (over the ensemble of states) that we find the spin to be up will be $|c_+|^2$.

Observables in quantum mechanics

Observables in quantum mechanics are represented by operators in Hilbert space. Operators (here denoted \hat{O}) transform a state vector into another state vector.

$$\hat{O}|a\rangle = |b\rangle \quad (2.22)$$

As a conventional definition, operators only act to the right on kets. Some states are eigenvectors or eigenstates of the operators; the action of the operator rescales the state:

$$\hat{A}|a\rangle \propto |a\rangle \quad (2.23)$$

We usually say

$$\hat{A}|a\rangle = a|a\rangle \quad a \in \mathbf{C} \quad (2.24)$$

where the number a (the prefactor) is the eigenvalue of the operator. Two operators \hat{A} and \hat{B} are equal if

$$\hat{A}|a\rangle = \hat{B}|a\rangle \quad (2.25)$$

for all ket vectors $|a\rangle$ in a space. Most operators we encounter are also linear; that is, they satisfy the following relation:

$$\hat{A}(\alpha|a\rangle + \beta|b\rangle) = \alpha\hat{A}|a\rangle + \beta\hat{A}|b\rangle. \quad (2.26)$$

An operator A is *null* if $A|a\rangle = |0\rangle$ for all $|a\rangle$. Operators generally do not commute:

$$\hat{A}\hat{B} \neq \hat{B}\hat{A} \quad (2.27)$$

(meaning that $\hat{A}\hat{B}|\psi\rangle \neq \hat{B}\hat{A}|\psi\rangle$). The commutator of two operators is often written as $[A, B] = AB - BA$.

The adjoint of an operator \hat{A}^\dagger is defined through the inner product

$$(|g\rangle, \hat{A}|f\rangle) = (\hat{A}^\dagger|g\rangle, |f\rangle) = (|f\rangle, \hat{A}^\dagger|g\rangle)^*. \quad (2.28)$$

Note that the adjoint also acts to the right. Alternatively, in “bra Hilbert space” the abstract relation of a bra to a ket $|a\rangle \leftrightarrow \langle a|$ means that if $X|a\rangle = |c\rangle$, $|c\rangle \leftrightarrow \langle c| = \langle a|X^\dagger$ (with the adjoint operator acting on the bra to its left).

The following relations are easily shown to hold for operators:

$$\begin{aligned} (\alpha\hat{A})^\dagger &= \alpha^*\hat{A}^\dagger \\ (\hat{A} + \hat{B})^\dagger &= \hat{A}^\dagger + \hat{B}^\dagger \\ (\hat{A}\hat{B})^\dagger &= \hat{B}^\dagger\hat{A}^\dagger \end{aligned} \quad (2.29)$$

Let us make a slightly sharper statement about the relation of operators and observables. An observable B “corresponds to” – or is identified with – an operator in the sense that any measurement of the observable B is postulated to yield one of the eigenstates of the operator \hat{B} ,

$$\hat{B}|b_j\rangle = b_j|b_j\rangle. \quad (2.30)$$

Furthermore, we postulate that, as a result of measurement, a system is localized into the state $|b_j\rangle$. (If you like jargon, I am describing the “collapse of the wave function” under measurement.)

If this is confusing, think of an example: let b_j be some value of the coordinate x . Measuring the coordinate of the particle to be at x at some time t simply means that we know that the particle is at x at time t – its wave function must be sharply peaked at x . This is the same statement as to say that the wave function is in an eigenfunction of the operator \hat{x} whose eigenvalue – at that instant of time – is x . Notice that this does not say anything about the location of the particle at later times.

To be honest, I have to say that the previous couple of paragraphs open a big can of worms called the “measurement problem in quantum mechanics.” The statement that a measurement collapses the wave function is a postulate associated with what is called the “Copenhagen interpretation” of quantum mechanics. There are other interpretations of

what is going on which I will leave for you to discover on your own. (Some of them are more interesting than other ones.) On the micro level, we can ask “what does it mean to measure something?” You have a quantum system, you do something to it, you read the result on a macroscopic instrument which you might want to think of as a classical object – or maybe not. The micro situation will always have some particular detailed answer which depends on what you are actually doing, precisely specified.

Let us try to avoid these issues for now and move on.

A consequence of this postulate is that, if the system is known to be in a state $|a\rangle$, then the probability that a measurement of \hat{B} gives the value b_j is $|\langle b_j|a\rangle|^2$. There could be a finite discrete set of possibilities for b_j ; in that case the Hilbert space is finite dimensional. (Think of spin up and spin down states for an electron, for a two dimensional Hilbert space.) However, if B has a continuous spectrum (like a coordinate), we have to say that the probability that a measurement of B lies in the range $b', b' + db$ will be $|\langle a|b'\rangle|^2 db$. Since there is a separate ray for every eigenvalue b_j , the basis in which a continuous variable is diagonal must have infinite dimensionality.

Note that if $X|a\rangle = a|a\rangle \equiv |c\rangle$, then $\langle c| = a^* \langle a| = \langle a|X^\dagger$, i.e., $X|a\rangle \leftrightarrow \langle a|X^\dagger$. For arbitrary X there is generally no connection between the ket vector $X|a\rangle$ and the bra vector $\langle a|X$.

Most of the operators we deal with in quantum mechanics are Hermitian operators, for which $\hat{A}^\dagger = \hat{A}$. The reason that we deal with Hermitian operators is that their eigenvalues are real. Physical observables are real numbers, so this identification is natural. The reality of the eigenvalues of a Hermitian operator follows from the equality

$$\langle a_2|\hat{A}|a_1\rangle = \langle a_1|\hat{A}|a_2\rangle^*, \quad (2.31)$$

which implies that

$$a_1 \langle a_2|a_1\rangle = a_2^* \langle a_1|a_2\rangle^* = a_2^* \langle a_2|a_1\rangle \quad (2.32)$$

or

$$(a_2^* - a_1) \langle a_2|a_1\rangle = 0. \quad (2.33)$$

This identity can be satisfied in two ways. Suppose that the bra $\langle a_2|$ is dual to the ket $|a_1\rangle$ – $\langle a_2| = \langle a_1|$. Then $\langle a_2|a_1\rangle$ is nonzero. However, this duality means that $a_2^* = a_1^* = a_1$, and of course this says that a_1 is real. The other possibility is that $a_2^* \neq a_1$. Then we can only satisfy the equality if the two states are orthogonal. This gives us a second very useful result – eigenstates of Hermitian operators with different eigenvalues are orthogonal.

It can happen that two or more states yield the same eigenvalue of an operator. We speak of the states as being “degenerate” or say that the operator has a degeneracy. In this case one cannot show that $\langle a_2|a_1\rangle = 0$. But one can construct a basis $|a'_1\rangle, |a'_2\rangle, \dots$ such that $\langle a'_2|a'_1\rangle = 0$, and recover our previous results. We can do this by Gram-Schmidt orthogonalization. We outline the first few steps of this process by the following:

$$\begin{aligned} |a'_1\rangle &= |a_1\rangle \\ |a'_2\rangle &= |a_2\rangle + c|a'_1\rangle \quad c = -\frac{\langle a_1|a_2\rangle}{\langle a_1|a_1\rangle} \end{aligned} \tag{2.34}$$

Clearly $\langle a'_2|a_1\rangle = 0$. A normalized second vector can be constructed by defining $|a''_2\rangle = |a'_2\rangle / \sqrt{\langle a'_2|a'_2\rangle}$. Thus, for all practical purposes, our statement that the eigenstates of Hermitian operators can be regarded as orthogonal still holds.

Note in passing that, because operators transform states into states, we can imagine constructing an operator from a bra and a ket:

$$\hat{O} = |b\rangle \langle a|. \tag{2.35}$$

This operator acts on any state to transform it into the state $|b\rangle$:

$$\begin{aligned} \hat{O}|\gamma\rangle &= |b\rangle \langle a|\gamma\rangle \\ &= |b\rangle \times (\text{complex } \#). \end{aligned} \tag{2.36}$$

If the possible number of eigenstates of an operator is finite in number, any state can be written as a linear superposition of the basis states. (If the dimension of the basis is infinite, one must make this statement as a postulate.) Then we can expand a state as

$$|\alpha\rangle = \sum_n c_n |n\rangle. \tag{2.37}$$

To find the values of the c_n 's we simply take the inner product on both sides of the equation with the ket $|m\rangle$ and use orthogonality,

$$\langle m|\alpha\rangle = \sum_n c_n \langle m|n\rangle = c_m \tag{2.38}$$

so

$$|\alpha\rangle = \sum_n (\langle n|\alpha\rangle) |n\rangle = \sum_n |n\rangle \langle n|\alpha\rangle. \tag{2.39}$$

We see that

$$|\alpha\rangle = \left\{ \sum_n |n\rangle \langle n| \right\} |\alpha\rangle. \quad (2.40)$$

We have discovered the very useful “identity operator”

$$\hat{\mathbf{1}} = \sum_n |n\rangle \langle n|. \quad (2.41)$$

The operator $\Lambda_a = |a\rangle \langle a|$ is called a projection operator. Applied to any state $|\psi\rangle$, it projects out the part of the state which is aligned along $|a\rangle$: $\Lambda_a |\psi\rangle = |a\rangle \langle a|\psi\rangle$. If the $|a\rangle$'s form a complete basis then summing over all projectors gives the state itself back again: a physical realization of our identity operator.

$$\hat{\mathbf{1}} = \sum \Lambda_a = \sum |a\rangle \langle a|. \quad (2.42)$$

From our definition of the projection operator it is fairly easy to see that a product of projection operators is also a projection operator,

$$\Lambda_a \Lambda_{a'} = \delta_{aa'} \Lambda_a. \quad (2.43)$$

Bras, kets, vectors, and matrices

It is useful to make an analogy between the rays of an abstract Hilbert space and their operators, and ordinary vectors and matrices. We can think of ψ as a column vector in an n -dimensional vector space.

$$|\psi\rangle = \sum_n c_n |n\rangle \quad (2.44)$$

where the $|n\rangle$'s are the rays along the axes of the space. Then we can think of writing the states as

$$|\psi\rangle \rightarrow \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{pmatrix}.$$

$\langle\psi|$ is associated with the row vector whose elements are the complex conjugates of $|\psi\rangle$'s:

$$\langle\psi| \rightarrow \left(c_1^* \ c_2^* \ c_3^* \ \dots \ c_n^* \right).$$

Thus $\langle\psi|\psi\rangle = c_1^*c_1 + c_2^*c_2 + \dots$ is an ordinary vector product. The projection operator is a matrix (a dyadic)

$$\Lambda_1 = |1\rangle\langle 1| = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdot & \cdot & 0 \\ 0 & \cdot & & & \\ \cdot & & \cdot & & \\ \cdot & & & \cdot & \\ 0 & & & & 0 \end{pmatrix}$$

The action of a projection operator on a state $|\psi\rangle$ is

$$\Lambda_1 |\psi\rangle = \begin{pmatrix} c_1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

Linear operators are then represented as matrices in the space: if $|\psi\rangle = \sum_n \alpha_n |n\rangle$ then

$$\hat{O} |\psi\rangle = |\psi'\rangle = \sum_n \alpha_n \hat{O} |n\rangle = \sum_m \beta_m |m\rangle \quad (2.45)$$

where the expansion coefficient is

$$\beta_m = \sum_n \langle m|\hat{O}|n\rangle \alpha_n = \sum_n O_{mn} \alpha_n \quad (2.46)$$

Thus we have a relation between the coefficients for $\hat{O} |\psi\rangle$ and the coefficients for $|\psi\rangle$.

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_n \end{pmatrix} = \begin{pmatrix} O_{11} & O_{12} & \cdot & \cdot & O_{1n} \\ O_{21} & \cdot & & & \\ \cdot & & \cdot & & \\ \cdot & & & \cdot & \\ O_{n1} & & & & O_{nn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \alpha_n \end{pmatrix}$$

We can interpret the equation $|\psi'\rangle = \hat{O} |\psi\rangle$ as one where \hat{O} is a matrix and the kets are vectors. The complex numbers $\hat{O}_{nm} = \langle n|\hat{O}|m\rangle$ are called “matrix elements.” Of course, the matrix elements of Hermitian operators form Hermitian matrices, $\hat{O}_{ij} = \hat{O}_{ji}^*$.

Transformation of basis

Suppose we have a state written in some orthonormal basis as $|j'\rangle$. We can write the state in another basis $|j\rangle$ as

$$|j'\rangle = \sum_j |j\rangle \langle j|j'\rangle \equiv \sum_j U_{jj'}^\dagger |j\rangle. \quad (2.47)$$

(Note $\langle j|j'\rangle = U_{jj'}^\dagger$ – the dagger is a convention.) This transformation must leave the norm of the state unchanged, and so U must obey the constraint

$$\langle i'|j'\rangle = \delta_{i'j'} = \sum_i \langle i'|i\rangle \langle i|j'\rangle = \sum_i U_{i'i} U_{ij'}^\dagger. \quad (2.48)$$

This implies that $U^\dagger = U^{-1}$. Such norm-preserving transformations are called unitary transformations, and they are realized as unitary matrices.

Now consider how an arbitrary state $|\psi\rangle$ is expanded in either basis:

$$\langle j|\psi\rangle = \sum_{j'} \langle j|j'\rangle \langle j'|\psi\rangle \quad (2.49)$$

or, writing $\psi_j = \langle j|\psi\rangle$,

$$\psi_j = \sum_{j'} U_{jj'}^\dagger \psi'_{j'}. \quad (2.50)$$

It is also useful to see how operators transform under unitary transformations:

$$\langle i|\hat{O}|j\rangle = \sum_{i'j'} \langle i|i'\rangle \langle i'|\hat{O}|j'\rangle \langle j'|j\rangle \quad (2.51)$$

or

$$\hat{O}_{ij} = \sum_{i'j'} U_{ii'}^\dagger \hat{O}_{i'j'} U_{j'j}. \quad (2.52)$$

Thus an operator represented in one basis is related to the same operator in another basis by a similarity transformation. (Note the identical order of U and U^\dagger – U^\dagger always to the left – for ψ and \hat{O} !)

Finally, suppose that we want to find the eigenfunctions and eigenvalues of an operator. Call the operator \hat{Q} . We want a vector $|a\rangle$ for which

$$\hat{Q}|a\rangle = \lambda|a\rangle. \quad (2.53)$$

We assume that we have some complete basis $|n\rangle$ and expand our state in it,

$$|a\rangle = \sum_n \alpha_n |n\rangle \quad \alpha_n \equiv \langle n|a\rangle, \quad (2.54)$$

and then the eigenvalue equation becomes

$$\hat{Q}|a\rangle = \sum_n \alpha_n \hat{Q}|n\rangle = \sum_n \alpha_n \lambda |n\rangle. \quad (2.55)$$

We now take the inner product of both sides of this equation with one member of the basis, $\langle i|$. Orthogonality forces $\langle i|m\rangle = \delta_{im}$, so

$$\sum_n \langle i|\hat{Q}|n\rangle \alpha_n = \sum_n \hat{Q}_{in} \alpha_n = \sum_n \alpha_n \lambda \delta_{in} \quad (2.56)$$

or

$$\sum_n (\hat{Q}_{in} - \delta_{in} \lambda) \alpha_n = 0. \quad (2.57)$$

This set of linear equations has a solution only if $\det(\hat{Q}_{in} - \lambda \delta_{in}) = 0$. This equation is called the “characteristic equation,” and if the dimension of the space is N , this is an N^{th} order polynomial equation. There will be N values for λ . The similarity transform which diagonalizes the matrix has a transformation matrix which is composed of the eigenvectors. (Note, though, the eigenvalues need not be distinct – this is the signal of degeneracy.)

Eigenvalues are basis-independent, though eigenvectors are obviously not. To see this fact, we begin with the characteristic equation

$$\det(\hat{Q} - \lambda \mathbf{1}) = 0 \quad (2.58)$$

We then transform our system to a new basis

$$\begin{aligned} \det(\hat{U}^\dagger(\hat{Q} - \lambda \mathbf{1})\hat{U}) &= \det(\hat{U}^\dagger) \det(\hat{Q} - \lambda \mathbf{1}) \det(\hat{U}) \\ &= \det(\hat{U}) \det(\hat{U}^\dagger) \det(\hat{Q} - \lambda \mathbf{1}) \\ &= \det(\hat{U}\hat{U}^\dagger) \det(\hat{Q} - \lambda \mathbf{1}) \\ &= \det(\hat{Q} - \lambda \mathbf{1}). \end{aligned} \quad (2.59)$$

We still obtain the same characteristic equation, with same eigenvalues as its solutions. Thus the two operators Q and $U^{-1}QU$, where U is any unitary transformation, have the same spectra.

More about operators

We appear to have gotten rather deeply into quantum mechanics without making any contact with classical mechanics. Let us rectify that omission, by writing down two last postulates.

Suppose we have two operators \hat{A} and \hat{B} which correspond to the classical variables A and B . We postulate that they will obey the following commutation relation.

$$\hat{A}\hat{B} - \hat{B}\hat{A} = [\hat{A}, \hat{B}] = i\hbar\{A, B\}_{op} \quad (2.60)$$

where $\{A, B\}_{op}$ is the operator corresponding to the classical Poisson bracket of the two dynamical variables

$$\{A, B\} = \sum_i \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial q_i} \frac{\partial A}{\partial p_i} \right). \quad (2.61)$$

q_i and p_i correspond to the canonically conjugate coordinates and momenta of the classical Hamiltonian, and thus

$$[\hat{q}_i, \hat{p}_j] = i\hbar\delta_{ij}, \quad (2.62)$$

that is, different components of \vec{p} and \vec{q} commute, but the p and q of identical components have a commutator which is equal to $i\hbar$.

Time dependence of states

We arrive at our final postulate of quantum mechanics, stated cautiously. It is possible to make a choice of states, called the Schrödinger Representation, in which a state changes with time according to the Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle. \quad (2.63)$$

In this expression, \hat{H} is the Hamiltonian operator. The reason for caution is that there are many equivalent ways of describing time evolution in quantum mechanics: the states may evolve in time, while the operators do not, or the operators may evolve in time while at least some of the states do not, or both states and operators may evolve in time. We will return to this point in the next chapter.

The two-state system

The two-state system is a very important object in quantum mechanics, possibly the simplest one which exhibits interesting quantum mechanical behavior. It is one of the few systems which we can almost always solve exactly. It also has many apparently quite different physical

applications. Let's consider it "on its own," before connecting it to some experimentally realizable situations. Suppose we have a Hamiltonian

$$\hat{H} = \begin{pmatrix} \epsilon_1 & \Delta \\ \Delta^* & \epsilon_2 \end{pmatrix}. \quad (2.64)$$

The off diagonal elements are complex conjugates of each other because \hat{H} is Hermitian. To find the eigenvalues of the Hamiltonian (i.e. the energy eigenvalues) we evaluate the determinant

$$\begin{vmatrix} \epsilon_1 - \lambda & \Delta \\ \Delta^* & \epsilon_2 - \lambda \end{vmatrix} = (\epsilon_1 - \lambda)(\epsilon_2 - \lambda) - |\Delta|^2 = 0 \quad (2.65)$$

and a little algebra gives us the following very useful formula, well worth memorizing:

$$\lambda = \frac{\epsilon_1 + \epsilon_2}{2} \pm \sqrt{\left(\frac{\epsilon_1 - \epsilon_2}{2}\right)^2 + |\Delta|^2} \equiv \frac{\epsilon_1 + \epsilon_2}{2} \pm \hbar\omega \quad (2.66)$$

The eigenvalues are equal to the average of the two energy levels (ϵ_1 and ϵ_2) plus a splitting term which is the sum of the squares of the variance in the diagonal term and of the norm of the off-diagonal term.

There is no universally useful formula for the eigenfunctions, and when \hat{H} is simple, it is often easiest just to proceed "from scratch." Let's look at the general case with language whose use will become more apparent shortly. To simplify our discussion, let us alter our notation slightly and replace the complex Δ in Eq. 2.64 by $\Delta \exp(i\chi)$ where now Δ is real. We wish to find eigenfunctions, solving

$$\hat{H}\psi_{\pm} = \lambda_{\pm}\psi_{\pm}. \quad (2.67)$$

We write a parameterization which automatically preserves unitarity

$$|\psi_{+}\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\phi} \end{pmatrix} \quad |\psi_{-}\rangle = \begin{pmatrix} -\sin \frac{\theta}{2} e^{-i\phi} \\ \cos \frac{\theta}{2} \end{pmatrix}, \quad (2.68)$$

and then the upper eigenvector equation corresponding to λ_{+} is

$$(\hat{H} - \lambda_{+})\psi_{+} = (\epsilon_1 - \lambda_{+}) \cos \frac{\theta}{2} + \Delta \sin \frac{\theta}{2} e^{i(\chi+\phi)} = 0. \quad (2.69)$$

Clearly, a solution to this equation is $\phi = -\chi$ and

$$\tan \frac{\theta}{2} = \frac{\lambda_{+} - \epsilon_1}{\Delta} = \frac{\frac{\epsilon_2 - \epsilon_1}{2} + \sqrt{\left(\frac{\epsilon_1 - \epsilon_2}{2}\right)^2 + \Delta^2}}{\Delta}. \quad (2.70)$$

Note that the amount of mixing between the two states only depends on the difference of the energies. We could have seen this more easily had we written

$$\hat{H} = \frac{\epsilon_1 + \epsilon_2}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \epsilon & \Delta \\ \Delta^* & -\epsilon \end{pmatrix} \quad (2.71)$$

where $\epsilon = (\epsilon_1 - \epsilon_2)/2$. This way of writing the Hamiltonian shows us that its first part only corresponds to an overall energy shift, while the second part gives us the interesting physics. Let us call $\delta = \Delta/\epsilon$. Then we can write

$$\tan \frac{\theta}{2} = -\frac{1}{\delta} + \sqrt{1 + \frac{1}{\delta^2}} \quad (2.72)$$

or, after invoking a trigonometric identity,

$$\tan \theta = \frac{2 \tan \frac{\theta}{2}}{1 - \tan^2 \frac{\theta}{2}} = \delta = \frac{\Delta}{\epsilon}. \quad (2.73)$$

We could have written

$$\hat{H} = \begin{pmatrix} \epsilon & \Delta \\ \Delta^* & -\epsilon \end{pmatrix} = \hbar\omega \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix} \quad (2.74)$$

if we wanted to encode this solution into the definition of \hat{H} from the beginning. (This is sometimes useful, but not always!) Now we are equipped with a general formula for the mixing as well one for the energies. Let's consider some limiting cases, further assuming the ordering $\epsilon_1 > \epsilon_2$, or $\epsilon > 0$.

1. $\Delta \ll \epsilon$: In this case we can Taylor expand the square root factor

$$\begin{aligned} \lambda_{\pm} &= \frac{\epsilon_1 + \epsilon_2}{2} \pm \frac{|\epsilon_1 - \epsilon_2|}{2} \left(1 + \frac{1}{2} \frac{\Delta^2}{\epsilon} + \dots \right) \\ &\simeq \frac{\epsilon_1 + \epsilon_2}{2} \pm \frac{|\epsilon_1 - \epsilon_2|}{2} \pm \frac{\Delta^2}{|\epsilon_1 - \epsilon_2|} \end{aligned} \quad (2.75)$$

and thus

$$\lambda_+ = \epsilon_1 + \frac{\Delta^2}{\epsilon_1 - \epsilon_2} \quad \lambda_- = \epsilon_2 - \frac{\Delta^2}{\epsilon_1 - \epsilon_2}. \quad (2.76)$$

If Δ/ϵ is small, $\lambda_+ \simeq \epsilon_1$ and $\lambda_- \simeq \epsilon_2$ with

$$\psi_+ \simeq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \psi_- \simeq \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.77)$$

When Δ is small the gap varies quadratically with Δ . The behavior is shown in Fig. 2.6.

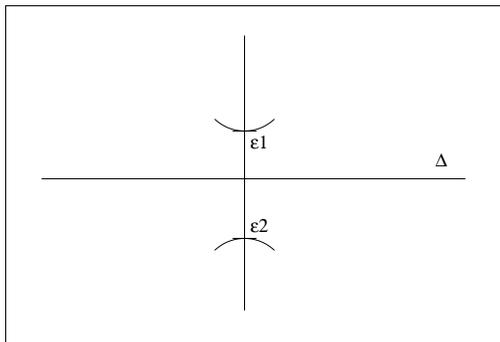
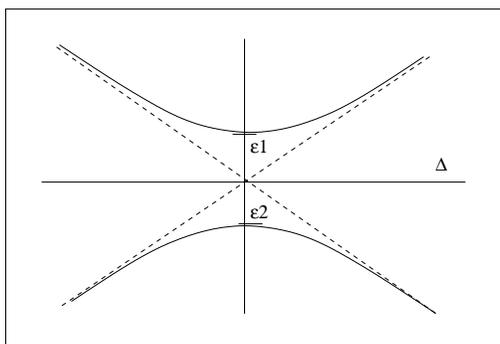
Figure 2.6: Eigenvalue spectrum for small Δ .

Figure 2.7: The full spectrum of the two-state system.

2. $\Delta \gg \epsilon$: In this case

$$\lambda_{\pm} \simeq \frac{\epsilon_1 + \epsilon_2}{2} \pm |\Delta|. \quad (2.78)$$

Thus the two energies are just the average of our two diagonal terms plus or minus the off diagonal term. See Fig. 2.7. In this case $\tan \theta = \Delta/\epsilon$ diverges, so the mixing angle $\theta \rightarrow \frac{\pi}{2}$. The eigenfunctions are as non-diagonal in our eigenstate basis as they could possibly be: if the phase of complex Δ is $\chi = 0$, then for large positive Δ

$$\psi_+ \rightarrow \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad \psi_- \rightarrow \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}. \quad (2.79)$$

Notice the situation for large negative Δ : up to overall phase factors

$$\psi_+ \rightarrow \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad \psi_- \rightarrow \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}. \quad (2.80)$$

That is, the wave functions have exchanged places. It is as if the system actually followed the dotted lines in Fig. 2.7 as Δ was tuned from a large positive value to a large negative value. But the dotted lines show the energy levels crossing; in reality, they do not. This is the iconic example of what is called an “avoided level crossing.” In the absence of a coupling between two levels (that is the ϵ term in \hat{H}) the levels could pass through each other and there would be a point of degeneracy at $\Delta = 0$. The coupling between the levels means that this degeneracy cannot occur. Avoided level crossings and the “swapping” of the wave functions of the would-be crossing pairs have many practical uses in quantum mechanics.

Any system with two states whose coupling to each other is large and whose coupling to other states is small can be approximated by a two-state system – independent of how complicated the states actually are. Many examples come to mind: neutrinos (in the context of neutrino oscillations) or pairs of atomic states, perhaps isolated from their neighbors by a clever experimentalist. But the “classic” two state system is the spin-1/2 particle in an external magnetic field. The Pauli matrices form a standard basis notation for the Hamiltonian:

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (2.81)$$

Any traceless 2×2 Hermitian matrix can be written as a linear combination of these three Pauli matrices, and if the matrix is not traceless, it can be written as a linear combination of the three Pauli matrices and the identity. By inspection, the eigenfunctions of σ_z , and their corresponding eigenvalues, are

$$\lambda = +1 \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \lambda = -1 \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.82)$$

(All the Pauli matrices have eigenvalues ± 1 ; stop reading and work out the eigenfunctions of σ_x and σ_y .) A spin- $\frac{1}{2}$ particle has a magnetic moment $\vec{\mu} = \frac{e}{mc}\vec{S}$ where $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$. (We will make this connection more formally, later.) The interaction Hamiltonian for a magnetic moment in an external magnetic field is

$$\hat{H} = -\vec{\mu} \cdot \vec{B}. \quad (2.83)$$

It has the matrix representation

$$\hat{H} = -\vec{\mu} \cdot \vec{B} = -\frac{e\hbar}{2mc} \begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix}. \quad (2.84)$$

The reader is encouraged to work out the eigenfunctions for the spin for various limiting orientations of the magnetic field. (We will return to this problem in the next chapter, if you want to look ahead.)

The Pauli matrices obey the following algebra ($\hat{1}$ is the identity matrix)

$$\hat{\sigma}_i \hat{\sigma}_j = \delta_{ij} \hat{1} + i\epsilon_{ijk} \hat{\sigma}_k \quad (2.85)$$

or equivalently

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2i\epsilon_{ijk} \hat{\sigma}_k. \quad (2.86)$$

The anti-commutator of two Pauli matrices is simple, too:

$$\hat{\sigma}_i \hat{\sigma}_j + \hat{\sigma}_j \hat{\sigma}_i \equiv \{\hat{\sigma}_i, \hat{\sigma}_j\} = 2\delta_{ij} \hat{1}. \quad (2.87)$$

Check these relations and remember them!

We often encounter complicated expressions made of several Pauli matrices. They can be simplified by repeatedly invoking appropriate commutation and anti-commutation relations. As an example let us consider

$$\begin{aligned} (\vec{\sigma} \cdot \mathbf{a})(\vec{\sigma} \cdot \mathbf{b}) &= \hat{\sigma}_i a_i \hat{\sigma}_j b_j \\ &= \hat{\sigma}_i \hat{\sigma}_j a_i b_j \\ &= (\delta_{ij} \hat{1} + i\epsilon_{ijk} \hat{\sigma}_k) a_i b_j \\ &= \hat{1} a_i b_j + i\epsilon_{ijk} a_i b_j \hat{\sigma}_k \\ &= \mathbf{a} \cdot \mathbf{b} \hat{1} + i\vec{\sigma} \cdot (\mathbf{a} \times \mathbf{b}). \end{aligned} \quad (2.88)$$

The final answer is a linear combination of terms proportional to the identity and to the three Pauli matrices, because that is all there is for a two dimensional Hilbert space.

Simultaneous observables and the uncertainty principle

Suppose we have a state $|a, b, \dots\rangle$ which is an eigenfunction of several operators, that is,

$$\hat{A}|a, b, \dots\rangle = a|a, b, \dots\rangle \quad \hat{B}|a, b, \dots\rangle = b|a, b, \dots\rangle. \quad (2.89)$$

This is a very desirable situation, since it means that we have precise information about the quantities represented by the operators for the state. But, can this happen for arbitrary

operators? No, it can only occur if the commutators of all pairs of operators are zero, for then (specialize to two operators)

$$(\hat{A}\hat{B} - \hat{B}\hat{A})|a, b, \dots\rangle = (ab - ba)|a, b, \dots\rangle = 0. \quad (2.90)$$

A mathematical way of expressing the physical question, “How completely can I characterize a state?” is to ask “What is the complete set of commuting operators for the system?”

We can invert the discussion and ask, if $A, B \neq 0$, how precisely can one simultaneously observe $\langle \hat{A} \rangle$ and $\langle \hat{B} \rangle$? We will restrict ourselves the case where both of the operators are Hermitian operators, since that is what is most commonly seen in practice. Let us consider an arbitrary state $|\psi\rangle$ and consider the shifted operator

$$\Delta\hat{A} = \hat{A} - \langle \hat{A} \rangle \quad (2.91)$$

where $\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle$. The expectation value of the square of this operator is

$$\begin{aligned} \langle (\Delta\hat{A})^2 \rangle &= \langle \hat{A}^2 - 2\hat{A}\langle \hat{A} \rangle + \langle \hat{A} \rangle^2 \rangle \\ &= \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2. \end{aligned} \quad (2.92)$$

It is known as the dispersion of the expectation value of the operator in the state, since that is the definition of the variance of a quantity in statistics. Now we shall show that

$$\langle (\Delta\hat{A})^2 \rangle \langle (\Delta\hat{B})^2 \rangle \geq \frac{1}{4} \left| \langle [\hat{A}, \hat{B}] \rangle \right|^2. \quad (2.93)$$

This result is called the Uncertainty Principle.

The norm of any vector is always greater than or equal to zero, so for any λ

$$(|\alpha\rangle + \lambda|\beta\rangle, |\alpha\rangle + \lambda|\beta\rangle) \geq 0. \quad (2.94)$$

A direct evaluation of this expression gives

$$\langle \alpha | \alpha \rangle + \lambda \langle \alpha | \beta \rangle + \lambda^* \langle \beta | \alpha \rangle + |\lambda|^2 \langle \beta | \beta \rangle \geq 0. \quad (2.95)$$

Setting $\lambda = i\epsilon$, $|\alpha\rangle = (\Delta\hat{A})|\psi\rangle$, and $|\beta\rangle = (\Delta\hat{B})|\psi\rangle$,

$$\langle \psi | (\Delta\hat{A})^2 | \psi \rangle + i\epsilon \langle \psi | \Delta\hat{A}\Delta\hat{B} - \Delta\hat{B}\Delta\hat{A} | \psi \rangle + \epsilon^2 \langle \psi | (\Delta\hat{B})^2 | \psi \rangle \geq 0. \quad (2.96)$$

We minimize this equation with respect to ϵ . The minimum occurs when $2\epsilon \langle (\Delta \hat{B})^2 \rangle = -i \langle [\hat{A}, \hat{B}] \rangle$. Notice, by the way, that if \hat{A} and \hat{B} are Hermitian then their commutator is anti-Hermitian and $i[\hat{A}, \hat{B}]$ is Hermitian:

$$[\hat{A}, \hat{B}]^\dagger = \hat{B}^\dagger \hat{A}^\dagger - \hat{A}^\dagger \hat{B}^\dagger = [\hat{B}^\dagger, \hat{A}^\dagger] = [\hat{B}, \hat{A}] = -[\hat{A}, \hat{B}]. \quad (2.97)$$

Inserting the result of the minimization into our original equation, we learn that

$$\langle (\Delta \hat{A})^2 \rangle - \frac{|\langle [\hat{A}, \hat{B}] \rangle|^2}{2 \langle (\Delta \hat{B})^2 \rangle} + \frac{|\langle [\hat{A}, \hat{B}] \rangle|^2}{4 \langle (\Delta \hat{B})^2 \rangle} \geq 0 \quad (2.98)$$

and so

$$\langle (\Delta \hat{A})^2 \rangle \langle (\Delta \hat{B})^2 \rangle \geq \frac{1}{4} \left| \langle [\hat{A}, \hat{B}] \rangle \right|^2. \quad (2.99)$$

The uncertainty principle is a deep statement about what can and cannot be measured in quantum mechanics. The most-seen case of the uncertainty principle is $A = x$, $B = p$, $C = \hbar$, yielding the familiar result $\Delta x \Delta p \geq \hbar/2$.

Again, if $[A, B] = 0$ one can in principle have simultaneous eigenstates of A and B , states in which we can determine expectation values of both operators to arbitrary precision.

We often encounter the uncertainty principle in casual situations, where it allows us to make back of the envelope calculations about the typical sizes of bound states and their energies. (In this context, it is basically a poor man's variational calculation.) As an example, consider the one body Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}(r) \quad (2.100)$$

and suppose that we want to guess the eigenvalues of this operator. We know that $\Delta p \geq \hbar/\Delta x$ and so we assume that in \hat{H} , $p \simeq \Delta p \simeq \hbar/\Delta x$, $r \simeq \Delta x$. Thus our Hamiltonian becomes

$$\hat{H} = \frac{\hbar^2}{2m(\Delta x)^2} + V(\Delta x). \quad (2.101)$$

To estimate the smallest possible bound state energy consistent with the uncertainty principle, minimize Hamiltonian with respect to Δx . For hydrogen we have

$$E(r) = \frac{\hbar^2}{2m\Delta x^2} - \frac{e^2}{\Delta x}. \quad (2.102)$$

Setting $\partial E/\partial \Delta x = 0$ gives $\Delta x = \hbar^2/(me^2)$ which upon insertion into the energy equation gives $E(\Delta x) = -\frac{1}{2} \left(\frac{e^2}{\hbar c} \right)^2 mc^2$ or -13.6 eV – a familiar result.

Operators with continuous spectra

Many operators – the coordinate x is a familiar example – have a continuous spectrum. We can treat operators with continuous spectra in a very similar fashion to those with discrete spectra, at the cost of generalizing the Kronecker delta $\langle A|A'\rangle = \delta_{A,A'}$ to the Dirac delta function $\delta(A-A')$. At the slender level of rigor seen in most physics books, the generalization is merely notational. We record it in this table:

discrete	continuous
$ \psi\rangle = \sum_A A\rangle \langle A \psi\rangle$	$ \psi\rangle = \int dA A\rangle \langle A \psi\rangle$
$\hat{1} = \sum_A A\rangle \langle A $	$\hat{1} = \int dA A\rangle \langle A $
$\langle A A'\rangle = \delta_{AA'}$	$\langle A A'\rangle = \delta(A-A')$
$\langle A_1 \hat{A} A_2\rangle = A_2\delta_{A_1A_2}$	$\langle A_1 \hat{A} A_2\rangle = A_2\delta(A_1-A_2)$

The Dirac delta function obeys

$$\int dx \delta(x) = 1; \quad \int dx \delta(x-x_0) f(x) = f(x_0). \quad (2.103)$$

There is, however, one important point, which requires caution: Dirac delta functions are not functions, but distributions. A distribution can be defined as a limit of a good function or defined as an operator under an integral. Working with delta functions in “stand-alone” mode can be quite hazardous.

Practical definitions of delta functions (as limits) are quite useful. A familiar one is

$$\begin{aligned} \delta(k_1 - k_2) &= \lim_{X \rightarrow \infty} \frac{1}{2\pi} \int_{-X}^X e^{i(k_1 - k_2)x} dx \\ &= \lim_{X \rightarrow \infty} \frac{2 \sin((k_1 - k_2)X)}{k_1 - k_2} \end{aligned} \quad (2.104)$$

Plotting this function, you will find that as $X \rightarrow \infty$ it has a peak at $x = 0$ that diverges as X/π , while its width falls to zero as π/X . Thus the area under the peak remains constant in the limit. The factor $1/2\pi$ assures that $\int dk_1 \delta(k_1 - k_2) = 1$.

Another example of a delta function is the limit of a very narrow Gaussian

$$\delta(x) = \lim_{a \rightarrow 0} \frac{1}{\sqrt{\pi}a} \exp\left(-\frac{x^2}{a^2}\right) \quad (2.105)$$

The prefactor normalizes the integral of the delta function to unity.

More useful properties of the delta function (all defined under an integral) are

$$\begin{aligned}
 \delta(x) &= \delta(-x) \\
 \delta'(x) &= -\delta'(-x) \\
 x\delta(x) &= 0 \\
 \delta(ax) &= \frac{1}{|a|}\delta(x) \\
 \delta(f(x)) &= \sum_j \frac{\delta(x - x_j)}{|f'(x_j)|}
 \end{aligned} \tag{2.106}$$

where the x_j 's are the zeros of the function in the argument of the delta function. Let us check the last of these identities, taking the limit of a Gaussian, Eq. 2.105, for our delta function:

$$\delta(f(x)) = \lim_{a \rightarrow 0} \frac{1}{\sqrt{\pi}a} \exp\left(-\frac{f(x)^2}{a^2}\right). \tag{2.107}$$

We Taylor expand $f(x)$ as

$$f(x) \simeq f(x_0) + (x - x_0)f'(x_0) \tag{2.108}$$

and if x_0 is the place where $f(x_0) = 0$, the delta function is

$$\delta(f(x)) = \frac{1}{\sqrt{\pi}a} \exp\left(-\frac{(x - x_0)^2 f'(x_0)^2}{a^2}\right). \tag{2.109}$$

To evaluate the integral $\int dx \delta(f(x))g(x)$, we Taylor expand both $f(x)$ and $g(x)$ about x_0 :

$$\begin{aligned}
 \int dx \delta(f(x))g(x) &= \int dx \frac{1}{\sqrt{\pi}a} \exp\left(-\frac{(x - x_0)^2 f'(x_0)^2}{a^2}\right) \\
 &\quad \times (g(x_0) + (x - x_0)g'(x_0) + \dots) \\
 &= \frac{g(x_0)}{\sqrt{\pi}a} \frac{\sqrt{\pi}a}{\sqrt{f'(x_0)^2}} + 0 + O(a^2) \\
 &= \frac{g(x_0)}{|f'(x_0)|},
 \end{aligned} \tag{2.110}$$

where we take the $a \rightarrow 0$ limit in the last step. This shows the origin of the extra factor in the denominator.

Parenthetically, this exercise illustrates that the simplest representation of a delta function as a tall, narrow step function, say of width ϵ and height $1/\epsilon$, is not optimal: it is not differentiable, and it is hard to retain the identities of delta functions of functions given above.

More about the coordinate operator

The coordinate operator \hat{x} gives the location of the particle in Cartesian space. A state which is diagonal in coordinate space is an eigenstate of \hat{x} with

$$\hat{x} |x'\rangle = x' |x'\rangle \quad (2.111)$$

and the matrix element of \hat{x} is

$$\langle x'' | \hat{x} | x' \rangle = x' \delta(x'' - x') = x' \langle x'' | x' \rangle. \quad (2.112)$$

We assume that the kets, $|x'\rangle$, form a complete basis. This means that the completeness relation allows us to write the state as

$$|\alpha\rangle = \int dx |x\rangle \langle x | \alpha \rangle. \quad (2.113)$$

Now suppose that we have an operator that records the location of a particle in an interval $x \in [-\Delta, \Delta]$. We can write this operator as

$$\hat{M}(x, \Delta) = \Theta(x + \Delta)\Theta(x - \Delta). \quad (2.114)$$

The expectation value of this operator in a state $|\alpha\rangle$ is

$$\langle \alpha | \hat{M} | \alpha \rangle = \int dx dx' \langle \alpha | x \rangle \langle x | \hat{M} | x' \rangle \langle x' | \alpha \rangle. \quad (2.115)$$

We insert the matrix element of the operator in states $|x\rangle$

$$\langle x | \hat{M} | x' \rangle = \delta(x - x')\Theta(x + \Delta)\Theta(x - \Delta) \quad (2.116)$$

to give

$$\begin{aligned} \langle \alpha | \hat{M} | \alpha \rangle &= \int dx dx' \langle \alpha | x \rangle \delta(x - x')\Theta(x + \Delta)\Theta(x - \Delta) \langle x' | \alpha \rangle \\ &= \int_{-\Delta}^{\Delta} dx \langle \alpha | x \rangle \langle x | \alpha \rangle \\ &= \int_{-\Delta}^{\Delta} dx |\langle \alpha | x \rangle|^2. \end{aligned} \quad (2.117)$$

If our states are normalized, this is just the probability to find a particle in the region $[-\Delta, \Delta]$. Thus the function $\langle x|\alpha\rangle$ is just what we would have identified using wave mechanics as the coordinate space wave function of the system, $\langle x|\alpha\rangle = \psi_\alpha(x)$. If we expand the region to $[-\infty, \infty]$ then the integral must equal unity – that is just the normalization constraint.

We can then write the following identities, which allow us to recover our wave-mechanical formulas as special cases of our Hilbert-space description,

$$\langle \beta|\alpha\rangle = \int dx \langle \beta|x\rangle \langle x|\alpha\rangle = \int dx \psi_\beta^*(x) \psi_\alpha(x) \quad (2.118)$$

and we have the further useful transcriptions

$$\begin{aligned} |\alpha\rangle &= \sum_j |j\rangle \langle j|\alpha\rangle = \sum_j |j\rangle c_j(\alpha), \\ \langle x|\alpha\rangle &= \sum_j \langle x|j\rangle \langle j|\alpha\rangle = \sum_j \psi_j(x) c_j(\alpha) = \psi_\alpha(x). \end{aligned} \quad (2.119)$$

Matrix elements can be constructed in coordinate basis:

$$\begin{aligned} \langle \beta|\hat{A}|\alpha\rangle &= \int dx dx' \langle \beta|x\rangle \langle x|\hat{A}|x'\rangle \langle x'|\alpha\rangle \\ &= \int dx dx' \psi_\beta^*(x) \langle x|\hat{A}|x'\rangle \psi_\alpha(x'). \end{aligned} \quad (2.120)$$

We must now analyze coordinate space matrix elements $\langle x|\hat{A}|x'\rangle$. If \hat{A} is just a function of the position operator, we merely have

$$\langle x|\hat{A}|x'\rangle = A(x')\delta(x - x'), \quad (2.121)$$

so

$$\langle \beta|\hat{A}|\alpha\rangle = \int dx \psi_\beta(x)^* A(x) \psi_\alpha(x). \quad (2.122)$$

More complicated operators require more thought. Consider the case of $\langle x|\hat{p}|x'\rangle$. To evaluate it, consider the fundamental commutation relation expanded in coordinate space states

$$\begin{aligned} \langle \alpha|[\hat{p}, \hat{x}]|\beta\rangle &= \frac{\hbar}{i} \langle \alpha|\beta\rangle \\ &= \int dx dx' \langle \alpha|x\rangle \langle x|\hat{p}\hat{x} - \hat{x}\hat{p}|x'\rangle \langle x'|\beta\rangle \\ &= \int dx dx' [\langle \alpha|x\rangle \langle x|\hat{p}|x'\rangle x' \langle x'|\beta\rangle - \langle \alpha|x\rangle x \langle x|\hat{p}|x'\rangle \langle x'|\beta\rangle] \end{aligned} \quad (2.123)$$

The reader can check that the following substitution solves the equation:

$$\langle x|\hat{p}|x'\rangle = \delta(x-x')\frac{\hbar}{i}\frac{\partial}{\partial x}. \quad (2.124)$$

It follows, then, that

$$\begin{aligned} \langle x|\hat{p}|\alpha\rangle &= \int dx' \langle x|\hat{p}|x'\rangle \langle x'|\alpha\rangle \\ &= \int dx' \delta(x-x')\frac{\hbar}{i}\frac{\partial}{\partial x} \langle x'|\alpha\rangle \\ &= \frac{\hbar}{i}\frac{\partial}{\partial x}\psi_\alpha(x). \end{aligned} \quad (2.125)$$

Moreover, higher powers of the momentum can also be expressed as differential operators:

$$\langle \beta|\hat{p}|\alpha\rangle = \int dx \psi_\beta^*(x)\frac{\hbar}{i}\frac{\partial}{\partial x}\psi_\alpha(x) \quad (2.126)$$

and

$$\langle \beta|\hat{p}^n|\alpha\rangle = \int dx \psi_\beta^*(x)\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^n \psi_\alpha(x) \quad (2.127)$$

and so on.

Let us complete our re-derivation of the equations of wave mechanics. The typical classical non-relativistic Hamiltonian of a particle in an external potential is

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x) \quad (2.128)$$

We consider the operator equation

$$\begin{aligned} \langle x|\hat{H}|\psi\rangle &= E \langle x|\psi\rangle = E\psi(x) \\ &= \int dx' \langle x|\hat{H}|x'\rangle \langle x'|\psi\rangle \\ &= \int dx' [\langle x|\frac{\hat{p}^2}{2m}|x'\rangle + \delta(x-x')V(x')]\psi(x') \\ &= \int dx' [\delta(x-x')\frac{1}{2m}\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^2 + \delta(x-x')V(x')]\psi(x') \end{aligned} \quad (2.129)$$

which gives us the usual Schrödinger differential equation

$$\left[\frac{1}{2m}\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^2 + V(x) \right] \psi(x) = E\psi(x). \quad (2.130)$$

The simple harmonic oscillator

Almost everything in Nature can be approximated by a simple harmonic oscillator. Its Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2. \quad (2.131)$$

We wish to find the energy eigenvalues and construct the eigenstates, by solving

$$\hat{H} |n\rangle = E_n |n\rangle. \quad (2.132)$$

We will do this algebraically, through the construction of so-called “ladder operators”

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} + \frac{i\hat{p}}{m\omega} \right) \quad \hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} - \frac{i\hat{p}}{m\omega} \right) \quad (2.133)$$

which are (non-Hermitian) combinations of the coordinate and the momentum,

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}} (\hat{a} + \hat{a}^\dagger) \quad \hat{p} = i\sqrt{\frac{\hbar m\omega}{2}} (\hat{a}^\dagger - \hat{a}). \quad (2.134)$$

For reasons which will shortly be apparent, we will call the operator \hat{a} the “lowering operator” and the operator \hat{a}^\dagger the “raising operator.” The commutator of these operators is

$$[\hat{a}, \hat{a}^\dagger] = 1; \quad [\hat{a}, \hat{a}] = 0 = [\hat{a}^\dagger, \hat{a}^\dagger]. \quad (2.135)$$

The arbitrary normalization in Eq. 2.134 is chosen to make the first commutator simple. We also introduce the (obviously Hermitian) “number operator” $N = \hat{a}^\dagger \hat{a}$. This is just the Hamiltonian, rescaled:

$$\begin{aligned} \hat{N} &= \frac{m\omega}{2\hbar} \left(\hat{x}^2 + \frac{\hat{p}^2}{m^2\omega^2} \right) + \frac{i}{2\hbar} [\hat{x}, \hat{p}] \\ &= \frac{\hat{H}}{\hbar\omega} - \frac{1}{2} \end{aligned} \quad (2.136)$$

and so

$$\hat{H} = \hbar\omega \left(\hat{N} + \frac{1}{2} \right). \quad (2.137)$$

Obviously, eigenstates of the Hamiltonian will also be eigenstates of \hat{N} .

$$\hat{N} |n\rangle = n |n\rangle \rightarrow \hat{H} |n\rangle = \hbar\omega \left(n + \frac{1}{2} \right) |n\rangle. \quad (2.138)$$

To proceed, we consider commutators of the number operator with the raising and lowering operators:

$$[\hat{N}, \hat{a}] = [\hat{a}^\dagger \hat{a}, \hat{a}] = \hat{a}^\dagger [\hat{a}, \hat{a}] + [\hat{a}^\dagger, \hat{a}] \hat{a} = -\hat{a} \quad (2.139)$$

and

$$[\hat{N}, \hat{a}^\dagger] = \hat{a}^\dagger. \quad (2.140)$$

What do raising and lowering operators do to eigenstates of \hat{N} ? Consider

$$\begin{aligned} \hat{N} \hat{a}^\dagger |n\rangle &= ([\hat{N}, \hat{a}^\dagger] + \hat{a}^\dagger \hat{N}) |n\rangle \\ &= (\hat{a}^\dagger + \hat{a}^\dagger \hat{N}) |n\rangle \\ &= (n+1) \hat{a}^\dagger |n\rangle \end{aligned} \quad (2.141)$$

This shows us that $\hat{a}^\dagger |n\rangle$ is an un-normalized eigenstate of \hat{N} with eigenvalue $n+1$,

$$\hat{a}^\dagger |n\rangle = C_{n+1}^{(+)} |n+1\rangle. \quad (2.142)$$

The lowering operator behaves similarly:

$$\hat{N} \hat{a} |n\rangle = ([\hat{N}, \hat{a}] + \hat{a} \hat{N}) |n\rangle = (n-1) \hat{a} |n\rangle \quad (2.143)$$

so

$$\hat{a} |n\rangle = C_{n-1}^{(-)} |n-1\rangle. \quad (2.144)$$

To find the constants, consider the norm of the state $\hat{a} |n\rangle$:

$$\begin{aligned} \langle n | \hat{a}^\dagger \hat{a} |n\rangle &= (\langle n | \hat{a}^\dagger) (\hat{a} |n\rangle) = |C_{n-1}^{(-)}|^2 \langle n-1 | n-1\rangle \\ &= n \langle n | n\rangle = n. \end{aligned} \quad (2.145)$$

Therefore, (choosing a phase) the normalization factor is

$$C_{n-1}^{(-)} = \sqrt{n}. \quad (2.146)$$

An analogous calculation for the raising operator gives

$$\hat{a} |n\rangle = \sqrt{n} |n-1\rangle \quad (2.147)$$

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle. \quad (2.148)$$

Observe that we cannot lower the state forever, because the norm of the state (the inner product of bra and ket) must always be greater than 0. From Eq. 2.146, this norm is n .

Therefore, the minimum n must then be 0, and all the n 's must be integers (otherwise one could lower states to negative values of n). This tells us that the spectrum of the oscillator is set of equally spaced levels labeled by a single integer n whose value is $n \geq 0$

$$\hat{H} |N\rangle = E_n |n\rangle \quad E_n = \hbar\omega \left(n + \frac{1}{2} \right) \quad n = 0, 1, 2, \dots \quad (2.149)$$

Given the ground state $|0\rangle$, we can construct the rest of the states using raising operators:

$$\begin{aligned} \hat{a}^\dagger |0\rangle &= |1\rangle \\ \hat{a}^\dagger |1\rangle &= \sqrt{2} |2\rangle \rightarrow \frac{(\hat{a}^\dagger)^2}{\sqrt{2 \cdot 1}} |0\rangle = |2\rangle \\ |3\rangle &= \frac{\hat{a}^\dagger}{\sqrt{3}} |2\rangle = \frac{(\hat{a}^\dagger)^3}{\sqrt{3 \cdot 2 \cdot 1}} |0\rangle \end{aligned} \quad (2.150)$$

and so

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}} |0\rangle. \quad (2.151)$$

This construction can be used to find the coordinate space wave functions $\langle x|n\rangle \equiv \psi_n(x)$. To do this, use the differential form of the lowering operator acting on the ground state,

$$\hat{a} |0\rangle = 0 \rightarrow \langle x|\hat{a}|0\rangle = \sqrt{\frac{m\omega}{2\hbar}} \langle x|\hat{x} + \frac{i\hat{p}}{m\omega}|0\rangle = 0 \quad (2.152)$$

or

$$\left(x + \frac{\hbar}{m\omega} \frac{\partial}{\partial x} \right) \langle x|0\rangle = 0 \quad (2.153)$$

This is just a differential equation for $\langle x|0\rangle$, whose solution is

$$\langle x|0\rangle = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \exp \left[-\frac{1}{2} \frac{m\omega}{\hbar} x^2 \right]. \quad (2.154)$$

The rest of the states can be constructed using the raising operator. For example,

$$\langle x|1\rangle = \langle x|\hat{a}^\dagger|0\rangle = \sqrt{\frac{m\omega}{2\hbar}} \left(x - \frac{\hbar}{m\omega} \frac{\partial}{\partial x} \right) \langle x|0\rangle. \quad (2.155)$$

This construction will give the usual coordinate space wave functions which arose from a direct attack on H as a differential operator.

We can find the matrix elements of \hat{x} and \hat{p} with the help of the raising and lowering operators. (This is of enormous practical utility!) We first note that

$$\langle n' | \hat{a} | n \rangle = \sqrt{n} \delta_{n', n-1} \quad (2.156)$$

$$\langle n' | \hat{a}^\dagger | n \rangle = \sqrt{n+1} \delta_{n', n+1} \quad (2.157)$$

Then we write

$$\langle n' | \hat{x} | n \rangle = \sqrt{\frac{\hbar}{2m\omega}} (\sqrt{n} \delta_{n', n-1} + \sqrt{n+1} \delta_{n', n+1}) \quad (2.158)$$

The matrix form of this expression is

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}} \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 1 & 0 & \sqrt{2} & 0 & \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \\ 0 & 0 & \sqrt{3} & 0 & \\ \vdots & & & & \ddots \end{pmatrix} \quad (2.159)$$

Powers of \hat{x} or \hat{p} are equally simple to construct.

Applications of oscillators

Many systems share this algebraic structure. Let us consider a few examples:

1. Two simple harmonic oscillators : The Hamiltonian is

$$\hat{H} = \frac{\hat{p}_1^2}{2m} + \frac{1}{2} m \omega^2 \hat{x}_1^2 + \frac{\hat{p}_2^2}{2m} + \frac{1}{2} m \omega^2 \hat{x}_2^2 \quad (2.160)$$

To solve this system, we note that coordinates and momenta for different oscillators commute, $[x_1, p_2] = 0$. This means that we can use two sets of commuting ladder operators, one for each oscillator: $[a_1^\dagger, a_2] = 0$ States are labeled by two n 's, one for each oscillator, and $E |n_1, n_2\rangle = (\hbar\omega_1(n_1 + \frac{1}{2}) + \hbar\omega_2(n_2 + \frac{1}{2})) |n_1, n_2\rangle$.

2. N uncoupled simple harmonic oscillators: Just let the subscript (x_i, p_i) be N -fold, and repeat the previous case.
3. Coupled oscillators: Suppose that

$$H = \frac{1}{2} \sum_k p_k^2 + \sum_{i,j} C_{ij} x_i x_j. \quad (2.161)$$

Perform a canonical transformation, or a simple change of variables, to diagonalize the potential: $V(\{x_i\}) = \sum_{i,j} C_{ij} x_i x_j \rightarrow \sum_k \frac{1}{2} \Omega_k z_k^2$. The canonically conjugate pairs will be $\{z_k, P_k\}$. The Hamiltonian is now diagonal in the new basis, $E = \sum_k \hbar \Omega_k (n_k + \frac{1}{2})$. We introduce ladder operators in the $\{z_k, P_k\}$ basis. States are labeled $|n_1, n_2 \dots\rangle$.

4. Classical systems obeying the differential equation

$$\frac{d^2 y_k}{dt^2} + \Omega_k^2 y_k = 0 : \quad (2.162)$$

These systems can be defined through a Hamiltonian $H = \frac{1}{2} \sum_k (p_k^2 + \Omega_k^2 y_k^2)$ and are obviously also oscillators, with states labeled by an integer for each oscillator.

5. Photons: Each mode of the classical electromagnetic field (working with the vector potential in Coulomb gauge) obeys an equation of motion

$$\frac{d^2 \vec{A}(\vec{k}, \vec{\epsilon})}{dt^2} + \omega_k^2 \vec{A}(\vec{k}, \vec{\epsilon}) = 0 \quad (2.163)$$

From what we have done so far, the quantum mechanical state of each mode of the electromagnetic field must be characterized by a ket $|n_{\vec{k}, \vec{\epsilon}}\rangle$. The extra letters \vec{k} and $\vec{\epsilon}$ label the wave number and polarization of the mode.

The energy of the mode is given by an integer multiple of $\hbar\omega_k$, $E_n = n\hbar\omega_k$ up to an overall constant. In the usual language of Planck and Einstein, the integer n counts the number of photons in the mode. We do not have to restrict ourselves to a single mode. In complete analogy with the case of coupled oscillators, we can also have a state with many photons in many modes.

$$|n_1, n_2, n_3, \dots\rangle \quad (2.164)$$

with energy $\sum_{\vec{k}, \vec{\epsilon}} \hbar\omega_k (n_{\vec{k}, \vec{\epsilon}} + \frac{1}{2})$. (In contrast to the two-oscillator problem, the labels \vec{k} and $\vec{\epsilon}$ can vary continuously.)

The quantum mechanical analogs of coordinates and momenta are operators. Coordinates and momenta are the canonically conjugate variables describing the classical system's Hamiltonian dynamics. One can construct a Hamiltonian description of the classical electromagnetic field. Not surprisingly, its canonical coordinates are the vector potential and the electric field. Thus in the quantum description of the electromagnetic field, the vector potential is an operator which acts on the many-photon wave function.

Again, it is no surprise, the vector potential is a linear superposition of raising and lowering operators. The raising operator increases the number of photons in the state

$$\hat{a}_{\vec{k},\vec{\epsilon}}^\dagger |n_{\vec{k},\vec{\epsilon}}\rangle = \sqrt{n_{\vec{k},\vec{\epsilon}} + 1} |n_{\vec{k},\vec{\epsilon}} + 1\rangle \quad (2.165)$$

while the lowering operator decreases the number of photons in the state,

$$\hat{a}_{\vec{k},\vec{\epsilon}} |n_{\vec{k},\vec{\epsilon}}\rangle = \sqrt{n_{\vec{k},\vec{\epsilon}}} |n_{\vec{k},\vec{\epsilon}} - 1\rangle \quad (2.166)$$

We will use these results when we consider the fully quantum mechanical problem of the emission and absorption of electromagnetic radiation from matter.

A few words about eigenfunctions of momentum

Our final example of the chapter is the free particle,

$$H = \frac{p^2}{2m}. \quad (2.167)$$

Obviously, the momentum operator itself commutes with H , so energy eigenstates will also be momentum eigenstates. Consider, then, eigenstates of the momentum operator (in one dimension, to simplify notation)

$$\hat{p} |p'\rangle = p' |p'\rangle. \quad (2.168)$$

They occupy, of course, an infinite-dimensional Hilbert space. In complete analogy with coordinate space expansions, any abstract state can be expanded in momentum eigenstates

$$\begin{aligned} |\alpha\rangle &= \int dp' |p'\rangle \langle p'|\alpha\rangle \\ &= \int dp' |p'\rangle \psi_\alpha(p') \end{aligned} \quad (2.169)$$

where $\psi_\alpha(p')$ is the momentum space wave function. Repeating an earlier construction, it is straightforward to show that the position operator in momentum space is

$$\langle p'|\hat{x}|p\rangle = -\frac{\hbar}{i}\delta(p' - p)\frac{\partial}{\partial p}. \quad (2.170)$$

Now we want to find the wave function $\langle x|p\rangle$, a momentum eigenfunction in coordinate space representation. We start with $\hat{p}|p\rangle = p'|p\rangle$ and take the inner product with $\langle x|$:

$$\langle x|\hat{p}|p\rangle = p \langle x|p\rangle. \quad (2.171)$$

We process the left hand side using

$$\int dx' \langle x|\hat{p}|x'\rangle \langle x'|p\rangle = \int dx' \delta(x-x') \frac{\hbar}{i} \frac{\partial}{\partial x} \langle x'|p\rangle. \quad (2.172)$$

This gives

$$\frac{\hbar}{i} \frac{\partial}{\partial x} \langle x|p\rangle = p \langle x|p\rangle \quad (2.173)$$

or

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \exp\left(\frac{ipx}{\hbar}\right). \quad (2.174)$$

The normalization factor is chosen so that $\langle x|x'\rangle = \delta(x-x') = \int dp \langle x|p\rangle \langle p|x'\rangle$. This result is used to pass between position space wave functions and momentum space ones. We find

$$\psi_\alpha(x) = \langle x|\alpha\rangle = \int dp \langle x|p\rangle \langle p|\alpha\rangle = \int dp \frac{1}{\sqrt{2\pi\hbar}} \exp\left(\frac{ipx}{\hbar}\right) \psi_\alpha(p) \quad (2.175)$$

The switch between $\psi_\alpha(x)$ and $\psi_\alpha(p)$ is a simple Fourier transformation.

Chapter 3

Time dependence in quantum mechanics

Quantum dynamics

So far, we have dealt with three kinds of quantum mechanical objects:

- Vectors, which specify the state of the system.
- Dynamical variables, which correspond to operators acting on the state vectors, transforming them into new state vectors
- Unitary transformations, which, at least as far as we have discussed them, do not change the state of the system, but change how we look at the state of the system.

So far, we have not discussed time evolution in quantum mechanics. While the motion of particles can only be described probabilistically, the evolution of the probability itself evolves causally. There are several equivalent ways of describing this causal evolution. We can equivalently regard the states as changing with time, or the dynamical variables as changing with time, or make a combination of the two pictures. In the case that we put all the time evolution onto the states, we can describe the evolution of the state directly, or we can encode the time evolution into the action of a unitary operator, acting on the states.

Schrödinger picture

We begin by assuming that the state vectors are the fundamental quantities which evolve in time. How they do so is governed by the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\alpha_s(t)\rangle = \hat{H} |\alpha_s(t)\rangle. \quad (3.1)$$

There is a dual equation for the bra vectors,

$$-i\hbar \frac{\partial}{\partial t} \langle \alpha_s(t)| = \langle \alpha_s(t)| \hat{H}^\dagger = \langle \alpha_s(t)| \hat{H}. \quad (3.2)$$

Again, the last equality is true only if the Hamiltonian is Hermitian.

Eq. 3.1 may be formally integrated to give

$$\psi(x, t) = \exp\left(-\frac{i\hat{H}}{\hbar}(t - t_0)\right)\psi(t_0). \quad (3.3)$$

However, the reader ought to treat this formal integration with great caution. The Hamiltonian is an operator, and the exponential of an operator is also an operator. In principle, the exponential is only well defined as a power series,

$$\exp -\frac{i\hat{H}}{\hbar}t = \hat{1} - \frac{i\hat{H}}{\hbar}t - \frac{1}{2!} \frac{\hat{H}^2}{\hbar^2}t^2 + \dots \quad (3.4)$$

It may be quite difficult to evaluate this series in a useful way, especially if H has explicit time dependence.

There is one extremely useful special case of time evolution. If the Hamiltonian is time-independent, and if $|\psi\rangle$ can be written as a superposition of eigenfunctions of the Hamiltonian

$$\hat{H}|\psi\rangle = \sum_n c_n E_n |E_n\rangle, \quad (3.5)$$

then the solution of Eq. 3.1 is very simple:

$$|\psi(t)\rangle = \sum_n c_n e^{-\frac{iE_n t}{\hbar}} |E_n\rangle \quad (3.6)$$

Each of the different energy eigenfunctions composing $|\psi\rangle$ evolves independently in time, and its evolution is a pure change of phase. Note that in this case, the probability of observing the state to be in any particular energy eigenstate is a time-independent constant. Projecting $|\psi(t)\rangle$ onto a state which is not an energy eigenstate will reveal interesting interference effects between the different energy eigenstates.

A general (but formal) way to rewrite the time-dependent Schrödinger equation can be given by introducing an “evolution operator”, $\hat{U}(t, t_0)$. The action of this operator is to move a state forward in time by a specified amount

$$|\alpha_s(t)\rangle = \hat{U}(t, t_0) |\alpha_s(t_0)\rangle \quad (3.7)$$

Inserting this into the Schrödinger equation, we have

$$i\hbar \frac{\partial}{\partial t} |\alpha_s(t)\rangle = \left[i\hbar \frac{\partial}{\partial t} \hat{U}(t, t_0) \right] |\alpha_s(t_0)\rangle = \hat{H} \hat{U}(t, t_0) |\alpha_s(t_0)\rangle \quad (3.8)$$

or

$$i\hbar \frac{\partial \hat{U}(t, t_0)}{\partial t} = \hat{H} \hat{U}(t, t_0). \quad (3.9)$$

The utility of this formalism is that if we can once evaluate $\hat{U}(t, t_0)$, any state may be carried forward in time. Formally, of course,

$$\hat{U}(t, t_0) = \exp\left(-\frac{i\hat{H}}{\hbar}(t - t_0)\right) \quad (3.10)$$

and if \hat{H} is Hermitian, \hat{U} is unitary.

Spin- $\frac{1}{2}$ particle in a magnetic field

Let us consider a simple example of time evolution, the two-state system with a time-independent Hamiltonian. We will specifically consider the spin-1/2 electron in an external magnetic field. The Hamiltonian is

$$\hat{H} = -\vec{\mu} \cdot \vec{B} = \frac{e\hbar}{2mc} \vec{\sigma} \cdot \vec{B} \quad (3.11)$$

and the negative charge of the electron flips the sign in the rightmost expression. The dot product has the explicit expansion

$$\vec{\sigma} \cdot \vec{B} = \begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix}. \quad (3.12)$$

Let us compute the time evolution of states under this Hamiltonian, in several different ways.

It is useful to introduce the unit vector $\vec{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ and to rewrite the Hamiltonian as

$$\hat{H} = \frac{e\hbar}{2mc} \vec{\sigma} \cdot \vec{B} = \hbar\omega \vec{\sigma} \cdot \vec{n}. \quad (3.13)$$

Because the Hamiltonian has no explicit time dependence, we can straightforwardly write the time evolution operator as

$$\hat{U} = e^{-\frac{i\hat{H}t}{\hbar}} = e^{-i\omega \vec{\sigma} \cdot \vec{n} t} \quad (3.14)$$

To evaluate this expression, we expand in a Taylor series

$$e^{-i\omega \vec{\sigma} \cdot \vec{n} t} = \hat{1} - i\omega t \vec{\sigma} \cdot \vec{n} + \frac{1}{2}(i\omega t)^2 \hat{\sigma}_i n_i \hat{\sigma}_j n_j + \dots \quad (3.15)$$

Recalling that $\hat{\sigma}_i \hat{\sigma}_j = i\epsilon_{ijk} \hat{\sigma}_k + \hat{1} \delta_{ij}$, we write

$$e^{-i\omega \vec{\sigma} \cdot \vec{n} t} = \hat{1} - i\omega t \vec{\sigma} \cdot \vec{n} + \frac{1}{2}(i\omega t)^2 (i\epsilon_{ijk} \hat{\sigma}_k + \hat{1} \delta_{ij}) n_i n_j + \dots \quad (3.16)$$

Since $\epsilon_{ijK} n_i n_j = 0$, this is

$$= \hat{1} - i\omega t \vec{\sigma} \cdot \vec{n} + \frac{1}{2}(i\omega t)^2 + \dots \quad (3.17)$$

The next term in the series as $\frac{1}{3!}(-i\omega t)^3 (\vec{\sigma} \cdot \vec{n})^3 = \frac{1}{3!}(-i\omega t)^3 (\vec{\sigma} \cdot \vec{n})$. Thus the time evolution operator is

$$\begin{aligned} \hat{U} &= \hat{1} \left(1 - \frac{1}{2}\omega^2 t^2 + \dots \right) + \vec{\sigma} \cdot \hat{n} \left(-i\omega t + \frac{(-i\omega t)^3}{3!} + \dots \right) \\ &= \hat{1} \cos(\omega t) - i\vec{\sigma} \cdot \hat{n} \sin(\omega t). \end{aligned} \quad (3.18)$$

Expressing our result in matrix form, we have

$$\begin{aligned}\hat{U} &= \begin{pmatrix} \cos \omega t - i n_z \sin \omega t & \sin \omega t (-i n_x - n_y) \\ \sin \omega t (-i n_x + n_y) & \cos \omega t + i n_z \sin \omega t \end{pmatrix} \\ &= \begin{pmatrix} \cos \omega t - i \cos \theta \sin \omega t & -i \sin \omega t \sin \theta e^{-i\phi} \\ -i \sin \omega t \sin \theta e^{i\phi} & \cos \omega t + i \cos \theta \sin \omega t \end{pmatrix}.\end{aligned}\quad (3.19)$$

This result can be used to compute the time evolution of any state. For example, suppose that at time $t = 0$ the electron is spin-up. Then $|\psi(t=0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and at a later time t

$$|\psi(t)\rangle = \begin{pmatrix} \cos \omega t - i \cos \theta \sin \omega t \\ -i \sin \omega t \sin \theta e^{i\phi} \end{pmatrix}.\quad (3.20)$$

The amplitude for the particle to be spin up at a later time t is

$$\langle \psi_+ | \psi(t) \rangle = \cos \omega t - i \cos \theta \sin \omega t \quad (3.21)$$

and the resulting probability is

$$|\langle \psi_+ | \psi(t) \rangle|^2 = \cos^2 \omega t + \cos^2 \theta \sin^2 \omega t = 1 - \sin^2 \theta \sin^2 \omega t.\quad (3.22)$$

Notice how the probability oscillates with time, if $\theta \neq 0$. This oscillation occurs because the state $|\psi(+)\rangle$ is not an energy eigenstate unless $\theta = 0$; it is a mixture of the two energy eigenstates. The fraction of the amplitude associated with each eigenstate is varying with time, and the two components' interference is time-dependent.

We can also evaluate the evolution operator a second way: use completeness twice and write the evolution operator as

$$\hat{U} = e^{\frac{-i\hat{H}t}{\hbar}} = \sum_{i,j} |i\rangle \langle i| e^{\frac{-i\hat{H}t}{\hbar}} |j\rangle \langle j|.\quad (3.23)$$

In the sums, pick the $|i\rangle$'s and the $|j\rangle$'s to be energy eigenfunctions. In that case,

$$\hat{U}(t) = \sum_i |i\rangle e^{\frac{-iE_i t}{\hbar}} \langle i|.\quad (3.24)$$

To proceed, we need the eigenfunctions and eigenvalues of the Hamiltonian. The Hamiltonian matrix is

$$\begin{aligned}H &= \hbar\omega \begin{pmatrix} n_z & n_x - i n_y \\ n_x + i n_y & -n_z \end{pmatrix} \\ &= \hbar\omega \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix}\end{aligned}\quad (3.25)$$

where on the second line we re-introduce the direction vector \vec{n} of the magnetic field in spherical coordinates. Recalling our discussion in the previous chapter, we discover quickly that our eigenvalues are $\lambda_{\pm} = \pm\hbar\omega$ and our eigenstates are

$$|\psi_{+}\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\phi} \end{pmatrix} \quad (3.26)$$

and

$$|\psi_{-}\rangle = \begin{pmatrix} -\sin \frac{\theta}{2} e^{-i\phi} \\ \cos \frac{\theta}{2} \end{pmatrix}. \quad (3.27)$$

A parenthetical remark: We used Eq. 3.26 with its funny half angle for θ in our discussion in Chapter 2. Now we see that this choice actually contained physics: The angles θ and ϕ parameterize the direction of the magnetic field – or, more generally, the “direction” of the Hamiltonian in our two dimensional Hilbert space. Eq. 3.26 is called the “Bloch sphere” parameterization of the spinor wave function. In fact, one can define a vector (the “Bloch vector”)

$$\vec{p} = \langle \psi_{+} | \vec{\sigma} | \psi_{+} \rangle = \hat{z} \cos \theta + \hat{x} \sin \theta \cos \phi + \hat{y} \sin \theta \sin \phi \quad (3.28)$$

to label the orientation of the state in the Bloch sphere. To visualize it, return to Eq. 3.25 and look at the direction vector \vec{n} . When \vec{n} points along the positive z direction (the north pole of the sphere) the positive energy eigenstate is

$$|\psi_{+}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (3.29)$$

and when \vec{n} points to the south pole, $\theta = \pi$ and

$$|\psi_{+}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (3.30)$$

When the field has no \hat{z} component the state is an equal (up to a phase) superposition of “up” and “down” states and the Bloch vector points along the equator.

We then reconstruct \hat{U} entry by entry. For example, we can recover the solution to our first example by taking the matrix element of \hat{U} between $|\psi(+)\rangle$ (the value of the state at $t = 0$) and $|\psi(+)\rangle$ (the observed state at time t):

$$\langle \uparrow | \hat{U} | \uparrow \rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \hat{U} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (3.31)$$

$$\begin{aligned}
&= \sum_i \langle \uparrow | i \rangle e^{\frac{iE_i t}{\hbar}} \langle i | \uparrow \rangle \\
&= \cos^2\left(\frac{\theta}{2}\right) e^{-i\omega t} + \sin^2\left(\frac{\theta}{2}\right) e^{i\omega t} \\
&= \cos \omega t - i \sin \omega t \cos \theta
\end{aligned}$$

as before.

Magnetic resonance

In many physical situations, the Hamiltonian is time dependent. In these cases it can be quite difficult to solve the equation of motion exactly. An exception is (again) the two state system. As an example of interesting dynamics, consider the situation of magnetic resonance. We have a spin-1/2 particle (we will call it an electron) in an external magnetic field which consists of a large constant term pointing in the \hat{z} direction and a small component which rotates in the $x - y$ plane, with an angular frequency ν . Explicitly, the magnetic field is

$$B_z = B \quad B_x = b \cos(\nu t) \quad B_y = b \sin(\nu t). \quad (3.32)$$

Then our Hamiltonian is

$$\hat{H} = \hbar\omega\hat{\sigma}_z + \hbar\Omega_0(\hat{\sigma}_x \cos(\nu t) + \hat{\sigma}_y \sin(\nu t)) = \hbar \begin{pmatrix} \omega & \Omega_0 e^{-i\nu t} \\ \Omega_0 e^{i\nu t} & -\omega \end{pmatrix} \quad (3.33)$$

where we have defined $\hbar\omega = eB\hbar/(2mc)$ and $\hbar\Omega_0 = eb\hbar/(2mc)$. In the magnetic resonance literature, Ω_0 is called the Rabi frequency. A general solution can be written as

$$|\psi(t)\rangle = \begin{pmatrix} a_+(t) \\ a_-(t) \end{pmatrix} \quad (3.34)$$

and the corresponding equations of motion are

$$i\dot{a}_+ = \omega a_+ + \Omega_0 e^{-i\omega t} a_- \quad (3.35)$$

$$i\dot{a}_- = \Omega_0 e^{i\omega t} a_+ - \omega a_- \quad (3.36)$$

To find a solution, we assume that

$$a_{\pm} = A_{\pm} e^{-i\lambda_{\pm} t} \quad (3.37)$$

so that

$$\begin{pmatrix} \omega - \lambda_+ & \Omega_0 e^{-i(\nu - \lambda_+ + \lambda_-)t} \\ \Omega_0 e^{i(\nu - \lambda_+ + \lambda_-)t} & -\omega - \lambda_- \end{pmatrix} \begin{pmatrix} A_+ \\ A_- \end{pmatrix} = 0. \quad (3.38)$$

We eliminate the time dependence entirely by making the ansatz that $\lambda_+ = \lambda_- + \nu$. Then we are left with a simple matrix equation

$$\begin{pmatrix} \omega - \lambda_- - \nu & \Omega_0 \\ \Omega_0 & -\omega - \lambda_- \end{pmatrix} \begin{pmatrix} A_+ \\ A_- \end{pmatrix} = 0. \quad (3.39)$$

Let us split this equation apart to write it in a more familiar way:

$$\begin{pmatrix} \omega - \nu & \Omega_0 \\ \Omega_0 & -\omega \end{pmatrix} \begin{pmatrix} A_+ \\ A_- \end{pmatrix} = \lambda_- \begin{pmatrix} A_+ \\ A_- \end{pmatrix} \quad (3.40)$$

or

$$H \begin{pmatrix} A_+ \\ A_- \end{pmatrix} = E \begin{pmatrix} A_+ \\ A_- \end{pmatrix} \quad (3.41)$$

with $E = \hbar\lambda_-$ and

$$H = -\frac{\hbar\nu}{2}\mathbf{1} + \hbar(\omega - \frac{1}{2}\nu)\sigma_z + \hbar\Omega_0\sigma_x. \quad (3.42)$$

This is the eigenvalue equation we found for a two state system when its Hamiltonian was constant in time. The solution is the same: the determinant of Eq. 3.39 must vanish, giving

$$\lambda_- = -\frac{\nu}{2} \pm \Delta \quad \Delta \equiv \sqrt{\left(\omega - \frac{1}{2}\nu\right)^2 + \Omega_0^2}. \quad (3.43)$$

Because there are two possible λ_- 's, the most general solution to the equation will be a linear combination of the energy eigenfunctions. Thus we have two unknown coefficients to determine,

$$a_-(t) = A_-^{(1)} e^{-i(-\frac{\nu}{2} + \Delta)t} + A_-^{(2)} e^{-i(-\frac{\nu}{2} - \Delta)t}. \quad (3.44)$$

We can fix $A_-^{(1)}$ and $A_-^{(2)}$ from the boundary conditions at $t = 0$ plus the normalization condition. Suppose, for example, that at $t = 0$ the electron has spin up:

$$|\psi(t=0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (3.45)$$

This implies

$$a_-(t=0) = 0 \rightarrow A_-^{(1)} = A_-^{(2)} \quad (3.46)$$

or

$$a_-(t) = A e^{\frac{i\nu t}{2}} \sin(\Delta t). \quad (3.47)$$

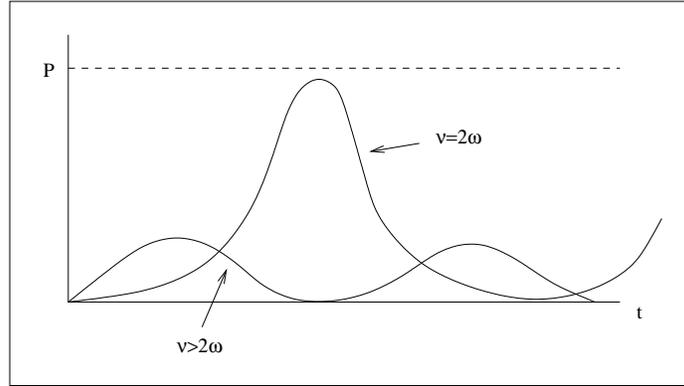


Figure 3.1: Variation of probability $P(\downarrow)$ with time in nuclear magnetic resonance.

An easy way to determine A is to return to our original equations and evaluate them at $t = 0$. We had

$$i\dot{a}_-(0) = \Omega_0 a_+(0) + i\omega a_-(0) \quad (3.48)$$

and since $a_-(t = 0) = 0$ and $a_+(t = 0) = 1$, $\Delta A = \Omega_0$ and

$$a_-(t) = \frac{\Omega_0}{\Delta} \sin(\Delta t) e^{\frac{i\nu t}{2}}. \quad (3.49)$$

Thus the probability that the spin of the electron is down at any time t is

$$P(\downarrow) = |a_-(t)|^2 = \frac{\Omega_0^2}{\Delta^2} \sin^2(\Delta t). \quad (3.50)$$

This result is usually written as

$$P(\downarrow) = \frac{\Omega_0^2}{\Omega^2} \sin(\Omega t), \quad \Omega = \sqrt{\left(\omega - \frac{1}{2}\nu\right)^2 + \Omega_0^2}. \quad (3.51)$$

Notice the behavior of $P(\downarrow)$ – especially its maximum value – on the frequency ν and strength of rotating field. See Fig. 3.1. The functional form of Ω_0^2/Ω^2 is called a “Lorentzian”

$$\frac{\Omega_0^2}{\Omega^2} = \frac{1}{(\nu - \nu_0)^2 + \Gamma^2/4} \quad (3.52)$$

and is shown in Fig. 3.2. The quantity Γ is called the “full width at half maximum” of the curve. Notice that as the strength of the rotating field falls to zero, the transition probability becomes very very tightly peaked as a function of frequency. It becomes vanishingly small unless the driving frequency corresponds to the frequency difference of the two states. This

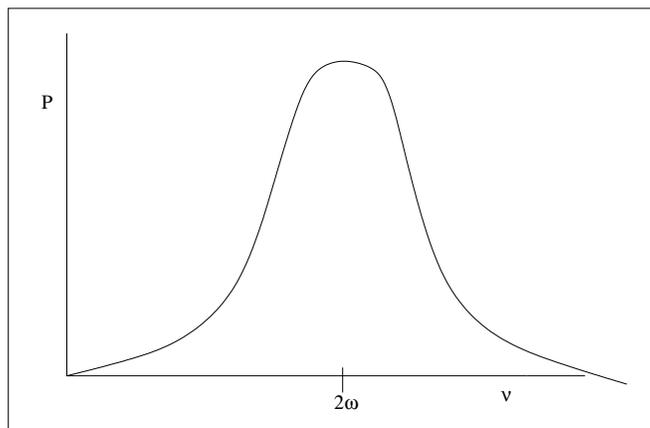


Figure 3.2: The Lorentzian line shape.

“resonance behavior” occurs at $\hbar\nu = 2\hbar\omega$. Physically, the rotating field must supply the right energy $\hbar\nu = \hbar\omega - (-\hbar\omega)$ to carry the electron from the lower state to the upper state, and then it can return the energy to the oscillating field as it falls back to the ground state.

Magnetic resonance has many analogs. One commonly-seen example is the case of a two-state atom in an external electromagnetic field. The electromagnetic field oscillates with frequency ν . The Hamiltonian has two parts. One term (\hat{H}_1) specifies the energy levels of the atom and the other parameterizes the interaction of the electron and the electromagnetic field (\hat{H}_2). Obviously, the first term can be written as

$$\hat{H}_1 = \begin{pmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_2 \end{pmatrix}. \quad (3.53)$$

The interaction between the electron and the external electric field is a dipole interaction $-\vec{p} \cdot \vec{E}$ between the atom and the field. Most systems do not have permanent dipole moments, so $\langle i|\vec{p}|i\rangle = 0$. However, the off-diagonal terms need not vanish. We can parametrize this situation as

$$\hat{H}_2 = \begin{pmatrix} 0 & \Delta e^{-i\nu t} \\ \Delta^* e^{i\nu t} & 0 \end{pmatrix} \quad (3.54)$$

This Hamiltonian is exactly what we have just analyzed, but with the substitutions

$$\Omega_0 \leftrightarrow \Delta \quad \omega \leftrightarrow \frac{\epsilon_1 - \epsilon_2}{2}. \quad (3.55)$$

Heisenberg picture

In the Schrödinger picture, states evolve as

$$|\alpha_s(t)\rangle = e^{-\frac{i\hat{H}t}{\hbar}} |\alpha_s(0)\rangle, \quad (3.56)$$

while operators without explicit time dependence remain constant in time. For example, $x(t)$ is just x . How then, do matrix elements of operators change in the Schrödinger picture? Just compute the time derivative of the matrix element of an arbitrary operator

$$\begin{aligned} \frac{\partial}{\partial t} \langle \alpha_s(t) | \hat{O} | \beta_s(t) \rangle &= \left(\frac{\partial \langle \alpha |}{\partial t} \right) | \hat{O} | \beta \rangle + \langle \alpha | \frac{\partial \hat{O}}{\partial t} | \beta \rangle + \langle \alpha | \hat{O} \left(\frac{\partial | \beta \rangle}{\partial t} \right) \\ &= -\langle \alpha | \frac{\hat{H} \hat{O}}{i\hbar} | \beta \rangle + \langle \alpha | \frac{\partial \hat{O}}{\partial t} | \beta \rangle + \langle \alpha | \frac{\hat{O} \hat{H}}{i\hbar} | \beta \rangle \\ &= -\frac{1}{i\hbar} \langle \alpha | [\hat{H}, \hat{O}] | \beta \rangle + \langle \alpha | \frac{\partial \hat{O}}{\partial t} | \beta \rangle \end{aligned} \quad (3.57)$$

The time dependence arises first because the states themselves evolve with time, and second, because the operator itself might have explicit time dependence. We observe from this calculation a very important special case: if \hat{O} has no explicit time dependence and if it commutes with the Hamiltonian, its matrix elements are time independent. The operator \hat{O} is a constant of motion. This has important practical applications in terms of how we characterize states: they can be simultaneous eigenstates of the energy and of all dynamical variables which commute with \hat{H} .

Notice that we can re-write the operator equation as

$$\begin{aligned} \frac{d}{dt} \langle \alpha_s(t) | \hat{O} | \beta_s(t) \rangle &= \frac{d}{dt} \langle \alpha_s(0) | e^{\frac{i\hat{H}t}{\hbar}} \hat{O} e^{-\frac{i\hat{H}t}{\hbar}} | \beta_s(0) \rangle \\ &= \langle \alpha_s(0) | e^{\frac{i\hat{H}t}{\hbar}} \frac{d\hat{O}}{dt} e^{-\frac{i\hat{H}t}{\hbar}} | \beta_s(0) \rangle + \frac{1}{i\hbar} \langle \alpha_s(0) | [e^{\frac{i\hat{H}t}{\hbar}} \hat{O} e^{-\frac{i\hat{H}t}{\hbar}}, \hat{H}] | \beta_s(0) \rangle. \end{aligned} \quad (3.58)$$

We regroup the expression by defining time independent states

$$|\alpha_H(t)\rangle = |\alpha_s(0)\rangle \quad (3.59)$$

and time dependent dynamical variables,

$$\hat{O}_H = e^{\frac{i\hat{H}t}{\hbar}} \hat{O} e^{-\frac{i\hat{H}t}{\hbar}} \quad (3.60)$$

These definitions express the “Heisenberg representation” of quantum mechanics in which the states are time-independent, but the operators are explicitly time-dependent. Then Eq. 3.58 becomes

$$\begin{aligned}\frac{d}{dt} \langle \alpha_H | \hat{O}_H | \beta_H \rangle &= \langle \alpha_H | \frac{d\hat{O}_H}{dt} | \beta_H \rangle \\ &= \langle \alpha_H | \frac{\partial \hat{O}_H}{\partial t} | \beta_H \rangle + \frac{1}{i\hbar} \langle \alpha_H | [\hat{O}_H, \hat{H}] | \beta_H \rangle\end{aligned}\quad (3.61)$$

This result holds for any states $|\alpha\rangle$, $|\beta\rangle$, so it must be true for the operators themselves that their time evolution is given by

$$\frac{d\hat{O}_H}{dt} = \frac{\partial \hat{O}_H}{\partial t} + \frac{1}{i\hbar} [\hat{O}_H, \hat{H}]. \quad (3.62)$$

Taking the states to be time-independent and pushing all the time evolution onto the operators produces the Heisenberg representation, and the operator equation of motion is called the “Heisenberg equation of motion.” It closely resembles the Poisson bracket formalism of classical mechanics, where dynamical variables evolve through

$$\frac{dA}{dt} = \{A, H\} \rightleftharpoons \frac{1}{i\hbar} [\hat{O}_H, \hat{H}] \quad (3.63)$$

(and we remind the reader of the transcription between Poisson bracket and commutator). Note, however, that Heisenberg equations of motion will also govern object without classical analogues (for example, spin).

As a useful example, suppose our Hamiltonian is that of a non-relativistic particle in an external potential

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \quad (3.64)$$

The Heisenberg equations of motion for the momentum and position operators are

$$\begin{aligned}\frac{d\hat{p}}{dt} &= \frac{1}{i\hbar} [\hat{p}, \hat{H}] = -\frac{\partial V}{\partial x} \\ \frac{d\hat{x}}{dt} &= \frac{1}{i\hbar} [\hat{x}, \hat{H}] = \frac{\hat{p}}{m} \\ \frac{d^2\hat{x}}{dt^2} &= \frac{1}{i\hbar} \left[\frac{d\hat{x}}{dt}, \hat{H} \right] = \frac{1}{m} \frac{d\hat{p}}{dt}.\end{aligned}\quad (3.65)$$

We can take an expectation value of the last expression. When we do so, we discover Ehrenfest’s theorem,

$$m \frac{d^2 \langle \hat{x} \rangle}{dt^2} = \frac{d \langle \hat{p} \rangle}{dt} = - \langle \nabla V \rangle \quad (3.66)$$

Expectation values of the operators must be picture independent, so this result will be true regardless of whether we evaluate it in Heisenberg or in Schrödinger representation.

Now we understand how the states and operators evolve in time in both pictures, but we need to determine the time evolution of the base kets (the axes in Hilbert space). Suppose we choose the base kets at $t = 0$ to be the eigenstates of some operator \hat{A} .

$$\hat{A} |a\rangle = a |a\rangle \quad (3.67)$$

In the Schrödinger picture, the operators do not change with time so the base kets must also not change with time. But in the Heisenberg picture the operators evolve in time according to

$$\hat{A}_H(t) = e^{\frac{i\hat{H}t}{\hbar}} \hat{A}_S e^{-\frac{i\hat{H}t}{\hbar}}. \quad (3.68)$$

Multiplying our eigenvalue equation on the left hand side by $e^{\frac{i\hat{H}t}{\hbar}}$, we observe that

$$\begin{aligned} e^{\frac{i\hat{H}t}{\hbar}} \hat{A} |a\rangle &= e^{\frac{i\hat{H}t}{\hbar}} a |a\rangle \\ &= e^{\frac{i\hat{H}t}{\hbar}} \hat{A} e^{-\frac{i\hat{H}t}{\hbar}} e^{\frac{i\hat{H}t}{\hbar}} |a\rangle = a e^{\frac{i\hat{H}t}{\hbar}} |a\rangle \\ \hat{A}_H \left(e^{\frac{i\hat{H}t}{\hbar}} |a\rangle \right) &= a \left(e^{\frac{i\hat{H}t}{\hbar}} |a\rangle \right) \end{aligned} \quad (3.69)$$

Thus in the Heisenberg picture the base kets must evolve “backward” in time as $e^{\frac{i\hat{H}t}{\hbar}} |a\rangle$.

Regardless of picture, we can expand a state in terms of its base kets as

$$|\psi(t)\rangle = \sum_n c_n(t) |n\rangle. \quad (3.70)$$

In the Schrödinger picture we can write this as

$$c_n(t) = \langle n | \left(e^{-\frac{i\hat{H}t}{\hbar}} |\psi(t=0)\rangle \right), \quad (3.71)$$

putting the time dependence onto the state $|\psi(t)\rangle$. To work in the Heisenberg picture, we group terms differently,

$$c_n(t) = \left(\langle n | e^{-\frac{i\hat{H}t}{\hbar}} \right) |\psi(t=0)\rangle \quad (3.72)$$

In the Heisenberg picture, the state does not change, $|\psi(t=0)\rangle = |\psi(t)\rangle$, but the base kets change. The expansion coefficients are the same in both the Schrödinger and Heisenberg pictures.

The calculation of matrix elements is also the same in either basis. Assume that the system is in state $|a\rangle$ at $t = 0$. What is the amplitude to be in an eigenstate of an operator \hat{B} with eigenvalue b at a later time? In the Schrödinger picture this is

$$\langle b | \left(e^{-\frac{i\hat{H}t}{\hbar}} |a\rangle \right) \quad (3.73)$$

while in the Heisenberg picture we have

$$\left(\langle b | e^{-\frac{i\hat{H}t}{\hbar}} \right) |a\rangle \quad (3.74)$$

We have merely regrouped terms.

A specific example has some historical context. Consider a matrix element of the coordinate

$$x(t)_{mn} = \langle m(t) | x | n(t) \rangle = \langle m | x(t) | n \rangle, \quad (3.75)$$

writing it first in Schrödinger and then in Heisenberg language. This matrix element drives electromagnetic electric dipole transitions (the dipole moment is $\vec{p} = e\vec{x}$). In classical electrodynamics we learn that if the dipole moment oscillates in time as $\vec{p}\exp(-i\omega t)$, the frequency of the emitted radiation is also ω . Quantum mechanically, however, we have the Bohr condition that the radiation emitted in a transition from state n to state m is $\omega = (E_n - E_m)/\hbar$. Taking $|n\rangle$ and $|m\rangle$ to be energy eigenstates, we see that the time dependence of our “coordinate in transition” is

$$x(t)_{mn} = \langle m | x | n \rangle \exp(i(E_m - E_n)t/\hbar) \quad (3.76)$$

which of course gives our desired result. But there is more: If a transition $n \rightarrow k$ is allowed ($E_k < E_n$), and if the transition $k \rightarrow m$ is also allowed, there is often a direct transition $n \rightarrow m$, although it is usually a weak “forbidden” line. (This is called the “Rydberg - Ritz combination principle” and it dates from 1908 long before quantum mechanics.) The matrix element would be

$$x^2(t)_{mn} = \langle m | x^2 | n \rangle \exp(i(E_m - E_n)t/\hbar) \quad (3.77)$$

so that the Bohr condition again obtains. In addition,

$$x_{mn}^2(t) = \sum_k \langle m | x(t) | k \rangle \langle k | x(t) | n \rangle = \sum_k x(t)_{mk} x(t)_{kn} \quad (3.78)$$

from the fundamental rule for products of operators. Classically, if $x(t) \sim \exp(i\omega t)$, then $x(t)^2 \sim \exp(2i\omega t)$ and there is no way to get the Bohr condition. Quantum mechanically, everything works: note the cancellation $\exp(-i(E_k - E_k)t/\hbar)$ in Eq. 3.78. Perhaps Heisenberg was led to the product rule by a desire to encode the Bohr condition in quantum dynamics in a natural way.

Example: Time evolution of the oscillator

The Heisenberg equations of motion of the oscillator are easily seen to mimic the classical equations of motion

$$\begin{aligned}\frac{d\hat{p}}{dt} &= -m\omega^2\hat{x} \\ \frac{d\hat{x}}{dt} &= \frac{\hat{p}}{m}\end{aligned}\tag{3.79}$$

This translates into equations for the time dependence of the raising and lowering operators of

$$\begin{aligned}\frac{d\hat{a}}{dt} &= \sqrt{\frac{m\omega}{2\hbar}} \left(\frac{\hat{p}}{m} - \frac{im\omega^2\hat{x}}{m\omega} \right) \\ &= -i\omega\hat{a} \\ \frac{d\hat{a}^\dagger}{dt} &= i\omega\hat{a}^\dagger.\end{aligned}\tag{3.80}$$

These can be easily integrated to find

$$\begin{aligned}\hat{a}(t) &= e^{-i\omega t}\hat{a}(0) \\ \hat{a}^\dagger(t) &= e^{i\omega t}\hat{a}^\dagger(0).\end{aligned}\tag{3.81}$$

Regrouping, we discover that the time evolution of the momentum and position operators is

$$\begin{aligned}\hat{x}(t) &= \hat{x}(0) \cos \omega t + \frac{\hat{p}(0)}{m\omega} \sin \omega t \\ \hat{p}(t) &= -m\omega\hat{x} \sin \omega t + \hat{p}(0) \cos \omega t.\end{aligned}\tag{3.82}$$

A second derivation of these quantities can be done beginning with the Heisenberg equations of motion

$$\begin{aligned}\hat{x}(t) &= e^{\frac{i\hat{H}t}{\hbar}}\hat{x}(0)e^{-\frac{i\hat{H}t}{\hbar}} \\ \hat{p}(t) &= e^{\frac{i\hat{H}t}{\hbar}}\hat{p}(0)e^{-\frac{i\hat{H}t}{\hbar}}\end{aligned}\tag{3.83}$$

by using the Baker-Hausdorff relation

$$e^{i\hat{G}\lambda} \hat{A} e^{-i\hat{G}\lambda} = \hat{A} + i\lambda[\hat{G}, \hat{A}] + \frac{(i\lambda)^2}{2!}[\hat{G}, [\hat{G}, \hat{A}]] + \dots \quad (3.84)$$

This gives

$$e^{\frac{i\hat{H}t}{\hbar}} \hat{x}(0) e^{-\frac{i\hat{H}t}{\hbar}} = \hat{x}(0) + \frac{it}{\hbar}[\hat{H}, \hat{x}(0)] + \frac{(it)^2}{2!\hbar^2}[\hat{H}, [\hat{H}, \hat{x}(0)]] + \dots \quad (3.85)$$

Fortunately, the algebra for the oscillator is simple:

$$[\hat{H}, \hat{x}(0)] = -\frac{i\hbar}{m}\hat{p}(0) \quad (3.86)$$

$$[\hat{H}, \hat{p}(0)] = i\hbar m\omega^2\hat{x}(0) \quad (3.87)$$

and so

$$\begin{aligned} e^{\frac{i\hat{H}t}{\hbar}} \hat{x}(0) e^{-\frac{i\hat{H}t}{\hbar}} &= \hat{x}(0) + \frac{\hat{p}(0)t\omega}{m} - \frac{1}{2}t^2\omega^2\hat{x}(0) - \frac{1}{3!}t^3\omega^2\frac{\hat{p}(0)}{m\omega} + \dots \\ &= \hat{x}(0) \cos \omega t + \frac{\hat{p}(0)}{m\omega} \sin \omega t. \end{aligned} \quad (3.88)$$

Notice that the operators show oscillatory time dependence. However, their expectation values in energy eigenstates will not show any time dependence. Only states which are superpositions of energy eigenstates will have time-varying expectation values of $\langle x(t) \rangle$ and $\langle p(t) \rangle$.

As a final example of Heisenberg equations of motion, we return to the two-state system. Here the Pauli matrices are the operators. (Recall that the spin $\vec{S} = (\hbar/2)\vec{\sigma}$.) The operator equation of motion will be

$$i\hbar \frac{\partial \hat{\sigma}_i}{\partial t} = [\hat{\sigma}_i, \hat{H}]. \quad (3.89)$$

Suppose our Hamiltonian is time independent: we write it as $\hat{H} = \hbar\omega\vec{\sigma} \cdot \hat{n} = \hbar\omega\hat{\sigma}_j\hat{n}_j$. Then the explicit equation of motion is

$$\begin{aligned} i\hbar \frac{\partial \hat{\sigma}_i}{\partial t} &= [\hat{\sigma}_i, \hbar\omega\vec{\sigma} \cdot \hat{n}] = \hbar\omega n_j [\hat{\sigma}_i, \hat{\sigma}_j] \\ &= 2\hbar\omega i\epsilon_{ijk} n_j \hat{\sigma}_k = 2i\hbar\omega \hat{n} \times \vec{\sigma}. \end{aligned} \quad (3.90)$$

This is

$$\frac{\partial \vec{\sigma}}{\partial t} = 2\omega \hat{n} \times \vec{\sigma}, \quad (3.91)$$

which is identical to the classical equation of motion for a magnetic dipole in an external magnetic field. This is an operator equation: sandwiching it between states turns it into an equation of motion for the Bloch vector introduced earlier in the chapter.

Let us complete the solution by splitting $\vec{\sigma}$ into a part parallel to \hat{n} and a part transverse to it, $\vec{\sigma} = \vec{\sigma}_{\parallel} + \vec{\sigma}_{\perp}$. The parallel part $\vec{\sigma}_{\parallel}$ will remain constant in time, because

$$\frac{\partial \vec{\sigma}_{\parallel}}{\partial t} = -2\omega(\vec{\sigma}_{\parallel} \times \hat{n}) = 0 \rightarrow \vec{\sigma}_{\parallel}(t) = \vec{\sigma}_{\parallel}(0). \quad (3.92)$$

We next decompose the transverse vector into two mutually orthogonal vectors σ_1 and σ_2 , where

$$\hat{n} \times \vec{\sigma}_1 = \vec{\sigma}_2 \quad \hat{n} \times \vec{\sigma}_2 = -\vec{\sigma}_1 \quad (3.93)$$

so that

$$\frac{\partial \vec{\sigma}_1}{\partial t} = -2\omega \vec{\sigma}_2 \quad \frac{\partial \vec{\sigma}_2}{\partial t} = 2\omega \vec{\sigma}_1 \quad (3.94)$$

Two new linear combinations of these vectors

$$\vec{\sigma}_+ = \vec{\sigma}_1 + i\vec{\sigma}_2 \quad \vec{\sigma}_- = \vec{\sigma}_1 - i\vec{\sigma}_2 \quad (3.95)$$

diagonalize the equations of motion. For σ_+ , this is

$$\frac{\partial}{\partial t}(\vec{\sigma}_1 + i\vec{\sigma}_2) = -2\omega(\vec{\sigma}_2 - i\vec{\sigma}_1) = 2i\omega(\vec{\sigma}_1 + i\vec{\sigma}_2) \quad (3.96)$$

and the equation for σ_- is similar. The result is

$$\begin{aligned} \vec{\sigma}_1(t) + i\vec{\sigma}_2(t) &= e^{2i\omega t}(\vec{\sigma}_1(0) + i\vec{\sigma}_2(0)) \\ \vec{\sigma}_1(t) - i\vec{\sigma}_2(t) &= e^{-2i\omega t}(\vec{\sigma}_1(0) - i\vec{\sigma}_2(0)) \end{aligned} \quad (3.97)$$

and we have completely specified all time dependence in the problem. The operator precesses around the direction specified by the Hamiltonian, essentially, this is cyclotron motion.

The interaction picture

Next we briefly consider a “mixed” picture, in which both the states and operators evolve in time. This picture will prove to be quite convenient later on when we study time-dependent perturbation theory. Let us suppose we can decompose the Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{V} \quad (3.98)$$

where \hat{H}_0 is time independent and simple, and \hat{V} is what is left over in the Hamiltonian. We define the interaction picture states as

$$|\alpha_I(t)\rangle = e^{\frac{i\hat{H}_0 t}{\hbar}} |\alpha_s(t)\rangle \quad (3.99)$$

and we define the operators in the interaction picture as

$$\hat{O}_I(t) = e^{\frac{i\hat{H}_0 t}{\hbar}} \hat{O} e^{-\frac{i\hat{H}_0 t}{\hbar}} \quad (3.100)$$

The idea is, that because H_0 is simple, the problem of how states evolve under H_0 alone can be completely specified. We wish remove its effect from the problem in order to continue. We find that $|\alpha_I(t)\rangle$ evolves with time only under the influence of the operator V (expressed in interaction picture)

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} |\alpha_I(t)\rangle &= -\hat{H}_0 e^{\frac{i\hat{H}_0 t}{\hbar}} |\alpha_s(t)\rangle + e^{\frac{i\hat{H}_0 t}{\hbar}} \times \left(i\hbar \frac{\partial}{\partial t} |\alpha_s(t)\rangle \right) \\ &= -\hat{H}_0 |\alpha_I(t)\rangle + e^{\frac{i\hat{H}_0 t}{\hbar}} (\hat{H}_0 + \hat{V}) |\alpha_s(t)\rangle \\ &= -\hat{H}_0 |\alpha_I(t)\rangle + \left(\hat{H}_0 + e^{\frac{i\hat{H}_0 t}{\hbar}} \hat{V} e^{-\frac{i\hat{H}_0 t}{\hbar}} \right) |\alpha_I(t)\rangle \\ &= \hat{V}_I |\alpha_I(t)\rangle. \end{aligned} \quad (3.101)$$

So much for the states. Operators evolve in time with a Heisenberg-like equation of motion, which only involves H_0 . It is easy to show that their interaction picture equation of motion is

$$\frac{d\hat{O}_I}{dt} = \frac{\partial \hat{O}_I}{\partial t} + \frac{1}{i\hbar} [\hat{O}_I, \hat{H}_0]. \quad (3.102)$$

The adiabatic theorem

Suppose you have a time dependent Hamiltonian. At $t = 0$ your system is in an energy eigenstate. As time goes on the eigenstates of the Hamiltonian evolve because $H(t)$ changes, but if the time dependence is very slow, the system will remain in the energy eigenstate. To illustrate the dynamics, glance back at Fig. 2.7, the two state system, and imagine that the off-diagonal term in \hat{H} , Δ , is large and positive at $t = 0$ and slowly decreases with time, eventually becoming negative. The energy of a system placed in the upper eigenvalue at $t = 0$ flows along the upper eigenvalue as Δ is varied. The state vector at any time remains the eigenstate at the appropriate value of $\Delta(t)$. Let's see how that happens.

We again solve

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = H(t) |\psi(t)\rangle. \quad (3.103)$$

If \hat{H} were time independent, we could expand in a complete set of energy eigenfunctions, $|\psi(t)\rangle = \sum_n c_n |n\rangle \exp(-iE_n t)$, and an eigenstate of \hat{H} would evolve as a pure phase, $|n, t\rangle = |n\rangle \exp(-iE_n t)$.

In an adiabatic process the Hamiltonian is time dependent, $H(t)$, but we assume it changes slowly. At any one time we can still solve the eigenvalue equation

$$H(t) |n, t\rangle = E_n(t) |n, t\rangle \quad (3.104)$$

and it is still the case that $\langle n, t | m, t \rangle = \delta_{mn}$. We guess a solution to the full equation

$$\psi(t) = \sum_n c_n(t) |n, t\rangle \exp(i\theta_n(t)). \quad (3.105)$$

Inserting our guess in the Schrödinger equation,

$$\begin{aligned} i\hbar \sum_n (\dot{c}_n(t) |n\rangle + c_n(t) \frac{d|n\rangle}{dt}) + i c_n(t) \dot{\theta}_n(t) e^{i\theta_n} &= \sum_n c_n(t) H(t) |n, t\rangle e^{i\theta_n} \\ &= \sum_n c_n(t) E_n(t) |n\rangle e^{i\theta_n} \end{aligned} \quad (3.106)$$

Setting

$$\theta_n(t) = -\frac{1}{\hbar} \int_0^t E_n(t') dt' \quad (3.107)$$

makes the right hand side equal to the last term on the left side, leaving

$$\sum_n \dot{c}_n(t) |n\rangle e^{i\theta_n} = - \sum_n c_n(t) \frac{d|n\rangle}{dt} e^{i\theta_n} \quad (3.108)$$

and taking an inner product with $\langle m |$ gives us

$$\dot{c}_m(t) = - \sum_n c_n(t) \langle m | \frac{d}{dt} |n\rangle e^{i(\theta_n - \theta_m)}. \quad (3.109)$$

Let's separate the $n = m$ term in the sum.:

$$\dot{c}_m(t) = c_m(t) \langle m | \frac{d}{dt} |m\rangle - \sum_{n \neq m} c_n(t) \langle m | \frac{d}{dt} |n\rangle e^{i(\theta_n - \theta_m)}. \quad (3.110)$$

To deal with the second term, we differentiate Eq: 3.104:

$$\frac{d}{dt}(H(t) |n, t\rangle) = \frac{d}{dt}(E_n(t) |n, t\rangle) \quad (3.111)$$

or

$$\dot{H} |n\rangle + H \frac{d}{dt} |n\rangle = \dot{E}_n(t) |n\rangle + \frac{d}{dt} |n\rangle \quad (3.112)$$

Again take an inner product with $\langle m(t)|$, and remember we will use this for the $n \neq m$ term

–

$$(E_m(t) - E_n(t) \langle m| \frac{d}{dt} |n\rangle) = \delta_{mn} \dot{E}_n(t) - \langle m|\dot{H}|n\rangle. \quad (3.113)$$

Inserting this result in Eq. 3.109 gives us

$$\dot{c}_m(t) = -c_m \langle m(t)| \frac{d}{dt} |m(t)\rangle - \sum_{n \neq m} c_n \frac{\langle m|\dot{H}|n\rangle}{E_m - E_n} e^{i(\theta_n - \theta_m)}. \quad (3.114)$$

This is exact. Now for adiabaticity. Suppose that $\dot{H} \sim \epsilon/\tau$ where ϵ is some typical energy which varies over a typical time τ . The second term scales as $\epsilon/(\tau(E_m - E_n))$. If this quantity is small compared to the $1/dt$ of the first term, that is, if $dt \ll \tau \Delta E/\epsilon$, we can neglect the second term compared to the first term. This says that the c_m 's evolve independently. We can integrate this expression:

$$c_m(t) = c_m(0) \exp\left(-\int_0^t dt' \langle m(t')| \frac{d}{dt} |m(t')\rangle\right) = c_m(0) \exp(i\gamma_m(t)). \quad (3.115)$$

(The integral is a pure phase since $\langle m(t')| \frac{d}{dt} |m(t')\rangle$ is purely imaginary – the time derivative is antihermitian.) If we start in the m th eigenstate of $H(0)$, we stay in the m th eigenstate of $H(t)$. All that happens is that the our wave function picks up an extra phase factor.

Pause for a second: this can have practical applications: if you control the time dependence of $H(t)$, you can tune it very slowly and “walk” your system, effectively, from one eigenstate of a time independent H to another one. In the case of the two state system, taking Δ from a large and positive value transforms the state into what it would be if there were no avoided level crossing, basically carrying you into the orthogonal spin state.

In this case the phase factor may not be important. But suppose the adiabatic transformation of the Hamiltonian takes it away from $H(0)$ and brings it back to $H(0)$ at a later time. The state which went away and came back picks up a phase factor compared to a state which stayed put. Interesting interference effects can result. In this case the phase factor is called a Berry phase.

Often we think of the transformation in $H(t)$ as resulting from moving the state around in space. We replace the time variable in Eq. 3.115 by a coordinate $\vec{R}(t)$. The chain rule tells us

$$\frac{d}{dt} = \frac{d\vec{R}}{dt} \cdot \vec{\nabla}_{\vec{R}}. \quad (3.116)$$

The phase becomes

$$\gamma = i \int_{R(0)}^{R(t)} d\vec{R}' \cdot \langle n(\vec{R}') | \vec{\nabla}_{\vec{R}'} | n(\vec{R}') \rangle \quad (3.117)$$

For a closed path, $\vec{R}(T) = \vec{R}(0)$ and the Berry phase is

$$\gamma = \oint d\vec{R}' \cdot \langle n(\vec{R}') | i \vec{\nabla}_{\vec{R}'} | n(\vec{R}') \rangle \quad (3.118)$$

We will see an application of this result in the next chapter when we look at the Arahamov-Bohm effect.

There is a simpler version of this derivation if we specialize to the two state system. Begin again with

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = H(t) |\psi(t)\rangle. \quad (3.119)$$

and assume that $|\psi(t)\rangle = V(t) |\phi_0(t)\rangle$ where $|\phi_0(t)\rangle$ is an energy eigenstate of $H(t)$. Rotate $|\psi(t)\rangle$ into a diagonal basis. The Schrödinger equation is

$$i\hbar \left(\frac{\partial V}{\partial t} |\phi_0\rangle + V(t) \frac{\partial |\phi_0\rangle}{\partial t} \right) = HV(t) |\phi_0\rangle. \quad (3.120)$$

This is

$$i\hbar \frac{\partial |\phi_0\rangle}{\partial t} = (V^\dagger HV) |\phi_0\rangle - i\hbar V^\dagger \frac{\partial V}{\partial t} |\phi_0\rangle \quad (3.121)$$

By construction, $V^\dagger HV$ is diagonal – V rotates the state $|\psi(t)\rangle$ into an energy eigenstate, so it rotates the Hamiltonian to be diagonal as well. This means that the first term keeps an energy eigenfunction an energy eigenfunction. The second term is a mixing term. To evaluate it, write

$$V = \begin{pmatrix} \cos \theta(t) & \sin \theta(t) \\ -\sin \theta(t) & \cos \theta(t) \end{pmatrix}. \quad (3.122)$$

(I am assuming that Δ , the off-diagonal entry in the 2×2 Hamiltonian, is real, and I am being careless about conventions for 2's in the definition of the angle.) A few lines of algebra give

$$-iV^\dagger \frac{\partial V}{\partial t} = -i \begin{pmatrix} 0 & \partial\theta(t)/\partial t \\ -\partial\theta(t)/\partial t & 0 \end{pmatrix} \quad (3.123)$$

so we recover the adiabatic theorem when $\partial\theta(t)/\partial t$ is small.

Neutrino oscillations

Neutrinos are nearly massless, nearly noninteracting particles that “pair” with charged leptons in the weak interactions. Technically, neutrinos and charged leptons form a doublet of weak isospin

$$l_e = \begin{pmatrix} \nu_e \\ e \end{pmatrix} \quad l_\mu = \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \quad l_\tau = \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \quad (3.124)$$

and the weak interactions raise or lower the third component of the weak isospin of the lepton. (It is actually a bit more complicated, but this is all we need to know for now). Anyway, it happens that the three neutrinos ν_e, ν_μ, ν_τ are not energy (“mass”) eigenstates; the mass eigenstates (which we will label as ν_1, ν_2, ν_3) are mixtures of these states. Thinking temporarily about two species of neutrinos, this is a situation we have seen before, there are two bases for our states, related by a unitary transformation,

$$\begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}. \quad (3.125)$$

The neutrino Hamiltonian is diagonal in the (ν_1, ν_2) basis. Neutrinos are very light so their energy is $E = \sqrt{p^2 + m^2}$ which is approximated as $E = p + \frac{1}{2}m^2/p$ (and $c = 1$ for now). We can write a rescaled H in the (ν_1, ν_2) basis as

$$H = \begin{pmatrix} m_1^2 & 0 \\ 0 & m_2^2 \end{pmatrix} = \frac{\Delta m^2}{2} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} + D \quad (3.126)$$

where D is a diagonal matrix. In the (ν_e, ν_μ) basis

$$H = \frac{\Delta m^2}{2} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} + D. \quad (3.127)$$

In a terrestrial experiment, neutrinos would be created in a ν_e, ν_μ, ν_τ basis, propagate in the ν_1, ν_2, ν_3 basis, and are detected in a ν_e, ν_μ, ν_τ basis. The physics – and formulas – describing the time evolution of a state are exactly those we saw at the start of this chapter; the probability that, say, an initial ν_e would be detected as a ν_e would oscillate in time. The situation is called “vacuum oscillation” in the literature.

However, the effects of neutrino oscillations were first noticed in the detection of neutrinos from the sun, and the physics is quite different. Fusion processes in the sun create ν_e 's. The luminosity of the sun tells us the absolute flux of ν_e 's expected to be seen on Earth. The

“solar neutrino problem” was that when the flux was observed (by Ray Davis and John Bahcall, in a large tank of cleaning fluid in a mine in South Dakota from 1970 to 1994), the observed detection rate was about a third of what was expected.

As they propagate through the sun, interactions with matter in the sun induce an additional term in the ν_e channel. For example, one could imagine a scattering process $\nu_e + e \rightarrow \nu_e + e$ due to exchange of a W boson. No such process occurs in the ν_μ channel because it would involve a muon and the energy scales in the sun are too low – KeV, far below the muon mass of about 100 MeV. So, in a (ν_e, ν_μ) basis we have an additional contribution

$$\Delta H = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} = \frac{A}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{A}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.128)$$

and the total Hamiltonian becomes

$$H = \frac{1}{2} \begin{pmatrix} A - \Delta m^2 \cos 2\theta & \Delta m^2 \sin 2\theta \\ \Delta m^2 \sin 2\theta & -A + \Delta m^2 \cos 2\theta \end{pmatrix} + \frac{1}{2}(A + m_1^2 + m_2^2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.129)$$

We have seen this Hamiltonian before, and we can write down its eigenvalues by inspection. – They are $(A + m_1^2 + m_2^2)/2 \pm E$ where

$$E = \frac{1}{2} \left((A - \Delta m^2 \cos 2\theta)^2 + (\Delta m^2 \sin 2\theta)^2 \right)^{1/2}. \quad (3.130)$$

Now for some physics: A is proportional to the density of material the neutrino can interact with, so it is largest at the center of the sun and decreases to zero at the sun’s surface. If A is large enough ($A \gg \Delta m^2 \cos 2\theta$), it dominates the Hamiltonian,

$$H = \frac{A}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.131)$$

and $|\nu_e\rangle$ is an eigenstate of H . A picture of the spectrum is shown in Fig. 3.3. In this idealized situation the initial $|\nu_e\rangle$ state is born on the far right of the picture, at large A , moves to the left as it exits the Sun, follows the upper eigenvalue’s branch as A varies, and emerges as the mass eigenstate $|\nu_2\rangle$. There is no oscillation en route from the sun to the earth since it is the mass eigenstate which propagates; the neutrino detector on the earth measures the weak eigenstate $|\nu_e\rangle = \cos \theta |\nu_1\rangle + \sin \theta |\nu_2\rangle$ and so the measured flux of electron neutrinos is reduced by a factor $|\langle \nu_e | \nu_2 \rangle|^2 = \sin^2 \theta$.

The Davis - Bahcall experiment detected relatively high energy neutrinos. It happens that A is energy dependent and proportional to E_ν . Without continuing the calculation we

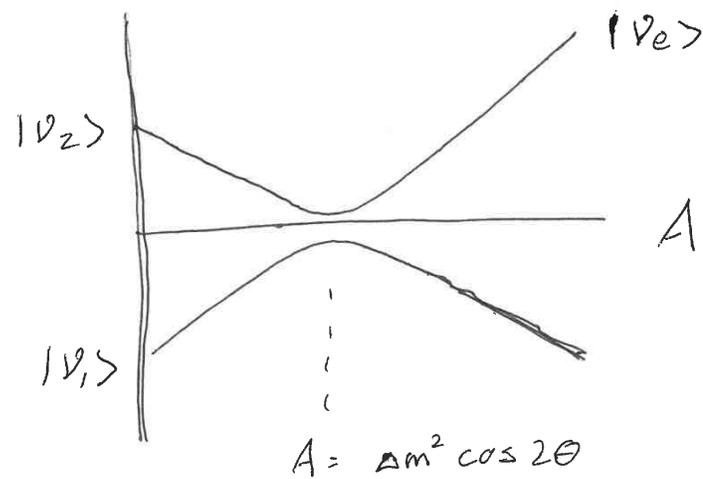


Figure 3.3: Neutrino energy spectrum in the Sun. Note the “resonance” or avoided level crossing at $A = \Delta m^2 \cos 2\theta$.

can observe that the reduction factor will be energy dependent (and indeed if $A = 0$ we are back to vacuum oscillations). Later experiments with different energy resolutions measured different dilution factors.

The situation is called “matter oscillation” or the MSW effect after its discoverers, Mikheyev, Smirnov, Wolfenstein. Useful references if you want to know more: check out the “Neutrino masses, mixing and oscillations” article in the Review of Particle Properties <https://pdg.lbl.gov/> and H. A. Bethe, “A Possible Explanation of the Solar Neutrino Puzzle,” Phys. Rev. Lett. **56**, 1305 (1986) doi:10.1103/PhysRevLett.56.1305.

Chapter 4

Propagators and path integrals

Propagators or Green's functions

It is often useful to be able to express the behavior of physical quantities in terms of a source function and a propagator or kernel. This can be done for the Schrödinger equation. We will first consider the case of a time-independent Hamiltonian. We expand our state as a superposition of energy eigenfunctions and write the time-dependent state as

$$\begin{aligned} |\alpha(t)\rangle &= e^{-\frac{i\hat{H}(t-t_0)}{\hbar}} |\alpha(t_0)\rangle \\ &= \sum_a |a\rangle \langle a|\alpha(t_0)\rangle e^{-\frac{iE_a(t-t_0)}{\hbar}}. \end{aligned} \quad (4.1)$$

The most useful way to present the result is to write it in coordinate space basis. We take the overlap of the state $|\alpha(t)\rangle$ with the bra $\langle x|$ and re-label $\langle x|\alpha(t)\rangle = \psi(x, t)$, $\langle x|a\rangle = u_a(x)$. This gives

$$\psi(x, t) = \sum_a u_a(x) \langle a|\alpha(t_0)\rangle e^{-\frac{iE_a(t-t_0)}{\hbar}} \quad (4.2)$$

where

$$\begin{aligned} \langle a|\alpha(t_0)\rangle &= \int d^3x' \langle a|x'\rangle \langle x'|\alpha(t_0)\rangle \\ &= \int d^3x' u_a^*(x') \psi(x', t_0) \end{aligned} \quad (4.3)$$

so

$$\begin{aligned} \psi(x, t) &= \sum_a u_a(x) \int d^3x' u_a^*(x') \psi(x', t_0) e^{-\frac{iE_a(t-t_0)}{\hbar}} \\ &= \int d^3x' \sum_a u_a(x) u_a^*(x') e^{-\frac{iE_a(t-t_0)}{\hbar}} \psi(x', t_0) \\ &= \int d^3x' K(x, t; x', t_0) \psi(x', t_0). \end{aligned} \quad (4.4)$$

We have introduced the propagator, $K(x, t; x', t_0)$.

The Schrödinger equation is a wave equation. You may recall a similar object in electromagnetism, with a distinction between a “retarded” propagator ($t > t_0$) and an “advanced” one ($t < t_0$). Generally, because of causality, we are interested in the retarded propagator,

$$K^+(x, t; x', t_0) = K(x, t; x', t_0) \theta(t - t_0) \quad (4.5)$$

where $\theta(t - t_0)$ is the step function (equal to unity for $t > t_0$ and zero otherwise). We can derive a familiar expression for the propagator by writing

$$\theta(t - t')\psi(x, t) = \int d^3x' a^+ K(x, t; x', t_0)\psi(x', t_0) \quad (4.6)$$

and multiplying through

$$\begin{aligned} (i\frac{\partial}{\partial t} - \frac{1}{\hbar}H)[\theta(t - t_0)\psi(x, t)] &= i\delta(t - t_0)\psi(x, t) \\ &= \int d^3x' (i\frac{\partial}{\partial t} - \frac{1}{\hbar}H)K^+(x, t; x', t_0)\psi(x', t_0). \end{aligned} \quad (4.7)$$

So it must be that

$$(i\hbar\frac{\partial}{\partial t} - H)K^+(x, t; x', t_0) = i\hbar\delta^3(x - x')\delta(t - t_0), \quad (4.8)$$

that is, like all propagators or Green's functions, $K^+(x, t; x', t_0)$ is the solution of an inhomogeneous Schrödinger equation with a delta function source. I will refer interchangeably to K and K^+ , always meaning the retarded propagator K^+ .

A compact expression for the propagator is given simply by expressing the time evolution operator in a coordinate space basis,

$$\begin{aligned} K(x, t; x', t_0) &= \sum_a \langle x|a\rangle e^{-\frac{iE_a(t-t_0)}{\hbar}} \langle a|x'\rangle \\ &= \langle x|e^{-\frac{i\hat{H}(t-t_0)}{\hbar}}|x'\rangle. \end{aligned} \quad (4.9)$$

A useful special case is the free particle, $\hat{H} = \hat{p}^2/2m$. It is easy to find the propagator for a free particle by passing to momentum space:

$$\begin{aligned} \langle x_1|e^{-\frac{i\hat{H}t}{\hbar}}|x_2\rangle &= \int d^3p \langle x_1|p\rangle e^{-\frac{i\hat{p}^2t}{2m\hbar}} \langle p|x_2\rangle \\ &= \int \frac{d^3p}{(2\pi)^3} e^{\frac{ipx_1}{\hbar}} e^{-\frac{i\hat{p}^2t}{2m\hbar}} e^{-\frac{ipx_2}{\hbar}} \end{aligned} \quad (4.10)$$

Performing the integral by completing the square gives us an explicit result for the free particle propagator,

$$K(x_1, t; x_2, 0) = \left(\frac{m}{2\pi i\hbar t}\right)^{\frac{3}{2}} \exp\left(\frac{i(x_1 - x_2)^2 m}{2\hbar t}\right) \theta(t). \quad (4.11)$$

Now imagine nesting propagators together, by splitting the time interval into two pieces and inserting a complete set of states $\int d^3x_3 |x_3\rangle \langle x_3|$ between the pieces,

$$K(x, t; x_2, t_2) = \int d^3x_3 \langle x_1 | e^{-i\frac{\hat{H}(t_1-t_3)}{\hbar}} |x_3\rangle \langle x_3 | e^{-i\frac{\hat{H}(t_3-t_2)}{\hbar}} |x_2\rangle. \quad (4.12)$$

Clearly, this is just the statement that

$$K(x, t; x_2, t_2) = \int d^3x_3 K(x_1, t_1; x_3, t_3) \times K(x_3, t_3; x_2, t_2). \quad (4.13)$$

This result can be iterated as needed. Variations on propagators are also often encountered. One of them is the quantity

$$G(t) = \int d^3x K(x, t; x, 0). \quad (4.14)$$

By a simple passage of steps, this is

$$\begin{aligned} G(t) &= \int d^3x \sum_a \langle x|a\rangle \langle a|x\rangle e^{-\frac{iE_a t}{\hbar}} \\ &= \sum_a \int d^3x \langle a|x\rangle \langle x|a\rangle e^{-\frac{iE_a t}{\hbar}} \\ &= \sum_s \langle a|a\rangle e^{-\frac{iE_a t}{\hbar}} = \sum_a e^{-\frac{iE_a t}{\hbar}}. \end{aligned} \quad (4.15)$$

A variation on $G(t)$ is used extensively in statistical mechanics, as the partition function

$$Z = \sum_a e^{-\beta E_a} \quad (4.16)$$

where $\beta = \frac{1}{kT}$. Clearly β plays the same role as $\frac{-it}{\hbar}$ in our equations. One can also Fourier transform of $G(t)$ into the energy domain

$$\begin{aligned} G(t) \rightarrow G(E) &= -i \int_0^\infty dt G(t) e^{\frac{iEt}{\hbar}} \\ &= -i \int_0^\infty dt \sum_a e^{-\frac{iE_a t}{\hbar}} e^{\frac{iEt}{\hbar}} \\ &= \lim_{\epsilon \rightarrow 0} -i \int_0^\infty dt \sum_a e^{i\left(\frac{E-E_a}{\hbar} + \frac{i\epsilon}{\hbar}\right)t} \\ &= -i\hbar \sum_a \frac{1}{E - E_a + i\epsilon}. \end{aligned} \quad (4.17)$$

Variations on this formula appear when one is considering scattering processes.

Path integrals

What if H varies with time? How can we make sense of the time evolution operator? Let us begin with the formal expression for the evolution operator

$$U_{\alpha\beta}(t) = \langle \alpha | \exp\left(-\frac{i\hat{H}t}{\hbar}\right) | \beta \rangle. \quad (4.18)$$

In this formula t is some finite quantity, so it might be hard to compute $U_{\alpha\beta}(t)$. We propose to evaluate the time evolution operator by slicing the time interval into a series of N steps of infinitesimal time interval Δt , arranging that $t = N\Delta t$. In each time interval Δt , the Hamiltonian will be regarded as constant, while of course it will be allowed to vary from time step to time step. That is,

$$U_{\alpha\beta}(t) = \lim_{N \rightarrow \infty; \Delta t \rightarrow 0} \langle \alpha | \prod_{i=1}^N \exp\left(-\frac{i\hat{H}(t_i)\Delta t}{\hbar}\right) | \beta \rangle. \quad (4.19)$$

Inserting a complete set of states between each exponential factor, we have

$$\begin{aligned} U_{\alpha\beta}(t) &= \sum_{j_1} \sum_{j_2} \cdots \sum_{j_N} \\ &\langle \alpha | j_N \rangle \langle j_N | \exp\left(-\frac{i\hat{H}(t_N)\Delta t}{\hbar}\right) | j_{N-1} \rangle \\ &\langle j_{N-1} | \exp\left(-\frac{i\hat{H}(t_{N-1})\Delta t}{\hbar}\right) | j_{N-2} \rangle \cdots \\ &\langle j_1 | \exp\left(-\frac{i\hat{H}(t_1)\Delta t}{\hbar}\right) | \beta \rangle. \end{aligned} \quad (4.20)$$

Rather than thinking of this expression time slice by time slice, let us link together a particular set of states $|j'_1\rangle, |j'_2\rangle, \dots, |j'_N\rangle$, and connect them together as a “path” in Hilbert space, as shown in Fig. 4.1. Each path contributes a complex number to $U_{\beta\alpha}$, the product of each particular matrix element $\langle j'_M | \exp(-iH(t_M)\Delta t/\hbar) | j'_{M-1} \rangle$. $U_{\beta\alpha}$ is the sum of contributions from all paths: hence the name “path integral” associated with this representation of the time evolution operator.

Let us suppose that our intermediate states are diagonal in coordinate space, so that we can interpret each path through Hilbert space as a “real” trajectory, $x(t)$. Then,

$$U_{\alpha\beta}(t) = \int dx_0 \int dx_N \cdots \int dx_1 \langle \alpha | x_N \rangle \langle x_N | \exp\left(-\frac{i\hat{H}(t_N)\Delta t}{\hbar}\right) | x_N \rangle \times \quad (4.21)$$

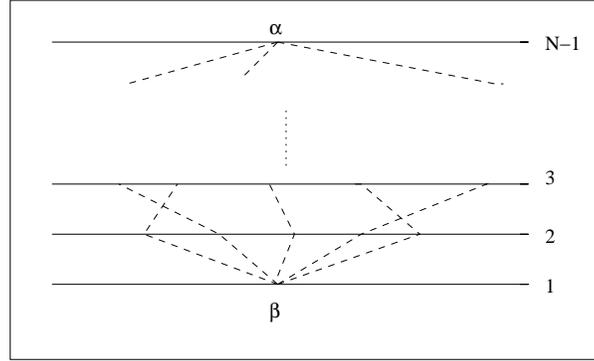


Figure 4.1: Paths in the time-sliced interval.

$$\langle x_N | \exp\left(-\frac{i\hat{H}(t_{N-1})}{\hbar}\Delta t\right) | x_{N-1} \rangle \cdots \langle x_2 | \exp\left(-\frac{i\hat{H}(t_1)}{\hbar}\Delta t\right) | x_0 \rangle \langle x_0 | \beta \rangle. \quad (4.22)$$

The time slicing helps us make sense of the evolution operator. Suppose next that the Hamiltonian is that of a particle in an external potential, $\hat{H} = \hat{H}_1 + \hat{H}_2$, where $\hat{H}_1 = \frac{\hat{p}^2}{2m}$ and $\hat{H}_2 = V(x)$. Then we can break the time evolution operator over a time Δt (which involves \hat{H} into a product of two terms

$$\exp\left(-\frac{i\hat{H}\Delta t}{\hbar}\right) = \exp\left(-\frac{i\hat{H}_1\Delta t}{\hbar}\right) \exp\left(-\frac{i\hat{H}_2\Delta t}{\hbar}\right) + O(\Delta t^2). \quad (4.23)$$

This means that

$$\begin{aligned} \langle x_1 | \exp\left(-\frac{i\hat{H}_1\Delta t}{\hbar}\right) \exp\left(-\frac{i\hat{H}_2\Delta t}{\hbar}\right) | x_2 \rangle &= \int dx_3 \langle x_1 | \exp\left(-\frac{i\hat{p}^2\Delta t}{2m\hbar}\right) | x_3 \rangle \\ &\quad \times \langle x_3 | \exp\left(-\frac{iV(x)\Delta t}{\hbar}\right) | x_2 \rangle \\ &= \int dx_3 \langle x_1 | \exp\left(-\frac{i\hat{p}^2\Delta t}{2m\hbar}\right) | x_3 \rangle \\ &\quad \times \delta(x_3 - x_2) \exp\left(-\frac{iV(x_2)\Delta t}{\hbar}\right). \end{aligned} \quad (4.24)$$

Notice that the term first term is just the propagator for a free particle, and so

$$\begin{aligned} \langle x_1 | \exp\left(-\frac{i\hat{H}\Delta t}{\hbar}\right) | x_2 \rangle &= \int dx_3 K(x_1, \Delta t; x_3, 0) \delta(x_3 - x_2) \exp\left(-\frac{iV(x_2)\Delta t}{\hbar}\right) \\ &= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{\frac{1}{2}} \exp\left[i\left(\frac{m(x_1 - x_2)^2}{2(\Delta t)^2} - V(x_2)\right) \frac{\Delta t}{\hbar}\right]. \end{aligned} \quad (4.25)$$

Thus the path integral will take the form

$$U_{\alpha\beta}(t) = \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{\frac{N}{2}} \int dx_N \cdots dx_1 \langle \alpha | x_N \rangle \langle x_1 | \beta \rangle \exp\left(\frac{i\Delta t}{\hbar} \sum_{j=1}^N \left(\frac{1}{2}m \left(\frac{x_{j+1} - x_j}{\Delta t}\right)^2 - V(x_j)\right)\right). \quad (4.26)$$

(We have factorized the Δt term in the expression in a useful way.) In the limit that $\Delta t \rightarrow 0$ the exponential factor associated with a single path $x(t')$ becomes

$$\exp\left(\frac{i}{\hbar} \int_0^t dt' \left(\frac{1}{2}m\dot{x}^2(t') - V(x(t'))\right)\right). \quad (4.27)$$

The reader should recognize the integrand as the classical Lagrangian for a nonrelativistic particle in an external potential. The integral is the classical action associated with its integral along a path specified by $x(t')$. This is the expression for the contribution to the path integral from a single path (a single term in the sum over intermediate states as defined in Eq. 4.20.) Each contribution is a phase, whose value is the ratio of the classical action for that particular path divided by \hbar . The time evolution operator is a sum of contributions over all possible paths the particle can take from the initial time to the final time.

We can now understand how classical dynamics can arise from quantum mechanical motion. Suppose that the action for every possible path the particle could take is large, much greater than \hbar . Compare two paths (which could correspond to paths which are similar in the classical sense). An action S_0 is associated with one path while the other has a different action, which we will write as $S_0 + \delta S$. These two paths combine to produce a contribution to the evolution operator of

$$\exp\frac{iS_0}{\hbar} + \exp\frac{i(S_0 + \delta S)}{\hbar} = \exp\frac{iS_0}{\hbar} \left(1 + \exp\frac{i\delta S}{\hbar}\right). \quad (4.28)$$

If $\delta S/\hbar$ is a large number, the phase difference between the paths, $\exp(i\delta S/\hbar)$, will not be small. There will be destructive interference between the two paths and their contribution to the transition amplitude will be small. Obviously, the only sets of paths which contribute will be the ones for which their contributions will add coherently. These will be the ones for which $\delta S \simeq 0$. This is the statement of the principle of least action: the dominant paths are the ones which extremize the classical action – that is, the particle follows the path which is the solution to the classical equations of motion.

As a contrast, it could be the case that all paths have a classical action which is on the order of, or smaller than \hbar . Then there are no sets of nearby paths which dominate the evolution operator. These systems are fully quantum mechanical; attempting to describe a particle's motion using the idea of a classical trajectory will not result in correct physics.

Electromagnetism in quantum mechanics

Now let us consider the interaction of a quantum mechanical system with the electromagnetic field, which we will treat as a classical object for the time being. Recall the connection between fields and potentials

$$\vec{B} = \vec{\nabla} \times \vec{A} \quad \vec{E} = -\vec{\nabla}\phi - \frac{1}{c} \frac{\partial \vec{A}}{\partial t}. \quad (4.29)$$

The classical Hamiltonian for a particle in an electromagnetic field involves the potentials ϕ and \vec{A} , not the fields \vec{E} and \vec{B} . The appropriate quantity for the kinetic part of the Hamiltonian involves the so-called ‘‘covariant derivative’’ (see Jackson Sec. 12.1 for a derivation),

$$H_{cl} = \frac{1}{2m} \left(\vec{p} - \frac{e\vec{A}}{c} \right)^2 + e\phi. \quad (4.30)$$

Promoting dynamical variables (\hat{p} and \hat{x}) to operators gives us the quantum Hamiltonian operator

$$\hat{H} = \frac{1}{2m} \left(\hat{p}^2 - \frac{e}{c} \left(\hat{p} \cdot \vec{A} + \vec{A} \cdot \hat{p} \right) + \frac{e^2 A^2}{c^2} \right) + e\hat{\phi}. \quad (4.31)$$

We have carefully kept both terms linear in \hat{p} because in coordinate representation $\hat{p} = (\hbar/i)\nabla$, and in principle, the derivative can act on the vector potential. The Heisenberg equation of motion is

$$i\hbar \frac{d\hat{x}}{dt} = [\hat{x}, \hat{H}] = \frac{1}{m} \left(\hat{p} - \frac{e\vec{A}}{c} \right) \quad (4.32)$$

and the quantum analog of the Lorentz force law is

$$m \frac{d^2 \hat{x}}{dt^2} = e\vec{E} + \frac{e}{2c} \left(\frac{d\hat{x}}{dt} \times \vec{B} - \vec{B} \times \frac{d\hat{x}}{dt} \right). \quad (4.33)$$

The appearance of \vec{A} rather than \vec{B} seems a bit odd when we remember that electrodynamics has a gauge symmetry: the laws of electrodynamics are invariant under gauge transformations

$$\vec{A} \rightarrow \vec{A} + \vec{\nabla}\chi; \quad \phi \rightarrow \phi - \frac{1}{c} \frac{\partial \chi}{\partial t} \quad (4.34)$$

where χ is any scalar function of x and t . This means that there are many possible choices for a vector potential which will give the same magnetic field. How do different gauge choices (choices for \vec{A}) affect the solution of the Schrödinger equation? Let's consider an example: a spinless particle moving in an external constant magnetic field, chosen to point in the \hat{z} direction for convenience. The classical situation is simple: circular motion around the field line, rotating at the cyclotron frequency. Let's make some choices for \vec{A} and see what we get:

1. $\vec{A} = (-yB, 0, 0)$: In this case the Hamiltonian is

$$\hat{H} = \frac{1}{2m} \left(\left(\hat{p}_x - \frac{eBy}{c} \right)^2 + \hat{p}_y^2 + \hat{p}_z^2 \right). \quad (4.35)$$

\hat{H} has no explicit dependence on x or z , so its commutator with p_x and p_z will be zero. These quantities are conserved or equivalently correspond to good quantum numbers of our states. (Or: states can be simultaneous eigenstates of H , p_x and p_z : we write $|E\rangle = |E, p'_z, p'_x\rangle$). The Hamiltonian is

$$\hat{H} = \frac{\hat{p}_y^2}{2m} + \frac{1}{2m} \left(p'_x - \frac{eBy}{c} \right)^2 + \frac{p_z'^2}{2m} \quad (4.36)$$

where p'_x and p'_z are c-numbers. Note that this means that the second term in the Hamiltonian is simply a shifted quadratic function of y , $e^2 B^2 / (2mc^2) \left(y - \frac{cp'_x}{eB} \right)^2$, and the Hamiltonian is that of a harmonic oscillator,

$$\hat{H} = \frac{p_z'^2}{2m} + \frac{\hat{p}_y^2}{2m} + \frac{m\omega_c^2}{2} (y - y_0)^2, \quad (4.37)$$

where ω_c is just the cyclotron frequency, $eB/(mc)$. Thus the energy levels are those of a harmonic oscillator superimposed on those of a free particle moving in the \hat{z} direction,

$$E_n = \hbar\omega_c \left(n + \frac{1}{2} \right) + \frac{p_z'^2}{2m}, \quad (4.38)$$

This is the quantum mechanical analog of circular motion at the cyclotron frequency, an orbit about a field line, combined with uniform motion along the field line. These levels are called “Landau levels.” Obviously, the states exhibit high degeneracy. When the quantum mechanical particles are fermions, there are interesting consequences associated with this degeneracy. This is a story for another day.

2. We could equally well make another gauge choice

$$A_x = -\frac{By}{2} \quad A_y = \frac{Bx}{2}. \quad (4.39)$$

In this gauge, the Hamiltonian becomes

$$\hat{H} = \frac{1}{2m} \left(\hat{p}_z^2 + \left(\hat{p}_x + \frac{eBy}{2c} \right)^2 + \left(\hat{p}_y - \frac{eBx}{2c} \right)^2 \right). \quad (4.40)$$

Let us now define $\hat{P}_x = \hat{p}_x + \frac{eBy}{2c}$ and $\hat{P}_y = \hat{p}_y - \frac{eBx}{2c}$. The commutator of these two quantities is

$$[\hat{P}_x, \hat{P}_y] = \frac{m\omega_x}{2} ([y, \hat{P}_y] - [\hat{P}_x, x]) = im\omega_c \hbar. \quad (4.41)$$

This is almost the commutator of a canonically conjugate pair of dynamical variables. We can introduce such variables by performing a rescaling, defining $\hat{Q} = -\hat{P}_y/(m\omega_c)$ and $\hat{P} = \hat{P}_x$. In that case we have the usual commutation relation $[\hat{Q}, \hat{P}] = i\hbar$. The Hamiltonian becomes

$$\hat{H} = \frac{p_z^2}{2m} + \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega_c^2 \hat{Q}^2. \quad (4.42)$$

Again, this is a harmonic oscillator, with the same spectrum we found before. The energy eigenvalues are gauge invariant. That seems like a sensible result! Physical results should be gauge invariant. The states, however, appear to be quite different. What is going on?

Let us answer this question by considering a different question: Where did the the $\vec{p} - \frac{e\vec{A}}{c}$ come from? Can we give an explanation which does not depend on a classical correspondence argument? Yes, we can.

Quantum mechanics is invariant under a transformation which is a global phase rotation, that is, if we multiply the wave function by an overall space and time independent phase factor, no physical predictions are altered. The transformation is

$$\psi(x, t)' = e^{i\theta} \psi(x, t). \quad (4.43)$$

Obviously, both the time dependent and time independent Schrödinger equations are unchanged by such a transformation. For example,

$$\hat{H}\psi'(x, t) = i\hbar \frac{\partial \psi'(x, t)}{\partial t} = e^{i\theta} \hat{H}\psi(x, t) = e^{i\theta} i\hbar \frac{\partial \psi(x, t)}{\partial t}. \quad (4.44)$$

Can we generalize this invariance to a form in which quantum mechanics is invariant under local phase transformations (i. e. can we imagine a symmetry transformation where the wave function can be changed by a different phase everywhere in space)? Such a transformation is

$$\psi(x, t) \rightarrow \psi(x, t)' = e^{i\theta(x)}\psi(x, t) \quad (4.45)$$

If the phase factor could be pulled back through the Hamiltonian and the $\partial/\partial t$ operators, this transformation would also be a symmetry.

The potential term in the Hamiltonian will allow this: $E\psi = V(x)\psi \rightarrow e^{i\theta(x)}(E\psi = V(x)\psi)$. However, the momentum term is not invariant under this transformation,

$$\hat{p}\psi' = \frac{\hbar}{i} \frac{\partial}{\partial x} (e^{i\theta(x)}\psi(x)) = e^{i\theta(x)} \left(\frac{\hbar}{i} \frac{\partial \psi}{\partial x} + \hbar \frac{\partial \theta}{\partial x} \psi \right). \quad (4.46)$$

To preserve the symmetry we must alter the dynamics. (Or, said differently, only a very particular kind of dynamics will be consistent with our proposed symmetry.) We need a new kind of momentum variable such that

$$\hat{p}'\psi' = e^{i\theta(x)}\hat{p}\psi \quad (4.47)$$

The phase variation will factor out of every term in the Schrödinger equation. Such a new momentum term is

$$\vec{p}' = \vec{p} - \frac{e\vec{A}}{c} \quad (4.48)$$

where $\vec{A}(x, t)$ is a new dynamical variable which must be transformed simultaneously with ψ . The combined transformation we need is

$$\psi' = e^{i\theta(x)}\psi; \quad \vec{A}' = \vec{A} - \frac{\hbar c}{e} \vec{\nabla} \theta(x). \quad (4.49)$$

This symmetry transformation is called a “local gauge transformation.” To check that it indeed is a symmetry, just work it out:

$$\begin{aligned} \hat{p}'\psi' &= e^{i\theta(x)} \left(\frac{\hbar}{i} \frac{\partial}{\partial x} + \hbar \frac{\partial \theta}{\partial x} \right) \psi - e^{i\theta(x)} \frac{e\vec{A}}{c} \psi - \frac{e\hbar c}{c} e^{i\theta(x)} \frac{\partial \theta}{\partial x} \psi \\ &= e^{i\theta(x)} \left(\frac{\hbar}{i} \frac{\partial \psi}{\partial x} - \frac{e\vec{A}'}{c} \right) \psi. \end{aligned} \quad (4.50)$$

The factor of $\exp(i\theta(x))$ factors out, so that that it can be eliminated from the Schrödinger equation, just as was done in Eq. 4.44. It is slightly more conventional to write the combined

gauge transformation as

$$\vec{A} \rightarrow \vec{A} - \vec{\nabla}\chi(x) \quad (4.51)$$

$$\psi(x) \rightarrow \exp\left(\frac{ie\chi(x)}{\hbar c}\right)\psi(x). \quad (4.52)$$

Note that without further dynamics our new theory is trivial. If A can take on any value we could wash away any interesting behavior in ψ by performing a local gauge transformation: A could even vary randomly in space. We have to assume some kind of dynamics for A . In order for the new dynamics to be gauge invariant, it must involve gauge invariant combinations of A (and ϕ , once we think about time dependent gauge transformations). These combinations are the electric and magnetic fields. Thus the new dynamics must involve Hamiltonians built out of \vec{E} and \vec{B} . Exactly what this dynamics is (i.e. whether there is a unique Hamiltonian for the electromagnetic field) is not answered by our construction.

The reader will notice that we did not answer the original question, why $p - eA/c$? The answer must come by running the argument backward: we have to assume from the beginning that Nature possesses a dynamics which preserves an invariance under local gauge invariance. Then the vector potential must enter into the Hamiltonian as we have written it. Notice that in doing so, we have specified the interaction between systems carrying charge, which obey Schrödinger dynamics, and the electromagnetic field.

Aharonov-Bohm effect

If you were a classically trained physicist, you might find the appearance of the vector potential in quantum mechanics disturbing. How can the vector potential be “real” when one can change it with a gauge transformation? Shouldn’t physics only depend on the electromagnetic fields, not on the potentials? The following remarkable phenomenon tells us that we should not be so dogmatic:

Imagine we have an infinitely long cylindrical solenoid of radius a , as shown in Fig. 4.2. Inside the cylinder, there is a magnetic field running along the axis of the cylinder, but outside the field is zero. Even though $B = 0$ outside the cylinder, the vector potential \vec{A} will not vanish there. Up to gauge transformations, it is

$$\vec{A} = \hat{\phi} \frac{Ba^2}{2r}. \quad (4.53)$$

Notice for future use that the integral of the vector potential around any closed path which

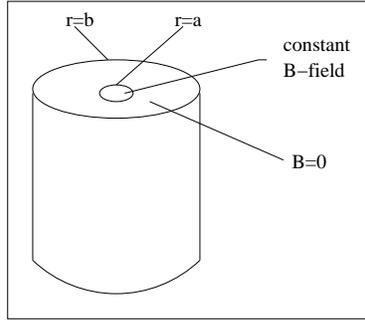


Figure 4.2: The solenoid.

encloses the solenoid is equal to the magnetic flux carried by the solenoid,

$$\oint \vec{A} \cdot d\vec{l} = \int \vec{B} \cdot \hat{n} dS = \Phi_B. \quad (4.54)$$

Imagine that we make the solenoid infinitesimally thin, so that classically a particle could never penetrate into a region of nonzero B , while holding the flux constant.

Schrödinger's equation involves the momentum operator as

$$\hat{p}_\phi \rightarrow \hat{p}_\phi - \frac{eA_\phi}{c} \quad (4.55)$$

or in coordinate basis

$$\frac{1}{r} \frac{\partial}{\partial \phi} \rightarrow \frac{1}{r} \frac{\partial}{\partial \phi} - \frac{ieBa^2}{2\hbar c \pi r}. \quad (4.56)$$

The Hamiltonian – and thus the energy eigenstates – must depend explicitly on B even though classically the Lorentz force is zero. This effect is completely quantum mechanical.

The real Aharanov-Bohm effect involves the influence of a solenoid on a diffraction pattern, say on two-slit diffraction. As shown in Fig. 4.3, the solenoid is placed between the screen with its slits and the plane with our particle detector. The fringes of the pattern will shift as a function of the B-field even though the B-field is confined within the solenoid. We can analyze the shift using path integrals. The classical Lagrangian for a particle in an external electromagnetic field is

$$L = \frac{1}{2} m \dot{x}^2 + \frac{e}{c} \dot{\vec{x}} \cdot \vec{A} \quad (4.57)$$

Consider the change in the action as we move from (x_{n-1}, t_{n-1}) to (x_n, t_n) along some path.

$$S = \int_{t_{n-1}}^{t_n} dt L = S_0 + \int_{t_{n-1}}^{t_n} dt \frac{e}{c} \frac{d\vec{x}}{dt} \cdot \vec{A}$$

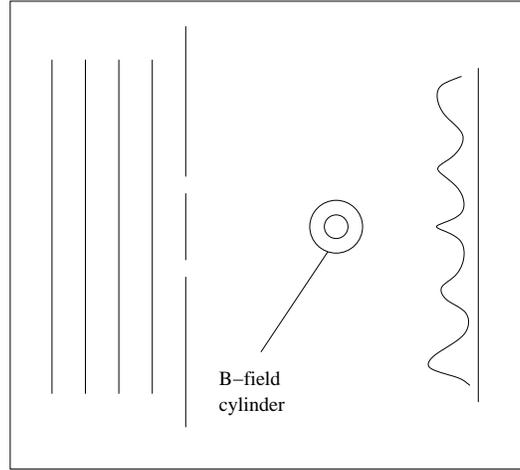


Figure 4.3: Experimental setup for the Aharanov-Bohm effect.

$$= S_0 + \frac{e}{c} \int_{x_{n-1}}^{x_n} \vec{A} \cdot d\vec{s}. \quad (4.58)$$

Adding up all the little intervals, we discover that there is a phase factor associated with this path

$$e^{\frac{iS}{\hbar}} = e^{\frac{iS_0}{\hbar}} e^{\frac{ie}{\hbar c} \int \vec{A} \cdot d\vec{s}}. \quad (4.59)$$

Now let us look at a few of the paths, as in Fig. 4.4. We will group them according to which side of the solenoid they pass. For the quantum mechanical transition amplitude we need to add up all the factors of $\exp \int \vec{A} \cdot d\vec{s}$ from all the paths. Group the paths two-by-two, one from each side of the solenoid, before counting them all up. That is, first combine terms as

$$\int_{path (a)} [dx] e^{\frac{iS_0}{\hbar}} e^{\frac{ie}{\hbar c} \int_a^b \vec{A} \cdot d\vec{s}} + \int_{path (b)} [dx] e^{\frac{iS_0}{\hbar}} e^{\frac{ie}{\hbar c} \int_a^b \vec{A} \cdot d\vec{s}}. \quad (4.60)$$

The intensity of the diffraction pattern is given by the square of the amplitude, whose contribution from these paths is

$$P(b) = |\psi|^2 = \left| \int_{path (a)} [dx] e^{\frac{iS_a}{\hbar}} \right|^2 + \left| \int_{path (b)} [dx] e^{\frac{iS_b}{\hbar}} \right|^2 + 2Re \left[\int [dx] e^{\frac{iS_a}{\hbar}} \int [dx] e^{\frac{-iS_b}{\hbar}} \right]. \quad (4.61)$$

In the direct terms, the A -dependent pieces square to unity and disappear. In the cross term, however, they contribute a term proportional to

$$\exp(i/\hbar \int_{(a)}^{(b)} \vec{A} \cdot d\vec{l}|_{P_1} + i/\hbar \int_{(b)}^{(a)} \vec{A} \cdot d\vec{l}|_{P_2}) \quad (4.62)$$

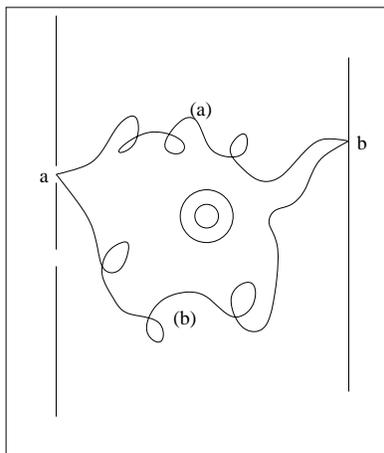


Figure 4.4: Representative paths in the Aharanov-Bohm apparatus.

which is the exponential of the line integral around the solenoid. Now recall Eq. 4.54: this is

$$\exp \frac{ie\Phi_B}{\hbar c}. \quad (4.63)$$

The diffraction pattern will be modified by the presence of the solenoid, even though the particle never traverses a region of nonzero magnetic field. The interference pattern could be observed by tuning the magnetic flux: the pattern would repeat as Φ_B changes by an amount $2\pi\hbar c/e$.

It seems that the vector potential is “real,” after all. How can we reconcile this fact with the notion of local gauge invariance? We must think of the vector potential as being a fundamental quantity, in the sense that it appears in the equations of motion, but it is a quantity which contains redundant information. The redundancy is expressed through the invariance of physical observables under local gauge transformations. As far as we know, it does not make sense to try to reduce this redundancy in any simple way (for example, it does not make sense to say that one gauge choice is more fundamental than another one).

Notice that the observable (the shift in the diffraction pattern) does depend on a gauge invariant quantity, the magnetic flux. We expect that physical observables will be gauge invariant. That is a different statement from the statement that only the fields should enter into the dynamics. In fact, the statement that only the fields appear in observables is special to electrodynamics. Only in electrodynamics are the fields gauge invariant while the potentials are gauge variant. Other interactions, such as the strong or weak interactions, are described by gauge theories with more complicated internal symmetry. The analogs

of the field variables in these theories are not gauge invariant; they change under gauge transformations. Only physically observable quantities, such as the energy density carried by the field, remain gauge invariant in these more general cases.

Let us conclude this chapter by recalling Max Born: “The world of our imagination is narrower and more special in its logical structure than the world of physical things.”

Chapter 5

Density matrices

Introduction to density matrices

We often encounter situations in physics where we are only interested in a subset of the degrees of freedom of a physical system, and we have to deal somehow with the degrees of freedom we want to leave out. We have already mentioned this in passing – truncating an atomic system down to a two state system. Sometimes we just ignore the unwanted degrees of freedom, but this is dangerous, if the part of the system we are studying can interact with them. What to do next is somewhat process - dependent, but the language is similar across multiple areas of physics: we talk about “integrating out” the unwanted degrees of freedom, leaving behind the ones we are interested in, and we aim to find some alteration in their dynamics due to the variables we have removed.

One of the techniques for doing this involves an object called the density matrix. To introduce it, we begin with some language: The variables we are interested in are called the “system,” which communicates with a “reservoir” (in statistical mechanics language), namely, everything else. We start by thinking about everything (system plus reservoir); we end with the system. We would like to set things up so that we can apply the formalism to cases where our knowledge of the reservoir is incomplete or lacking.

We start by assuming that we have a complete set of states of the system, $|\phi_i\rangle$, and a complete set of states for everything else, $|\theta_j\rangle$. The most general wave function we can construct, which includes the system and reservoir, can be written as a superposition of product states

$$|\psi\rangle = \sum_{ij} c_{ij} |\phi_i\rangle |\theta_j\rangle. \quad (5.1)$$

We can let x be a coordinate of the system, and y label the coordinates of the reservoir variables, and write the combined wave function in coordinate space as

$$\begin{aligned} \psi(x, y) &= (\langle y| \langle x|) |\psi\rangle \\ &= \sum_{ij} c_{ij} \langle x|\phi_i\rangle \langle y|\theta_j\rangle \\ &= \sum_i c_i(y) \phi_i(x) \end{aligned} \quad (5.2)$$

where the expansion coefficient is a sum over the reservoir wave functions

$$c_i(y) = \sum_j c_{ij} \langle y|\theta_j\rangle. \quad (5.3)$$

This expresses the coordinate space wave function of both the system and reservoir as a superposition of wave functions of the system alone.

Next, suppose that we have a set of operators \hat{A} which act only on the system, $\hat{A} |\phi_i\rangle |\theta_j\rangle = (\hat{A} |\phi_i\rangle) |\theta_j\rangle$. We can write these operators as projectors acting on the combined wave function

$$\begin{aligned}\hat{A} &= \sum_{i,i',j} (|\phi_i\rangle |\theta_j\rangle) \langle \phi_i | \hat{A} | \phi_{i'} \rangle (\langle \phi_{i'} | \langle \theta_j |) \\ &= \sum_{i,i',j} A_{ii'} |\phi_i\rangle |\theta_j\rangle \langle \theta_j | \langle \phi_{i'} |.\end{aligned}\tag{5.4}$$

An expectation value of the operator can then be expressed as its system expectation value, via

$$\begin{aligned}\langle \psi | \hat{A} | \psi \rangle \equiv \langle \hat{A} \rangle &= \sum_{i,j,i',j'} c_{ij}^* c_{i'j'} \langle \theta_j | \langle \phi_i | \hat{A} | \phi_{i'} \rangle | \theta_{j'} \rangle \\ &= \sum_{i,i',j} c_{ij}^* c_{i'j} \langle \phi_i | \hat{A} | \phi_{i'} \rangle,\end{aligned}\tag{5.5}$$

the last line arising because of the orthogonality of the $|\theta_j\rangle$'s. The final expression can be written compactly as

$$\langle \hat{A} \rangle = \sum_{i,i'} \langle \phi_i | \hat{A} | \phi_{i'} \rangle \rho_{i',i}\tag{5.6}$$

where

$$\rho_{i',i} = \sum_j c_{ij}^* c_{i'j}\tag{5.7}$$

is called the “density matrix.” It carries all the information about the reservoir, needed to evaluate matrix elements of the system. Note that $\rho_{i',i} = \rho_{i,i'}^*$, which means that ρ is Hermitian. We can define an operator $\hat{\rho}$, which only acts on the system or x -variables, as

$$\rho_{i,i'} = \langle \phi_i | \hat{\rho} | \phi_{i'} \rangle.\tag{5.8}$$

Then, using completeness, the expectation value of \hat{A} becomes

$$\begin{aligned}\langle \psi | \hat{A} | \psi \rangle &= \sum_{i,i'} \rho_{i',i} \langle \phi_i | \hat{A} | \phi_{i'} \rangle \\ &= \sum_{i,i'} \langle \phi_i | \hat{A} | \phi_{i'} \rangle \langle \phi_{i'} | \hat{\rho} | \phi_i \rangle\end{aligned}$$

$$\begin{aligned}
&= \sum_i \langle \phi_i | \hat{\rho} \hat{A} | \phi_i \rangle \\
&= \text{Tr } \hat{\rho} \hat{A}.
\end{aligned} \tag{5.9}$$

Because $\hat{\rho}$ is Hermitian, its eigenvalues w_i are real and it can be diagonalized with a complete orthonormal set of eigenvectors

$$\hat{\rho} = \sum_i w_i |i\rangle \langle i|. \tag{5.10}$$

The w_i 's have two useful properties: they are all positive, and they sum to unity. We can discover these properties by considering two examples.

First, if \hat{A} is equal to the identity operator

$$1 = \langle \psi | 1 | \psi \rangle = \text{Tr } \hat{\rho} 1 = \text{Tr } \hat{\rho} = \sum_i w_i. \tag{5.11}$$

Next, if \hat{A} is a projector, $\hat{A} = |j\rangle \langle j|$, then

$$\text{Tr } \hat{\rho} \hat{A} = w_j = \langle \psi | \hat{A} | \psi \rangle. \tag{5.12}$$

We recall that this quantity is the norm of a state, and hence it is always positive or zero:

$$\begin{aligned}
\langle \psi | \hat{A} | \psi \rangle &= \sum_k [\langle \psi | j \rangle \langle \theta_k |] \langle \theta_k | j \rangle \langle \psi | \\
&= \sum_k |\langle \psi | j \rangle \langle \theta_k | j \rangle|^2 \geq 0.
\end{aligned} \tag{5.13}$$

That is, $w_j \geq 0$ and $\sum_j w_j = 1$.

Quantum mechanics using density matrices

Let us think about quantum mechanics in terms of $\hat{\rho}$. Any system can be described by a density matrix $\hat{\rho} = \sum_i w_i |i\rangle \langle i|$, where

- $|i\rangle$ is complete and orthonormal
- $w_i \geq 0$

- $\sum_i w_i = 1$
- $\langle \hat{A} \rangle = \text{Tr } \hat{\rho} \hat{A}$

We can expand $\langle \hat{A} \rangle$ in eigenstates of $\hat{\rho}$. This gives $\langle \hat{A} \rangle = \sum_j w_j \langle j | \hat{A} | j \rangle$. Because $\langle j | \hat{A} | j \rangle$ is the expectation value of the operator \hat{A} in state $|j\rangle$, we can interpret w_j as the probability that the system is in the density matrix eigenstate $|j\rangle$.

If all $w_i = 0$, except for one $w_j = 1$ (recall the sum rule), then we say that the system is in a “pure state.” In all other cases, we say that the system is in a “mixed state.” In a pure state, $\hat{\rho} = |i_{\text{pure}}\rangle \langle i_{\text{pure}}|$, and its matrix element in any other basis is

$$\rho_{ij} = \langle \phi_i | i_{\text{pure}} \rangle \langle i_{\text{pure}} | \phi_j \rangle = \langle \phi_i | i_{\text{pure}} \rangle \langle \phi_j | i_{\text{pure}} \rangle^*. \quad (5.14)$$

Otherwise, we must use the full expression

$$\rho_{ij} = \sum_k w_k \langle \phi_i | k \rangle \langle \phi_j | k \rangle^*. \quad (5.15)$$

Notice that for a pure state $\hat{\rho}^2 = |i_{\text{pure}}\rangle \langle i_{\text{pure}}| |i_{\text{pure}}\rangle \langle i_{\text{pure}}| |i_{\text{pure}}\rangle \langle i_{\text{pure}}| = \hat{\rho}$. In situations where you are not sure whether you have a mixed state or a pure one, $\hat{\rho}^2 = \hat{\rho}$ is a useful diagnostic.

When we have a pure state, expectation values are the usual quantum mechanical ones,

$$\langle A \rangle = \langle i_{\text{pure}} | A | i_{\text{pure}} \rangle. \quad (5.16)$$

Furthermore, recall Eq. 5.1. If we have a pure state, the general formula

$$\langle \psi | \hat{A} | \psi \rangle = \sum_{i,i',j} c_{ij}^* c_{i'j} \langle \phi_i | \hat{A} | \phi_{i'} \rangle, \quad (5.17)$$

collapses to a single term. The coefficient c_{ij} must vanish unless $i = i_{\text{pure}}$. The combined state of system and reservoir becomes

$$|\psi\rangle = \left[\sum_j c_{i_{\text{pure}}j} |\theta_j\rangle \right] |i_{\text{pure}}\rangle. \quad (5.18)$$

That is, we only have a pure state when the system and reservoir are completely decoupled.

If it makes sense to talk about our system in a coordinate basis, then the density matrix becomes

$$\rho(x', x) \equiv \langle x' | \hat{\rho} | x \rangle = \sum_k w_k \langle x' | k \rangle \langle k | x \rangle = \sum_k k(x) k(x')^*, \quad (5.19)$$

(this is just a change of notation, but potentially an evocative one), and a matrix element is

$$\langle \hat{A} \rangle = \text{Tr } \hat{\rho} \hat{A} = \int dx \langle x | \hat{\rho} \hat{A} | x \rangle. \quad (5.20)$$

The second term is

$$\begin{aligned} \langle x | \hat{\rho} \hat{A} | x \rangle &= \int dx' \langle x | \hat{\rho} | x' \rangle \langle x' | \hat{A} | x \rangle \\ &= \int dx' \rho(x, x') A(x', x) \end{aligned} \quad (5.21)$$

and so

$$\langle \hat{A} \rangle = \int dx dx' \rho(x, x') A(x', x). \quad (5.22)$$

Returning to our earlier discussion, the expectation value of the operator is taken between states of the entire universe. Recall that we labeled the system's coordinates as x and the reservoir's, as y . Recall also that the operator \hat{A} only acted on the system. Then we can write directly

$$\langle \hat{A} \rangle = \int dx dx' dy \psi^*(x', y) A(x', x) \psi(x, y). \quad (5.23)$$

This means that yet another definition for the density matrix is

$$\rho(x, x') = \int dy \psi^*(x', y) \psi(x, y). \quad (5.24)$$

Again, the density matrix is the integral over the unseen variables of the combined probability.

Some examples

The simplest example of a density matrix is for a two-state system

$$\begin{aligned} \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\equiv |+\rangle \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} &\equiv |-\rangle. \end{aligned} \quad (5.25)$$

$|+\rangle$ and $|-\rangle$ could label spins, up or down, or two orthogonal polarization states of light (x and y for linear polarization). Any pure state is

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (5.26)$$

where $|\alpha|^2 + |\beta|^2 = 1$. For a pure state, $\hat{\rho} = |i_{\text{pure}}\rangle \langle i_{\text{pure}}|$, or

$$\rho = \begin{pmatrix} \alpha\alpha^* & \beta\alpha^* \\ \alpha\beta^* & \beta\beta^* \end{pmatrix}. \quad (5.27)$$

Note $\text{Tr } \rho = 1$, $\rho_{ii} > 0$, and $\rho^2 = \rho$, as required. Various examples of pure states, expressed physically as polarization states, are

- x -polarized light: $\alpha = 1, \beta = 0, \rho_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
- y -polarized light: $\alpha = 0, \beta = 1, \rho_y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
- 45° polarization: $\alpha = \beta = \frac{1}{\sqrt{2}}, \rho_{45} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
- 135° polarization: $\alpha = -\beta = -\frac{1}{\sqrt{2}}, \rho_{135} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

Examples of mixed states are

- 50% x , 50% y : $\rho = \frac{1}{2}(\rho_x + \rho_y) = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$
- 50% 45° , 50% 135° : $\rho = \frac{1}{2}(\rho_{45} + \rho_{135}) = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$

These two mixed states have the same density matrix and correspond to the same physical effect. They are both realizations of an unpolarized beam of light.

A slightly more complicated example is a mixed spin state for electrons, in which a fraction $1 - x$ is in a state where the spin is in the $+z$ state and a fraction x where the spin

is in the $+x$ -state.

$$\begin{aligned}\rho &= (1-x) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + x \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} 1 - \frac{x}{2} & \frac{x}{2} \\ \frac{x}{2} & \frac{x}{2} \end{pmatrix}.\end{aligned}\tag{5.28}$$

The second term is built using the pure state

$$|x\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.\tag{5.29}$$

Can you verify that

$$\langle\sigma_z\rangle = \text{Tr } \rho\sigma_z = 1 - x,\tag{5.30}$$

$$\langle\sigma_x\rangle = \text{Tr } \rho\sigma_x = x,\tag{5.31}$$

and

$$\langle\sigma_y\rangle = \text{Tr } \rho\sigma_y = 0?\tag{5.32}$$

Notice how the usual quantum mechanical behavior appears in the limits when the density matrix becomes that of a pure state.

As a final comment, notice the distinction between a pure state which is a product state such as

$$|\psi\rangle = |s\rangle |E\rangle\tag{5.33}$$

and one which is a superposition, such as

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|s_1\rangle |E_1\rangle + |s_2\rangle |E_2\rangle).\tag{5.34}$$

Tracing over the environment ($|E\rangle$) in Eq. 5.33 obviously leaves a pure state behind. But tracing over the environment for the state of Eq. 5.34 leaves behind a mixed state. (We did the algebra for two-state systems in our discussion of polarized and unpolarized photon beams.) States like Eq. 5.34 (more generally, any state which is not a product state) are called an “entangled states.” Generally, system states which are entangled with environment states show a loss of quantum coherence, equivalently a loss of unitary time evolution. This would be the first sentence in a discussion of quantum information theory, were I competent enough to do so. We will, however, talk a bit more about entangled states when we describe the addition of angular momentum.

Time evolution and related questions

If $|i\rangle$ changes with time, so does the density matrix. Explicitly,

$$\hat{\rho}(t) = \sum_i w_i |i(t)\rangle \langle i(t)|. \quad (5.35)$$

A useful set of relations based on this expression can be had if we expand $|i(t)\rangle$ in energy eigenstates. Then, if at time $t = 0$ the state is

$$|i(0)\rangle = \sum_n |E_n\rangle \langle E_n|i\rangle, \quad (5.36)$$

at later times it becomes

$$|i(t)\rangle = \sum_n |E_n\rangle e^{-iE_n t/\hbar} \langle E_n|i\rangle = e^{-iHt/\hbar} |i(0)\rangle \quad (5.37)$$

and the density matrix is

$$\begin{aligned} \hat{\rho}(t) &= \sum_i w_i e^{-iHt/\hbar} |i(0)\rangle |i(t)\rangle \langle i(t)| e^{iHt/\hbar} |i(0)\rangle \\ &= e^{-iHt/\hbar} \hat{\rho}(0) e^{iHt/\hbar}. \end{aligned} \quad (5.38)$$

Differentiating this expression gives an equation for the time evolution of $\hat{\rho}$:

$$i\hbar \frac{d\hat{\rho}}{dt} = H\hat{\rho} - \hat{\rho}H. \quad (5.39)$$

Notice that this has the opposite sign to the Heisenberg equation of motion for an observable \hat{A} ,

$$i\hbar \frac{d\hat{A}}{dt} = \hat{A}H - H\hat{A}, \quad (5.40)$$

a curiosity, but not a problem – the density matrix is just not a Heisenberg representation observable. The trace is time independent – $\text{Tr} \hat{\rho}(t) = \text{Tr} e^{-iHt/\hbar} \hat{\rho}(0) e^{iHt/\hbar} = \text{Tr} \hat{\rho}(0)$. Does this seem sensible?

Let us look at a few more examples using density matrices. Recall that a coordinate space expectation value is

$$\langle \hat{A} \rangle = \int dx dx' \rho(x, x') A(x', x). \quad (5.41)$$

Consider first $A = x$, the coordinate operator. Since $\langle x|x|x'\rangle = x\delta(x - x')$,

$$\langle x \rangle = \int dx x \rho(x, x). \quad (5.42)$$

Next is the momentum operator. In coordinate space, $\langle x|p|x'\rangle = \frac{\hbar}{i} \frac{\partial}{\partial x} \delta(x - x')$. Inserting this result in our standard formula and integrating by parts,

$$\langle p \rangle = - \int dx dx' \frac{\hbar}{i} \frac{\partial}{\partial x} \rho(x, x') = - \frac{\hbar}{i} \int dx \left[\frac{\partial}{\partial x} \rho(x, x') \right]_{x=x'}. \quad (5.43)$$

For our last example, we compute the probability that we make a measurement and observe the system to be in a state $|\chi\rangle$. When we do this, we are measuring the expectation value of the projector $\hat{A} = |\chi\rangle \langle \chi|$. This is

$$\begin{aligned} \langle \hat{A} \rangle &= \text{Tr} \hat{\rho} |\chi\rangle \langle \chi| \\ &= \text{Tr} \left(\sum_i w_i |i\rangle \langle i| \right) |\chi\rangle \langle \chi| \\ &= w_i \langle \chi|o\rangle|^2 \\ &= \langle \chi| \sum_i w_i |i\rangle \langle i| \chi \rangle \\ &= \langle \chi| \hat{\rho} | \chi \rangle. \end{aligned} \quad (5.44)$$

The interpretation is clear – we measure the expectation value of the density matrix in our desired state.

To illustrate this result, let us return to our complicated example, with the density matrix of Eq. 5.28. What is the probability to measure $|+\rangle$? It is

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 - \frac{x}{2} & \frac{x}{2} \\ \frac{x}{2} & \frac{x}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 - \frac{x}{2}. \quad (5.45)$$

The $|+y\rangle$ state is $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$. Can you show that the probability of observing this state, given the density matrix, is $\frac{1}{2}$?

Quantum canonical ensemble

To formulate quantum statistical mechanics requires the use of density matrices. That need is most obvious for the canonical ensemble, where the system is in contact with a reservoir of temperature T . Here, the probability that the system is found in a state $|\phi_i\rangle$ is

$$P(|\phi_i\rangle) = \frac{e^{-\beta E_i}}{Z} \quad (5.46)$$

where k is Boltzmann's constant, T is the temperature, $\beta^{-1} = kT$, and the normalizing factor Z is the partition function, to be defined below. The density matrix, then, is diagonal in an energy basis (labeled by $|\phi_n\rangle$),

$$\hat{\rho} = \sum_n w_n |\phi_n\rangle \langle \phi_n| \quad (5.47)$$

where $w_n = \exp(-\beta E_n)/Z$. Thus

$$\hat{\rho} = \frac{1}{Z} \sum_n e^{-\beta H} |\phi_n\rangle \langle \phi_n| = \frac{e^{-\beta H}}{Z} \quad (5.48)$$

(writing the density matrix as an operator) and the partition function is

$$Z = \sum_n e^{-\beta E_n} = \text{Tr} e^{-\beta H} \equiv e^{-\beta F} \quad (5.49)$$

where F is the Helmholtz free energy. Thus the normalized density matrix is

$$\hat{\rho} = \frac{e^{-\beta H}}{\text{Tr} e^{-\beta H}}. \quad (5.50)$$

An obvious example to check is to find the partition function for a free particle. To do that, let us pause to consider the un-normalized ρ as a function of β ,

$$\rho(\beta) = \exp(-\beta H). \quad (5.51)$$

Note that $\rho(0) = 1$ and

$$-\frac{\partial}{\partial \beta} \rho(\beta) = H \rho(\beta). \quad (5.52)$$

Can you confirm this result, by expanding in energy eigenfunctions? In coordinate space, relabeling $H \rightarrow H_x$, the equation of motion is

$$-\frac{\partial}{\partial \beta} \rho(x, x', \beta) = H_x \rho(x, x', \beta), \quad (5.53)$$

with $\rho(x, x', 0) = \delta(x - x')$ expressing the identity operator in coordinate basis.

Let's use this expression to solve for a single one-dimensional particle in a heat bath. Here $H = p^2/(2m)$ and Eq. 5.53 becomes

$$-\frac{\partial}{\partial \beta} \rho(x, x', \beta) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \rho(x, x', \beta). \quad (5.54)$$

With the boundary condition. $\rho(x, x', 0) = \delta(x - x')$, this is the equation for the Green's function for diffusion, and we write down the answer by inspection,

$$\rho(x, x', \beta) = \sqrt{\frac{m}{2\pi\hbar^2\beta}} \exp\left(-\frac{m}{2\beta\hbar^2}(x - x')^2\right). \quad (5.55)$$

Confine the particle on a line of length L to evaluate the partition function, and

$$Z_1 = e^{-\beta F} = \int dx \rho(x, x) = L \sqrt{\frac{m}{2\pi\hbar^2\beta}}. \quad (5.56)$$

The three-dimensional system's partition function is the cube of this expression (L^3 gives the familiar volume factor) and naively generalizing to N particles raises Z_1^3 to the power N to give the usual free particle partition function. Since we have not considered the effects of identical particles, our answer does not treat the entropy of mixing properly, but that can be accounted for simply dividing by $N!$. We are going too far afield, into statistical mechanics, to continue the discussion any farther.

Euclidean time path integral

Notice that the un-normalized thermal density matrix, Eq. 5.51, is identical to the time evolution operator, with the substitution of $-i/\hbar$ for $-\beta$. This means that we can use our calculation of the path integral formulation of the time evolution operator to construct a path integral expression for the partition function. Before we copy expressions, it is useful to make a change of variables, introducing $u = \beta\hbar$. Then the evolution equation becomes

$$\hbar \frac{\partial}{\partial u} \rho = -H\rho. \quad (5.57)$$

Note that u has dimensions of time, so that we can interpret our equation as one for evolution in imaginary time, with a formal solution $\rho = \exp(-Hu/\hbar)$. The free particle propagator is replaced by the diffusion Green's function, Eq. 5.54, and the expression for the evolution operator, Eq. 4.26, becomes

$$Z = \text{Tr } \rho = \left(\frac{m}{2\pi i \hbar \Delta t}\right)^{\frac{N}{2}} \int dx_{N-1} \cdots dx_1 \exp\left(\frac{-\Delta t}{\hbar} \sum_{j=1}^N \left(\frac{1}{2} m \left(\frac{x_{j+1} - x_j}{\Delta t}\right)^2 + V(x_j)\right)\right). \quad (5.58)$$

Differences in the two expressions (the absence of matrix elements at the beginning and end times, the fact that there are only $N - 1$ integrals over dx , and the so-far-unstated constraint

that $x_N = x_1$) are a consequence of the fact that Z is a trace. In the limit that $\Delta t \rightarrow 0$ the exponential becomes

$$\exp\left(-\frac{1}{\hbar} \int_0^{\tau=\beta/\hbar} dt' \left(\frac{1}{2}m\dot{x}^2(t') + V(x(t'))\right)\right). \quad (5.59)$$

The integral is called the “Euclidean action” S_E associated with a path $x(t)$, periodic so $x(0) = x(T)$ – again the trace at work. Note that it is minus the sum of kinetic and potential energies.

Just as in the case of the ordinary path integral, we can ask, “Which paths are important?” The answer, here, is that the paths for which S_E/\hbar are smallest are the important ones.

There might be many of them! However, consider the classical limit: $\beta \rightarrow 0$ or $T \rightarrow \infty$. The range of the integral, $\tau = \beta/\hbar$, becomes small. If the important paths are those for which x does not change, the density matrix becomes

$$\rho(x, x, T) = \left[\int \prod dx_j \exp\left(-\frac{1}{\hbar} \int_0^\tau dt' \frac{1}{2}m\dot{x}^2(t')\right) \right] \exp(-\beta V(x)). \quad (5.60)$$

The object in square brackets is the free particle propagator, Eq. 5.55, evaluated at $x = x'$. The partition function is

$$Z = \int dx \sqrt{\frac{m}{2\pi\hbar^2\beta}} \exp(-\beta V(x)), \quad (5.61)$$

the familiar classical result for a particle in a potential at finite temperature.

This formalism is particularly useful for carrying out numerical studies of statistical mechanics systems. Many general properties of quantum mechanical systems can also be obtained through its generalizations. Note that all non-commutativity has disappeared from Eq. 5.58. It is just an N -dimensional integral over a set of classical variables x_k . Computers can do this integral using what are called “importance sampling” techniques. If we can generate a set of N “typical configurations,” which are sets of variables $x_i = \{x_k^i\}$, whose density in the ensemble is proportional to $\exp(-S_E(x_i)/\hbar)$, then thermal averages are just averages over the ensemble of configurations, and the thermal expectation value of any function of x is

$$\langle O(x) \rangle = \frac{1}{N} \sum_{i=1}^N O(x_i) + O\left(\frac{1}{\sqrt{N}}\right). \quad (5.62)$$

Because the thermal weighting probability is real, these expressions are generally quite stable. Contrast this behavior to the sum of complex amplitudes resulting from the real-time path

integral! The energies of low lying states, and expectation values of matrix elements, can be computed using products of operators at different Euclidean times, such as $x(t_1)x(t_2)$.

Chapter 6

Wave mechanics

In this chapter we return to what is (hopefully) more familiar ground for the reader: the Schrödinger equation as a wave equation. We will explore its behavior in a variety of simple cases, attempting to extract behavior which is common to all similar situations even though the differential equation itself may not have such simple solutions.

Wave mechanics in one dimension

The time independent Schrödinger equation is

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V(x)\psi = E\psi. \quad (6.1)$$

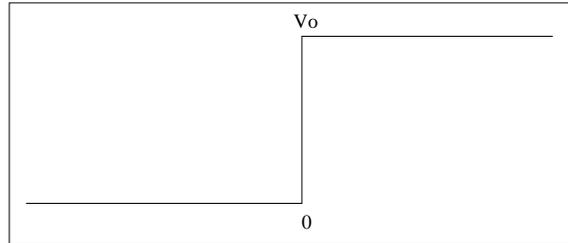
For constant $V(x) = V$, and if $E > V$, the most general solution is

$$\psi_E(x) = Ae^{\frac{ipx}{\hbar}} + Be^{-\frac{ipx}{\hbar}}, \quad \frac{p^2}{2m} = E - V. \quad (6.2)$$

If $E < V$, in contrast,

$$\psi_E(x) = A'e^{\frac{px}{\hbar}} + B'e^{-\frac{px}{\hbar}}, \quad \frac{p^2}{2m} = V - E. \quad (6.3)$$

Now let us consider a case where the potential has a step, $V(x) = 0$ for $x < 0$; $V(x) = V_0$ for $x > 0$.



We assume $E > 0$. On the left hand side of the boundary

$$\psi_L = Ae^{\frac{ip_L x}{\hbar}} + Be^{-\frac{ip_L x}{\hbar}}, \quad p_L = \sqrt{2mE} \quad (6.4)$$

and on the right hand side, if $E > V_0$ we have

$$\psi_R = Ce^{\frac{ip_R x}{\hbar}} + De^{-\frac{ip_R x}{\hbar}}, \quad p_R = \sqrt{2m(E - V_0)}. \quad (6.5)$$

We know that the wave equation itself must be continuous at the boundary between region I and region II, so

$$A + B = C + D \quad (6.6)$$

What about the first derivative? Let us integrate the Schrödinger equation from $x = -\epsilon$ to $x = \epsilon$, across the boundary,

$$\frac{-\hbar^2}{2m} \frac{\partial \Psi}{\partial x} \Big|_{-\epsilon}^{\epsilon} + \int_{-\epsilon}^{\epsilon} (V(x) - E) \Psi(x) dx = 0. \quad (6.7)$$

Making ϵ an infinitesimal, the second term in the expression vanishes. (Note that this would NOT be the case if $V(x)$ were a delta function.) Thus the first derivative must also be continuous across the boundary,

$$\frac{-\hbar^2}{2m} \frac{\partial \Psi}{\partial x} \Big|_{-\epsilon}^{\epsilon} = 0. \quad (6.8)$$

In the example we are considering, this gives

$$p(A - B) = p'(C - D). \quad (6.9)$$

We still have a problem, though. We have four unknowns in our wave equations, but we only have two constraint equations. We need a way to resolve this difficulty. We cannot use normalization as a constraint, because plane waves are not normalizable functions. We must, instead, carefully restate the physical situation. Let us suppose that we are considering the case of a beam coming in from the right and either scattering from, or reflecting off, the barrier. To the right of the barrier, we only have an outgoing wave. This means that we can set $D = 0$. Next, if we imagine performing this experiment, we can see that we will have control over the coefficient A , the amplitude of the incoming particles. We do not have control over B and C because they are the amplitude of the reflected and transmitted waves and those will in principle depend on the potential barrier. However, we expect that B and C will depend linearly on A . (After all, the rate at which particles are reflected from the barrier depends on the rate that they strike it.) Thus there are two relevant quantities, the relative reflection amplitude B/A and relative transmission amplitude C/A . (One often drops the word “relative.”) The squared modulus $\frac{|C|^2}{|A|^2}$ is known as the transmission coefficient, T , and the squared modulus $\frac{|B|^2}{|A|^2}$ is known as the reflection coefficient, R , in analogy with optics usage.

This discussion has been a bit informal, so let us pause to define a particle current. Currents obey the usual relation to the time rate of change of a density

$$\vec{\nabla} \cdot \vec{J} + \frac{\partial \rho}{\partial t} = 0, \quad (6.10)$$

where the function $\rho = \Psi^*(\vec{x}, t)\Psi(\vec{x}, t)$ is just the modulus squared of the wave function. The first time derivative of the function ρ is

$$\frac{\partial \rho}{\partial t} = \frac{\partial \Psi^*}{\partial t} \Psi(x, t) + \Psi^*(x, t) \frac{\partial \Psi}{\partial t}. \quad (6.11)$$

Invoking the time-dependent Schrödinger equation,

$$i\hbar \frac{\partial \Psi}{\partial t} = \frac{-\hbar^2}{2m} \nabla^2 \Psi + V\Psi \quad (6.12)$$

we obtain by substitution

$$\frac{\partial \rho}{\partial t} = \frac{1}{i\hbar} \left[-\left(\frac{-\hbar^2}{2m} \nabla^2 \Psi^* + V\Psi^* \right) \Psi + \Psi^* \left(\frac{-\hbar^2}{2m} \nabla^2 \Psi + V\Psi \right) \right] \quad (6.13)$$

or

$$\frac{\partial \rho}{\partial t} = -\vec{\nabla} \cdot \left[\Psi^* \frac{\hbar}{2mi} \vec{\nabla} \Psi + \frac{1}{2m} \left(\frac{\hbar}{i} \vec{\nabla} \Psi \right)^* \Psi \right]. \quad (6.14)$$

So we identify the probability current as

$$\vec{J} = \frac{\hbar}{2mi} [\Psi^* \vec{\nabla} \Psi - (\vec{\nabla} \Psi)^* \Psi]. \quad (6.15)$$

As an example let us consider a plane wave of the form $\Psi(x) = \exp i\vec{k}\vec{r}$. Its current is

$$\vec{J} = \frac{1}{2m} [e^{-i\vec{k}\vec{r}} (\hbar\vec{k}) e^{i\vec{k}\vec{r}} - \frac{\hbar}{i} (-i\vec{k}) e^{-i\vec{k}\vec{r}} e^{i\vec{k}\vec{r}}] = \frac{\hbar\vec{k}}{m}. \quad (6.16)$$

Note $\hbar\vec{k}/m$ is, as expected, just the velocity.

Now we return to the barrier problem. Current conservation should hold at the barrier, so

$$|A|^2 \frac{\hbar\vec{k}}{m} - |B|^2 \frac{\hbar\vec{k}}{m} = |C|^2 \frac{\hbar\vec{k}'}{m}. \quad (6.17)$$

From this we can again see what matters physically is the proportionality between A and B and between A and C , not the absolute values. Solving the two equations (continuity of the wave function and continuity of the derivative), we find the following values for $\frac{B}{A}$ and $\frac{C}{A}$.

$$\frac{B}{A} = \frac{p - p'}{p + p'} \quad (6.18)$$

$$\frac{C}{A} = \frac{2p}{p + p'} \quad (6.19)$$

The reader should check that current conservation is obeyed.

So much for the case where $E > V_0$. When $E < V_0$ the solutions on the right hand side, Ψ_R , will be real exponentials, and

$$\Psi_R = C e^{-p'x/\hbar} + D e^{p'x/\hbar} \quad (6.20)$$

where $p'^2 = 2m(V_0 - E)$. Because the potential barrier extends all the way to positive infinity, the exponential with the positive argument must vanish; otherwise the wave function in that region is not normalizable. This forces $D = 0$. We can recycle our work from the $E > V_0$ case by making the substitution $p' \rightarrow ip'$. Then

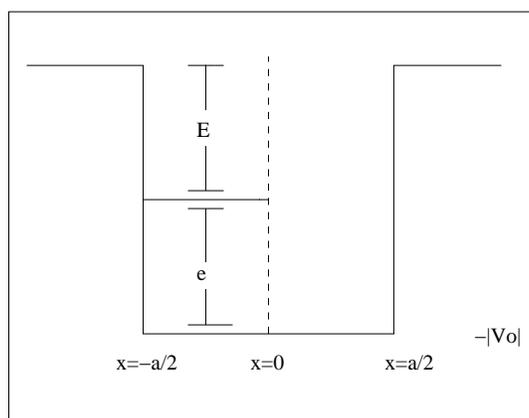
$$\frac{B}{A} = \frac{p - ip'}{p + ip'} \quad (6.21)$$

$$\frac{C}{A} = \frac{2p}{p + ip'} \quad (6.22)$$

Note that the squared modulus of the factor, $\frac{B}{A}$, is just equal to unity. From this one can deduce that then the quantity, $\frac{B}{A}$, must be equal to a pure phase. Thus our reflected wave is just a phase shifted reversal of our original wave.

Particle in a finite well

Next we consider a potential well $V(x) = V_0$ for $-a/2 < x < a/2$, $V(x) = 0$ otherwise. We let the potential be attractive, with well depth $-|V_0|$, as illustrated in the figure.



If this potential possesses any bound states, they must have negative energy. We will label the bound state energy by $-E$. Equivalently, and more usefully for the algebra, any bound state will have energy ϵ above the bottom of the well. Inside the box, wave functions have the following form

$$\Psi(x) = Ae^{ikx} + Be^{-ikx} \quad (6.23)$$

with

$$\frac{\hbar^2 k^2}{2m} = \epsilon \quad (6.24)$$

where

$$|E| + \epsilon = |V_0|. \quad (6.25)$$

Outside the box, the wave function dies exponentially, $\Psi(x) \propto e^{-\alpha|x|}$ with

$$\frac{\hbar^2 \alpha^2}{2m} = |V_0| - \epsilon = |E|. \quad (6.26)$$

To continue, we must match boundary conditions for the wave function and its derivative at $x = \pm a/2$: there are four such equations, all told.

Parity

It is worthwhile, once again, to pause and think formally, this time about mirror symmetry or parity. We will assume that we are in three spatial dimensions for this section. The parity operator \hat{P} performs the transformation $\Psi(\vec{x}) \rightarrow \Psi(-\vec{x})$. Thus

$$\hat{P}\Psi(\vec{x}) = \Psi(-\vec{x}). \quad (6.27)$$

Since it flips the sign of the coordinate, the parity operator also obeys the operator equation

$$\hat{P}^\dagger \hat{\vec{r}} \hat{P} = -\hat{\vec{r}}. \quad (6.28)$$

This relation should be true for all vector operators, not just \vec{r} . The parity operator is unitary. To see this, apply it twice:

$$\hat{P}^\dagger \hat{\vec{r}} \hat{P} \cdot \hat{P}^\dagger \hat{\vec{r}} \hat{P} = \vec{r} \cdot \vec{r}. \quad (6.29)$$

The quantity r^2 is obviously rotationally invariant, so the only way the two sides will be equal is if $\hat{P}^\dagger = \hat{P}^{-1}$, which is the definition of unitarity. The operator \hat{P} can act on state vectors as well as on operators.

$$\langle \vec{r} | \hat{P} | \Psi \rangle = \langle -\vec{r} | \Psi \rangle = \Psi(-\vec{r}). \quad (6.30)$$

Applied to the Hamiltonian, it gives

$$\hat{P}^\dagger \hat{H}(\vec{r}) \hat{P} = \hat{H}(-\vec{r}). \quad (6.31)$$

We know that if \hat{P} commutes with \hat{H} then we can construct simultaneous eigenvectors of \hat{H} and \hat{P} . Commutivity requires that $\hat{P}\hat{H} = \hat{H}\hat{P}$, and, multiplying both sides by \hat{P}^\dagger ,

$$\hat{P}^\dagger \hat{P} \hat{H} = \hat{P}^\dagger \hat{H} \hat{P} \quad (6.32)$$

or, because \hat{P} is unitary,

$$\hat{H} = \hat{P}^\dagger \hat{H} \hat{P} \quad (6.33)$$

or

$$\hat{H}(\vec{r}) = \hat{H}(-\vec{r}). \quad (6.34)$$

The squared momentum is obviously rotationally invariant, so whether or not parity commutes with the Hamiltonian depends on the form of the potential. For example, in three dimensions, parity is a good quantum number for potentials depending only on the radial coordinate $V(r)$. Obviously, we say that a state has “positive parity” when

$$\Psi(\vec{x}) = \Psi(-\vec{x}) \quad (6.35)$$

and a “negative parity” state obeys

$$\Psi(\vec{x}) = -\Psi(-\vec{x}). \quad (6.36)$$

Parity appears in many physical applications. For example, integrals of the form

$$\int d^3x \Psi_i(\vec{x}) \vec{r} \Psi_j(\vec{x}) \quad (6.37)$$

vanish if Ψ_i and Ψ_j have the same parity. (This is the mathematical origin for selection rules for radiative transitions in atoms.) We can see this result formally using

$$\langle i | \vec{r} | j \rangle = -\langle i | \hat{P}^\dagger \vec{r} \hat{P} | j \rangle. \quad (6.38)$$

If

$$\hat{P} | i \rangle = (-1)^{P_i} | i \rangle \quad (6.39)$$

and

$$\hat{P} | j \rangle = (-1)^{P_j} | j \rangle, \quad (6.40)$$

then

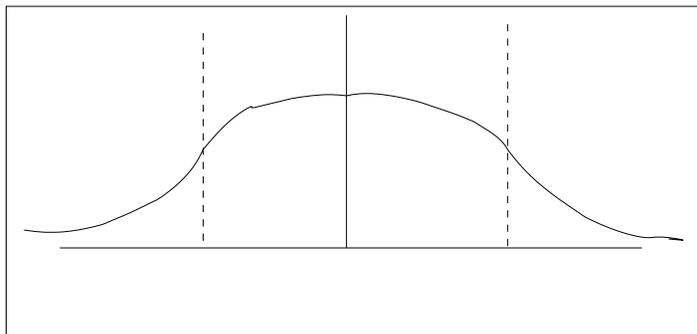
$$\langle i | \vec{r} | j \rangle = -(-1)^{(P_i + P_j)} \langle i | \vec{r} | j \rangle. \quad (6.41)$$

Now back to the one-dimensional square well. The potential is symmetric about its midpoint, so energy eigenfunctions will also be parity eigenfunctions. For even parity states, the interior solution is

$$\Psi = A \cos kx \quad (6.42)$$

and outside the well

$$\Psi = B e^{-\alpha|x|}. \quad (6.43)$$

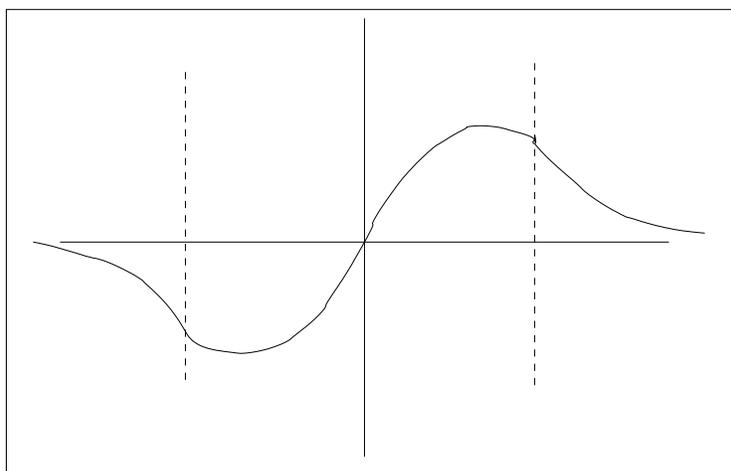


Similarly, odd parity wave functions are

$$\Psi = A' \sin kx \quad (6.44)$$

inside the well and outside the well

$$\Psi = \pm B' e^{-\alpha|x|}. \quad (6.45)$$



For each parity, we have three unknowns (E,A,B), and we have three constraints on these equations: continuity of the wave function, as well as the continuity of its first spatial derivative, at $x = a/2$, and a normalization condition on the wave function. For the even parity states, the first two boundary conditions give

$$A \cos\left(\frac{ka}{2}\right) = B e^{-\alpha\frac{a}{2}} \quad (6.46)$$

$$-Ak \sin\left(\frac{ka}{2}\right) = -B\alpha e^{-\alpha\frac{a}{2}}. \quad (6.47)$$

Dividing the second equation by the first equation, and inserting an 'a' (the width of the barrier) on both sides for ease of interpretation, we get

$$ka \tan\left(\frac{ka}{2}\right) = a\alpha. \quad (6.48)$$

At $x = \frac{a}{2}$ the odd parity states obey

$$A' \sin\left(\frac{ka}{2}\right) = B' e^{-\alpha|\frac{a}{2}|} \quad (6.49)$$

and

$$A' k \cos\left(\frac{ka}{2}\right) = -B' \alpha e^{-\alpha|\frac{a}{2}|} \quad (6.50)$$

or

$$ka \cot\left(\frac{ka}{2}\right) = -a\alpha. \quad (6.51)$$

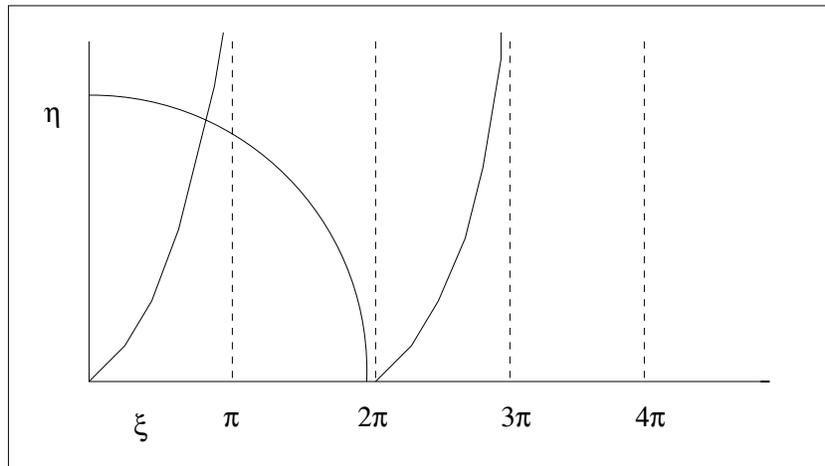
Introducing $\xi \equiv ka$ and $\eta \equiv \alpha a$, the defining relations between V_0 , E , α and k become

$$\xi^2 + \eta^2 = \frac{2ma^2}{\hbar^2}(E + V_0 - E) = \frac{2ma^2}{\hbar^2}V_0. \quad (6.52)$$

The continuity equation for the even parity solutions becomes

$$\xi \tan\left(\frac{\xi}{2}\right) = \eta. \quad (6.53)$$

Eqs. 6.52 and 6.53 may be solved graphically. Allowed values of ξ and the η will form a circle of radius $\sqrt{\frac{2ma^2V_0}{\hbar^2}}$, which must intersect the function $\xi \tan\left(\frac{\xi}{2}\right)$. The situation for even parity is shown in the following figure:



Notice that there is always at least one positive parity bound state, because no matter how small is the circle formed by ξ and η , it will always intersect the $\xi \tan\left(\frac{\xi}{2}\right)$ function. The number of bound states is given by the number of intersections. Clearly, as V_0 becomes bigger and bigger, the number of intersections rises, and there are more bound states. For example, two solutions will exist when

$$2\pi \leq \sqrt{\frac{2ma^2V_0}{\hbar^2}} \leq 4\pi, \quad (6.54)$$

three solutions will exist when

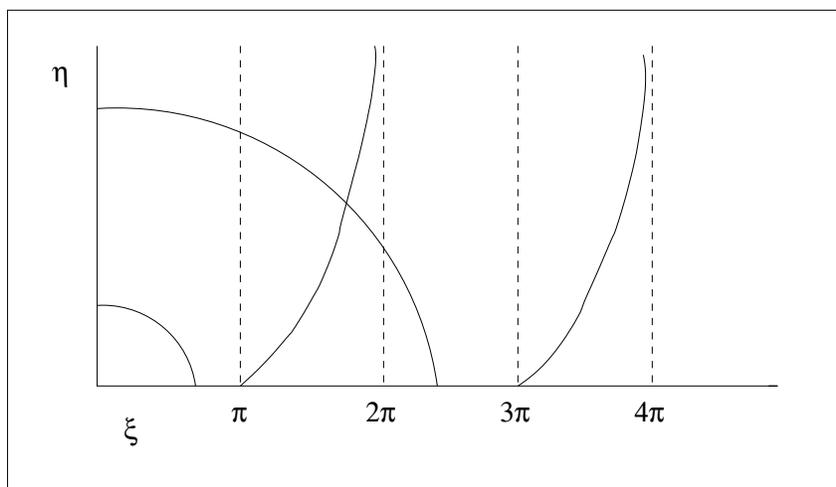
$$4\pi \leq \sqrt{\frac{2ma^2V_0}{\hbar^2}} \leq 6\pi, \quad (6.55)$$

and so on. As $V_0 \rightarrow \infty$ the number of solutions will also approach infinity. The eigenvalue equation becomes $\xi = (2n + 1)\pi$ for $n = 0, 1, 2, \dots$, the expected result for the particle in the infinitely deep well.

The analysis for the odd parity states is similar. The second eigenvalue equation becomes

$$-\xi \cot\left(\frac{\xi}{2}\right) = \eta. \quad (6.56)$$

The situation is shown in the following figure. Again, solutions exist where the curves intersect.

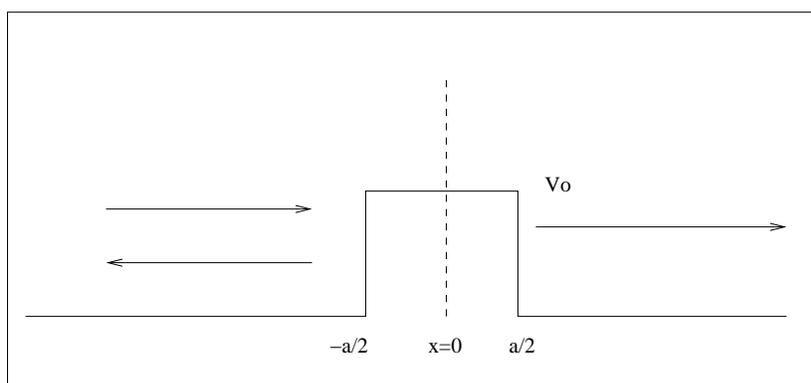


In contrast to the positive parity case, if the radius of the circle, $\sqrt{\frac{2ma^2V_0}{\hbar^2}}$, is less than π there will be no bound state solutions. Thus if the finite well is not deep enough, no odd parity bound states will exist. As the well deepens, the number of solutions for odd parity

states will behave qualitatively like even parity case. Again, if $V_0 \rightarrow \infty$ solutions will be of the form $\xi = 2\pi n$ where $n = 1, 2, \dots$. Thus we recover the familiar case of the infinitely deep well.

Barrier penetration

Next, consider scattering from a rectangular barrier of constant potential and finite width. We assume the beam is entering from the left. The potential is $V(x) = V_0$ for $-a/2 < x < a/2$, $V(x) = 0$ otherwise.



Outside of the barrier we have our familiar plane wave solution; on the left, there are incoming and reflected waves,

$$u(x) = Ae^{ikx} + B^{-ikx} \quad (6.57)$$

and only a transmitted wave on the right,

$$u(x) = Ce^{ikx}, \quad (6.58)$$

where $\frac{\hbar^2 k^2}{2m} = E$. The transmission coefficient is $\frac{|C|^2}{|A|^2}$ and the reflection coefficient is $\frac{|B|^2}{|A|^2}$. To simplify all our results, we will simply scale A to 1. Let's begin with the case where the energy of the wave, $E > V_0$. In this case the solution in the interior of the barrier oscillates,

$$u(x) = Fe^{i\alpha x} + Ge^{-i\alpha x} \quad (6.59)$$

Applying our continuity boundary conditions at $x = -\frac{a}{2}$, we have

$$e^{-ika/2} + Be^{ika/2} = Fe^{-i\alpha a/2} + Ge^{i\alpha a/2} \quad (6.60)$$

and

$$ik[e^{-ika/2} - Be^{ika/2}] = i\alpha[Fe^{-i\alpha a/2} - Ge^{i\alpha a/2}]. \quad (6.61)$$

Similarly, at $x = \frac{a}{2}$ we have

$$Ce^{ika/2} = Fe^{i\alpha a/2} + Ge^{-i\alpha a/2} \quad (6.62)$$

and

$$ik[Ce^{ika/2}] = i\alpha[Fe^{i\alpha a/2} - Ge^{-i\alpha a/2}]. \quad (6.63)$$

This constitutes a 4×4 matrix of equations whose solution gives B , C , F , and G . The general solution is not illuminating, and so we consider interesting special cases. First, suppose $\alpha a = 2\pi n$. The matching equations reduce to

$$e^{-ika/2} + Be^{ika/2} = e^{i\pi n}[F + G] \quad (6.64)$$

$$ik[e^{-ika/2} - Be^{ika/2}] = i\alpha e^{i\pi n}[F - G] \quad (6.65)$$

$$Ce^{ika/2} = e^{i\pi n}[F + G] \quad (6.66)$$

$$ik[Ce^{ika/2}] = i\alpha e^{i\pi n}[F - G]. \quad (6.67)$$

A few lines of algebra quickly reveal that $B = 0$, $C = \exp(-ika)$. The interpretation is easy: The reflection coefficient is 0 and thus there is no reflected wave at the potential barrier. In this case we have what is called a transmission resonance. It occurs only for particular values of the wavelength and energy, when $\alpha^2 = 2m(E - V_0) = 4\pi^2/a^2$. This is qualitatively similar to the situation of Brewster's angle in an optical system, where at a particular angle of incidence one polarization component will be transmitted completely.

The complete solution for the transmission and reflection amplitudes are

$$\frac{B}{A} = \frac{(k^2 - \alpha^2)(1 - e^{2i\alpha a})e^{-ika}}{(k + \alpha)^2 - (k - \alpha)^2 e^{2i\alpha a}} \quad (6.68)$$

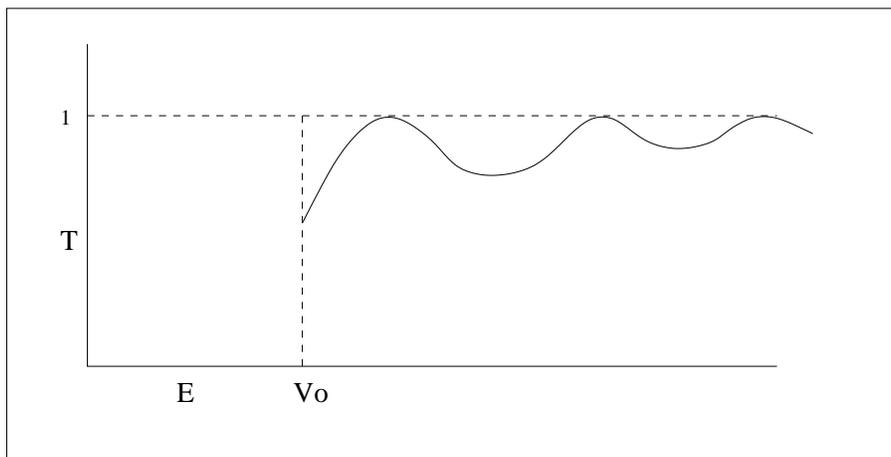
$$\frac{C}{A} = \frac{4k\alpha e^{i(\alpha - k)a}}{(k + \alpha)^2 - (k - \alpha)^2 e^{2i\alpha a}}. \quad (6.69)$$

The transmission coefficient is

$$T = \frac{|C|^2}{|A|^2} = \left[1 + \frac{(k^2 - \alpha^2)^2 \sin^2(\alpha a)}{4k^2 \alpha^2} \right]^{-1} \quad (6.70)$$

$$= \left[1 + \frac{V_0^2 \sin^2(\alpha a)}{4E(E - V_0)} \right]^{-1}. \quad (6.71)$$

Recall that this is for the case $E \geq V_0$. The graph of this transmission coefficient is shown in the figure; note the transmission resonances:



When $E \leq V_0$ the solution of the Schrödinger equation inside the barrier is

$$u(x) = Fe^{\beta x} + Ge^{-\beta x} \quad (6.72)$$

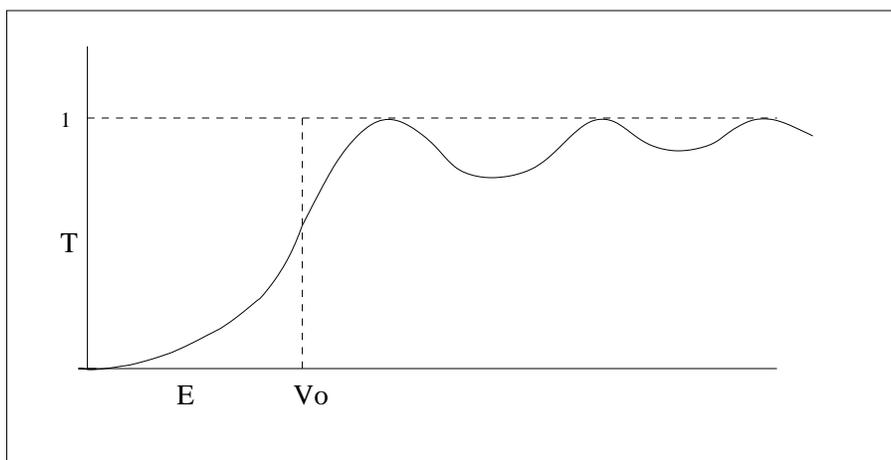
where $\hbar^2\beta^2 = 2m(V_0 - E)$. The easiest way to find the transmission amplitude in this case is by analytic continuation $\alpha \rightarrow i\beta$ from our previous result. This gives

$$T = \left[1 - \frac{V_0^2 \sinh^2(\beta a)}{4E(E - V_0)} \right]^{-1}. \quad (6.73)$$

At small E/V_0 the transmission falls to zero exponentially with increasing barrier width or height,

$$T \simeq \frac{1}{\frac{V_0^2}{16E(V_0 - E)} e^{2\beta a}} \simeq e^{-2\beta a}. \quad (6.74)$$

The following figure displays this behavior.



It is easy to obtain an approximate solution for a very tall ($V_0 \gg E$) and very wide $e^{\beta a} \gg 1$ barrier. For ease of calculation, let us shift the location of the barrier from $-\frac{a}{2} < x < \frac{a}{2}$ to the range $0 < x < a$. The interior wave function is

$$u(x) = Fe^{\beta x} + Ge^{-\beta x}. \quad (6.75)$$

In order to keep this solution finite, the coefficient F , must be very tiny. More precisely, at $x = a$ we need to keep $F \exp(\beta a) \propto O(1)$. Thus for the equations to be consistent, we need to retain terms involving $F \exp(\beta a)$, but can discard terms that only involve F . This amounts to replacing the exact boundary conditions by approximate ones: At $x = 0$,

$$1 + B = F + G \simeq G \quad (6.76)$$

and

$$ik(1 - B) = \beta(F - G) \simeq -\beta G. \quad (6.77)$$

They imply

$$G = \frac{2ik}{ik - \beta}. \quad (6.78)$$

Similarly, at $x = a$ we have

$$Ce^{ika} \equiv C' = Fe^{\beta a} + Ge^{-\beta a} \quad (6.79)$$

and

$$ikC' = \beta(Fe^{\beta a} - Ge^{-\beta a}) = \beta(C' - 2Ge^{-\beta a}) \quad (6.80)$$

because $Fe^{\beta a} = (C' - Ge^{-\beta a})$. Thus

$$(ik - \beta)C' = -2Ge^{-\beta a} \quad (6.81)$$

and so

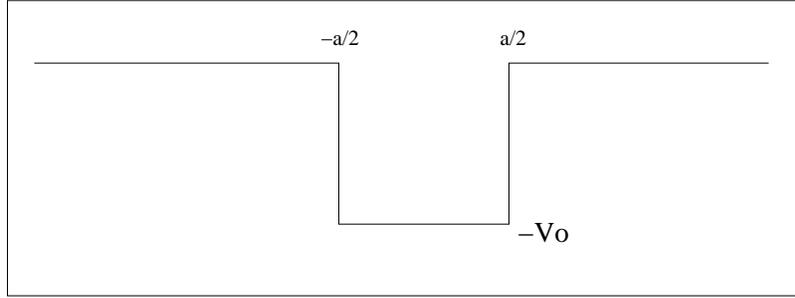
$$C' = \frac{-2\beta Ge^{-\beta a}}{ik - \beta} = \frac{-4ik\beta e^{-\beta a}}{(ik - \beta)^2}. \quad (6.82)$$

Thus we have found the transmission coefficient in the limit of a tall, wide barrier to be

$$T = \frac{16k^2\beta^2 e^{-2\beta a}}{(k^2 + \beta^2)^2}, \quad (6.83)$$

which agrees with Eq. 6.74.

To study scattering from an attractive potential well,



we can take all of our formulas and simply change the sign of V_0 . Then for $E \geq V_0$ the transmission coefficient is

$$T(E) = \left[1 + \frac{V_0^2 \sin^2(\alpha a)}{4E(E + |V_0|)} \right]^{-1}. \quad (6.84)$$

We still have transmission resonances, whenever $\alpha a = \pi n$. In that case

$$E + V_0 = \frac{4\pi^2 \hbar^2 n^2}{2ma^2}. \quad (6.85)$$

Stepping back for a moment, we recall that the transmission amplitude is a complex function of the energy, or the wave number, k .

$$\frac{C}{A} = S(E) = \frac{4k\alpha e^{i(\alpha-k)a}}{(k+\alpha)^2 - (k-\alpha)^2 e^{2i\alpha a}} \equiv e^{-ika} \tilde{S}(E) \quad (6.86)$$

with

$$\tilde{S}(E) = \frac{1}{\cos(\alpha a) - \frac{i}{2} \left(\frac{k}{\alpha} + \frac{\alpha}{k} \right) \sin(\alpha a)}. \quad (6.87)$$

We observe that $\tilde{S}(E)$ will have a pole whenever

$$\cos(\alpha a) = \frac{i}{2} \left(\frac{k}{\alpha} + \frac{\alpha}{k} \right) \sin(\alpha a) \quad (6.88)$$

The improbable trigonometric identity $\tan(2x) = 2/(\cot(x) - \tan(x))$ transforms the pole condition into

$$\cot\left(\frac{\alpha a}{2}\right) - \tan\left(\frac{\alpha a}{2}\right) = i \left[\frac{k}{\alpha} + \frac{\alpha}{k} \right] \quad (6.89)$$

so that a pole appears whenever either of the following two relations holds.

$$\tan\left(\frac{\alpha a}{2}\right) = \frac{-ik}{\alpha} \quad (6.90)$$

or

$$\cot\left(\frac{\alpha a}{2}\right) = \frac{ik}{\alpha}, \quad (6.91)$$

To convert these relations into statements about energy, we recall $\hbar^2 k^2 = 2mE$ and $\hbar^2 \alpha^2 = 2m(E + V_0)$. A pole will only occur if $E < 0$ and $\alpha > 0$. In that case the \sqrt{E} must be treated with care. It has a branch cut in the complex energy plane. To locate the branch cut we define

$$E = |E|e^{i\phi} \quad (6.92)$$

so that

$$\sqrt{E} = |E|^{\frac{1}{2}} e^{\frac{i\phi}{2}}. \quad (6.93)$$

Then Eqs. 6.90 and 6.91 become

$$\alpha \tan\left(\frac{\alpha a}{2}\right) = K \quad (6.94)$$

and

$$\alpha \cot\left(\frac{\alpha a}{2}\right) = -K \quad (6.95)$$

where

$$iK \equiv \frac{i\sqrt{2m|E|}}{\hbar} = k. \quad (6.96)$$

These are the equations which gave the values of bound state energies, Eqs. 6.53 and 6.56. This is a general result: the scattering amplitude has a pole at all energies corresponding to the locations of bound states in the potential. (Note that these energies may not be directly accessible to experiment: for this potential, the bound states have negative energy but only positive energy beams can be produced.) Physically, the pole corresponds to the presence of a solution near the potential which is very large compared to the value of the solution far away from the potential. This can only occur when the energy associated with the solution is in fact the energy of a bound state.

Now we examine how $\tilde{S}(E)$ behaves near a transmission resonance. We assume that we have $E > 0$ and that a transmission resonance occurs when $\alpha a = \pi n$ or $E = E_0$. We can rewrite the transmission coefficient as

$$\tilde{S}(E) = \frac{1}{\cos(\alpha a) \left[1 - \frac{i}{2} \left(\frac{k}{\alpha} + \frac{\alpha}{k} \right) \tan(\alpha a) \right]}. \quad (6.97)$$

A transmission resonance occurs when $\tan(\alpha a) \simeq 0$. Let us Taylor expand the denominator in a power series about E_0 . Writing the result as

$$\left(\frac{k}{\alpha} + \frac{\alpha}{k} \right) \tan(\alpha a) \equiv \frac{4}{\Gamma} (E - E_0) \quad (6.98)$$

where

$$\frac{4}{\Gamma} \equiv \frac{d}{dE} \left(\left(\frac{k}{\alpha} + \frac{\alpha}{k} \right) \tan(\alpha a) \right) \Big|_{\alpha a = \pi n}, \quad (6.99)$$

we discover that near the resonance the scattering amplitude is

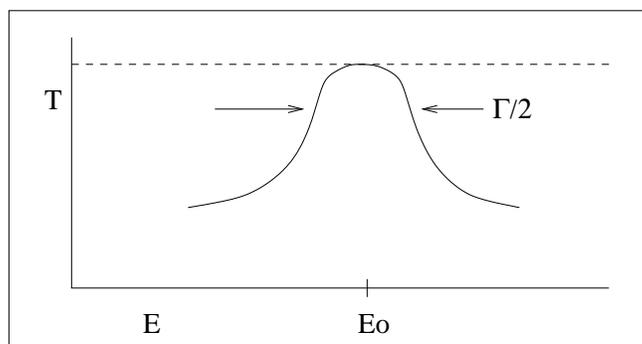
$$\tilde{S}(E) \simeq \frac{1}{\cos(\alpha a)} \left(\frac{\frac{i\Gamma}{2}}{E - E_0 + \frac{i\Gamma}{2}} \right). \quad (6.100)$$

At the resonance $\cos(\alpha a) \rightarrow 1$, so the only important term is that inside the parentheses.

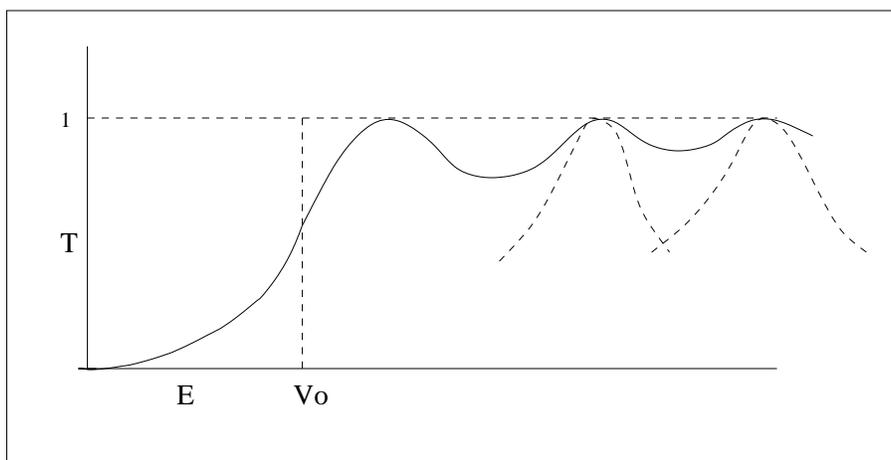
The transmission coefficient near a resonance is

$$T(E) \propto |\tilde{S}(E)|^2 = \frac{\frac{\Gamma^2}{4}}{(E - E_0)^2 + \frac{\Gamma^2}{4}}. \quad (6.101)$$

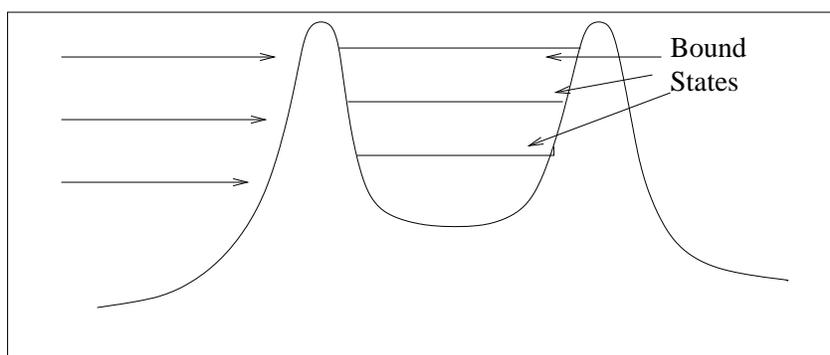
This function is the same Lorentzian lineshape that we had discussed earlier (in the example of magnetic resonance). It is also called a “Breit-Wigner” lineshape.



Note that the approximation we have made is only valid near the transmission resonance. In fact, looking at the graph of the transmission coefficient, we see that in this case, it is not a very good approximation away from a resonance.



Again, we observe that the transmission amplitude has a pole at $E = E_0 - \frac{i\Gamma}{2}$. When the energy is near this value in the complex energy plane, the nearby singularity completely dominates the function. Of course, at physically accessible energies, the scattering amplitude cannot diverge, for that would lead to a violation of unitarity. The largest it can be is a pure phase, in this case, $\tilde{S}(E) = 1$. The scattering amplitude has poles at $E < 0$, but these energies are not accessible in a scattering experiment. If we had a potential of the following form,



we could probe the bound states by scattering $E > 0$ particles and we would see resonant behavior whenever we were near a bound state energy. Examples of such behavior occur frequently in Nature; we will encounter it again when we study three dimensional scattering.

Scattering and the one-dimensional delta function potential

The delta function potential

$$V(x) = V_0\delta(x) \quad (6.102)$$

is a useful approximation to a situation where the range of the potential is very small compared to the distances over which we are able to probe in a scattering experiment. There are “conventional” ways to solve the Schrödinger equation (relegated to homework problems). Let us consider an unusual one: The Schrödinger equation.

$$\left(E + \frac{\hbar^2}{2m} \frac{d^2}{dx^2}\right)\Psi(x) = V_0\delta(x)\Psi(x) = V_0\delta(x)\Psi(0) \quad (6.103)$$

can be replaced by an integral equation

$$\Psi(x) = \Psi_i(x) + \int_{-\infty}^{\infty} \frac{dp}{2\pi\hbar} \frac{e^{ipx/\hbar}}{E - \frac{p^2}{2m}} V_0\Psi(0), \quad (6.104)$$

where the homogeneous solution is that $\Psi_i(x) = e^{ipx/\hbar}$. To check the validity of this solution we apply the operator $\left(E + \frac{\hbar^2}{2m} \frac{d^2}{dx^2}\right)$ to both sides, producing

$$\begin{aligned} \left(E + \frac{\hbar^2}{2m} \frac{d^2}{dx^2}\right) \Psi(x) &= 0 + \int_{-\infty}^{\infty} \frac{dp}{2\pi\hbar} \frac{E - \frac{p^2}{2m}}{E - \frac{p^2}{2m}} e^{ipx/\hbar} V_0 \Psi(0) \\ &= V_0 \Psi(0) \int_{-\infty}^{\infty} \frac{dp}{2\pi\hbar} e^{ipx/\hbar} \\ &= V_0 \Psi(0) \delta(x). \end{aligned} \tag{6.105}$$

Unfortunately, our “solution” is not complete – the integral is singular, due to the pole at $E = p^2/(2m)$.

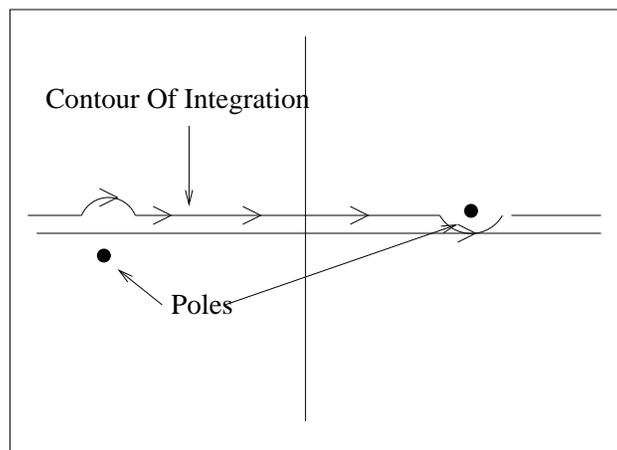
We must combine a physical problem – what boundary conditions do we want for $\Psi(x)$? – with a mathematical problem – how can we distort the integration contour to avoid the singularity? We can approach the solution by thinking of the potential as a source of waves. Then, we want an inhomogeneous solution that has outgoing waves (away from the potential) in both directions, positive and negative x . That is, we must require

$$\Psi(x) \propto e^{ikx} \quad x > 0 \tag{6.106}$$

and

$$\Psi(x) \propto e^{-ikx} \quad x < 0. \tag{6.107}$$

Mathematically, we can achieve this result by picking a contour of integration in the complex momentum plane as shown:



We can close the contour in the upper or lower half plane, picking a contour which gives zero contribution to the integral. To determine how to do this, we must consider separately the cases of x positive or negative. For $x > 0$ we need to close the contour in the upper half of the plane so that its positive imaginary momentum will give a negative exponential. The residue theorem gives the following for $x > 0$.

$$\oint \frac{dp}{2\pi\hbar} \frac{e^{ipx/\hbar}}{E - \frac{p^2}{2m}} = \frac{m \exp[i\sqrt{2mE}x/\hbar]}{i\hbar \sqrt{2mE}} \quad (6.108)$$

For $x < 0$ we close the contour in the lower half plane, giving

$$\oint \frac{dp}{2\pi\hbar} \frac{e^{ipx/\hbar}}{E - \frac{p^2}{2m}} = \frac{m \exp[-i\sqrt{2mE}x/\hbar]}{i\hbar \sqrt{2mE}} \quad (6.109)$$

Thus our solution for $\Psi(x)$ is

$$\Psi(x) = \Psi_i(x) + \frac{mV_0}{ip\hbar} e^{ip|x|/\hbar} \Psi(0) \quad (6.110)$$

or, solving for $\Psi(0)$ and collecting terms,

$$\Psi(x) = e^{ipx/\hbar} + \frac{mV_0}{ip\hbar - mV_0} e^{ip|x|/\hbar}. \quad (6.111)$$

The $x > 0$ solution gives the transmitted wave

$$\Psi_{trans} = \frac{ip\hbar}{ip\hbar - mV_0} e^{ipx/\hbar}, \quad (6.112)$$

from which, the scattering amplitude is

$$\tilde{S}(E) = \frac{ip\hbar}{ip\hbar - mV_0} \quad (6.113)$$

The function $\tilde{S}(E)$ has a pole whose location tells us the value of the bound state energy. The pole is at

$$p\hbar = \hbar\sqrt{2mE} = -imV_0. \quad (6.114)$$

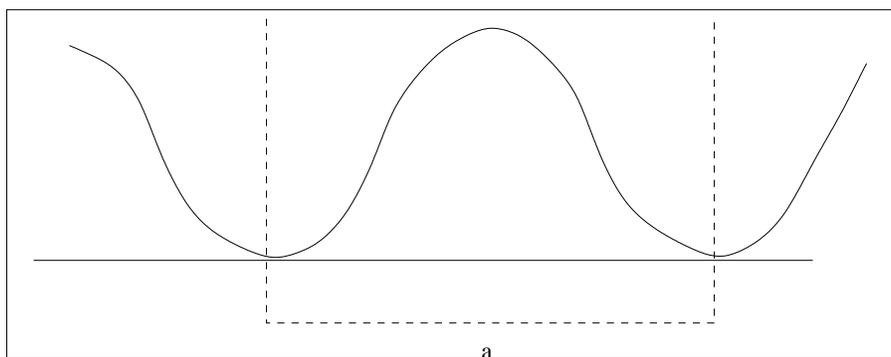
To complete the calculation, the bound state energy must be negative, and

$$E = \frac{-mV_0^2}{2\hbar^2}. \quad (6.115)$$

Note that the one dimensional delta function potential has only a single bound state.

Bloch waves and periodic potentials (in one dimension)

Imagine a particle in a one-dimensional periodic potential $V(a) = V(x + a)$, as illustrated in the figure,



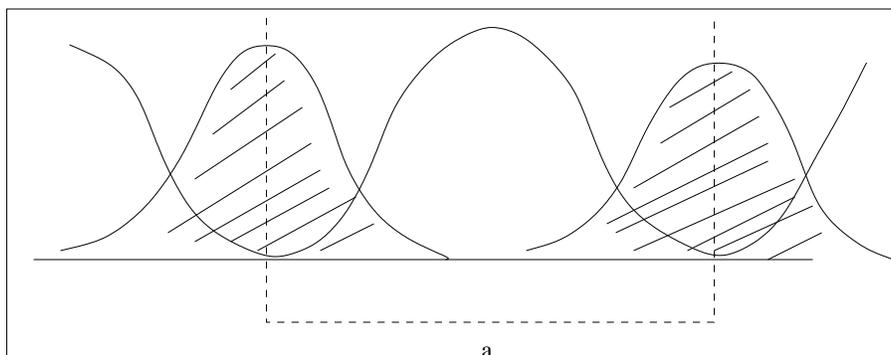
where a is the length of one period of the potential. Associated with this situation we can define an operator $\hat{\tau}$, called the translation operator, with the property that

$$\hat{\tau}x\hat{\tau}^\dagger = x + a \quad (6.116)$$

This operator will transform potential $\hat{\tau}V(x)\hat{\tau}^\dagger = V(x + a)$. For a periodic potential this is equal to $V(x)$. This means that the Hamiltonian itself also obeys $\hat{\tau}\hat{H}\hat{\tau}^\dagger = \hat{H}$, and, since $\hat{\tau}$ is unitary (prove it!), $[\hat{H}, \hat{\tau}] = 0$. We should be able to find simultaneous eigenfunctions of \hat{H} and $\hat{\tau}$.

The tight binding approximation

To proceed, let us suppose we have a set of eigenfunctions $|n\rangle$ such that each of them is localized around one of the wells of the potential. We label the one in the j th potential well as $|j\rangle$.



Now imagine that the wells of $V(x)$ are sufficiently deep and sufficiently far apart, that there is no overlap between the wave functions of neighboring wells. This means that $\langle j|j'\rangle = \delta_{jj'}$. Because of the translational invariance of the Hamiltonian, the expectation value of the Hamiltonian for a state in any particular well should equal that of any other well. We will call this energy $\langle j|\hat{H}|j\rangle = E_0$. Next, let us assume that the only nonvanishing matrix elements of the Hamiltonian connect neighboring states. We parameterise this matrix element as $\langle j'|\hat{H}|j\rangle = -\Delta\delta_{j'j\pm 1}$. These equations compose what is referred to in solid state physics as the “tight binding assumption.”

The translation operator shifts states: $\hat{\tau}|j\rangle = |j+1\rangle$. What are its eigenstates? Let us consider the following state.

$$|\theta\rangle = \sum_{n=-\infty}^{\infty} e^{in\theta} |n\rangle. \quad (6.117)$$

Applying the translation operator to it gives

$$\hat{\tau}|\theta\rangle = \sum_{n=-\infty}^{\infty} e^{in\theta} |n+1\rangle \quad (6.118)$$

or, shifting the index of the sum,

$$\hat{\tau}|\theta\rangle = \sum_{n=-\infty}^{\infty} e^{i(n-1)\theta} |n\rangle = e^{-i\theta} |\theta\rangle. \quad (6.119)$$

Therefore

$$\hat{\tau}|\theta\rangle = e^{-i\theta} |\theta\rangle. \quad (6.120)$$

The $|\theta\rangle$ states are eigenstates of the translation operator with eigenvalues of pure phase. These states are eigenstates of H . To see this, begin with the action of H on a localized state

$$\hat{H}|n\rangle = E_0|n\rangle - \Delta|n+1\rangle - \Delta|n-1\rangle. \quad (6.121)$$

Use this result to apply the Hamiltonian to the states $|\theta\rangle$.

$$\hat{H}|\theta\rangle = E_0|\theta\rangle - \Delta\left(\sum_{n=-\infty}^{\infty} e^{in\theta} |n+1\rangle\right) - \Delta\left(\sum_{n=-\infty}^{\infty} e^{in\theta} |n-1\rangle\right) \quad (6.122)$$

$$= E_0|\theta\rangle - \Delta(e^{-\theta} + e^{i\theta})|\theta\rangle \quad (6.123)$$

or

$$\hat{H}|\theta\rangle = (E_0 - 2\Delta\cos(\theta))|\theta\rangle. \quad (6.124)$$

Indeed, the $|\theta\rangle$ state is also an energy eigenstate. Notice that the energy lies in the range

$$E \in [E_0 - 2\Delta, E_0 + 2\Delta]. \quad (6.125)$$

There is a band of allowed energies. But physically, what is the parameter θ ? The position space wave function is $\langle x|\theta\rangle$, so let us consider $\langle x|\hat{\tau}|\theta\rangle$. Letting $\hat{\tau}$ act to the left gives

$$\langle x|\hat{\tau}|\theta\rangle = \langle x - a|\theta\rangle, \quad (6.126)$$

while letting $\hat{\tau}$ act to the right gives

$$\langle x|\hat{\tau}|\theta\rangle = \langle x|\theta\rangle e^{-i\theta}. \quad (6.127)$$

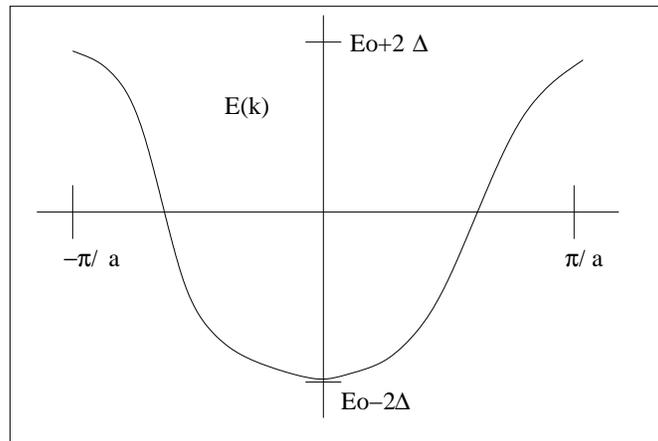
The quantities in equations 6.126 and 6.127 must be equal. To solve this constraint, define

$$\langle x|\theta\rangle = e^{ikx} u_k(x) \quad \theta = ka \quad (6.128)$$

where $u_k(x)$ is a periodic function with period a , i.e., $u_k(x + a) = u_k(x)$. Note that

$$\langle x - a|\theta\rangle = e^{ik(x-a)} u_k(x - a) = e^{ik(x-a)} u_k(x) = u_k(x) e^{ikx} e^{-ika}, \quad (6.129)$$

which is Eq. 6.127. These solutions are called Bloch functions, and k is a wave number. Note the form of the solution – the wave function “in the well” is multiplied by a plane wave. The “in the well” wave function picks up a phase as we move from well to well. Putting all the pieces together, we see that $-\pi \leq \theta \leq \pi$ is equivalent to $-\pi/a \leq k \leq \pi/a$ and the energy is $E(k) = E_0 - 2\Delta \cos ka$. This behavior is sketched in the figure:



For an infinite chain, k is continuous. However, for a finite chain of length N with periodic boundary conditions, N applications of τ returns us to our starting point,

$$\langle x|\hat{\tau}^N|\theta\rangle = \langle x|\theta\rangle = \langle x - Na|\theta\rangle = \langle x|\theta\rangle e^{iN\theta} \quad (6.130)$$

so that

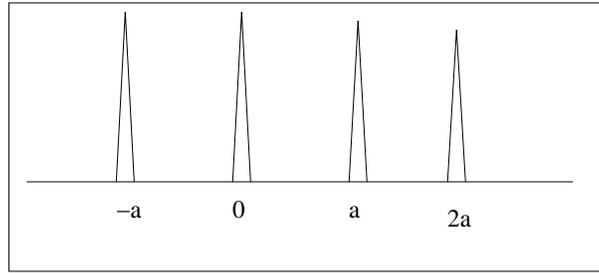
$$\theta = \frac{2\pi}{N}j. \quad (6.131)$$

where j is an integer in the range $-N/2 < j \leq N/2$. This gives a set of discrete levels, or a set of discrete momenta, as many momenta as there are sites.

A more complicated example – a periodic array of delta functions

Now we take a potential function which is a periodic set of delta functions

$$V(x) = \sum_{n=-\infty}^{\infty} V_0 \delta(x - na). \quad (6.132)$$



We again label the Bloch state as $\Psi(x) = e^{ikx}u_k(x)$, with periodicity condition $u_k(x + a) = u_k(x)$.

Between the delta functions, the solution to the Schrödinger equation is

$$\Psi(x) = Ae^{iqx} + Be^{-iqx} \quad (6.133)$$

where $\frac{\hbar^2 q^2}{2m} = E$. The Bloch function, then, is

$$u_k(x) = Ae^{i(q-k)x} + Be^{-i(q+k)x}. \quad (6.134)$$

A and B may be determined from boundary conditions. The first boundary condition is the periodicity of $u_k(x)$:

$$u_k(0) = u_k(a) \quad (6.135)$$

which demands

$$A + B = Ae^{i(q-k)a} + Be^{-i(q+k)a}. \quad (6.136)$$

The second boundary condition arises from the behavior of the derivative of the wave function at the boundary. To find it, we integrate the Schrödinger equation across the boundary,

$$\int_{-\epsilon}^{\epsilon} dx \left(E + \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right) \Psi(x) = \int_{-\epsilon}^{\epsilon} dx V_0 \delta(x) \Psi(x). \quad (6.137)$$

which simplifies to

$$\left. \frac{\partial \Psi}{\partial x} \right|_{-\epsilon}^{\epsilon} = \frac{2m}{\hbar^2} V_0 \Psi(0) = \frac{2m}{\hbar^2} V_0 (A + B). \quad (6.138)$$

We know that just inside the unit cell

$$\left(\frac{\partial \Psi}{\partial x} \right)_{\epsilon} = iq(A - B), \quad (6.139)$$

but we must be careful when evaluating the first derivative at the other side of the potential, at $x = -\epsilon$. This location is actually at the edge of the neighboring cell, or the far edge of our cell:

$$\Psi(-\epsilon) = e^{-ika} \Psi(a - \epsilon) \quad (6.140)$$

and so

$$\left(\frac{\partial \Psi}{\partial x} \right)_{-\epsilon} = \left(\frac{\partial \Psi}{\partial x} \right)_{a-\epsilon} = iq(Ae^{iqa} - Be^{-iqa})e^{-ika}. \quad (6.141)$$

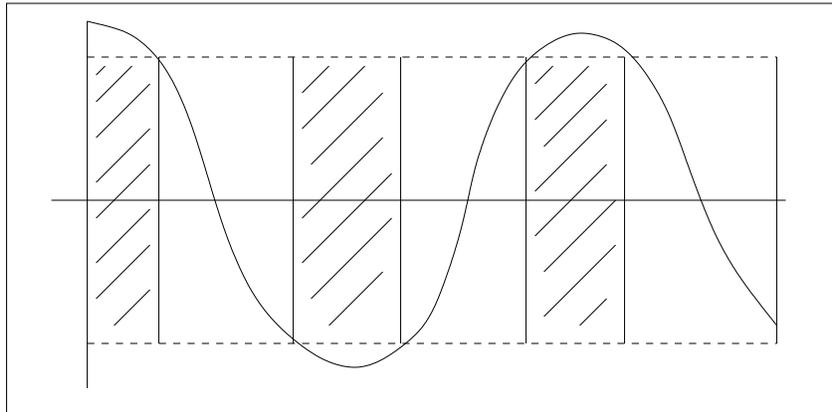
The second boundary condition is therefore

$$\frac{2mV_0}{\hbar^2} (A + B) = iq(A - B - Ae^{i(q-k)a} + Be^{-i(q+k)a}). \quad (6.142)$$

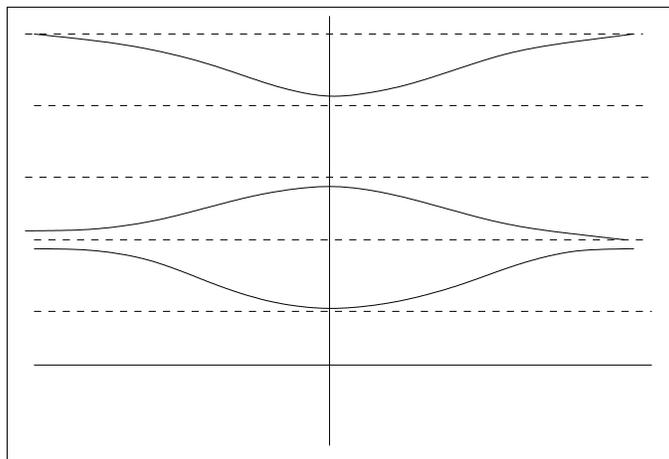
Let us restrict the discussion to $V_0 > 0$. Solving for A and B yields the following relation on the momenta k and q ,

$$\cos(ka) = \cos(qa) + \frac{mV_0}{\hbar^2} \frac{\sin(qa)}{qa} \quad (6.143)$$

The right hand side of this equation is shown as a function of qa in the figure:



The value of the left hand side of the equation is constrained by the requirement $-1 \leq \cos(ka) \leq 1$. This means that only particular values of q , and hence of the energy $E = \hbar^2 q^2 / (2m)$, are allowed. These regions are shaded in the figure. Within the allowed region there are a continuum of energy states – a “band” of energies. We can see the behavior from the next figure, showing allowed values of $E(k)$ vs. k , where again $-\pi/a \leq k \leq \pi/a$.



The bands are separated by “band gaps,” values of the energy for which no solutions of the Schrödinger equation exist.

We are on the edge of being able to explain the difference between insulators and metals. Crystalline solids are obviously examples of systems with periodic potentials. Electrons obey the Pauli principle and hence there can be only one electron (per spin) per energy level. If we assume that electrons in a solid do not interact (never mind that they have charge!), then the electronic ground state of the solid has one electron (per spin) in every single particle state, beginning with the lowest. If we run out of electrons while in the middle of a band, then the ground state is separated from the lowest excited state by only an infinitesimal amount of energy. It is easy to excite an electron from a filled state into an empty one. This is the situation for a conductor. However, if the supply of electrons is finished just at a band filling, then exciting an electron requires supplying a finite energy (the distance to the bottom of the next band) to the system. Then the system is an insulator.

Schrödinger equation in three dimensions

We turn now to three dimensional problems with central potentials: $V(\vec{r}) = V(|r|)$. The most useful coordinate choice for these systems is spherical coordinates. We recall the definition

$$\begin{aligned} z &= r \cos \theta \\ x &= r \sin \theta \cos \phi \\ y &= r \sin \theta \sin \phi \end{aligned}$$

with $-1 < \cos \theta < 1$, $0 < \phi < 2\pi$ and $0 < r < \infty$. The Laplacian can be split into a radial and an angular piece

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} - \frac{\vec{L}^2}{\hbar^2 r^2}, \quad (6.144)$$

where \vec{L}^2 is the squared angular momentum operator,

$$\vec{L}^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \quad (6.145)$$

While one might imagine that this is just a convenient shorthand labeling the angular part of the Laplacian, it is none the less true that when we define the angular momentum operator as

$$\vec{L} = \vec{r} \times \vec{p} = \frac{\hbar}{i} (\vec{r} \times \vec{\nabla}) \quad (6.146)$$

we discover that its components, expressed in spherical coordinates, are

$$\begin{aligned} L_z &= \frac{\hbar}{i} \frac{\partial}{\partial \phi} \\ L_x &= i\hbar \left[\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right] \\ L_y &= i\hbar \left[-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right] \end{aligned} \quad (6.147)$$

and its square is given by Eq. 6.145.

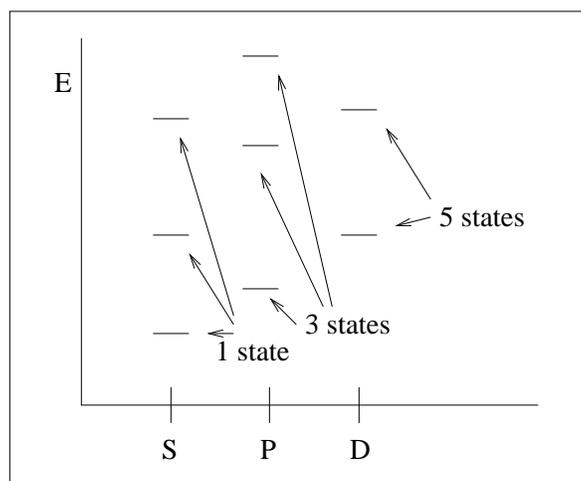
The wave function can be written in separated form, $\psi(\vec{r}) = R(r)Y_l^m(\theta, \phi)$. where the functions $Y_l^m(\theta, \phi)$, called the spherical harmonics, obey the eigenvalue equations

$$\begin{aligned} \vec{L}^2 Y_l^m(\theta, \phi) &= \hbar^2 l(l+1) Y_l^m(\theta, \phi). \\ L_z Y_l^m(\theta, \phi) &= \hbar m Y_l^m(\theta, \phi). \end{aligned} \quad (6.148)$$

That is, they are eigenfunctions of L^2 and L_z with eigenvalues $\hbar^2 l(l+1)$ and $\hbar m$, respectively. We state without proof that the l 's are restricted to be integers, $l = 1, 2, \dots$, and $-l \leq m \leq l$. Then we are left with an ordinary differential equation for the radial dependence of the wave function.

$$\left[-\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + V(r) + \frac{\hbar^2 l(l+1)}{2mr^2} \right] R(r) = ER(r). \quad (6.149)$$

Notice that the angular dependence of a wave function for a central potential problem is independent of the particular form of the potential; this part of the wave function is universal. Notice also that the radial equation depends on $V(r)$ and on l . Solutions of the radial equation will in general involve a distinct set of energy levels for each value of l . Thus, all wave functions for central potentials in three dimensions are labeled by three quantum numbers (not including spin), one for E , l , and m . The radial solution (and the energy eigenvalues) depend on l but not on m . Because there is no dependence on m there will be a $2l + 1$ degeneracy among the energy levels described by E and l .



Review of spherical harmonics

Angular momentum is so important that it deserves (and will get) its own chapter. For the time being, we remark that commutation relations for the L_i 's are simple: $[L_i, L_j] = i\hbar\epsilon_{ijk}L_k$ and $[\vec{L}^2, L_i] = 0$. For central potentials, $[H, \vec{L}^2] = 0$, and $[H, L_i] = 0$, which, together with the commutation relations for the L_i 's, implies that one can find simultaneous eigenstates of H , L^2 and one component of \vec{L} . We label such a state as $|\psi\rangle = |E, \vec{L}^2, L_i\rangle$. It is conventional to pick this to be eigenstates of L_z . In Dirac notation, $Y_l^m(\theta, \phi) \equiv \langle \theta, \phi | \vec{L}^2, L_z \rangle$. All of

these statements (and much more) can be derived using pure operator methods. Before we do that, however, we pause to fill in some details about the spherical harmonics as special functions.

The explicit form of Eq. 6.148 is

$$0 = \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] Y_l^m(\theta, \phi) + l(l+1) Y_l^m(\theta, \phi). \quad (6.150)$$

This partial differential equation can be completely separated by writing $Y_l^m(\theta, \phi) = P(\theta)Q(\phi)$. The resulting two equations are

$$\begin{aligned} \left[\frac{\partial^2}{\partial \phi^2} + m^2 \right] Q(\phi) &= 0 \\ \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \left(l(l+1) - \frac{m^2}{\sin^2 \theta} \right) \right] P(\theta) &= 0. \end{aligned} \quad (6.151)$$

We immediately find solutions of the form $Q(\phi) = e^{im\phi}$, and by demanding $Q(0) = Q(2\pi)$, we impose the quantization condition $m = \{0, \pm 1, \pm 2, \dots\}$. This is quantization of the z -component of angular momentum,

$$L_z Y_l^m(\theta, \phi) = \hbar m Y_l^m(\theta, \phi). \quad (6.152)$$

The equation for θ is called the associated Legendre equation in the mathematics literature. When $m = 0$ its solutions are the usual Legendre polynomials. Away from $m = 0$, it has well behaved solutions only for $l = \{0, 1, 2, \dots\}$ and $|m| \leq l$. The first condition amounts to quantization of the squared angular momentum, equal to $\hbar^2 l(l+1)$ where l is an integer. (As a shorthand, one often says, “the total angular momentum takes on integer values.”) The m constraint makes physical sense because one would expect the magnitude of the z -component of angular momentum to be less the magnitude of the total angular momentum, $\sqrt{l(l+1)}$. Also notice that since Eq. 6.151 only depends on m^2 , there is freedom in the definition of the $Y_l^m(\theta, \phi)$'s. The standard convention is to define $Y_l^{-m}(\theta, \phi) \equiv (-1)^m Y_l^m(\theta, \phi)^*$.

We state a few of the relevant properties of spherical harmonics in Dirac notation: Orthonormality:

$$\langle lm | l' m' \rangle = \delta_{ll'} \delta_{mm'}; \quad (6.153)$$

orthonormality in coordinate space:

$$\int d\Omega \langle lm | \theta, \phi \rangle \langle \theta, \phi | l' m' \rangle = \int d\Omega Y_l^m(\theta, \phi) Y_{l'}^{m'}(\theta, \phi)^* = \delta_{ll'} \delta_{mm'}; \quad (6.154)$$

Completeness:

$$\hat{\mathbf{1}} = \sum_{lm} |lm\rangle \langle lm| = \int d\Omega |\theta, \phi\rangle \langle \theta, \phi|; \quad (6.155)$$

and completeness in coordinate space

$$\delta(\cos \theta - \cos \theta') \delta(\phi - \phi') = \sum_{lm} \langle \theta', \phi' | lm \rangle \langle lm | \theta, \phi \rangle = \sum_{l=0}^{\infty} \sum_{m=-l}^l Y_l^m(\theta', \phi') Y_l^m(\theta, \phi)^*. \quad (6.156)$$

Finally, we tabulate the first few spherical harmonics.

$$\begin{aligned} Y_0^0 &= \frac{1}{\sqrt{4\pi}} \\ Y_1^0 &= \sqrt{\frac{3}{4\pi}} \cos \theta \\ Y_1^1 &= -\sqrt{\frac{3}{8\pi}} \sin \theta e^{i\phi} \\ Y_2^2 &= \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{2i\phi} \\ Y_2^1 &= -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{i\phi} \\ Y_2^0 &= \sqrt{\frac{5}{4\pi}} \left(\frac{3 \cos^2 \theta - 1}{2} \right). \end{aligned} \quad (6.157)$$

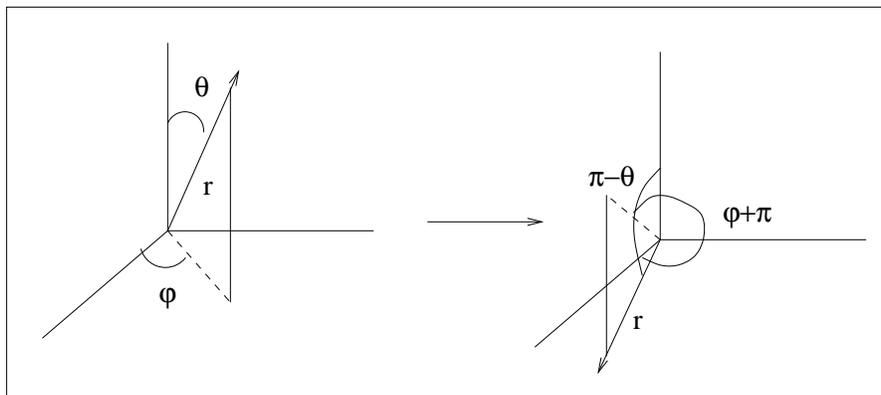
The first three l spherical harmonics are important enough to commit to memory. For historical reasons the different l levels have been given names.

l	level	m
0	s	0
1	p	$0, \pm 1$
2	d	$0, \pm 1, \pm 2$
3	f	\dots
\vdots	\vdots	\vdots

s , p , d , and f stand for “sharp,” “principal,” “diffuse,” and “fine,” respectively. This is called “*spectroscopic notation*” and continues down the alphabet for subsequent l ’s (g , h , \dots), without names, as far as I know.

Parity in three dimensions

Parity comes into its own in three dimensions. Let us first investigate the parity properties of the spherical harmonics. The transformation $\vec{r} \rightarrow -\vec{r}$ in spherical coordinates is shown in the figure:



Specifically, it is

$$r \rightarrow r \quad \theta \rightarrow \pi - \theta \quad \phi \rightarrow \phi + \pi. \quad (6.158)$$

One can also express the parity operation on θ as $\cos \theta \rightarrow -\cos \theta$.

To analyze the parity operator's effect on the spherical harmonics, we recall their explicit form

$$Y_l^m(\theta, \phi) = P_l^m(\theta) e^{im\phi} \quad (6.159)$$

where the P_l^m 's are the associated Legendre polynomials. Consider first $l = 0$, where these functions are the usual Legendre polynomials. They are l -th order polynomials in $\cos \theta$ containing only even or odd powers of $\cos \theta$. Therefore

$$\hat{P}P_l(\cos \theta) = (-1)^l P_l(\cos \theta). \quad (6.160)$$

Thus, the Legendre polynomials are eigenfunctions of the parity operator with eigenvalues ± 1 . The associated Legendre polynomials transform under parity as

$$\hat{P}P_l^m(\cos \theta) = (-1)^{l-|m|} P_l^m(\cos \theta) \quad (6.161)$$

This behavior comes from the fact that the associated Legendre polynomials are products of a term $\propto (1 - \cos^2 \theta)^{\frac{|m|}{2}}$ with an even or odd polynomial in $\cos \theta$ of highest order $l - |m|$. The first term in the product is also even under a parity transformation; thus the eigenvalue

comes from the order of the polynomial. The azimuthal part of the spherical harmonic transforms under parity as

$$e^{im\phi} \rightarrow e^{im(\phi+\pi)} = (-1)^m e^{im\phi}. \quad (6.162)$$

Thus the behavior of the spherical harmonics under a parity transformation is

$$Y_l^m(\theta - \pi, \phi + \pi) = (-1)^{l-|m|} (-1)^m Y_l^m(\theta, \phi) = (-1)^l Y_l^m(\theta, \phi). \quad (6.163)$$

As m is an integer, $(-1)^{l-|m|}(-1)^m = (-1)^l$. Thus the spherical harmonics are even or odd under parity depending on whether l is even or odd.

The spherical square well

Perhaps it is time for an example. The spherical square well is defined by

$$V(r) = -V_0 \quad r < a \quad (6.164)$$

$$V(r) = 0 \quad r > a \quad (6.165)$$

The radial equation obeys the differential equation

$$\left\{ \frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} - \frac{l(l+1)}{r^2} - \frac{2m(V-E)}{\hbar^2} \right\} f(r) = 0. \quad (6.166)$$

For later use, let us define $k^2 \equiv -\frac{2m(V_0-E)}{\hbar^2}$. Eq. 6.166 is a form of the Bessel's equation and its solutions are the spherical Bessel functions. Expressed in terms of the cylindrical Bessel functions, they are

$$j_l(x) = \sqrt{\frac{\pi}{2x}} J_{l+1/2}(x) \quad (6.167)$$

$$n_l(x) = (-1)^{l+1} \sqrt{\frac{\pi}{2x}} J_{-l-1/2}(x). \quad (6.168)$$

(The $n(x)$'s are sometimes called the spherical Neumann functions.) The first few are

$$j_0(x) = \frac{\sin(x)}{x}, \quad j_1(x) = \frac{\sin(x)}{x^2} - \frac{\cos(x)}{x} \quad (6.169)$$

$$n_0(x) = \frac{\cos(x)}{x}, \quad n_1(x) = \frac{-\cos(x)}{x^2} - \frac{\sin(x)}{x}. \quad (6.170)$$

The limiting forms of these functions are quite useful. As $x \rightarrow 0$

$$j_l \simeq \frac{x^l}{1 \cdot 3 \cdot 5 \cdots (2l+1)} = \frac{x^l}{(2l+1)!!} \quad n_l \simeq \frac{1 \cdot 3 \cdot 5 \cdots (2l+1)}{x^{l+1}} = \frac{(2l+1)!!}{x^{l+1}} \quad (6.171)$$

and as $x \rightarrow \infty$ we have

$$j_l \simeq \frac{1}{x} \sin\left(x - \frac{l\pi}{2}\right) \quad n_l \simeq \frac{1}{x} \cos\left(x - \frac{l\pi}{2}\right). \quad (6.172)$$

The requirement that the wave function be nonsingular at the origin eliminates the Neumann functions as solutions of the square well. Other boundary conditions require the use of linear combinations of these two sets of functions. One such useful linear combination is called the Hankel function.

$$h_l(x) = j_l(x) + in_l(x) \simeq \frac{1}{x} \exp\left(i\left(x - \frac{(l+1)\pi}{2}\right)\right). \quad (6.173)$$

It represents a wave moving radially outward, and its complex conjugate represents an incoming wave. We will need these functions when we consider scattering solutions in three dimensions.

Everything we have done so far assumes that $E > V_0$, and consequently that $k^2 > 0$. In the classically forbidden region we define

$$\hbar K = \sqrt{2m(V_0 - E)} \quad (6.174)$$

which amounts to replacing kr with iKr as the independent variable in the spherical Bessel and Neumann functions. In the classically forbidden region the Hankel functions become

$$h_0(iKr) = \frac{1}{Kr} e^{-Kr} \quad h_1(iKr) = i\left(\frac{1}{Kr} + \frac{1}{K^2 r^2}\right) e^{-Kr}. \quad (6.175)$$

Infinitely deep spherical well

In this case it is better to shift the origin of potential and write

$$V = 0 \quad r < a \quad (6.176)$$

$$V = \infty \quad r > a. \quad (6.177)$$

The solutions of the Schrodinger equation are

$$\psi_{E,l,m} = j_l(kr) Y_l^m(\Omega). \quad (6.178)$$

The vanishing of the wave function at $r = a$ gives the quantization condition

$$j_l(ka) = 0. \quad (6.179)$$

If $l = 0$ this is merely $ka = n\pi$.

General results for arbitrary potentials

A useful change of variables converts the three dimensional radial equation into a one dimensional Schrödinger equation.. To do this, we define $f(r) \equiv u(r)/r$. Then

$$\frac{d}{dr} r^2 \frac{d}{dr} \frac{u}{r} = \frac{d}{dr} \left(r^2 \left(\frac{u'}{r} - \frac{u}{r^2} \right) \right) = \frac{d}{dr} (ru' - u) \quad (6.180)$$

$$= u' + ru'' - u' = ru''. \quad (6.181)$$

Inserting this into the radial equation, we have

$$\frac{-\hbar^2}{2m} \frac{u''}{r} + \left[V(r) + \frac{l(l+1)\hbar^2}{2mr^2} \right] \frac{u}{r} = E \frac{u}{r} \quad (6.182)$$

or

$$\frac{-\hbar^2}{2m} u'' + \left[V(r) + \frac{l(l+1)\hbar^2}{2mr^2} \right] u = Eu. \quad (6.183)$$

This is just the usual one-dimensional Schrodinger equation. The only differences with a true one-dimensional problem are that the range of r is $0 < r < \infty$, and that the boundary condition at the origin is $u(r=0) = 0$ so that $f(r)$ remains a finite function. The normalization condition also becomes very one-dimensional,

$$\int_0^\infty dr |u(r)|^2 = 1. \quad (6.184)$$

Defining

$$V_{eff} = \left[V(r) + \frac{l(l+1)\hbar^2}{2mr^2} \right] \quad (6.185)$$

we can rewrite the radial equation as

$$\frac{-\hbar^2}{2m} u'' + V_{eff} u = Eu. \quad (6.186)$$

The potential that truly matters to the radial degrees of freedom is $V_{eff}(r)$, not simply $V(r)$. The extra term is the centrifugal barrier, or the energy cost associated with having nonzero angular momentum.

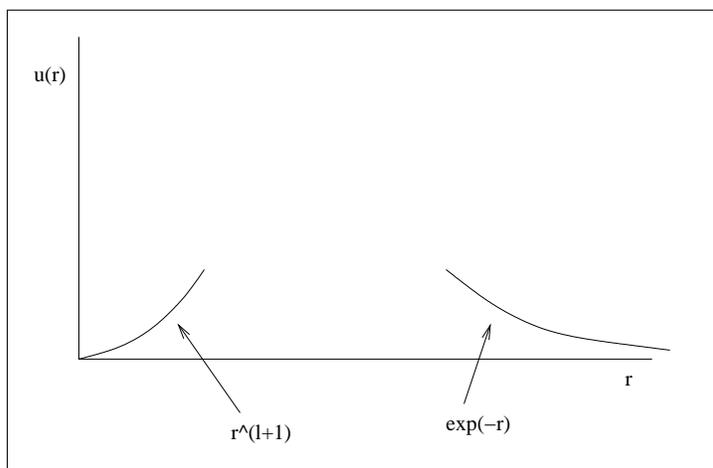
It is useful to consider some limiting cases. Suppose that we take r to be small and also assume that $V(r)$ is less singular than $1/r^2$. Then the centrifugal term $\hbar^2 l(l+1)/(2mr^2)$ dominates $V_{eff}(r)$ and we must solve

$$\frac{d^2 u}{dr^2} + \frac{l(l+1)}{r^2} u = 0. \quad (6.187)$$

The solution of this differential equation is $u \propto r^{l+1}$. Physically,, the centrifugal barrier forces the wave function away from the origin. Next suppose that $V(r) \rightarrow 0$ as $r \rightarrow \infty$. Then, in that limit,

$$\frac{d^2u}{dr^2} = \frac{-2mE}{\hbar^2}u(r) \quad (6.188)$$

At large distances, if $E < 0$, as we would expect for a bound state, we find $u \propto \exp\left(-\sqrt{\frac{2m|E|}{\hbar^2}}r\right)$. Of course, E must be known before this is useful. The two limiting forms are sketched in the figure. No simple argument can tell us anything about $u(r)$ at intermediate distances.



The two-body central potential

Two particles interacting by a central potential have a Hamiltonian

$$\hat{H} = \frac{\hat{p}_1^2}{2m_1} + \frac{\hat{p}_2^2}{2m_2} + V(|\vec{r}_2 - \vec{r}_1|). \quad (6.189)$$

As the Hamiltonian is invariant under an overall translation, we expect that the momentum of the entire system will be conserved. We see that with a change of variables into relative (\vec{r}) and center of mass (\vec{R}) coordinates:

$$\psi(\vec{r}_1, \vec{r}_2) \rightarrow \psi(\vec{r}, \vec{R}) \quad (6.190)$$

$$\vec{R} \equiv \frac{m_1\vec{r}_1 + m_2\vec{r}_2}{m_1 + m_2} \quad \vec{r} \equiv \vec{r}_1 - \vec{r}_2. \quad (6.191)$$

Below, we will label

$$\vec{R} = (X, Y, Z) \quad \vec{r} = (x, y, z). \quad (6.192)$$

Let us sketch the simple steps in the change of variables, looking only at one coordinate.

$$\frac{\hat{p}_1^2}{2m_1} + \frac{\hat{p}_2^2}{2m_2} = \frac{-\hbar^2}{2m_1} \frac{\partial^2}{\partial x_1^2} + \frac{-\hbar^2}{2m_2} \frac{\partial^2}{\partial x_2^2}. \quad (6.193)$$

From the chain rule, and using $M = m_1 + m_2$, we find

$$\frac{\partial \psi(x, X)}{\partial x_1} = \frac{\partial x}{\partial x_1} \frac{\partial \psi}{\partial x} + \frac{\partial X}{\partial x_1} \frac{\partial \psi}{\partial X} = \frac{\partial \psi}{\partial x} + \frac{m_1}{M} \frac{\partial \psi}{\partial X}. \quad (6.194)$$

The second derivatives are

$$\frac{\partial^2 \psi}{\partial x_1^2} = \frac{\partial^2 \psi}{\partial x^2} + \frac{2m_1}{M} \frac{\partial^2 \psi}{\partial x \partial X} + \frac{m_1^2}{M^2} \frac{\partial^2 \psi}{\partial X^2} \quad (6.195)$$

and

$$\frac{\partial^2 \psi}{\partial x_2^2} = \frac{\partial^2 \psi}{\partial x^2} - \frac{2m_2}{M} \frac{\partial^2 \psi}{\partial x \partial X} + \frac{m_2^2}{M^2} \frac{\partial^2 \psi}{\partial X^2} \quad (6.196)$$

so that the kinetic term is

$$-\hbar^2 \left\{ \frac{1}{2} \left(\frac{1}{m_1} + \frac{1}{m_2} \right) \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial x \partial X} \left(\frac{2}{2M} - \frac{2}{2M} \right) + \frac{\partial^2 \psi}{\partial X^2} \left(\frac{m_1 + m_2}{2M^2} \right) \right\}. \quad (6.197)$$

Thus the the Schrodinger equation becomes

$$\left[\frac{-\hbar^2}{2\mu} \nabla_r^2 + \frac{-\hbar^2}{2M} \nabla_R^2 + V(|\vec{r}|) \right] \psi = E\psi. \quad (6.198)$$

The quantity μ is the ‘reduced mass,’ $1/\mu = 1/m_1 + 1/m_2$. It is no surprise that we can assume a separated solution

$$\psi(\vec{R}, \vec{r}) = \Phi(\vec{R})\phi(\vec{r}) \quad (6.199)$$

where the energy

$$E = E_1 + E_2 \quad (6.200)$$

consists of a relative-coordinate piece

$$E_1 \phi(\vec{r}) = \left(\frac{-\hbar^2}{2\mu} \nabla_r^2 + V(r) \right) \phi(\vec{r}) \quad (6.201)$$

and a center of mass piece

$$E_2 \Phi(\vec{R}) = \frac{-\hbar^2}{2M} \nabla_R^2 \Phi(\vec{R}). \quad (6.202)$$

The center of mass motion is obviously that of a free particle, $\Phi(\vec{R}) = e^{i\vec{k}\cdot\vec{R}}$ with $E_2 = \hbar^2 k^2 / (2M)$, and the relative piece is our one-body central potential problem again.

Instead of hydrogen

The author lacks the strength to present the full solution of the radial equation for hydrogen, a problem which appears in most textbooks. Let us merely assume that the student has worked through it previously, and summarize relevant information in (I admit) insufficient detail to be of much use.

The Schrödinger equations for the hydrogen atom is

$$\frac{-\hbar^2}{2\mu}u'' + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{Ze^2}{r} \right] u = Eu. \quad (6.203)$$

To remove numerical factors we define rescaled energy and radial variables

$$E = -\frac{1}{2}\mu c^2 Z^2 \alpha^2 \epsilon \quad r = \frac{\hbar c}{\mu c^2 \alpha Z} \rho = (a_0/Z)\rho. \quad (6.204)$$

(a_0 is the familiar Bohr radius.) With these substitutions the Schrödinger equation reduces to

$$-\frac{1}{2} \frac{d^2 u}{d\rho^2} + \frac{l(l+1)}{2\rho^2} u - \frac{u}{\rho} = \frac{-\epsilon\mu}{2}. \quad (6.205)$$

At small ρ we again see that $u \propto \rho^{l+1}$ and at large ρ , $u \propto e^{-\sqrt{\epsilon}\rho}$. Omitting a solution to this equation, I will simply remark that the quantization condition is $\epsilon = \frac{1}{n^2}$. The radial solutions are

$$u_{n,l}(\rho) = N_{n,l} e^{-\frac{\rho}{n}} P_{n-l-1}(\rho) \quad (6.206)$$

where the $P_{n-l-1}(\rho)$'s are polynomials of degree $n-l-1$. Since this degree must be greater than zero, there are n such polynomials for each n ($l = 0$ to $n-1$) and since the energy magically does not depend on l , these states are all degenerate: counting a factor of 2 for spin, hydrogenic levels are $2 \sum_{l=0}^{n-1} (2l+1) = 2n^2$ degenerate. This degeneracy makes hydrogen an atypical atom. Our limited discussion makes it seem rather contrived. Its true origin is the existence of a set of conserved quantities (the Runge-Lenz vector) associated with a $1/r$ potential. The first solution of hydrogen, before Schrödinger's wave solution, was performed by Pauli using these quantities.

Table 6.1: Relation between n and l

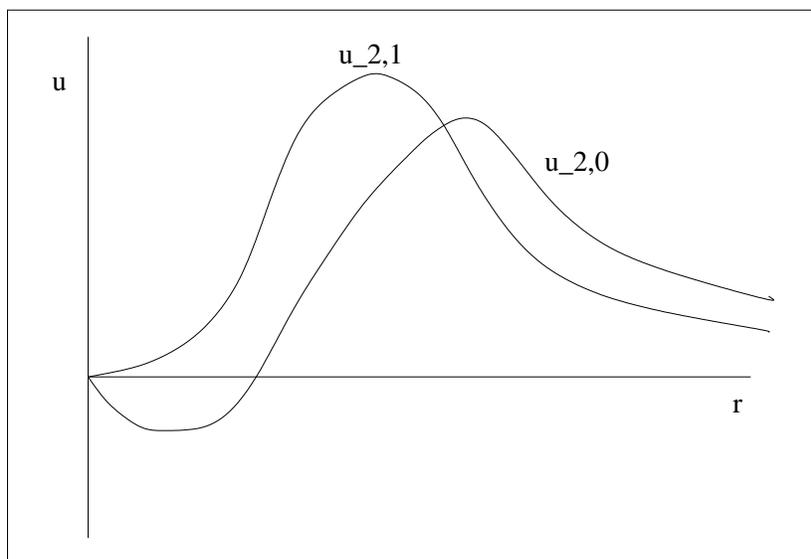
$n = 1$	$l = 0$	$1S$
$n = 2$	$l = 0, 1$	$2S, 2P$
$n = 3$	$l = 0, 1, 2$	$3S, 3P, 3D$

It is useful to have a functional form for at least part of the family of hydrogenic wave functions. The $n = l + 1$ solutions are simple:

$$u(r) = Cr^{l+1} \exp\left(-\frac{Zr}{na_0}\right) \quad (6.207)$$

The reader can readily check (by substitution) that it is a solution.

From the solution we see that the size of the wave function is about a_0/Z . The $1/Z$ simply reflects the increased Coulomb attraction of a multiply charged nucleus. At larger n the wave function pushes out from the nucleus. This is also unsurprising – the electron has more energy and sits higher in the potential well. Because the radial solution is a polynomial of degree $n - l - 1$, the wave function of the (n, l) state will have $n - l$ maxima. The graph illustrates this for $u_{2,1}$ and $u_{2,0}$.



We see that $\langle r \rangle$, the average value of the radius of the electron will be a little bit bigger for smaller l . In fact, one can show that

$$\langle r \rangle = \frac{a_0}{2Z} (3n^2 - l(l+1)). \quad (6.208)$$

A few of these values are shown in the table.

So, to summarize, the energy levels of a hydrogenic bound state

$$E_n = \frac{1}{n^2} \cdot \frac{1}{2} \alpha^2 m_e c^2 \quad (6.209)$$

Table 6.2: Values of $\langle r \rangle$ coefficient

n	l	coeff.
2	0	12
2	1	10
3	0	27
3	1	25
3	2	21

where the fine structure constant, $\alpha \equiv \frac{e^2}{\hbar c}$ is a dimensionless constant equal to about $1/137$. The quantity $\frac{1}{2}\alpha^2 m_e c^2$ is defined as the Rydberg and is equal to 13.6eV. m_e is the mass of an electron, $m_e c^2 = 511$ KeV. (To be more accurate, it is the reduced mass of the electron and the proton.) If the nucleus has charge Z , the energy levels scale as

$$E_n = Z^2 E_n(H). \quad (6.210)$$

The solution of the Schrödinger equation is far from the whole story for hydrogenic atoms. The electron's velocity is not vanishingly small (in fact $\langle \frac{v}{c} \rangle \simeq \frac{1}{137}$), so that while it is a good assumption to assume that the non-relativistic, this is not completely correct. Because the electron has spin, it has a magnetic moment. Because the electron moves in the Coulomb field of the nucleus, it sees (by Lorentz transformation) a magnetic field. Its interaction with the field, plus relativistic corrections to the kinetic energy, give a "fine structure" contribution to the energy.

$$\Delta E \propto Z^4 \alpha^4 m_e c^2 \quad (6.211)$$

which breaks the degeneracy of the different levels of common n . The electron also experiences a direct "hyperfine" interaction with the spin of the nucleus

$$\Delta E \propto Z^3 \alpha^4 m_e c^2 \left(\frac{m_e}{m_p} \right). \quad (6.212)$$

The ground state of hydrogen is split with a frequency difference of about 1420 MHz. Finally, quantum fluctuations in the electromagnetic field induce a splitting between the (otherwise degenerate) $2S$ and one of the $2P$ levels of hydrogen, with a frequency of about 1000 MHz. This is called the Lamb shift.

Because all these effects are small, they are best studied using perturbative techniques. We will return later to all of them except the Lamb shift.

The semi-classical method of Wentzel, Kramers, and Brillouin

The WKB (Wentzel - Kramers - Brillouin) method is a set of tools for solving the Schrödinger equation which are applicable when the action of typical configurations is large compared to \hbar . These methods are relevant for systems of particles which are in potentials whose variation in space is many de Broglie wavelengths in extent. The potential itself could be very broad, or (as in the case of highly excited states) the de Broglie wavelength could simply be small compared to the scale of variation in the potential $V(r)$. These methods actually predate quantum mechanics. They are basically a quantum mechanical transcription of the method of stationary phase in optics. This is an enormous field, including topics such as the passage from wave optics to geometric optics, the transmission of light through regions of spatially varying index of refraction, and gravitational lensing. As a byproduct they also give most of the old Bohr-Sommerfeld quantization rules for bound states.

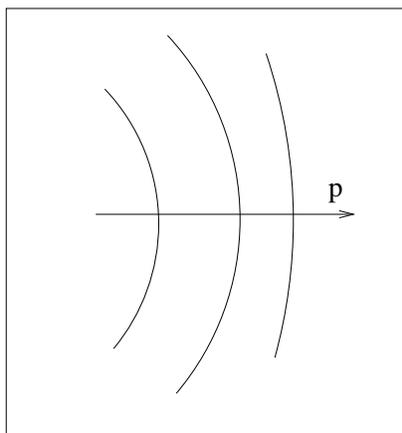
We begin by writing the coordinate space wave function as

$$\psi(\vec{r}, t) = A \exp \frac{iW(\vec{r}, t)}{\hbar}. \quad (6.213)$$

This equation assigns $W(x, t)/\hbar$ as a spatially and temporally varying phase of the wave function. At fixed time, the phase varies along the direction of the wave number, so we immediately identify the (local) momentum and energy as

$$\vec{p}(x, t) = \hbar \vec{k}(x, t) = \vec{\nabla} W; \quad \hbar \omega(x, t) = \frac{\partial W}{\partial t}. \quad (6.214)$$

Particles move on trajectories orthogonal to the surfaces of constant W , or to the surfaces of constant phase in ψ .



Substituting this functional form into the time-dependent Schrödinger equation gives

$$\left[\frac{\partial W}{\partial t} + \frac{1}{2m}(\nabla W)^2 + V(x) - \frac{i\hbar}{2m}\nabla^2 W \right] \psi = 0. \quad (6.215)$$

Note that if we set $\hbar = 0$, or, more precisely, if we assume that the last term in this expression is small compared to the other two terms, Eq. 6.215 simplifies to

$$\frac{\partial W}{\partial t} + H(x, p)W = 0 \quad (6.216)$$

where we have defined $H(x, p) \equiv \frac{1}{2m}(\nabla W)^2 + V(x)$ and $\vec{p} \equiv \vec{\nabla}W$. This happens to be the equation for Hamilton's principal function W in classical mechanics.

If ψ is an energy eigenfunction, we can write it as

$$\psi = u(r)e^{\frac{-iEt}{\hbar}} \quad (6.217)$$

and from Eq. 6.213, $W = S(\vec{r}) - Et$ and

$$u(r) = A \exp\left(\frac{i}{\hbar}S(\vec{r})\right). \quad (6.218)$$

The full equation of motion is

$$\frac{1}{2m}(\vec{\nabla}S)^2 - (E - V(r)) - \frac{i\hbar\nabla^2 S}{2m} = 0. \quad (6.219)$$

This is still exact. We will work with the one-dimensional version of this equation (although there is a large three - dimensional literature of WKB). For future use, let us define $\hbar k(x) \equiv \sqrt{2m(E - V(x))}$ so that the Schrödinger equation reduces to

$$\frac{d^2 u}{dx^2} + k^2(x)u = 0. \quad (6.220)$$

In three dimensions, working in spherical coordinates and keeping only the radial dependence, it is useful to include the centrifugal barrier in $V(r)$ and to set

$$\hbar k(r) = \sqrt{2m\left(E - V(r) - \frac{l(l+1)}{r^2}\right)}. \quad (6.221)$$

Now we are in one dimension: define $S' = dS/dx$ and $S'' = d^2S/dx^2$. We are interested in a solution where the de Broglie wavelength of the particle is short compared to the variation in $k(x, t)$. We can encode this situation in the requirement that $k'/k \ll 1/\lambda$. To get this we (improbably) use \hbar as a proxy, and we expand $S(r)$ in a power series in \hbar ,

$$S = S_0 + \hbar S_1 + \hbar^2 S_2 + \dots \quad (6.222)$$

Noting that $\hbar^2 k^2$ is actually $O(\hbar^0)$ (it is equal to $E - V$), the zeroth order expression is

$$(S'_0)^2 = \hbar^2 k^2 = 2m(E - V) \quad (6.223)$$

whose solution is

$$S_0 = \pm \hbar \int^x k(x') dx'. \quad (6.224)$$

The order \hbar^1 equation is

$$iS''_0 - 2S'_0 S'_1 = 0. \quad (6.225)$$

Substituting for S_0 gives

$$i\hbar k'(x) - 2\hbar k(x)S'_1 = 0 \quad (6.226)$$

or

$$S'_1 = \frac{ik'(x)}{2k(x)} \quad (6.227)$$

which integrates to

$$S_1(x) = \frac{i}{2} \ln(k(x)). \quad (6.228)$$

Thus to order \hbar^1 the complete solution is

$$u(x) = \frac{A}{\sqrt{2m(E - V(x))}} \exp\left[\pm \frac{i}{\hbar} \int^x dx' \sqrt{2m(E - V(x'))}\right]. \quad (6.229)$$

This is for $E > V$. For $E < V$ we have, defining $\hbar\kappa = \sqrt{2m(V(x) - E)}$,

$$u(x) = \frac{B}{\sqrt{\kappa(x)}} \exp\left[\pm \int^x dx' \kappa(x')\right], \quad (6.230)$$

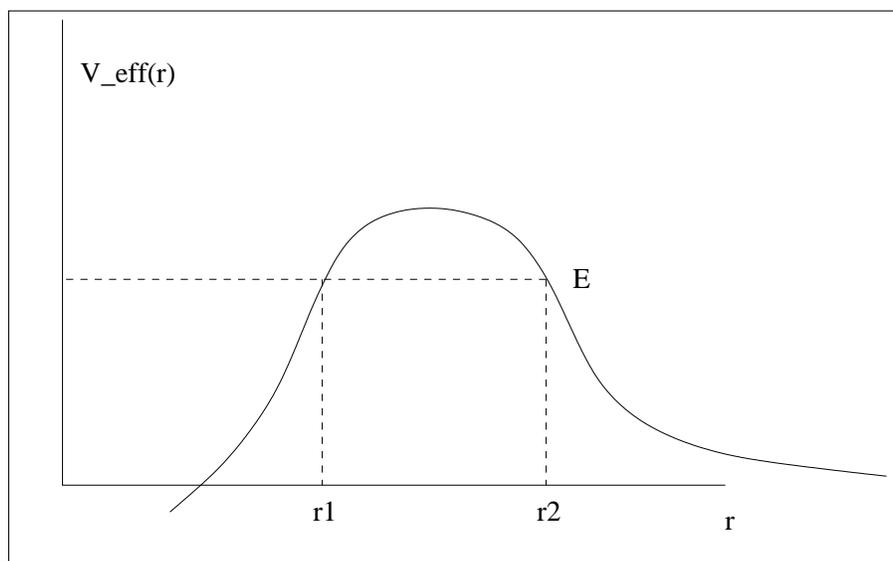
and we choose the sign from the appropriate boundary condition.

For an iterative solution to make sense, we must require $\hbar S'_1/S'_0 \ll 1$, or $|k'/(2k^2)| \ll 1$, because $S'_0 = k$ and $S'_1 = k'/k$. The de Broglie wavelength is $\lambda = 2\pi/k$ and so this condition is equivalent to the statement that $(\lambda/(4\pi))|dk/dx| \ll k$. Physically, this means that the fractional change in the wavelength in a distance $\lambda/(4\pi)$ should be small compared to k . The WKB approximation is a good one when the potential energy changes slowly over many wavelengths of the particle.

Before we continue the story, let's pause to discuss a particularly important application – barrier penetration.

Barrier Penetration and WKB

Suppose there is a potential barrier $V(r)$ separating two classically allowed regions. $V(r)$ could include the centrifugal term if we are considering a radial coordinate in a three dimensional problem.



Suppose also that for $r_1 < r < r_2$, $E < V(r)$. We imagine that we begin with a particle confined to the region $0 < r < r_1$. In the classically forbidden region the tail of the wave function behaves as

$$u(r) \propto \exp \left[- \int_{r_1}^r \kappa(r) dr \right], \quad (6.231)$$

where as usual $\hbar^2 \kappa^2(r) = 2m(V(r) - E)$. The ratio of the value of the wave function at location r_2 , where it leaves the forbidden region, to its value at r_1 , where it enters, is an interesting quantity. It measures how $u(r)$ falls off across the barrier. The probability ratio, its square, is called the penetration factor P .

$$P = \frac{|u(r_2)|^2}{|u(r_1)|^2} \propto \exp \left[- 2 \int_{r_1}^{r_2} \kappa(r) dr \right] \quad (6.232)$$

where explicitly

$$\kappa(r) = \frac{1}{\hbar} \sqrt{2m \left(V(r) + \frac{l(l+1)}{2mr^2} - E \right)}. \quad (6.233)$$

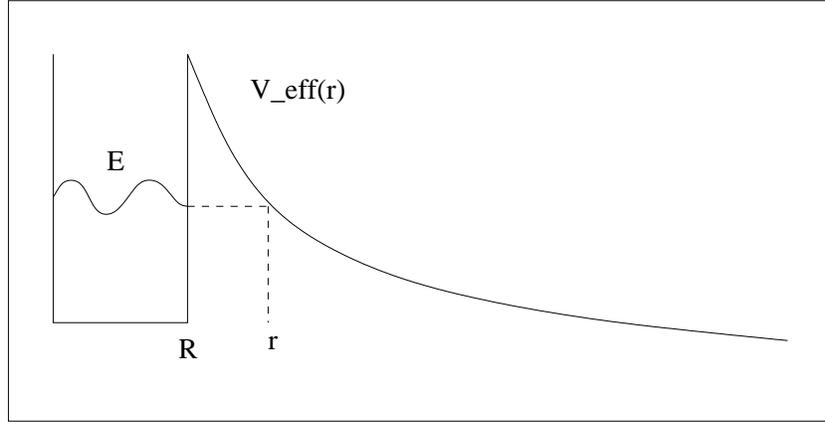


Figure 6.1: The potential felt by an alpha particle inside and outside a nucleus.

Tunneling through a square barrier is a special case of this formula. Recall that we found that the transmission coefficient for a barrier of height V_0 and width a was

$$T = \left[1 + \frac{V_0^2 \sinh^2(\beta a)}{4E(V_0 - E)} \right]^{-1} \quad (6.234)$$

where $\hbar^2 \beta^2 = 2m(V_0 - E)$. For large βa , T is approximately

$$T \simeq \left(\frac{16E(V_0 - E)}{V_0^2} \right) e^{-2\beta a} = \left(\right) \exp\left(-2 \int_0^a dx \sqrt{\frac{2m}{\hbar^2}(V_0 - E)}\right). \quad (6.235)$$

WKB gives the dominant exponential factor in this expression.

The classic example of barrier penetration is the alpha-decay of a nucleus. Inside the nucleus, protons and neutrons feel the strong attractive nuclear force. This force is short range, so outside the nucleus, nucleons only sense the Coulomb repulsion $ZZ'e^2/r$ of the nucleus. (The nucleon has charge $Z'e$ and the nucleus has charge Ze .) In alpha-decay, a nucleus emits an alpha particle – a bound state of two protons with $Z' = 2$. We follow Gamow, Condon and Gurney and assume that the alpha particle exists as a separate entity inside the nucleus, where it feels an attractive potential: the decay occurs as the alpha particle tunnels through the classically forbidden region to emerge as a free particle. The potential it feels is shown in Fig. 6.1. WKB gives the ($l = 0$) penetration factor as

$$-\frac{1}{2} \log P = \int_R^r \kappa(r) dr = \frac{\sqrt{2m}}{\hbar} \int_R^{r=\frac{ZZ'e^2}{E}} \left(\frac{ZZ'e^2}{r} - E \right)^{\frac{1}{2}} dr. \quad (6.236)$$

Taking $E = \frac{1}{2}mv^2$, $ZZ'e^2/R = \frac{1}{2}mv^2$, we evaluate the integral, finding

$$-\frac{1}{2} \log P = \frac{\pi ZZ'e^2}{\hbar v} \left(1 - \frac{2}{\pi} \sin^{-1} \frac{1}{\sqrt{ZZ'e^2/(ER)}}\right) - \frac{mvR}{\hbar} \left(\sqrt{\frac{ZZ'e^2}{ER}} - 1\right)^2. \quad (6.237)$$

We model the lifetime for alpha-decay as τ/P where τ is the rate at which the alpha particle strikes the barrier. In terms of the alpha's velocity v and the nuclear radius R , τ would be $2R/v$. With nuclear radii ranging from $10^{-12} - 10^{-13}$ cm, and known values of Z and E for alpha decays, the right hand side of Eq. 6.237 varies by about a factor of 25. It reproduces experimental data for lifetimes over 23 orders of magnitude.

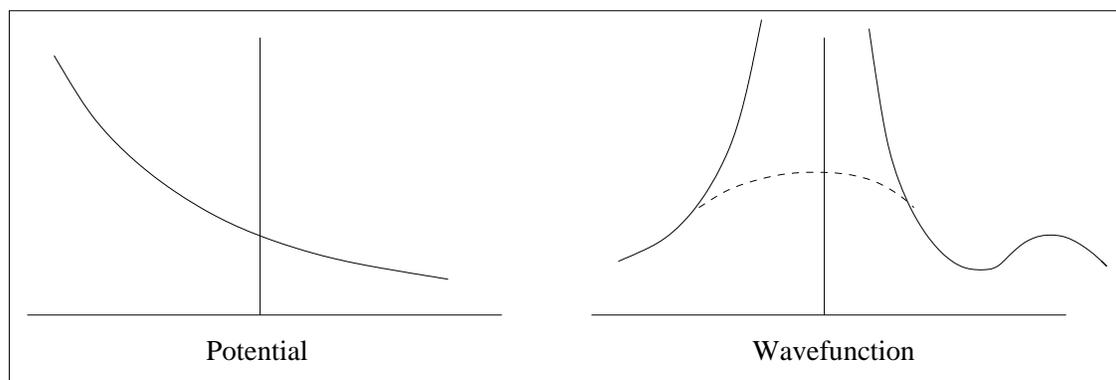
In low energy nuclear reactions one can run the problem backward: what is the probability that the alpha particle gets into the nucleus? It is

$$P \simeq \exp \left[- \frac{2\pi ZZ'e^2}{\hbar v} \right] \quad (6.238)$$

because of the Coulomb barrier. This quantity is called the “Gamow factor,” and is the dominant term in the rate for many low energy nuclear reactions. Fusion reactions in the sun are a particularly important example – hence the meme “The sun shines by quantum tunneling.”

There are many other more exotic variations of this calculation, including the decay of the early universe from a “false vacuum” to its current ground state.

WKB for energy eigenvalues



The condition $(\lambda/(4\pi))|dk/dx| \ll k$ is violated when $V(x) = E$ or $k = 0$, i.e., near the classical turning points for the orbits. How can we use WKB in that case? We need what

are called “connection formulas.” To get them, and their consequences for bound state, we need to make a mathematical digression. Introduce the variable

$$\xi = \int^x k(x') dx' \quad (6.239)$$

and rewrite

$$\frac{d^2 u_0(x)}{dx^2} + k(x)^2 u_0(x) = 0 \quad (6.240)$$

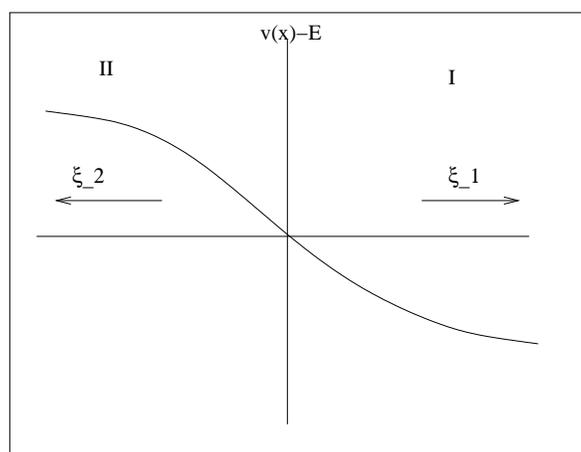
as

$$\frac{d^2 u_0}{d\xi^2} + \left[\frac{1}{k(x)^2} \frac{dk(x)}{dx} \right] \frac{du_0}{d\xi} + u_0 = 0. \quad (6.241)$$

In the good WKB limit, the object in square brackets vanishes, and our solutions are $u_0 \simeq \exp(\pm\xi)$, as expected. To derive the connection formula, we assume that $V(x) - E \rightarrow 0$ vanishes with a linear dependence on x . (This is almost always the case in practical applications.) Choose coordinates so that this zero occurs at $x = 0$. We write $V(x) - E = Cx$, where C is a constant. We introduce the variables for use in regions I and II (see figure)

$$\xi_1 \equiv \int_0^x k dx \quad (6.242)$$

$$\xi_2 \equiv \int_x^0 \kappa dx, \quad (6.243)$$



which both increase as we go away from zero. In region I we have

$$\frac{d^2 u_0}{d\xi^2} + \frac{1}{3\xi} \frac{du_0}{d\xi} + u_0 = 0. \quad (6.244)$$

The solutions to this textbook differential equation are called Airy functions. They are ordinary Bessel functions order of $\pm 1/3$. Explicitly, they are

$$u_1^\pm(x) = A_\pm \left(\frac{\xi_1}{k}\right)^{\frac{1}{2}} J_{\pm\frac{1}{3}}(\xi_1). \quad (6.245)$$

In region II the solutions are also Airy functions, this time of imaginary order:

$$u_2^\pm = B_\pm \left(\frac{\xi_2}{k}\right)^{\frac{1}{2}} I_{\pm\frac{1}{3}}(\xi_2). \quad (6.246)$$

As $x \rightarrow 0$ the Bessel functions behave like

$$\begin{aligned} J_{\pm\frac{1}{3}}(\xi_1) &= \frac{(\frac{1}{2}\xi_1)^{\pm\frac{1}{3}}}{\Gamma(1 \pm \frac{1}{3})} \\ I_{\pm\frac{1}{3}}(\xi_2) &= \frac{(\frac{1}{2}\xi_2)^{\pm\frac{1}{3}}}{\Gamma(1 \pm \frac{1}{3})} \end{aligned} \quad (6.247)$$

and as $x \rightarrow \infty$ they approach

$$J_{\pm\frac{1}{3}}(\xi_1) = \frac{1}{\sqrt{\frac{1}{2}\pi\xi_1}} \cos\left(\xi_1 \mp \frac{\pi}{6} - \frac{\pi}{4}\right) \quad (6.248)$$

and

$$I_{\pm\frac{1}{3}}(\xi_2) = \frac{1}{\sqrt{\frac{1}{2}\pi\xi_2}} \left[e^{\xi_2} + e^{-\xi_2} e^{-(\frac{1}{2}\pm\frac{1}{3})i\pi} \right]. \quad (6.249)$$

At the turning point $k^2 = Cx$ or $k = c\sqrt{x}$ and $\kappa^2 = C|x|^2$ or $\kappa = c\sqrt{|x|}$, so the limiting form of the solutions to the Schrödinger equation are

$$u_1^+ = \frac{A_+ \left(\frac{2}{3}\right)^{\frac{1}{2}} \left(\frac{1}{3}c\right)^{\frac{1}{3}} x}{\Gamma\left(\frac{4}{3}\right)} \quad u_1^- = \frac{A_- \left(\frac{2}{3}\right)^{\frac{1}{2}} \left(\frac{1}{3}c\right)^{\frac{1}{3}}}{\Gamma\left(\frac{2}{3}\right)} \quad (6.250)$$

and

$$u_2^+ = \frac{B_+ \left(\frac{2}{3}\right)^{\frac{1}{2}} \left(\frac{1}{3}c\right)^{\frac{1}{3}} |x|}{\Gamma\left(\frac{4}{3}\right)} \quad u_2^- = \frac{B_- \left(\frac{2}{3}\right)^{\frac{1}{2}} \left(\frac{1}{3}c\right)^{\frac{1}{3}}}{\Gamma\left(\frac{2}{3}\right)}. \quad (6.251)$$

If we set $B_+ = -A_+$ then u_1^+ joins u_2^+ , and similarly, to match u_1^- and u_2^- we require $A_- = B_-$. This achieves a match near the classical turning point. However, we also need to check that our solutions obey their boundary conditions at $\pm\infty$. The limiting forms are

$$u_1^+ \rightarrow \frac{1}{\sqrt{\frac{1}{2}\pi k}} \cos\left(\xi_1 - \frac{5\pi}{12}\right), \quad (6.252)$$

$$u_1^- \rightarrow \frac{1}{\sqrt{\frac{1}{2}\pi k}} \cos\left(\xi_1 - \frac{\pi}{12}\right), \quad (6.253)$$

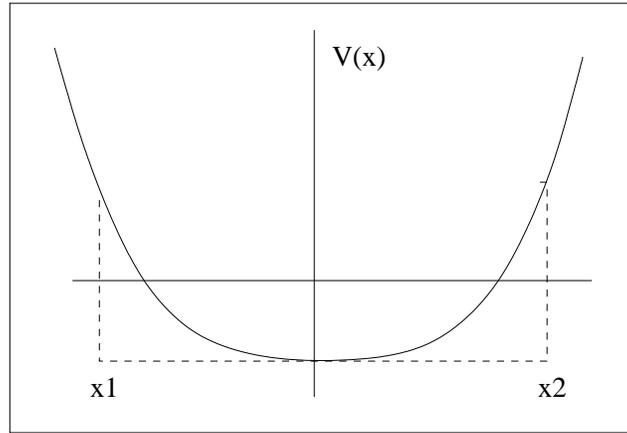
$$u_2^+ \rightarrow \frac{-1}{\sqrt{2\pi\kappa}} \left(e^{\xi_2} + e^{-\xi_2} e^{-\frac{5\pi i}{6}} \right) \quad (6.254)$$

and

$$u_2^- \rightarrow \frac{1}{\sqrt{2\pi\kappa}} \left(e^{\xi_2} + e^{-\xi_2} e^{-\frac{\pi i}{6}} \right). \quad (6.255)$$

In order that ψ vanishes as $x \rightarrow -\infty$, we must choose the linear combination $u_2^+ + u_2^-$. This oscillates as $u_1^+ + u_1^- \simeq \cos(\xi_1 - \pi/4)$ in the classically allowed region.

Now imagine we have a confining potential well with two turning points, as shown in the figure:



Inside the well (in the classically allowed region) we construct a solution by starting our integration at x_1 , one of the classical turning points. There

$$u(x) \simeq \frac{1}{\sqrt{k}} \cos \left[\int_{x_1}^x k dx' - \frac{\pi}{4} \right]. \quad (6.256)$$

We can also begin by integrating from x_2 , so

$$u(x) \simeq \frac{1}{\sqrt{k}} \cos \left[\int_x^{x_2} k dx' - \frac{\pi}{4} \right] \quad (6.257)$$

$$= \frac{1}{\sqrt{k}} \cos \left[\int_x^{x_1} k dx' + \int_{x_1}^{x_2} k dx' - \frac{\pi}{4} \right]. \quad (6.258)$$

These solutions must be identical. This forces the equality

$$- \int_{x_1}^{x_2} k dx + \frac{\pi}{4} = -\frac{\pi}{4} - n\pi, \quad (6.259)$$

which, upon re-introducing the potential, is the equation

$$\int_{x_1}^{x_2} dx \sqrt{\frac{2m(E - V(x))}{\hbar^2}} = \left(n + \frac{1}{2}\right)\pi \quad (6.260)$$

where $n \in \{0, 1, 2, \dots\}$. For historical reasons we replace \hbar by $h/(2\pi)$, and write $p(x) = \sqrt{2m(E - V(x))}$ to get

$$2 \int_{x_1}^{x_2} dx p(x) = \left(n + \frac{1}{2}\right)h. \quad (6.261)$$

We regard the integral from x_1 to x_2 as half of an integral for x_1 (a classical turning point) to x_2 (the other classical turning point) and back to the start: an integral of the momentum about the closed classical trajectory. Then the consistency condition is

$$\oint p dx = \left(n + \frac{1}{2}\right)h. \quad (6.262)$$

Equation 6.262 is known as the Bohr-Sommerfeld quantization condition, a 1916 generalization of Bohr's famous 1913 guess of $L = n\hbar$ for a circular orbit. We have derived it from the Schrödinger equation in the WKB limit. It is a quantization condition because it will only be satisfied for particular values of the energy E . In this formula, n is the number of nodes in the wave function. WKB is only a good approximation if the number of nodes is large, or (equivalently) if the classical turning points are many wavelengths apart.

In a few cases, WKB is exact. The simple harmonic oscillator is the most important such case. The connection formula gives the $1/2$ in the energy formula $E = \hbar\omega(n + \frac{1}{2})$. Clearly, if n is large, keeping the $1/2$ or leaving it out will not be very important.

Connection of the WKB method and the path integral

Recall that the equation for the time evolution of a wave function is

$$\psi(x, t) = \int U(x, t; x', t') \psi(x', t') dx' \quad (6.263)$$

where the time evolution operator is

$$U(x, t; x', t') = \int \prod_i dx_i \exp \left[\frac{i}{\hbar} \int_{t'}^t \mathcal{L}(x, \dot{x}) dt'' \right]. \quad (6.264)$$

Suppose that the motion of the particle is nearly classical, so that in evaluating the path integral, x and dx/dt remain very close to their classical values (the ones found by solving Lagrange's equation). Then

$$\begin{aligned} U(x, t; x', t') &= C \exp \left[\frac{i}{\hbar} \int_{t'}^t dt'' \mathcal{L}(x_{cl}, \dot{x}_{cl}) \right] \\ &= C \exp \left[\frac{i}{\hbar} \int_{t'}^t dt'' \left[\frac{p^2}{2m} - V(x_{cl}) \right] \right], \end{aligned} \quad (6.265)$$

replacing \dot{x} by p/m . The sum over paths disappears because the particle is confined to its classical path. On the classical path, energy is conserved, so we can replace $V(x)$ by $E - p^2/(2m)$, giving

$$U(x, t; x', t') = C \exp \left(\frac{i}{\hbar} \int_{t'}^t \frac{p^2}{m} dt'' \right) \exp \left(- \frac{i}{\hbar} E(t - t') \right). \quad (6.266)$$

Finally, the first term simplifies to

$$\int_{t'}^t p \frac{p}{m} dt'' = \int_{t'}^t p \frac{dx}{dt''} dt'' = \int_{x'}^x p dx'', \quad (6.267)$$

leaving an evolution equation

$$\psi(x, t) = \phi(x) \exp - \frac{iEt}{\hbar} \quad (6.268)$$

with

$$\phi(x) = A \int \left[\exp \left(\frac{i}{\hbar} \int_{x'}^x p(x'') dx'' \right) \right] \phi(x') dx'. \quad (6.269)$$

If we replace the integral $\int_{x'}^x$ by $\int_{x'}^{x_0} + \int_{x_0}^x$, then we can write this expression as

$$\phi(x) = \exp \left(\frac{i}{\hbar} \int_{x_0}^x p(x'') dx'' \right) \left[\int dx' \phi(x') A \exp \left(\frac{i}{\hbar} \int_{x'}^{x_0} p(x'') dx'' \right) \right]. \quad (6.270)$$

The quantity in the brackets is independent of x . We call it $f(x_0)$ and rewrite Eq. 6.270 as

$$\phi(x) = \exp \left[\frac{i}{\hbar} \int_{x_0}^x p(x'') dx'' \right] f(x_0). \quad (6.271)$$

Look closely at this expression – it tells us that $f(x_0)$ is actually $\phi(x_0)$. Thus

$$\phi(x) = \exp \left[\frac{i}{\hbar} \int_{x_0}^x p(x'') dx'' \right] \phi(x_0). \quad (6.272)$$

This is again the WKB formula for $\phi(x)$ in terms of $\phi(x_0)$.

Chapter 7

Angular momentum

Angular momentum and the consequences of its conservation appear in so many practical considerations that it is very useful to devote some time to its study. We will do this in the abstract, rather than through the use of coordinate space wave functions, like the spherical harmonics. Electrons (and other elementary particles) carry an intrinsic spin angular momentum of $\frac{1}{2}\hbar$, which, in contrast to orbital angular momentum, has no simple connection with either coordinate or momentum. An operator language is the only way to deal with spin. Accordingly, we assume the existence of a set of abstract operators $(\hat{J}_1, \hat{J}_2, \hat{J}_3)$ and their square, which obey the commutation relations

$$[\hat{J}^2, \hat{J}_i] = 0 \quad [\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k. \quad (7.1)$$

We defer any explanation of why such commutation relations might be expected or “natural.” (It will come.) We also assume the existence of states $|jm\rangle$ such that

$$\begin{aligned} \hat{J}^2 |j, m\rangle &= \hbar^2 f(j) |j, m\rangle \\ \hat{J}_z |j, m\rangle &= \hbar m |j, m\rangle. \end{aligned} \quad (7.2)$$

All the properties of the states and operators can be constructed from the commutation relations. To simplify the process, let us define two more operators, called raising and lowering operators,

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \quad \hat{J}_- = \hat{J}_x - i\hat{J}_y \quad (7.3)$$

which have an algebra

$$[\hat{J}^2, \hat{J}_\pm] = 0 \quad [\hat{J}_z, \hat{J}_\pm] = \pm\hbar\hat{J}_\pm \quad (7.4)$$

$$[\hat{J}_+, \hat{J}_-] = 2\hbar\hat{J}_z. \quad (7.5)$$

To see how they act on states, consider the matrix element

$$\langle j, m | [\hat{J}_z, \hat{J}_+] | j, m' \rangle = \langle j, m | \hat{J}_z \hat{J}_+ | j, m' \rangle - \langle j, m | \hat{J}_+ \hat{J}_z | j, m' \rangle = \hbar \langle j, m | \hat{J}_+ | j, m' \rangle. \quad (7.6)$$

Letting \hat{J}_z act to the left and right gives

$$0 = (m - m' - 1) \langle j, m | \hat{J}_+ | j, m' \rangle. \quad (7.7)$$

This means that either $m = m' + 1$, or the matrix element must vanish. Therefore, \hat{J}_+ changes the state $|j, m\rangle$ into a state proportional to $|j, m + 1\rangle$. There is a similar calculation for the \hat{J}_- operator. To summarize,

$$\hat{J}_+ |j, m\rangle = \hbar\lambda_m |j, m + 1\rangle \quad (7.8)$$

and

$$\hat{J}_- |j, m\rangle = \hbar\lambda_{m-1}^* |j, m-1\rangle. \quad (7.9)$$

We may determine the λ_m 's from

$$\langle jm | [\hat{J}_+, \hat{J}_-] |jm\rangle = \langle jm | 2\hbar\hat{J}_z |jm\rangle = 2\hbar^2 m. \quad (7.10)$$

Acting left and right with \hat{J}_+ and \hat{J}_- , we find

$$= \hbar^2 \{ |\lambda_{m-1}|^2 - |\lambda_m|^2 \} = 2m\hbar^2. \quad (7.11)$$

The solution to this equation is $|\lambda_m|^2 = C - m(m+1)$. Because $|\lambda_m|^2 \geq 0$, the spectrum of m 's in the states $|j, m\rangle$ is finite and bounded. There is a largest $m = m_1$, for which $\lambda_{m_1} = 0$, and

$$\hat{J}_+ |j, m_1\rangle = 0. \quad (7.12)$$

There is also a smallest m_2 with $\hat{J}_- |j, m_2+1\rangle = \lambda_{m_2} |j, m_2\rangle = 0$. Starting from $|j, m_2+1\rangle$, we may construct states $|j, m_2+2\rangle, |j, m_2+3\rangle \dots$ by repeated application of \hat{J}_+ . Eventually, we must hit the state $|j, m_1\rangle$ and the series must terminate. Otherwise, we would find a λ_m with $|\lambda_m|^2 < 0$. Thus the number of states must be an integer.

We may find m_1 and m_2 by solving $C - m(m+1) = 0$. The solutions to this quadratic equation are $m_{1,2} = -\frac{1}{2} \pm \frac{1}{2}\sqrt{4C+1}$. Note that $m_1 = -(m_2+1)$. Then either m_1 must be an integer (for example, $m = (-3, -2, -1, 0, 1, 2, 3)$) or m_1 must a half integer $m = (3/2, 1/2, -1/2, -3/2)$. Either way, the minimum m_1 is the negative of its maximum. The algebra of the commutation relation forces this result.

Finally, what is $f(j)$, the eigenvalue of \hat{J}^2 ? To find it, consider the matrix element

$$\begin{aligned} \langle jm | \hat{J}^2 |jm\rangle &= \langle jm | \hat{J}_z^2 + \frac{1}{2}(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+) |jm\rangle \\ &= \hbar^2 m^2 + \frac{1}{2}(|\lambda_{m-1}|^2 + |\lambda_m|^2) \\ &= \hbar^2 C \\ &= \hbar^2 m_1(m_1+1) \end{aligned} \quad (7.13)$$

We rename m_1 as j , so the eigenvalues of \hat{J}^2 are $\hbar^2 j(j+1)$. Thus the quantity j is either an integer or a half-integer, and the allowed values of m run in integer steps $m = -j, -j+1, -j+2, \dots, j-1, j$.

We summarize all the results for matrix elements:

$$\begin{aligned}\langle j, m+1 | \hat{J}_+ | jm \rangle &= \sqrt{j(j+1) - m(m+1)} \hbar \\ &= \sqrt{(j-m)(j+m+1)} \hbar\end{aligned}\quad (7.14)$$

$$\begin{aligned}\langle j, m-1 | \hat{J}_- | jm \rangle &= \sqrt{j(j+1) - (m-1)m} \hbar \\ &= \sqrt{(j+m)(j-m+1)} \hbar\end{aligned}\quad (7.15)$$

$$\hat{J}_z |j, m\rangle = \hbar m |j, m\rangle \quad (7.16)$$

$$\hat{J}^2 |j, m\rangle = \hbar^2(j(j+1)) |j, m\rangle. \quad (7.17)$$

Let us consider some examples of angular momentum matrix elements.

- $j = 0$: There is only one state, the $|j = 0, m = 0\rangle$ state. The operators \hat{J}_z and \hat{J}^2 may be regarded as 1×1 matrices (with the value zero).
- $j = \frac{1}{2}$: There are two states $|\frac{1}{2}, +\frac{1}{2}\rangle$ and $|\frac{1}{2}, -\frac{1}{2}\rangle$ corresponding to the vectors

$$|\frac{1}{2}, +\frac{1}{2}\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |\frac{1}{2}, -\frac{1}{2}\rangle \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (7.18)$$

The matrix expressions for the operators are $\hat{\vec{J}} = \frac{\hbar}{2} \vec{\sigma}$ where the $\vec{\sigma}$ are the Pauli matrices.

$$\hat{J}_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \hat{J}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \hat{J}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (7.19)$$

$$\hat{J}_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \hat{J}_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (7.20)$$

- $j = 1$: With the three m - states, we associate vectors

$$|1, +1\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad |1, 0\rangle \rightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad |1, -1\rangle \rightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (7.21)$$

and operators

$$\hat{j}^2 = 2\hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \hat{J}_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (7.22)$$

$$\hat{J}_+ = \sqrt{2}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \hat{J}_- = \sqrt{2}\hbar \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad (7.23)$$

- $j = \frac{3}{2}$: The angular momentum operators are

$$\hat{J}_z = \hbar \begin{pmatrix} -\frac{3}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{3}{2} \end{pmatrix} \quad \hat{J}_+ = \hbar \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (7.24)$$

$$\hat{J}_- = \hbar \begin{pmatrix} 0 & 0 & 0 & 0 \\ \sqrt{3} & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}. \quad (7.25)$$

Combining angular momenta

Let us imagine that we have a system with two kinds of angular momenta, \hat{J}_1 and \hat{J}_2 , and a Hamiltonian \hat{H} which commutes with their sum, though not with either individual angular momentum:

$$[\hat{\vec{J}}_1 + \hat{\vec{J}}_2, \hat{H}] = 0, \quad [\hat{\vec{J}}_1, \hat{H}] \neq 0, \quad [\hat{\vec{J}}_2, \hat{H}] \neq 0. \quad (7.26)$$

This situation often occurs in Nature. A simple example is a one-electron atom. The two angular momenta are the electron's spin \vec{S} and its orbital angular momentum \vec{L} . We know that the electron experiences a spin-orbital interaction, $H \simeq \vec{L} \cdot \vec{S}$. Clearly, $[\hat{\vec{S}}, \hat{\vec{L}} \cdot \hat{\vec{S}}] \neq 0$. The electron's total angular momentum is a constant of the motion, simply because there are no torques acting on the system as a whole. Thus $\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$ is conserved. We want to be able to construct states which are eigenstates of all conserved quantum numbers. In this case, these quantities can only be \hat{J} and one of its components, which we select to be \hat{J}_z .

However, we only know how to construct states which are (separately) eigenfunctions of $(\hat{J}_1^2, \hat{J}_{1z})$ and $(\hat{J}_2^2, \hat{J}_{2z})$. As long as \hat{J}_1 and \hat{J}_2 involve different sectors of Hilbert space, we

can write these combined states as product states

$$|J_1 m_1\rangle |J_2 m_2\rangle \equiv |J_1 J_2 m_1 m_2\rangle \quad (7.27)$$

where, for example,

$$\hat{J}_1^2 |J_1 J_2 m_1 m_2\rangle = \hbar^2(j_1(j_1 + 1)) |J_1 J_2 m_1 m_2\rangle \quad (7.28)$$

and so on.

Since the total angular momentum obeys a commutation relation $\vec{J} \times \vec{J} = i\hbar\vec{J}$, we can try to build eigenstates of \hat{J}^2 and \hat{J}_z from a linear combination of $|J_1 J_2 m_1 m_2\rangle$'s. \hat{J}_1^2 and \hat{J}_2^2 are scalars, so they will commute with \hat{J} ; hence, these states will be labeled $|j_1 j_2 j m\rangle$. These states will be a linear combination of the product states:

$$|j_1 j_2 j m\rangle = \sum_{m_1 m_2} |j_1 j_2 m_1 m_2\rangle \langle j_1 j_2 m_1 m_2 | j_1 j_2 j m\rangle. \quad (7.29)$$

The object $\langle j_1 j_2 m_1 m_2 | j_1 j_2 j m\rangle$ is a number: it is called a Clebsch-Gordon coefficient. The sum runs over all values of m_1 and m_2 ; however, since $\hat{J}_z = \hat{J}_{1z} + \hat{J}_{2z}$, there is a constraint that $m_1 + m_2 = m$, and the sum is restricted to all values of m_1 and m_2 such that $m_1 + m_2 = m$.

Example: Spin $\frac{1}{2}$ electrons in a P-wave orbital

Recoupling angular momentum is a task which can simplify many physical problems. It is worthwhile to spend time on several examples, introducing along the way a method (not the best nor fastest, but certainly the most widely known) for finding Clebsch-Gordon coefficients. Recall why we are doing this: the basis $|j_1 j_2 j m\rangle$ may be more useful than the product state $|j_1 j_2 m_1 m_2\rangle$.

We begin imagining a spin-1/2 electron in a P -orbital. There are six possible states corresponding to the two spin orientations and three L_z states. We can label these states as

$$|11\rangle |\uparrow\rangle \quad |11\rangle |\downarrow\rangle \quad |10\rangle |\uparrow\rangle \quad (7.30)$$

$$|10\rangle |\downarrow\rangle \quad |1-1\rangle |\uparrow\rangle \quad |1-1\rangle |\downarrow\rangle. \quad (7.31)$$

We expect

$$|J.M\rangle = \sum_{s,m} |1, \frac{1}{2}, m, s\rangle \langle 1, \frac{1}{2}, m, s | J, M\rangle \quad (7.32)$$

with $s + m = M$. Note that if $M = \frac{3}{2}$, only one term contributes to the sum. For $M = \frac{1}{2}$ there are 2 states, for $M = -\frac{1}{2}$ there are also 2 states, and for $M = -\frac{3}{2}$ there is one state. Thus

$$|J, M = \frac{3}{2}\rangle = |11\rangle |\uparrow\rangle. \quad (7.33)$$

The Clebsch-Gordon coefficient must be unity to preserve the normalization of the state. To build other states, write the raising and lowering operators as $\hat{J}_{\pm} = \hat{L}_{\pm} + \hat{S}_{\pm}$ and then examine the action of an operator on the state

$$\hat{J}_+ |J, M = \frac{3}{2}\rangle = (\hat{L}_+ + \hat{S}_+) |11\rangle |\uparrow\rangle. \quad (7.34)$$

Because this state is the highest state in the hierarchy, the raising operator must yield zero.

$$\hat{J}_+ |J, M = \frac{3}{2}\rangle = \sqrt{j(j+1) - \frac{3}{2}(\frac{3}{2} + 1)} |J, M = \frac{3}{2} + 1\rangle = 0. \quad (7.35)$$

This constraint tells us that $J = \frac{3}{2}$. A similar calculation can be done for the $|J, M = -\frac{3}{2}\rangle$ state. To calculate the Clebsch-Gordon coefficients for the full spectrum of states, we use raising and lowering operators starting from states we know about, namely, the highest and lowest state in the spectrum.

$$\hat{J}_- |\frac{3}{2}, \frac{3}{2}\rangle \propto |\frac{3}{2}, \frac{1}{2}\rangle \propto (\hat{L}_- + \hat{S}_-) |11\rangle |\uparrow\rangle \propto |10\rangle |\uparrow\rangle + |11\rangle |\downarrow\rangle. \quad (7.36)$$

We can also raise the lowest state,

$$\hat{J}_+ |\frac{3}{2}, -\frac{3}{2}\rangle = (\hat{L}_+ + \hat{S}_+) |1-1\rangle |\downarrow\rangle \propto |10\rangle |\downarrow\rangle + |1-1\rangle |\uparrow\rangle. \quad (7.37)$$

There are six states in all; four of them make up the $J = \frac{3}{2}$ multiplet. The two remaining states comprise the $J = \frac{1}{2}$ multiplet. As a rule one can show that the allowed values of J range from $J = J_1 + J_2$ down to $|J_1 - J_2|$, and here, with $J_1 = 1$ and $J_2 = \frac{1}{2}$, the $J = \frac{3}{2}$ and $J = \frac{1}{2}$ multiplets are the only ones present in the system. Symbolically, we write

$$1 \otimes \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2}. \quad (7.38)$$

We read this equation as “3 states times 2 states gives a decomposition of 4 states plus 2 states.” The same sort of direct product and direct sum can be done for other systems. As an example, consider the coupling of a $J = 1$ system to a $J = 1$ system.

$$1 \otimes 1 = 2 \oplus 1 \oplus 0. \quad (7.39)$$

State counting works: $3 \times 3 = 5 + 3 + 1$.

We return to the construction of the full set of states of the $J = \frac{3}{2}$ system. We act on the $|\frac{3}{2}, \frac{3}{2}\rangle$ state with the \hat{J}_- operator, to find

$$\begin{aligned}\hat{J}_- |\frac{3}{2}, \frac{3}{2}\rangle &= \hat{L}_- |11\rangle |\uparrow\rangle + \hat{S}_- |11\rangle |\uparrow\rangle \\ &= \sqrt{1 \cdot 2 - 1 \cdot 0} |10\rangle |\uparrow\rangle + \sqrt{\frac{1}{2} \cdot \frac{3}{2} - \frac{1}{2} \cdot (-\frac{1}{2})} |11\rangle |\downarrow\rangle \\ \hat{J}_- |\frac{3}{2}, \frac{3}{2}\rangle &= \sqrt{j(j+1) - m(m-1)} |\frac{3}{2}, \frac{1}{2}\rangle = \sqrt{\frac{3}{2} \cdot \frac{5}{2} - \frac{3}{2} \cdot \frac{1}{2}} |\frac{3}{2}, \frac{1}{2}\rangle.\end{aligned}\tag{7.40}$$

Therefore

$$|\frac{3}{2}, \frac{1}{2}\rangle = \sqrt{\frac{2}{3}} |10\rangle |\uparrow\rangle + \sqrt{\frac{1}{3}} |11\rangle |\downarrow\rangle.\tag{7.41}$$

We can build the $|\frac{3}{2}, -\frac{1}{2}\rangle$ state by either of two methods. We can either act on the $|\frac{3}{2}, \frac{1}{2}\rangle$ states with a lowering operator, or we can act on the $|\frac{3}{2}, -\frac{3}{2}\rangle$ state with a raising operator. Making the second choice yields

$$\begin{aligned}\hat{J}_+ |\frac{3}{2}, -\frac{3}{2}\rangle &= \sqrt{\frac{3}{2} \cdot \frac{5}{2} + \frac{3}{2} \cdot (-\frac{3}{2})} |\frac{3}{2}, -\frac{1}{2}\rangle \\ &= \sqrt{2} |10\rangle |\downarrow\rangle + |1-1\rangle |\uparrow\rangle\end{aligned}\tag{7.42}$$

and so

$$|\frac{3}{2}, -\frac{1}{2}\rangle = \sqrt{\frac{2}{3}} |10\rangle |\downarrow\rangle + \sqrt{\frac{1}{3}} |1-1\rangle |\uparrow\rangle.\tag{7.43}$$

The $J = \frac{1}{2}$ states remain. There will be a $|\frac{1}{2}, +\frac{1}{2}\rangle$ and a $|\frac{1}{2}, -\frac{1}{2}\rangle$ state. These states should be orthogonal to all the other states. Because the states can only be composed of (sums of) direct products of states where the z-component of angular momentum equals the appropriate value of $\pm\frac{1}{2}$,

$$\begin{aligned}|\frac{1}{2}, +\frac{1}{2}\rangle &\sim |10\rangle |\uparrow\rangle + |11\rangle |\downarrow\rangle \\ |\frac{1}{2}, -\frac{1}{2}\rangle &\sim |10\rangle |\downarrow\rangle + |1-1\rangle |\uparrow\rangle.\end{aligned}\tag{7.44}$$

The Clebsch-Gordon coefficients can be found using orthogonality, $\langle \frac{3}{2}, \frac{1}{2} | \frac{1}{2}, \frac{1}{2} \rangle = 0$ and $\langle \frac{3}{2}, -\frac{1}{2} | \frac{1}{2}, -\frac{1}{2} \rangle = 0$. We recall

$$|\frac{3}{2}, \frac{1}{2}\rangle = \sqrt{\frac{2}{3}} |10\rangle |\uparrow\rangle + \sqrt{\frac{1}{3}} |11\rangle |\downarrow\rangle\tag{7.45}$$

and parameterize

$$|\frac{1}{2}, \frac{1}{2}\rangle = \alpha |10\rangle |\uparrow\rangle + \beta |11\rangle |\downarrow\rangle, \quad (7.46)$$

and then orthogonality gives

$$\alpha\sqrt{\frac{2}{3}} + \beta\sqrt{\frac{1}{3}} = 0 \rightarrow \alpha = -\sqrt{\frac{1}{3}} \quad \beta = \sqrt{\frac{2}{3}} \quad (7.47)$$

A similar operation could be used to construct the $|\frac{1}{2}, -\frac{1}{2}\rangle$ state. Alternatively, could also apply the lowering operator to the $|\frac{1}{2}, \frac{1}{2}\rangle$ state.

Two spin $\frac{1}{2}$ particles

The total angular momentum of a system of two spin-1/2 states is $\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$. A simple application of the techniques of the last section yields

$$\begin{aligned} |11\rangle &= |\uparrow\rangle |\uparrow\rangle \\ |10\rangle &= \sqrt{\frac{1}{2}}(|\uparrow\rangle |\downarrow\rangle + |\downarrow\rangle |\uparrow\rangle) \\ |1-1\rangle &= |\downarrow\rangle |\downarrow\rangle \\ |00\rangle &= \sqrt{\frac{1}{2}}(|\uparrow\rangle |\downarrow\rangle - |\downarrow\rangle |\uparrow\rangle) \end{aligned} \quad (7.48)$$

This combination appears often in fermionic systems. For example, a meson is a bound state of two spin- $\frac{1}{2}$ quarks. One can parameterize the interaction Hamiltonian between the quarks as a spin coupling term $\hat{\vec{S}}_1 \cdot \hat{\vec{S}}_2$. What are the eigenstates of the interaction Hamiltonian, and their degeneracies? One way to find them would be to write the interaction term as

$$\hat{\vec{S}}_1 \cdot \hat{\vec{S}}_2 = \hat{S}_{1z}\hat{S}_{2z} + \frac{1}{2}(\hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+}) \quad (7.49)$$

and diagonalize it in the four-dimensional basis. An easier approach is to write

$$\hat{S}^2 = \hat{S}_1^2 + \hat{S}_2^2 + 2\hat{\vec{S}}_1 \cdot \hat{\vec{S}}_2 \quad (7.50)$$

or

$$\hat{\vec{S}}_1 \cdot \hat{\vec{S}}_2 = \frac{1}{2}(\hat{S}^2 - \hat{S}_1^2 - \hat{S}_2^2) \quad (7.51)$$

from which

$$\hat{\vec{S}}_1 \cdot \hat{\vec{S}}_2 = \frac{1}{2}\hat{S}^2 - \frac{3}{4}\hbar^2. \quad (7.52)$$

The eigenvalue of this operator is $+\frac{1}{4}\hbar^2$ for any of the three $S = 1$ states and $-\frac{3}{4}\hbar^2$ for the $S = 0$ state.

Example: $J = l \pm \frac{1}{2}$ decompositions

This example is particularly relevant to single-electron atoms. The eigenstate $|j, m\rangle$ is

$$|j, m\rangle = \alpha |l, m - \frac{1}{2}\rangle |\uparrow\rangle + \beta |l, m + \frac{1}{2}\rangle |\downarrow\rangle \equiv \alpha |-\rangle |\uparrow\rangle + \beta |+\rangle |\downarrow\rangle. \quad (7.53)$$

It is an eigenstate of $\hat{J}_z = \hat{L}_z + \hat{S}_z$ with eigenvalue $m\hbar$. We would like it also to be an eigenvalue of \hat{J}^2 , and we can force it to be so by picking suitable α and β . Recall that

$$\hat{J}^2 = \hat{L}^2 + \hat{S}^2 + 2\hat{L}_z\hat{S}_z + \hat{L}_+\hat{S}_- + \hat{L}_-\hat{S}_+. \quad (7.54)$$

Acting with \hat{J}^2 on the $|j = l + 1/2, m\rangle$ state produces

$$\begin{aligned} \hat{J}^2 |j, m\rangle &= (l + \frac{1}{2})(l + \frac{3}{2})[\alpha |-\rangle |\uparrow\rangle + \beta |+\rangle |\downarrow\rangle] \\ &= \{(l(l+1)) + \frac{3}{4}\}(\alpha |-\rangle |\uparrow\rangle + \beta |-\rangle |\downarrow\rangle) + (m - \frac{1}{2})\alpha |-\rangle |\uparrow\rangle - (m + \frac{1}{2})\beta |+\rangle |\downarrow\rangle \\ &\quad + \alpha \sqrt{(l+m+\frac{1}{2})(l-m+\frac{1}{2})} |+\rangle |\downarrow\rangle + \beta \sqrt{(l-m+\frac{1}{2})(l+m+\frac{1}{2})} |-\rangle |\uparrow\rangle. \end{aligned} \quad (7.55)$$

Taking the inner product on both sides with either $\langle\downarrow|\langle+|$ or $\langle\uparrow|\langle-|$ gives a relation between α and β . With the latter choice,

$$\alpha(l + \frac{1}{2})(l + \frac{3}{2}) = \{l(l+1) + \frac{3}{4} + m - \frac{1}{2}\}\alpha + \beta \sqrt{(l-m+\frac{1}{2})(l+m+\frac{1}{2})}. \quad (7.56)$$

Equating like terms on both sides,

$$\alpha(l-m+\frac{1}{2}) = \beta \sqrt{(l-m+\frac{1}{2})(l+m+\frac{1}{2})} \quad (7.57)$$

and so

$$\frac{\alpha}{\beta} = \sqrt{\frac{l+m+\frac{1}{2}}{l-m+\frac{1}{2}}}. \quad (7.58)$$

Requiring $\alpha^2 + \beta^2 = 1$ gives

$$\alpha^2 = \frac{l+m+\frac{1}{2}}{2l+1} \quad \beta^2 = \frac{l-m+\frac{1}{2}}{2l+1}. \quad (7.59)$$

By orthogonality,

$$|l, \frac{1}{2}, J = l - \frac{1}{2}, m\rangle = -\beta |l, m - \frac{1}{2}\rangle + \alpha |l, m + \frac{1}{2}\rangle. \quad (7.60)$$

This completes the construction of the mixing matrix for the system,

$$\begin{pmatrix} |l, \frac{1}{2}, l + \frac{1}{2}, m\rangle \\ |l, \frac{1}{2}, l - \frac{1}{2}, m\rangle \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{l+m+\frac{1}{2}}{2l+1}} & \sqrt{\frac{l-m+\frac{1}{2}}{2l+1}} \\ -\sqrt{\frac{l-m+\frac{1}{2}}{2l+1}} & \sqrt{\frac{l+m+\frac{1}{2}}{2l+1}} \end{pmatrix} \begin{pmatrix} |m - \frac{1}{2}\rangle |\uparrow\rangle \\ |m + \frac{1}{2}\rangle |\downarrow\rangle \end{pmatrix}. \quad (7.61)$$

Entanglement, again

Let's recopy the states of two spin-1/2 particles:

$$\begin{aligned}
 |11\rangle &= |\uparrow\rangle |\uparrow\rangle \\
 |10\rangle &= \sqrt{\frac{1}{2}}(|\uparrow\rangle |\downarrow\rangle + |\downarrow\rangle |\uparrow\rangle) \\
 |1-1\rangle &= |\downarrow\rangle |\downarrow\rangle \\
 |00\rangle &= \sqrt{\frac{1}{2}}(|\uparrow\rangle |\downarrow\rangle - |\downarrow\rangle |\uparrow\rangle)
 \end{aligned}
 \tag{7.62}$$

States like $|11\rangle$ and $|1-1\rangle$ are examples of “pure states” – they are just product states. States like $|10\rangle$ or $|00\rangle$ are examples of what are called “entangled states” – basically, any many-body state which is not a product state is entangled. Entangled states have some very peculiar properties – peculiar, at least, from a conceptual point of view, not a mathematical one.

Imagine that you have prepared a $|00\rangle$ state, and you measure the z-component of one of the spins, say $S_z(1)$. The average value is zero, made up of an equal number of measurements giving $\langle S_z(1) \rangle = 1/2$ or $-1/2$. And of course, if you measure $\langle S_z(1) \rangle = 1/2$, then it must be that $\langle S_z(2) \rangle = -1/2$. Physically, of course, this follows from angular momentum conservation, but it also follows from the form of the state, involving only $|\uparrow\rangle |\downarrow\rangle$ or $|\downarrow\rangle |\uparrow\rangle$.

This correlation would persist even if the two spins were physically separated. An example of such a situation would be the decay of a spinless particle into two spin - 1/2 fermions which are produced back to back in the rest frame of the parent particle. Now the detection of the orientation of spin 1 can take place at a macroscopic separation from the detection of the orientation of spin 2. Nevertheless, there is a perfect correlation in the value of $S_z(1)$ and $S_z(2)$.

This is not just classical physics. The observers could measure $S_x(1)$ and $S_x(2)$: they would be perfectly correlated. $S_z(1)$ and $S_x(2)$ are, however, not correlated.

Such nonlocal correlations bother some people. Einstein was one of them. To pursue that direction would take us into a discussion of the “Einstein - Rosen - Podolsky” paradox and Bell's theorem as a test of quantum mechanics as compared to “hidden variable” alternatives. Let us leave that for the reader to explore on his/her own and simply remark that one can sometimes use these long - distance correlations to do physics. In my own field, this has been

done on an industrial scale, in what are called B factories. B mesons and their antiparticles are bound states of a heavy b quark and a light antiquark (or vice versa). It happens that the neutral ones can oscillate into each other, $B^0 \rightarrow \bar{B}^0 \rightarrow B^0 \dots$ (Think of a two state system in an external magnetic field not aligned along the spin measurement direction.) They are produced in electron positron annihilation through an intermediate resonance, so that the two mesons oscillate coherently; until one particle decays, the system consists of one B and one \bar{B} . But these particles can decay, and when one decays, its clock stops. The nature of the decay tags the particle as B or \bar{B} . The other one goes along until it decays, so there is a possibility of events with two B 's or two \bar{B} 's, with a probability which is governed by the time between the two decays. The interesting physics is in the properties of the B mesons; fundamental quantum mechanics through the entangled state allows the experimentalist to study it. The BaBar experiment collected about 4.7×10^8 pairs of B mesons – industrial quantum mechanics, indeed.

Symmetries and transformations

In classical mechanics, a symmetry is a transformation of the variables which leaves the Lagrangian unchanged. In general, the existence of a symmetry implies a conservation law. For example, consider a non-relativistic particle in a potential which is a function of the magnitude of its distance from the origin. The system is rotationally invariant and the particle's angular momentum will be conserved. To investigate the general form of this connection, let us perturb a coordinate in the Lagrangian by a small amount.

$$\mathcal{L}(q + \delta q, \dot{q}, \dots) = \mathcal{L}(q) + \delta q \frac{\partial \mathcal{L}}{\partial q}. \quad (7.63)$$

If this operation is a symmetry transformation, then

$$\frac{\partial \mathcal{L}}{\partial q} = 0. \quad (7.64)$$

The Lagrange equations of motion

$$\frac{\partial \mathcal{L}}{\partial q} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \quad (7.65)$$

then require that $\partial \mathcal{L} / \partial \dot{q}$ must remain constant through the time evolution of the system. In Hamiltonian dynamics this quantity is the momentum conjugate to the coordinate,

$$p = \frac{\partial \mathcal{L}}{\partial \dot{q}} \quad (7.66)$$

and (equivalently) the presence of the symmetry requires that it be a conserved quantity.

Now consider the quantum mechanical version of these statements. Consider shifting the “center of the wave function” by an infinitesimal variation in q ,

$$\begin{aligned}\psi(q - (q_0 + \delta q)) &= \psi(q - q_0) - \delta q \frac{\partial \psi}{\partial q} \psi(q - q_0) + \dots \\ &\equiv \hat{S}(\delta q) \psi(q - q_0).\end{aligned}\tag{7.67}$$

If δq is small, the transformation is

$$\hat{S} = \hat{1} - \delta q \frac{\partial}{\partial q}.\tag{7.68}$$

The variation $\partial/\partial q$ can be written in terms of p , the operator conjugate to q (working in a basis where q is diagonal). Because of the canonical commutation relation $[p, q] = -i\hbar$, $\partial/\partial q = (i/\hbar)p$ and

$$\hat{S} = \hat{1} - \frac{i}{\hbar} \delta q \hat{p}.\tag{7.69}$$

A mathematician writes a general transformation by an amount δq as

$$\hat{S} = \hat{1} - \frac{i}{\hbar} \delta q \hat{G}\tag{7.70}$$

where \hat{G} is called the “generator of the transformation.” If the transformation $\psi \rightarrow \hat{S}\psi$ is unitary (and so far it nearly is, $S^\dagger S \sim (\delta q)^2$), then the Hamiltonian transforms as

$$\hat{H} \rightarrow \hat{S} \hat{H} \hat{S}^\dagger.\tag{7.71}$$

If \hat{H} is invariant under this transformation then $\hat{H} = \hat{S} \hat{H} \hat{S}^\dagger$ or $\hat{H} \hat{S} = \hat{S} \hat{H}$ and $[\hat{H}, \hat{S}] = 0$. Equivalently, the generator and the Hamiltonian commute, $[\hat{H}, \hat{G}] = 0$. The Heisenberg equations of motion then tell us that the physical observable corresponding to the generator of the transformation is conserved, because

$$[\hat{H}, \hat{S}] = 0 \rightarrow \frac{d\hat{G}}{dt} = 0.\tag{7.72}$$

Further, if the transformation is unitary and if $|\psi\rangle$ is an eigenstate of the Hamiltonian, $\hat{S}|\psi\rangle$ will also be an eigenstate with the same eigenvalue

$$\hat{H} \hat{S} |\psi\rangle = \hat{S} \hat{H} |\psi\rangle = \hat{S}(E |\psi\rangle) = E(\hat{S} |\psi\rangle).\tag{7.73}$$

Thus, in addition to leading to a conservation law, symmetry also implies a degeneracy of states.

The discussion so far involved only an infinitesimal transformation. The finite transformation can be constructed by linking together a set of N infinitesimal transformations one after another. We can define a finite transformation Δq as $N\delta q$, and

$$\hat{S}(\Delta q) = \left(1 - \frac{i}{\hbar} \frac{\Delta q}{N} \hat{p}\right)^N. \quad (7.74)$$

The limit $N \rightarrow \infty$ gives

$$\hat{S}(\Delta q) = \exp\left(-\frac{i}{\hbar} \Delta q \hat{p}\right). \quad (7.75)$$

Observe that if \hat{p} is Hermitian, then the operator \hat{S} is unitary.

The translation operator

The translation operator is

$$\hat{S} = \hat{1} - \frac{i}{\hbar} \delta \vec{x} \cdot \vec{p} = \hat{1} - \delta \vec{x} \cdot \vec{\nabla}. \quad (7.76)$$

Specialize to one dimension and imagine applying this operator to a Gaussian wave packet centered at x_0 , $\psi \propto \exp(-(x - x_0)^2)$. It should push the state's center to $x_0 + \delta x$, $\psi' \propto \exp(-(x - x_0 - \delta x)^2)$. If δx is small,

$$\begin{aligned} \psi' &\propto (1 + 2\delta x(x - x_0)) \exp(-(x - x_0)^2) \\ &= \left(1 - \delta x \frac{\partial}{\partial x}\right) \exp(-(x - x_0)^2). \end{aligned} \quad (7.77)$$

This operation only represents a symmetry if $[\hat{H}, \hat{S}] = 0$. This in turn can only happen if the potential is zero everywhere in space. Thus the translation operator only represents a symmetry for a free particle. The conserved quantity is, of course, the linear momentum.

Rotations

We consider the situation when the symmetry corresponds to a rotation of coordinates. An ordinary vector \vec{v} is transformed into a new vector by multiplication by a rotation matrix

$$\vec{v}' = R\vec{v} \quad (7.78)$$

where R is a 3×3 orthogonal matrix (i.e. $R^\dagger R = RR^\dagger = \mathbf{1}$). The analogous transformation for a state is operation by a unitary rotation operator $\mathcal{D}(R)$, converting a state $|\alpha\rangle$ into a new state $|\alpha_R\rangle$. We write

$$|\alpha_R\rangle = \mathcal{D}(R) |\alpha\rangle. \quad (7.79)$$

One can in principle consider two kinds of transformations. One can rotate the state, or one can obtain the same final state by rotation of the coordinate axis by the negative of the angle by which we rotated the state. This is called a “passive rotation.” We will deal with “active rotations,” rotations where the state is explicitly rotated, not the coordinate axes.

The three rotation matrices for rotations about the usual Cartesian axes are

$$R_z = \begin{pmatrix} \cos \phi_z & -\sin \phi_z & 0 \\ \sin \phi_z & \cos \phi_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi_x & -\sin \phi_x \\ 0 & \sin \phi_x & \cos \phi_x \end{pmatrix} \quad R_y = \begin{pmatrix} \cos \phi_y & 0 & -\sin \phi_y \\ 0 & 1 & 0 \\ \sin \phi_y & 0 & \cos \phi_y \end{pmatrix}. \quad (7.80)$$

Their infinitesimal equivalents, up to second order, are

$$R_z = \begin{pmatrix} 1 - \frac{\epsilon_z^2}{2} & -\epsilon_z & 0 \\ \epsilon_z & 1 - \frac{\epsilon_z^2}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \frac{\epsilon_x^2}{2} & -\epsilon_x \\ 0 & \epsilon_x & 1 - \frac{\epsilon_x^2}{2} \end{pmatrix} \quad R_y = \begin{pmatrix} 1 - \frac{\epsilon_y^2}{2} & 0 & -\epsilon_y \\ 0 & 1 & 0 \\ \epsilon_y & 0 & 1 - \frac{\epsilon_y^2}{2} \end{pmatrix}. \quad (7.81)$$

By multiplying out the infinitesimal matrices and keeping terms of second order in ϵ (set all the ϵ_i 's equal), we observe that

$$R_x(\epsilon)R_y(\epsilon) - R_y(\epsilon)R_x(\epsilon) = R_z(\epsilon^2) - \mathbf{1}. \quad (7.82)$$

Since angular momentum is the momentum conjugate to an angle, we can define the generator of rotations as the angular momentum operator \hat{J} . Then we can write the rotation operator for infinitesimal rotations as

$$\mathcal{D}(\vec{n}, d\phi) = \hat{1} - \frac{i}{\hbar} \delta\phi (\vec{n} \cdot \hat{\vec{J}}) \quad (7.83)$$

and the operator for finite rotations is then

$$\mathcal{D}(\vec{n}, \phi) = \exp\left(-\frac{i}{\hbar} \phi (\vec{n} \cdot \hat{\vec{J}})\right). \quad (7.84)$$

Since the rotation operators should obey the same relations as their classical analogues (rotations of vectors), the following relations should hold:

$$\mathcal{D}(R)\mathcal{D}(\mathbf{1}) = \mathcal{D}(R)$$

$$\begin{aligned}
\mathcal{D}(R_1)\mathcal{D}(R_2) &= \mathcal{D}(R_3) = \mathcal{D}(R_1R_2) \\
\mathcal{D}(R)\mathcal{D}(R^{-1}) &= \mathcal{D}(\mathbf{1}).
\end{aligned}
\tag{7.85}$$

These relations imply that

$$\begin{aligned}
\mathcal{D}(\mathbf{1}) &= \hat{\mathbf{1}} \\
\mathcal{D}(R^{-1}) &= \mathcal{D}^{-1}(R).
\end{aligned}
\tag{7.86}$$

Return to Eq. 7.82 and consider the operator analogue for this equation. We use the infinitesimal form of the rotation operator

$$\mathcal{D}(R_z(\epsilon)) = 1 - \frac{i}{\hbar} \hat{J}_z \epsilon - \frac{1}{2\hbar^2} \hat{J}_z^2 \epsilon^2 + \dots
\tag{7.87}$$

and require the same relation as for the rotation matrix,

$$[\mathcal{D}(R_x(\epsilon)), \mathcal{D}(R_y(\epsilon))] = \mathcal{D}(R_z(\epsilon^2)) - \hat{\mathbf{1}}.
\tag{7.88}$$

Expanding this to second order in ϵ gives

$$\left[\hat{\mathbf{1}} - \frac{i}{\hbar} \hat{J}_x \epsilon - \frac{1}{2\hbar^2} \hat{J}_x^2 \epsilon^2, \hat{\mathbf{1}} - \frac{i}{\hbar} \hat{J}_y \epsilon - \frac{1}{2\hbar^2} \hat{J}_y^2 \epsilon^2 \right] = \hat{\mathbf{1}} - \frac{i}{\hbar} \hat{J}_z \epsilon^2 - \hat{\mathbf{1}}.
\tag{7.89}$$

The second order terms

$$\hat{\mathbf{1}} - \hat{\mathbf{1}} - \frac{1}{\hbar^2} \epsilon^2 [\hat{J}_x, \hat{J}_y] = \hat{\mathbf{1}} - \hat{\mathbf{1}} - \frac{i}{\hbar} \hat{J}_z \epsilon^2
\tag{7.90}$$

give the familiar commutation relations for the angular momentum operators,

$$[\hat{J}_x, \hat{J}_y] = i\hbar \hat{J}_z.
\tag{7.91}$$

From this, we see that the commutation relations for angular momentum operators arise because of the geometrical properties of rotations.

Let us consider the case $J = 1/2$. The rotation matrix is

$$\mathcal{D}_{mm'}^{\frac{1}{2}}(\phi, \vec{n}) = \left\langle \frac{1}{2}, m \left| \exp\left(-i\frac{1}{2}\vec{\sigma} \cdot \vec{n}\phi\right) \right| \frac{1}{2}, m' \right\rangle.
\tag{7.92}$$

This is identical to the evolution operator for a spin-1/2 particle in a magnetic field, $\hat{H} = \frac{e\hbar}{2mc} \vec{B} \cdot \vec{\sigma}$,

$$U(t) = \exp \left[-\frac{ieB\vec{n} \cdot \vec{\sigma}t}{2mc} \right]
\tag{7.93}$$

with the substitution $\phi/2 \rightarrow eBt/(2mc)$. Explicitly, the rotation matrix is

$$\mathcal{D}_{mm'}^{\frac{1}{2}}(\phi, \vec{n}) = \begin{pmatrix} \cos(\frac{\phi}{2}) - in_z \sin(\frac{\phi}{2}) & (in_x + n_y) \sin(\frac{\phi}{2}) \\ (-in_x + n_y) \sin(\frac{\phi}{2}) & \cos(\frac{\phi}{2}) + in_z \sin(\frac{\phi}{2}) \end{pmatrix}. \quad (7.94)$$

Let us investigate a curious property of the $J = 1/2$ rotation matrix by imagining a rotation about the z axis. The rotation matrix simplifies to

$$\mathcal{D} = \begin{pmatrix} e^{-\frac{i\phi}{2}} & 0 \\ 0 & e^{\frac{i\phi}{2}} \end{pmatrix}. \quad (7.95)$$

A rotation by 360° corresponds to a transformation

$$\mathcal{D} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (7.96)$$

that is, to a change of sign in the wave function. When $J = \frac{1}{2}$ it is necessary to rotate the state by 4π to return the state to its original configuration.

Euler angles

Rotations of rigid bodies are conventionally characterized by three Euler angles (α, β, γ) . The first rotation is about the z -axis by an amount α . The second rotation is about the (new) y' -axis by an amount β . Finally, we rotate by an angle γ amount the (new) z' -axis. The full rotation can be written in terms of rotation matrices,

$$R(\alpha, \beta, \gamma) = R_{z'}(\gamma)R_{y'}(\beta)R_z(\alpha). \quad (7.97)$$

This is a bit too complicated for practical applications, so let us simplify it. We can decompose two of the rotations as follows

$$\begin{aligned} R_{y'}(\beta) &= R_z(\alpha)R_y(\beta)R_x^{-1}(\alpha) \\ R_{z'}(\gamma) &= R_{y'}(\beta)R_z(\gamma)R_{y'}^{-1}(\beta). \end{aligned} \quad (7.98)$$

Thus we have

$$\begin{aligned} R_{z'}(\gamma)R_{y'}(\beta)R_z(\alpha) &= R_{y'}(\beta)R_z(\gamma)R_{y'}^{-1}(\beta)R_{y'}(\beta)R_z(\alpha) \\ &= R_z(\alpha)R_y(\beta)R_z^{-1}(\alpha)R_z(\gamma)R_z(\alpha). \end{aligned} \quad (7.99)$$

Rotations about the same axis commute, so as

$$R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma), \quad (7.100)$$

which involves only rotations about the original axes. Thus the rotation operator is

$$\mathcal{D}(\alpha, \beta, \gamma) = \mathcal{D}_z(\alpha)\mathcal{D}_y(\beta)\mathcal{D}_z(\gamma). \quad (7.101)$$

Let us work in a basis where \hat{J}_z is diagonal, so rotations about the z-axis are simple. In that case it is conventional to define

$$\begin{aligned} \mathcal{D}_{mm'}^j &= \langle j, m' | e^{-\frac{i\hat{J}_z}{\hbar}\alpha} e^{-\frac{i\hat{J}_y}{\hbar}\beta} e^{-\frac{i\hat{J}_z}{\hbar}\gamma} | j, m \rangle \\ &= e^{-i(m'\alpha+m\gamma)} \langle j, m' | e^{-\frac{i\hat{J}_y}{\hbar}\beta} | j, m \rangle \\ &= e^{-i(m'\alpha+m\gamma)} d_{mm'}^j(\beta). \end{aligned} \quad (7.102)$$

As an example, consider again $J = \frac{1}{2}$. The rotation operator is

$$\mathcal{D}(\alpha, \beta, \gamma) = \begin{pmatrix} e^{-\frac{i\alpha}{2}} & 0 \\ 0 & e^{\frac{i\alpha}{2}} \end{pmatrix} \begin{pmatrix} \cos(\frac{\beta}{2}) & -\sin(\frac{\beta}{2}) \\ \sin(\frac{\beta}{2}) & \cos(\frac{\beta}{2}) \end{pmatrix} \begin{pmatrix} e^{-\frac{i\gamma}{2}} & 0 \\ 0 & e^{\frac{i\gamma}{2}} \end{pmatrix} \quad (7.103)$$

or

$$\mathcal{D}(\alpha, \beta, \gamma) = \begin{pmatrix} e^{-\frac{i(\alpha+\gamma)}{2}} \cos(\frac{\beta}{2}) & -e^{-\frac{i(\alpha-\gamma)}{2}} \sin(\frac{\beta}{2}) \\ e^{\frac{i(\alpha-\gamma)}{2}} \sin(\frac{\beta}{2}) & e^{\frac{i(\alpha+\gamma)}{2}} \cos(\frac{\beta}{2}) \end{pmatrix}. \quad (7.104)$$

Considered as an operator acting in the space of all states, the rotation operator is reducible in J . This happens because $[\hat{J}^2, \hat{J}] = 0$ and $[\hat{J}^2, \mathcal{D}] = 0$. As a consequence,

$$\langle j', m' | \mathcal{D} | j, m' \rangle \propto \delta_{jj'}. \quad (7.105)$$

Thus the entire rotation matrix for all values of J is block diagonal

$$\mathcal{D} = \begin{pmatrix} \mathcal{D}^{(j_1)} & & & \\ & \mathcal{D}^{(j_2)} & & \\ & & \mathcal{D}^{(j_3)} & \\ & & & \ddots \end{pmatrix} \quad (7.106)$$

where

$$\mathcal{D}_{m_1, m_2}^{(J)} = \langle J, m_1 | \exp[-\frac{i}{\hbar} \hat{J} \cdot \vec{n} \phi] | J, m_2 \rangle, \quad (7.107)$$

a $(2j+1) \times (2j+1)$ matrix, is called the “ J th irreducible representation of the rotation matrix.”

Relation of Y_l^m 's and \mathcal{D} 's

The spherical harmonics are coordinate space representations of angular momentum eigenkets. The arguments of the spherical harmonic are the values of the unit vector parameterizing the orientation of the coordinates.

$$\langle \vec{n} | lm \rangle = Y_l^m(\vec{n} = (\theta, \phi)). \quad (7.108)$$

Let us exploit this fact to find explicit expressions for a class of rotation operator matrix elements. A ket oriented in the \vec{n} direction can be constructed from one along the z-axis using a rotation matrix,

$$|\vec{n}\rangle = \mathcal{D}(R) |\vec{z}\rangle. \quad (7.109)$$

We project this relation into coordinate space. We can parameterize the rotation with Euler angles $\alpha = \phi, \beta = \theta, \gamma = 0$. Then

$$\begin{aligned} \langle l, m' | \vec{n} \rangle &= \sum_{m''} \langle l, m' | \mathcal{D}(R) | l, m'' \rangle \langle l, m'' | \vec{z} \rangle \\ &= \sum_{m''} \mathcal{D}_{m' m''}^l(\phi, \theta, 0) \langle l, m'' | \vec{z} \rangle. \end{aligned} \quad (7.110)$$

$\langle l, m' | \vec{z} \rangle$ is just the complex conjugate of a spherical harmonic, evaluated at $\theta = 0$. At $\theta = 0$ only the Y_l^0 's are non-vanishing, so the ϕ dependence disappears. The Y_l^0 's are proportional to the Legendre polynomials. At $\theta = 0$ the Legendre polynomials are equal to the identity, so

$$\langle l, m | \vec{z} \rangle = \sqrt{\frac{2l+1}{4\pi}} \delta_{m,0}. \quad (7.111)$$

Thus Eq. 7.110 becomes

$$Y_l^m(\theta, \phi)^* = \sqrt{\frac{2l+1}{4\pi}} \mathcal{D}_{m,0}^{(l)}(\theta, \phi, 0). \quad (7.112)$$

We will use this improbable identity momentarily.

Clebsch-Gordon coefficients, rotation matrices, Y_l^m 's, ...

Do you recall how we could write states involving two independent angular momenta in terms of eigenstates of total angular momentum? We may need to know how these states

transform under rotations. This, of course involves the rotation matrix in a reducible basis. Its expectation value is

$$\langle J_1, J_2, m_1, m_2 | \mathcal{D}(R) | J_1, J_2, m'_1, m'_2 \rangle = \mathcal{D}_{m_1, m'_1}^{J_1} \mathcal{D}_{m_2, m'_2}^{J_2} \quad (7.113)$$

We can also write the reducible state as a superposition of irreducible states, and express the result in terms of a sum of matrix elements involving a single rotation matrix. Not surprisingly, this involves a pair of Clebsch-Gordon coefficients:

$$\begin{aligned} \langle J_1, J_2, m_1, m_2 | \mathcal{D}(R) | J_1, J_2, m'_1, m'_2 \rangle &= \mathcal{D}_{m_1, m'_1}^{J_1} \mathcal{D}_{m_2, m'_2}^{J_2} \\ &= \sum_{J, M, M'} \langle J_1, J_2, m_1, m_2 | J, M, J_1, J_2 \rangle \times \\ &\quad \langle J, M | \mathcal{D}(R) | J, M' \rangle \langle J_1, J_2, m'_1, m'_2 | J, M', J_1, J_2 \rangle \\ &= \sum_{J, M, M'} \langle J_1, J_2, m_1, m_2 | J, M, J_1, J_2 \rangle \times \\ &\quad \langle J_1, J_2, m'_1, m'_2 | J, M', J_1, J_2 \rangle \mathcal{D}_{M M'}^J(R). \end{aligned} \quad (7.114)$$

This is a completely general result. However, suppose that all the J 's are integers. Recalling Eq. 7.112, we learn that Eq. 7.114 can be written as

$$\begin{aligned} Y_{l_1}^{m_1}(\theta, \phi) Y_{l_2}^{m_2}(\theta, \phi) &= \sqrt{\frac{(2l_1 + 1)(2l_2 + 1)}{(4\pi)^2}} \sum_{l', m'} \langle l_1, l_2, m_1, m_2 | l_1, l_2, l', m' \rangle \times \\ &\quad \langle l_1, l_2, 0, 0 | l_1, l_2, l', 0 \rangle \sqrt{\frac{4\pi}{2l' + 1}} Y_{l'}^{m'}(\theta, \phi). \end{aligned} \quad (7.115)$$

Multiply both sides by $Y_{l'}^{m'}(\theta, \phi)^*$ and integrate over the solid angle. By orthogonality, this gives

$$\begin{aligned} \int d\Omega Y_{l'}^{m'}(\theta, \phi)^* Y_{l_1}^{m_1}(\theta, \phi) Y_{l_2}^{m_2}(\theta, \phi) &= \\ \sqrt{\frac{(2l_1 + 1)(2l_2 + 1)}{4\pi(2l' + 1)}} \langle l_1, l_2, m_1, m_2 | l_1, l_2, l', m' \rangle \langle l_1, l_2, 0, 0 | l_1, l_2, l', 0 \rangle. \end{aligned} \quad (7.116)$$

This certainly looks horrible! On the contrary, it is a very beautiful (and useful) result. Matrix elements like the left hand side commonly arise in practical calculations, where the initial and final state wave functions are eigenfunctions (or superpositions of eigenfunctions)

of orbital angular momentum – spherical harmonics. If the object whose matrix element is desired can also be written as a superposition of spherical harmonics, we have the integral on the left side of Eq. 7.116 to evaluate. Eq. 7.115 makes this evaluation trivial. The point of course is that the reducible state (a product of two spherical harmonics) can be written as a superposition of irreducible states (a single spherical harmonic, in coordinate representation). The result simply involves a product of Clebsch-Gordon coefficients. These are tabulated. (The reader is encouraged to perform some of these integrals without using this result, to appreciate its utility.)

Notice that the m -dependence is contained entirely in one Clebsch-Gordon coefficient. This immediately gives a selection rule: the integral is zero unless $m' = m_1 + m_2$. (This is just the z -component of angular momentum conservation, between the initial state, operator, and final state.) The remainder of the expression (which is often called the “reduced matrix element”) is independent of m . To look ahead, the integral is a special case of the Wigner-Eckhart Theorem, which we will derive in a few pages.

Note that we can use of the fact that the Clebsch-Gordon coefficients are real numbers, to write expressions more symmetrically without regard to order.

Vector and tensor operators

We have been using vector operators like $\hat{\mathbf{L}}, \hat{\mathbf{S}}, \hat{\mathbf{r}}, \dots$ for some time, without thinking very deeply about their vector nature. Classical vector quantities transform under rotations as

$$v_i = \sum_j R_{ij} v_j \quad (7.117)$$

where R is the rotation matrix. In quantum mechanics, a state transforms as $|\alpha\rangle \rightarrow \mathcal{D}(R)|\alpha\rangle$. The components of the expectation values of a vector operator are also c-numbers, and they will transform as a vector under rotations:

$$\langle\alpha|\hat{v}_i|\alpha\rangle_R = \sum_j R_{ij} \langle\alpha|\hat{v}_j|\alpha\rangle = \langle\alpha|\sum_j R_{ij}\hat{v}_j|\alpha\rangle. \quad (7.118)$$

In quantum mechanics, a state transforms as $|\alpha\rangle \rightarrow \mathcal{D}(R)|\alpha\rangle$. We could equally well evaluate the rotated matrix elements by rotating the states,

$$\langle\alpha|\hat{v}_i|\alpha\rangle_R = \langle\alpha|\mathcal{D}^\dagger(R)\hat{v}_i\mathcal{D}(R)|\alpha\rangle. \quad (7.119)$$

Equating these expressions tells us that a vector operator is an operator that transform under rotations as

$$\mathcal{D}^\dagger(R)\hat{v}_i\mathcal{D}(R) = \sum_j R_{ij}\hat{v}_j. \quad (7.120)$$

In the case of an infinitesimal rotation, where

$$\mathcal{D}(R) = \hat{1} - \frac{i}{\hbar}\epsilon\hat{\mathbf{J}} \cdot \hat{\mathbf{n}}, \quad (7.121)$$

Eq. 7.120 demands that

$$\hat{v}_i - \frac{i\epsilon}{\hbar}[\hat{v}_i, \hat{\mathbf{J}} \cdot \hat{\mathbf{n}}] = \sum_j R_{ij}\hat{v}_j \quad (7.122)$$

To interpret this expression, assume that $\hat{\mathbf{n}}$ points along the z-axis. The rotation matrix is

$$R \rightarrow \begin{pmatrix} 1 & -\epsilon & 0 \\ \epsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7.123)$$

and the vector transforms as

$$\begin{aligned} \hat{v}_x - \frac{i\epsilon}{\hbar}[\hat{v}_x, \hat{J}_z] &= \hat{v}_x - \epsilon\hat{v}_y \\ \hat{v}_y - \frac{i\epsilon}{\hbar}[\hat{v}_y, \hat{J}_z] &= \hat{v}_y + \epsilon\hat{v}_x \\ \hat{v}_z &= \hat{v}_z. \end{aligned} \quad (7.124)$$

This is a special case of the general requirement for vector operators that

$$[\hat{v}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{v}_k. \quad (7.125)$$

We can turn the discussion around: any operator which obeys this relation is, by definition, a vector operator.

The generalization of a vector is a tensor, specifically, a Cartesian tensor of rank N . It transforms under rotations as

$$T_{ijk\dots N} \rightarrow \sum_{i'j'k'\dots N'} R_{ii'}R_{jj'}R_{kk'}\cdots R_{NN'}T_{i'j'k'\dots N'}. \quad (7.126)$$

Cartesian tensors transform reducibly under rotations, and so are not the most convenient objects for study. To see how to proceed, consider the simplest example of a rank $N = 2$ Cartesian tensor, a dyadic. It is defined as a direct product of two vector operators

$$T_{ij} = u_iv_j. \quad (7.127)$$

This is a collection of 9 numbers, the T_{ij} 's. We can decompose it as follows:

$$u_i v_j = \frac{1}{3} \delta_{ij} (\mathbf{u} \cdot \mathbf{v}) + \frac{1}{2} (u_i v_j - u_j v_i) + \left(\frac{u_i v_j + u_j v_i}{2} - \frac{1}{3} \delta_{ij} (\mathbf{u} \cdot \mathbf{v}) \right). \quad (7.128)$$

The first term in the decomposition is a scalar under rotations, the second is an axial vector (a cross product) and the last term is a symmetric traceless tensor. This decomposition preserves the number of components ($3 \times 3 = 9$) in the the dyadic. There is one scalar term, three axial vector terms, and five terms from the traceless, symmetric tensor.

“Spherical tensors” transform simply (and irreducibly) under rotations. Their elements are linear superpositions of Cartesian tensor elements. A simple way to define a spherical tensor from a Cartesian vector \mathbf{v} is via the replacement

$$T_q^{(k)} = Y_{l=k}^{m=q}(\mathbf{v}). \quad (7.129)$$

The rank-1 spherical tensor (a vector) is very useful. To build it, we make the substitutions $z/r = v_z$, $y/r = v_y$ and $x/r = v_x$, to give

$$\begin{aligned} Y_1^0 &= \sqrt{\frac{3}{4\pi}} \cos(\theta) = \sqrt{\frac{3}{4\pi}} \frac{z}{r} \rightarrow T_0^{(1)} = \sqrt{\frac{3}{4\pi}} \hat{v}_z \\ Y_1^{\pm 1} &= \mp \sqrt{\frac{3}{4\pi}} \left(\frac{x \pm iy}{\sqrt{2}r} \right) \rightarrow T_{\pm 1}^{(1)} = \mp \sqrt{\frac{3}{4\pi}} \left(\frac{\hat{v}_x \pm i\hat{v}_y}{\sqrt{2}} \right) \\ Y_2^{\pm 2} &= \sqrt{\frac{15}{32\pi}} \left(\frac{x \pm iy}{r} \right)^2 \rightarrow T_{\pm 2}^{(2)} = \sqrt{\frac{15}{32\pi}} (\hat{v}_x \pm i\hat{v}_y)^2 \end{aligned} \quad (7.130)$$

and so on.

How do spherical tensors transform under rotations? We know that the spherical harmonics can be written as $Y_l^m(\hat{\mathbf{n}}) = \langle \hat{\mathbf{n}} | l, m \rangle$. Furthermore, we recall the transformation law $|\hat{\mathbf{n}}'\rangle = \mathcal{D}(R) |\hat{\mathbf{n}}\rangle$. We also know that

$$\mathcal{D}^\dagger(R) |l, m'\rangle = \sum_{m''} \mathcal{D}_{mm''}^{(l)}(R)^* |l, m''\rangle. \quad (7.131)$$

Taking the inner product of both sides with $\langle \hat{\mathbf{n}} |$ gives

$$\langle \hat{\mathbf{n}} | \mathcal{D}^\dagger(R) |l, m'\rangle = \langle \hat{\mathbf{n}}' | l, m'\rangle = Y_l^{m'}(\hat{\mathbf{n}}') = \sum_{m''} \mathcal{D}_{mm''}^{(l)}(R)^* Y_l^{m''}(\hat{\mathbf{n}}). \quad (7.132)$$

This is the transformation law for states. Noting that $\mathcal{D}_{mm'}^{(l)}(R)^* = \mathcal{D}_{mm'}^{(l)}(R^{-1})$. We expect that operators ought to transform as

$$\mathcal{D}^\dagger(R) \hat{T}_q^{(k)} \mathcal{D}(R) = \sum_{q'} T_{q'}^{(k)} \mathcal{D}_{mm'}^{(k)}(R^{-1}) \quad (7.133)$$

and this is the definition for the transformation property of a spherical tensor operator. Alternatively, we can consider the defining relation for a spherical tensor operator through commutators of the operator with \hat{J} . These are

$$[\hat{J}_z, \hat{T}_q^{(k)}] = \hbar q \hat{T}_q^{(k)} \quad (7.134)$$

and

$$[\hat{J}_\pm, \hat{T}_q^{(k)}] = \hbar \sqrt{(k \mp q)(k \pm q + 1)} T_{q\pm 1}^{(k)}. \quad (7.135)$$

Combining tensor operators

As we saw in the last section, we can write the components of a vector operator \mathbf{u} in spherical basis

$$u_0 = u_z \quad u_{\pm 1} = \mp \left(\frac{u_x \pm i u_y}{\sqrt{2}} \right). \quad (7.136)$$

Let us assume that we have two vector operators \mathbf{u} and \mathbf{v} . A few special cases of their combination written as tensor operators are

$$T_0^{(0)} = -\frac{1}{\sqrt{3}} \mathbf{u} \cdot \mathbf{v} \quad (7.137)$$

$$T_q^{(1)} = \frac{i}{\sqrt{2}} (\mathbf{u} \times \mathbf{v})_q \quad T_{\pm 2}^{(2)} = u_{\pm 1} v_{\pm 1}. \quad (7.138)$$

In general, we combine spherical tensor operators just like we combined angular momenta, using Clebsch-Gordon coefficients:

$$\hat{T}_q^{(k)} = \sum_{q_1 q_2} \langle k_1, k_2, q_1, q_2 | k_1, k_2, k, q \rangle X_{q_1}^{(k_1)} Z_{q_2}^{(k_2)} \quad (7.139)$$

Let us examine how this transforms under a rotation.

$$\mathcal{D}^\dagger(R) \hat{T}_q^{(k)} \mathcal{D}(R) = \sum_{q_1 q_2} \langle k_1, k_2, q_1, q_2 | k_1, k_2, k, q \rangle \mathcal{D}^\dagger(R) X_{q_1}^{(k_1)} \mathcal{D}(R) \mathcal{D}^\dagger(R) Z_{q_2}^{(k_2)} \mathcal{D}(R) \quad (7.140)$$

or

$$\mathcal{D}^\dagger(R) \hat{T}_q^{(k)} \mathcal{D}(R) = \sum_{q_1 q_2} \langle k_1, k_2, q_1, q_2 | k_1, k_2, k, q \rangle \sum_{q'_1} X_{q'_1}^{(k_1)} \mathcal{D}_{q'_1 q_1}^{(k_1)}(R^{-1}) \sum_{q'_2} Z_{q'_2}^{(k_2)} \mathcal{D}_{q'_2 q_2}^{(k_2)}(R^{-1}). \quad (7.141)$$

We can also write

$$\mathcal{D}_{q'_1 q_1}^{(k_1)}(R^{-1}) \mathcal{D}_{q'_2 q_2}^{(k_2)}(R^{-1}) = \sum_{k'' q'' q'} \langle k_1, k_2, q_1, q_2 | k_1, k_2, k'', q' \rangle \langle k_1, k_2, q_1, q_2 | k_1, k_2, k'', q'' \rangle \mathcal{D}_{q' q''}^{(k'')}(R^{-1}). \quad (7.142)$$

Equating all of these long expressions, we have

$$\mathcal{D}^\dagger(R)\hat{T}_q^{(k)}\mathcal{D}(R) = \sum_{k''} \sum_{q_1 q_2} \sum_{q'_1} \sum_{q'_2} \sum_{q' q''} \langle k_1, k_2, q_1, q_2 | k_1, k_2, k, q \rangle \times \quad (7.143)$$

$$\langle k_1, k_2, q_1, q_2 | k_1, k_2, k'', q' \rangle \langle k_1, k_2, q_1, q_2 | k_1, k_2, k'', q'' \rangle \mathcal{D}_{q' q''}^{(k'')}(R^{-1}) X_{q'_1}^{(k_1)} Z_{q'_2}^{(k_2)}.$$

Recall that Clebsch-Gordon coefficients are real. Performing the sum over $q_1 q_2$ will yield the identity operator, which will introduce a $\delta_{k k''}$ and $\delta_{q q''}$. They collapse the expression to the more manageable

$$\mathcal{D}^\dagger(R)\hat{T}_q^{(k)}\mathcal{D}(R) = \sum_{q'} \left(\sum_{q'_1 q'_2} \langle k_1, k_2, q'_1, q'_2 | k_1, k_2, k, q' \rangle X_{q'_1}^{(k_1)} Z_{q'_2}^{(k_2)} \right) \mathcal{D}_{q' q}^{(k)}(R^{-1}) \quad (7.144)$$

Since the object in the large parentheses is $T_{q'}^{(k)}$ from Eq. 7.139, this is our defining transformation for a tensor operator, Eq. 7.133.

Let us consider three useful results for matrix elements of tensor operators, written as $\langle \alpha', j', m' | \hat{T}_q^{(k)} | \alpha, j, m \rangle$. (The alphas stand for the non-angular momentum quantum numbers.) The first is the m-selection rule,

$$\langle \alpha', j', m' | \hat{T}_q^{(k)} | \alpha, j, m \rangle = \delta_{m', q+m} \quad (7.145)$$

The matrix element of the commutator of \hat{J}_z and $\hat{T}_q^{(k)}$

$$[\hat{J}_z, \hat{T}_q^{(k)}] - \hbar q \hat{T}_q^{(k)} = 0 \quad (7.146)$$

is

$$\langle \alpha', j', m' | [\hat{J}_z, \hat{T}_q^{(k)}] - \hbar q \hat{T}_q^{(k)} | \alpha, j, m \rangle = \hbar(m' - m - q) \langle \alpha', j', m' | \hat{T}_q^{(k)} | \alpha, j, m \rangle = 0. \quad (7.147)$$

Either $m' = m + q$ or the matrix element itself must vanish.

Next is the Wigner-Eckhart Theorem:

$$\langle \alpha', j', m' | \hat{T}_q^{(k)} | \alpha, j, m \rangle = \langle j, k, m, q | j, k, j', m' \rangle \frac{\langle \alpha', j' || T^{(k)} || \alpha, j \rangle}{\sqrt{2j+1}} \quad (7.148)$$

The quantity $\langle \alpha', j' || T^{(k)} || \alpha, j \rangle$ is called the “reduced matrix element.” It is fully independent of m, m' , and q . Thus all of the geometrical dependence is contained in the Clebsch-Gordon coefficients. From the m-selection rule we know that the only non-vanishing elements occur when $|j - k| \leq j' \leq j + k$.

To prove the Wigner-Eckhart theorem, we return to the relation

$$[\hat{J}_\pm, \hat{T}_q^{(k)}] - \hbar\sqrt{(k \mp q)(k \pm q + 1)}T_{q\pm 1}^{(k)} = 0 \quad (7.149)$$

We evaluate the equality between angular momentum states:

$$\langle \alpha', j', m' | [\hat{J}_\pm, \hat{T}_q^{(k)}] - \hbar\sqrt{(k \mp q)(k \pm q + 1)}T_{q\pm 1}^{(k)} | \alpha, j, m \rangle = 0. \quad (7.150)$$

Letting the angular momentum operators act on the states, we have

$$\begin{aligned} \sqrt{(j' \pm m')(j' \mp m' + 1)} \langle \alpha', j', m' + 1 | \hat{T}_q^{(k)} | \alpha, j, m \rangle = \\ \sqrt{(k \mp q)(k \pm q + 1)} \langle \alpha', j', m' | T_{q\pm 1}^{(k)} | \alpha, j, m \rangle + \\ \sqrt{(j \mp m)(j \pm m + 1)} \langle \alpha', j', m' | \hat{T}_q^{(k)} | \alpha, j, m + 1 \rangle. \end{aligned} \quad (7.151)$$

The coefficients of this equation are identical to the ones relating the Clebsch-Gordon coefficients themselves; there is only a translation of variables, given in Table 7. Because there two

CG		WE
J	\leftrightarrow	J'
m	\leftrightarrow	m'
J_1	\leftrightarrow	J
m_1	\leftrightarrow	m
J_2	\leftrightarrow	k
m_2	\leftrightarrow	q

sums of three terms with the same coefficients are equal to zero, they must be proportional to each other. This means that

$$\langle J_1, J_2, m_1, m_2 \pm 1 | J_1, J_2, J, M \rangle = \text{const} \times \langle \alpha', j', m' | T_{q\pm 1}^{(k)} | \alpha, j, m \rangle \quad (7.152)$$

or

$$\langle \alpha', j', m' | \hat{T}_q^{(k)} | \alpha, j, m \rangle \equiv C \langle j, k, m, q | j, k, j', m' \rangle \quad (7.153)$$

where

$$C = \frac{\langle \alpha', j' | | T^{(k)} | | \alpha, j \rangle}{\sqrt{2j + 1}}. \quad (7.154)$$

This theorem is a very useful tool for the evaluation of matrix elements and for finding selection rules. Let us consider a few examples.

First, consider matrix elements of a scalar operator (tensor of rank 0). They are

$$\langle \alpha', j', m' | \hat{S} | \alpha, j, m \rangle = \delta_{jj'} \delta_{mm'} \frac{\langle \alpha', j | |T^{(k)}| | \alpha, j \rangle}{\sqrt{2j+1}} \quad (7.155)$$

Scalar operators have no directional dependence and thus cannot change the angular momentum of a state.

The projection theorem states that matrix elements of a vector operator can be written as

$$\langle \alpha', j, m' | \hat{V}_q | \alpha, j, m \rangle = \frac{\langle \alpha', j, m' | \hat{\mathbf{J}} \cdot \hat{\mathbf{V}} | \alpha, j, m \rangle}{\hbar^2 J(J+1)} \langle j, m' | \hat{J}_q | j, m \rangle. \quad (7.156)$$

This follows from the Wigner-Eckhart theorem, because matrix elements of \hat{V}_q and \hat{J}_q can be written as

$$\begin{aligned} \langle \hat{V}_q \rangle &= C_1 \langle ||V|| \rangle \\ \langle \hat{J}_q \rangle &= C_1 \langle ||J|| \rangle \end{aligned} \quad (7.157)$$

where C_1 is a shorthand for all except the reduced matrix element. Two other related matrix elements are

$$\begin{aligned} \langle \hat{\mathbf{J}} \cdot \hat{\mathbf{V}} \rangle &= \langle \hat{J}_0 \hat{V}_0 - \hat{J}_1 \hat{V}_{-1} - \hat{J}_{-1} \hat{V}_1 \rangle \\ &= C_2 \langle ||V|| \rangle \\ \langle \hat{\mathbf{J}} \cdot \hat{\mathbf{J}} \rangle &= C_2 \langle ||J|| \rangle = \hbar^2 J(J+1) \end{aligned} \quad (7.158)$$

where C_2 is another linear combination of coefficients. Combining these results yields

$$\langle \hat{V}_q \rangle = \langle \hat{J}_q \rangle \frac{\langle ||V|| \rangle}{\langle ||J|| \rangle} = \langle \hat{J}_q \rangle \frac{\langle \hat{\mathbf{J}} \cdot \hat{\mathbf{V}} \rangle}{\langle \hat{\mathbf{J}} \cdot \hat{\mathbf{J}} \rangle} = \langle \hat{J}_q \rangle \frac{\langle \hat{\mathbf{J}} \cdot \hat{\mathbf{V}} \rangle}{\hbar^2 J(J+1)} \quad (7.159)$$

which is Eq. 7.156 when written less schematically.

The projection theorem's most famous application is for the Zeeman effect in light atoms. In a small magnetic field, the energy levels of the atom split due to the interaction of the electrons' magnetic moments with the field,

$$\hat{H}_B = \frac{e\vec{B}}{2mc} \cdot (\mathbf{L} + 2\mathbf{S}) \quad (7.160)$$

where \mathbf{L} and \mathbf{S} are the total orbital and spin angular momenta of the valence electrons. This splitting is called the Zeeman effect. The projection theorem gives us the splitting. (We defer discussions about why \mathbf{L} and \mathbf{S} are good quantum numbers for multi-electron light atoms, and why diagonal matrix elements of \hat{H}_B give the energy shifts.) Choose the axis of quantization along \vec{B} and define this as the \hat{z} axis. We rewrite $L_z + 2S_z = J_z + S_z$ and consider the matrix element

$$\langle L, S, J, M | J + S | L, S, J, M \rangle = \frac{\langle L, S, J, M | \hat{J}^2 + \hat{\mathbf{J}} \cdot \hat{\mathbf{S}} | L, S, J, M \rangle}{\hbar^2 J(J+1)} \hbar M \equiv g_J \hbar M \quad (7.161)$$

The quantity g_J is called the Lande g -factor. $2\hat{\mathbf{J}} \cdot \hat{\mathbf{S}} = -(\hat{\mathbf{J}} - \hat{\mathbf{S}})^2 + \hat{\mathbf{J}}^2 + \hat{\mathbf{S}}^2 = -\hat{\mathbf{L}}^2 + \hat{\mathbf{J}}^2 + \hat{\mathbf{S}}^2$, and the g -factor is

$$g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}. \quad (7.162)$$

The magnitude of the Zeeman splitting reveals information about the quantum numbers of the state.

Our third example involves the electric quadrupole moment of the deuteron, and illustrates the strengths and weaknesses of the Wigner-Eckart theorem. The deuteron is a bound state of a proton and a neutron. It is not a typical nucleon – it is one of only a few stable nuclei with an odd number of protons and an odd number of neutrons. It is weakly bound ($E_b \sim 2.2$ MeV, about 1/10 of typical nuclear binding energies). It was discovered by Urey in 1931, before the discovery of the neutron. It has a nonzero magnetic moment and a nonzero electric quadrupole moment.

Nuclear physics is complicated! The neutron and proton are nearly degenerate in mass, both have spin $\frac{1}{2}$, and have their own spin magnetic moments. However, they are not structureless particles like the electron – they themselves are bound states of quarks. The nuclear potential is short range, but it is not simple, either (unlike the Coulomb potential).

It's worth pausing to describe the phenomenology of magnetic moments. The orbital magnetic moment is simple: from

$$\hat{H}_B = \frac{e\vec{B}}{2mc} \cdot \mathbf{L}, \quad (7.163)$$

$\vec{\mu} = e/(2mc)\vec{L}$ (or $\mu = e/(2mc)$ or $e\hbar/(2mc)$ for the Bohr magneton). Now spin: for an electron, it is conventional to introduce the g -factor and write

$$\vec{\mu} = g \frac{e}{2mc} \vec{S}. \quad (7.164)$$

For the structureless electron described by the Dirac equation, $g = 2$. The value was known before Dirac, from analysis of the spectroscopy of light atoms. It's a rather confusing story,

with several competing factors of two – $S = \frac{1}{2}$, $g = 2$, and Thomas precession, so $g = 2$ was immediately realized to be a major success for Dirac. The real g-factor is a bit more complicated –

$$g = 2\left[1 + \frac{\alpha}{\pi} + \dots\right] \quad (7.165)$$

where the extra terms come from processes where the electron emits and absorbs virtual photons. You need quantum field theory to describe what is going on.

Back to the proton and the neutron. The literature is presented in several slightly inconsistent notations, but the bottom line is that $g \neq 2$. If we write $\vec{\mu}_i^{(s)} = g_i e / (2Mc) \vec{S}$, then $g_p^{(s)} = 2 \times 2.79$ and $g_n^{(s)} = -2 \times 2.79 \times 0.68$. (particle physicists remember the -0.68 as $-2/3$). These values were completely unexpected when Otto Stern measured the proton's magnetic moment in the 1930's. (The theorists all predicted $g = 2$ because of Dirac.)

Now for nuclei: it is conventional to define the magnetic moment of a nucleus with total angular momentum J as

$$\vec{\mu} = g \frac{e}{2m_p} \vec{J} \quad (7.166)$$

setting the scale with the proton mass.

And finally, the deuteron. The deuteron has $J = 1$ (from observing hyperfine splitting in deuterium), and $S = 1$ (as a thought experiment, scatter a neutron off a proton: the total spin is $\frac{1}{2} + \frac{1}{2} = 1,0$ and the deuteron is a nearby resonance in the $S = 1$ channel). It's reasonable to expect that $L = 0$ just because our intuition suggests that the lowest bound states are S-waves. However, the nonzero quadrupole moment (measured by Rabi) tells us that there is a small admixture of $L = 2$ in the wave function.

We take the magnetic moment from the magnetic energy ($H = -\vec{\mu} \cdot \vec{B}$) in the $M = J$ state:

$$\mu \equiv \langle LSJ, M = J | \mu_z | LSJ, M = J \rangle. \quad (7.167)$$

The projection theorem turns this into

$$\begin{aligned} \mu &= \frac{\langle LSJ, M = J | \vec{\mu} \cdot \vec{J} | LSJ, M = J \rangle}{\hbar^2 J(J+1)} \langle J, M = J | J_z | J, M = J \rangle \\ &= \frac{\langle LSJ, M = J | \vec{\mu} \cdot \vec{J} | LSJ, M = J \rangle}{j+1}. \end{aligned} \quad (7.168)$$

We need several ingredients to evaluate this formula. First,

$$\vec{\mu} = g^{(l)}\vec{L} + g^{(s)}\vec{S}. \quad (7.169)$$

Next, $g_p^{(l)} = 1$ and $g_n^{(l)} = 0$ (the neutron is neutral!). And we have $g_p^{(s)}$ and $g_n^{(s)}$, given above. There are a couple of two's associated with the conversion of individual spins to total spins. For \vec{S} , consider the $|11\rangle$ state, which is $|\uparrow\uparrow\rangle$. The expectation value of the magnetic moment in this state is

$$\langle\mu_z\rangle = \frac{e\hbar}{2Mc}(g_p^{(s)} + g_n^{(s)}) = \frac{e}{Mc} \frac{(g_p^{(s)} + g_n^{(s)})}{2} S_z. \quad (7.170)$$

There is an identical factor of two for the orbital angular momentum, using $|22\rangle = |11\rangle|11\rangle$. Putting the pieces together, we have

$$\mu = \frac{1}{2(j+1)}((g_p^{(s)} + g_n^{(s)})\vec{S} \cdot \vec{J} + \vec{L} \cdot \vec{J}) \frac{e}{2Mc}. \quad (7.171)$$

And, as we saw before, $\vec{L}^2 = (\vec{J} - \vec{S})^2$ and $\vec{S}^2 = (\vec{J} - \vec{L})^2$, so $\vec{J} \cdot \vec{S} = (J^2 + S^2 - L^2)/2$ and $\vec{J} \cdot \vec{L} = (J^2 - S^2 + L^2)/2$. this gives

$$\begin{aligned} \mu = & \frac{e}{2Mc} \frac{1}{4(j+1)} [(g_p^{(s)} + g_n^{(s)})(j(j+1) - l(l+1) + s(s+1)) \\ & + (j(j+1) - l(l+1) + s(s+1))]. \end{aligned} \quad (7.172)$$

We can now consider two cases: First, $S = 1, L = 0, J = 1$:

$$\mu = \frac{g_p^{(s)} + g_n^{(s)}}{2} \frac{e}{2Mc} \quad (7.173)$$

or $g = 0.879$. The second case is $S = 1, L = 2, J = 1$, and

$$g = \frac{3}{4} - \frac{1}{4}(g_p^{(s)} + g_n^{(s)}) \sim 0.3. \quad (7.174)$$

The observed value of g is 0.857, so the deuteron is mostly $S = 1, L = 0, J = 2$. The mismatch between 0.857 and 0.879 tells us that the state is not completely $L = 0$.

Now we return to the quadrupole moment. It is

$$Q = e(3z^2 - r^2). \quad (7.175)$$

It is one component of the rank-2 quadrupole tensor Q_{ij} . Its expectation value in the $J = 1$ deuteron state is (keeping only the labels J and m)

$$\begin{aligned} \langle Q \rangle &= \langle 1, m | T_0^{(2)} | 1, m \rangle \\ &= \frac{e \langle 1, m, 2, 0 | 1, 2, 1, m \rangle}{\sqrt{3}} \|Q\|. \end{aligned} \quad (7.176)$$

We can determine Q , or to be more precise, the reduced matrix element, by doing the full calculation any convenient combination of states (m) and component of Q_{ij} and dividing out the known Clebsch-Gordon coefficient and other factors.

We need the Clebsch - Gordon coefficients for $1 + 2 = 1$. Tables give us

$$\begin{aligned}\langle 1120|1211\rangle &= \sqrt{1/10} \\ \langle 1020|1210\rangle &= -\sqrt{2/15} \\ \langle 1-120|121-1\rangle &= \sqrt{1/10}.\end{aligned}\tag{7.177}$$

What is $\langle 1||T||1\rangle$? $T_0^{(2)} = r^2(3\cos^2\theta - 1)$. Evaluate $\langle Q\rangle$ in the $|10\rangle$ state:

$$-\sqrt{\frac{2}{15}} \frac{\langle 1||T||1\rangle}{\sqrt{3}} = \int d^3r \psi_{j=1,m=0}^* r^2 (3\cos^2\theta - 1) \psi_{1,0}.\tag{7.178}$$

Notice that if the deuteron were pure $l = 0$, Q would vanish,

$$\langle 1||T||1\rangle \sim \int d\cos\theta P_2(\cos\theta) = 0.\tag{7.179}$$

But Q is not zero (it is 0.28 e-fm² where one fm is 10⁻¹³ cm). One fm is about the size of a proton, so Q is a small quantity.

If we know that $S = 1$ and $J = 1$, then we could in principle have $L = 0, 1$ or 2 . If the deuteron is a parity eigenstate, then either it is positive parity ($L = 0$ or 2 , S or D - 3S_1 or 3D_1 in the standard ${}^{2S+1}L_J$ notation). If it is negative parity, it is 3P_1 . It seems unlikely to have a P-wave bound state and no S-wave one, which favors a mixture of S and D. However, that is the end of the story we can tell. All we can do to go further is to parameterize the wave function as a superposition of S and D wave,

$$\psi = \cos\omega f_0(r) |00\rangle + \sin\omega f_2(r) |20\rangle,\tag{7.180}$$

with a model for the potential and the two radial wave functions, and then use Q to determine ω .

A little group theory

It's probably useful to conclude this chapter with a brief discussion (little more than a dictionary) of aspects of group theory which have found their way into quantum mechanics.

A group is a set of elements G with a multiplication rule connecting the elements of the group to each other. Group properties include

- Closure: For any $a, b \in G$, $ab = c \in G$
- Associativity: $a(bc) = (ab)c$
- An identity element: $a \cdot 1 = 1 \cdot a$
- An inverse: $aa^{-1} = a^{-1}a = 1$

A representation of a group is a set of square non-singular matrices M_1, M_2, \dots , one M for each group element g , which multiply like the group elements. If the matrices are $n \times n$ then we speak of an n -dimensional representation of the group. If the matrices corresponding to different group elements g_i are different, the representation is called “faithful” and the group is isomorphic to the matrix group. Otherwise we have an “unfaithful representation” with a “homeomorphic” relation of the group to the set of matrices.

In physics we typically encounter two types of groups

- point groups, which consist of discrete sets of elements
- Lie groups, which are continuously connected groups whose parameters of products of elements are continuous, differentiable functions of the products of the elements. An example of a Lie group is the rotation group.

There are two types of Lie groups as well, compact and non-compact groups. A compact group is one in which the parameters of the group have a finite range of possible values (the rotation group is an example). In contrast, non-compact groups do not have a finite range for their parameters. An example of a non-compact group is the part of special-relativistic transformations corresponding to boosts.

In physics we often speak of “symmetry groups.” By that, we mean a set of transformations which leave the system unchanged. For example, rotations, which leave the length of vectors unchanged, form a compact symmetry group. The symmetry is the transformation $\mathbf{x} \rightarrow \mathbf{x}'$ which leaves $|\mathbf{x}|^2 = |\mathbf{x}'|^2$. The transformation is, of course,

$$\mathbf{x}' = R\mathbf{x} \tag{7.181}$$

where $|\mathbf{x}'|^2 = \mathbf{x}^T R^\dagger R \mathbf{x} = |\mathbf{x}|^2$. The vectors \mathbf{x} are assumed to be 3-dimensional, and so the matrices R must be 3×3 symmetric matrices with $R^\dagger R = 1$ in order to preserve the norm. These matrices have $9 - 6 = 3$ free parameters (since there are 6 constraints imposed by $R^\dagger R = 1$). Hence three parameters characterize R . (They could be the three Euler angles, or the angles of rotations about three axes.)

The maximum number of independent parameters which label the group parameters is referred to as the dimension of the Lie algebra, and mathematicians call this the dimension of the adjoint representation of the group. The rotation group is called $O(3)$ – the group of orthogonal transformations in three dimensions. (Clearly, one can repeat the story in N dimensions, and arrive at the group of transformations $O(N)$.)

As a second set of transformations, consider the matrices which rotate spinor states

$$U = e^{-i\hat{\sigma} \cdot \hat{\mathbf{n}} \phi / 2} \quad U |\psi\rangle \rightarrow |\psi'\rangle. \quad (7.182)$$

The matrix U must be unitary to conserve probability. U is also unimodular (i.e. the determinant must be equal to 1). The representation is two-dimensional because there are two possibilities for the spin. The most general unitary unimodular matrix can be written as

$$U(a, b) = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}. \quad (7.183)$$

where a and b are complex and the unimodularity condition gives the relation $|a|^2 + |b|^2 = 1$. Because a and b are complex, four real numbers label the transformations. Unimodularity gives one constraint, leaving three free parameters. Of course, we knew that already, from constructing the $j = 1/2$ rotation matrix: we count one angle and two (of three constrained) components of a unit vector, recalling

$$\begin{aligned} \operatorname{Re}(a) &= \cos\left(\frac{\phi}{2}\right) \\ \operatorname{Im}(a) &= -n_z \sin\left(\frac{\phi}{2}\right) \\ \operatorname{Re}(b) &= -n_y \sin\left(\frac{\phi}{2}\right) \\ \operatorname{Im}(b) &= -n_x \sin\left(\frac{\phi}{2}\right). \end{aligned} \quad (7.184)$$

(Because $\sum_i n_i^2 = 1$, only two of the n_i 's are independent.) These parameters are called the ‘‘Cayley-Klein’’ parameters and are used in classical mechanics to study rigid body rotations. The product of the U 's is itself a U , and

$$U^{-1}(a, b) = U(a^*, -b). \quad (7.185)$$

The matrices constituting the U 's form a group called $SU(2)$, the “special unitary group of two dimensions.”

Although $O(3)$ and $SU(2)$ both characterize rotations, the groups $SU(2)$ and $O(3)$ are not isomorphic groups, that is, there is no one-to-one correspondence between the group elements. We can see that by comparing rotations by 2π and 4π . In $O(3)$ both rotations are 3×3 identity matrices. In $SU(2)$ one is $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ and the other is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. In fact, $U(a, b)$ and $U(-a, -b)$ correspond to the same 3×3 rotation in $O(3)$: for a given R , U is double-valued.

If one did not require that $\det U = 1$, one would just have a “unitary group,” $U(2)$. The group of unitary transformations in one dimension, $U(1)$, is

$$U = e^{i\phi}. \quad (7.186)$$

This is the group of gauge transformations of electrodynamics. It is an “internal symmetry” group, in that the dynamical variables contain information which has nothing direct to do with space or time. (Of course, ϕ could vary with x and t , if we are considering a local gauge transformation.) One could also consider special unitary groups of rank n . The group $SU(3)$ has been particularly important in particle physics. For example, the theory of the strong interactions, Quantum Chromodynamics, declares that quarks come in three “colors,” or that a fermion spinor carries a three-fold internal index (red, green, blue?). This symmetry is an $SU(3)$ gauge symmetry; the transformation converts one color into a mixture of other ones.

Another example of an internal symmetry could apply to a magnetic material. We could imagine treating the spins, or locally averaged magnetic moments, as classical variables which could rotate in three dimensions (with an internal $O(3)$ symmetry), or could be confined to a plane (an $O(2)$ symmetry), or perhaps we could simply divorce the dimensionality of the internal degree of freedom from a geometric picture and think about an $O(N)$ internal symmetry. Restricting the magnetization to the real line, or perhaps to the values ± 1 , would take us to a discrete group $Z(2)$. Of course, in all of these cases we have a symmetry group only if the underlying dynamics respects the transformation of variables. This considerably restricts the possible kinds of terms present in the Hamiltonian.

Generators of Lie groups

Look more closely at the group elements infinitesimally close to the identity operator. They may be written

$$U \simeq 1 + i \sum_{j=1}^S \lambda_j \phi_j \rightarrow \exp(i \sum_{j=1}^S \lambda_j \phi_j). \quad (7.187)$$

If the group has S ϕ_j parameters, then there will be S generators λ_j where $j = 1 \dots S$. If the group is to remain a group (products of group rotations remain in the group), the generators of the group must obey an algebra which is called a ‘‘Lie algebra.’’ The algebra is conventionally given through the commutators

$$[\lambda_i, \lambda_j] = i c_{ijk} \lambda_k. \quad (7.188)$$

The coefficients c_{ijk} are called ‘‘structure constants.’’ In $SU(2)$ the generators are proportional to the Pauli matrices, $\lambda_i = \frac{1}{2} \sigma_i$, and $c_{ijk} = \frac{1}{2} \epsilon_{ijk}$. In $SU(N)$ it happens that there are $N^2 - 1$ generators and more complicated structure constants. (To count: there are N^2 linearly independent Hermitian matrices with N rows and columns, and one constraint: $\det U = 1$. The magic identity $\det e^A = \exp \text{Tr } A$ applied to $A = i \sum_j \lambda_j \phi_j$ says that the λ 's must be traceless, which is our constraint lowering N^2 to $N^2 - 1$.)

It may happen that some of the c_{ijk} 's are zero, so that the generators commute. The ‘‘rank’’ of a group is the maximum number of mutually commuting generators. For $SU(2)$, the rank is one.

If Hamiltonians are to be invariant under the group transformations, they will involve combinations of the operators which can be constructed from the generators, which commute with all the generators. These are called the Casimir operators. They are very useful because their value in a representation is independent of which generator happens to be diagonal. It turns out that the number of Casimirs is equal to the rank of the group. Since $SU(2)$ is rank 1, it has one Casimir, \hat{J}^2 , the square of the angular momentum.

Exactly as in the case of angular momentum, we want to construct states which are eigenstates of as many generators as possible, plus all the Casimirs. The quantity which is important in the present discussion is the dimensionality of a representation: it is the size of a matrix in which the Casimirs are degenerate. This dimensionality is $2j + 1$ for $SU(2)$ or $O(3)$. For some more vocabulary, we call the smallest nontrivial representation the ‘‘fundamental representation.’’ For $SU(2)$ this is $j = 1/2$. The ‘‘adjoint representation’’ has a dimensionality equal to the number of generators, $j = 1$ or $d = 3$ for $SU(2)$ or $O(3)$.

In most physics applications, we encounter $SU(2)$ and $O(3)$ over and over again, recycling results for our needs.

Isospin

Isospin is an interesting case of a symmetry group with wide applications to nuclear and particle physics.

Its introduction (by Heisenberg, in the 1930's) was inspired by the fact that the masses of the proton and the neutron are roughly equal, and that nuclear forces seemed to be independent of whether the interacting particles were neutrons or protons. Let us think of the proton and the neutron as being different components of the same object, the nucleon, which forms a two-component representation

$$p = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad n = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (7.189)$$

This is a familiar construction. We do not think of the spin-up and spin-down electrons as separate entities, just different versions of one underlying state, the electron. The quantum number which differentiates the neutron from the proton is called “isospin,” and we have chosen the states to be eigenstates of I_3 , in complete analogy with S_z . Notice that there is a relation between charge and isospin of $Q = I_3 + \frac{1}{2}B$, with B , the baryon number, equal to unity. In complete analogy to spin we have

$$I_3 |p\rangle = \frac{1}{2} |p\rangle \quad I_3 |n\rangle = -\frac{1}{2} |n\rangle \quad (7.190)$$

and isospin raising and lowering operators

$$I_+ |n\rangle = |p\rangle \quad I_- |p\rangle = |n\rangle. \quad (7.191)$$

These define the Lie algebra for isospin to be that of the group $SU(2)$, via $[I_i, I_j] = i\epsilon_{ijk}I_k$. The picture can be tested by assuming some form for the Hamiltonian in an isospin basis and by computing relations between different processes, basically, making predictions of ratios using the Wigner-Eckhart theorem. Isospin symmetry can be tested in nuclear spectroscopy by asking whether nuclear levels form isospin multiplets. Isospin will not be an exact symmetry because of electromagnetism: the proton is charged, the neutron is neutral.

Relations among processes can best be described by jumping ahead twenty years, to the early 1950's. There happen to be three pi-mesons, π^+ , π^- , and π^0 . The π^+ and the π^- are

a particle-antiparticle pair, and are therefore degenerate in mass.. The mass of the π^0 is not very different from this value (135 MeV versus 140 MeV), so to good approximation we can consider them all to be of equal mass. The three states smell like an isotriplet state, $|I = 1, I_3 = 1, 0, -1\rangle$. Mesons have $B = 0$, hence $Q = I_3$ for this multiplet.

Pions have a distinguished place in the zoo of strongly interacting particles: they are the lightest mesons, and they are responsible for the long distance part of the nuclear potential (the Yukawa potential).

Shortly after the discovery of the mesons, they were used in scattering experiments which observed the lightest excited states of the nucleon. These are the four states Δ^{++} , Δ^+ , Δ^0 and Δ^- , all with mass of about 1230 MeV. With $Q = I_3 + \frac{1}{2}B$ these states plausibly form an isospin quartet

$$|I = \frac{3}{2}, I_3 = \frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}\rangle. \quad (7.192)$$

Again, $Q = I_3 + B/2$. One can check the hypothesis (and make predictions) by computing relative rates for low energy pion-nucleon scattering at energies where the cross section is dominated by the Δ resonances. Ratios of the amplitudes agree with the appropriate products of $I = \frac{1}{2} \oplus I = 1$ Clebsch-Gordon coefficients.

Here is the classic example: at energies close to the mass of the Δ , pion - nucleon scattering is resonance scattering, that is, the reaction $\pi + N \rightarrow \pi + N$ proceeds through the intermediate state $\pi + N \rightarrow \Delta \rightarrow \pi + N$. The scattering amplitude in a particular channel (π^+p , π^-n , and so on) is proportional to the product of the amplitude for the initial state to couple to the particular isospin 3/2 state, times an amplitude for the intermediate state to couple to the final state. These amplitudes are proportional to the Clebsch -Gordon coefficients for $1 + 1/2 = 3/2$. We can decompose the Delta states into (for example)

$$|\Delta^{++}\rangle = \left| \frac{3}{2} \frac{3}{2} \right\rangle = |11\rangle \left| \frac{1}{2} \frac{1}{2} \right\rangle = |\pi^+p\rangle, \quad (7.193)$$

$$\begin{aligned} |\Delta^+\rangle = \left| \frac{3}{2} \frac{1}{2} \right\rangle &= \sqrt{\frac{1}{3}} |11\rangle \left| \frac{1}{2} -\frac{1}{2} \right\rangle + \sqrt{\frac{2}{3}} |10\rangle \left| \frac{1}{2} \frac{1}{2} \right\rangle \\ &= \sqrt{\frac{1}{3}} |\pi^+n\rangle + \sqrt{\frac{2}{3}} |\pi^0p\rangle \end{aligned} \quad (7.194)$$

and so on. Then we can write

$$A(\pi^+p \rightarrow \Delta^{++} \rightarrow \pi^+p) = A$$

$$\begin{aligned}
 A(\pi^0 p \rightarrow \Delta^+ \rightarrow \pi^- n) &= A\sqrt{\frac{2}{3}}\sqrt{\frac{1}{3}} \\
 A(\pi^0 p \rightarrow \Delta^+ \rightarrow \pi^0 p) &= A\sqrt{\frac{2}{3}}\sqrt{\frac{2}{3}}.
 \end{aligned}
 \tag{7.195}$$

The three cross sections would be in the ratio $1 : 2/9 : 4/9$.

The 1960's story is that the up and down quarks are $I_3 = \pm\frac{1}{2}$ members of an $I = 1/2$ multiplet;

$$u = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad d = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.
 \tag{7.196}$$

Now the charge assignment is $Q = I_3 + 1/6$: the $1/2$ for B is now $1/6$ because a quark carries baryon number $1/3$. Work out the (fractional) charges! And the proton is $(q, q, q)_{I=1/2, I_3=1/2}$, a (u, u, d) bound state.

All this discussion has an old fashioned flavor. The modern story is also worth mentioning. Isospin is an accidental symmetry which arises because the up and down quark masses are both tiny, much smaller than the typical scale of strong interaction dynamics. If they were degenerate, and if one could ignore electromagnetism, isospin would be an exact symmetry of the Standard Model. It is not exact in Nature because the masses are not precisely equal and because the quarks have different charges. The analog situation for the atomic physicist would be to do spectroscopy in the presence of a tiny external magnetic field, which slightly breaks rotational invariance, but still to recognize the near existence of the symmetry.

And one more decade: the 1970's saw the introduction of "weak isospin:" the electron and neutrino are bundled into one leptonic doublet of weak isospin

$$l = \begin{pmatrix} \nu_e \\ e \end{pmatrix}
 \tag{7.197}$$

while the W boson, which induces transitions from charged to neutral leptons, is proportional to an isospin raising or lowering operator $W^\pm \propto I_\pm$. (In fact, all the families of quarks and leptons – at least the ones which interact with W 's – are members of doublets of weak isospin.) Readers who know their particle physics will recognize all of this as a wildly oversimplified story, but hopefully everyone will appreciate the many applications of $SU(2)$ to physics.

Chapter 8

Identical particles

The classical mechanics of a collection of particles assumes that the particles are distinguishable, that one can somehow label all the particles at some point in time and follow their individual motion forever after. This assumption is untenable in the real world of quantum mechanics: identical particles really are identical. All electrons are alike. The consequences of that fact are deep and vast. Let us introduce them slowly, and by example.

We consider first a two electron system, like the Helium atom. Temporarily ignoring spin, we may write its Hamiltonian as a sum of kinetic terms, potentials for the interactions of the individual electrons with the nucleus, and an interaction term for the two electrons,

$$\hat{H} = \frac{\hat{p}_1^2}{2m} + \frac{\hat{p}_2^2}{2m} + V(x_1 - x_2) + V(x_2) + V(x_1). \quad (8.1)$$

Electrons are identical particles and so the Hamiltonian should be (and is) invariant under an exchange of the coordinates of the two electrons, $\hat{H}(1, 2) = \hat{H}(2, 1)$. The wave function of the system depends on the coordinates of the two electrons. Call it $\psi(x_1, x_2)$ or $\psi(1, 2)$, where we extend the idea of exchange to include possible spin degrees of freedom. One can invent an exchange operator \hat{P}_{12} , which acts on the wave function to interchange the coordinates of the two particles

$$\hat{P}_{12}\psi(1, 2) = \psi(2, 1). \quad (8.2)$$

On physical grounds, a glance at Eq. 8.1 tells us that \hat{H} and \hat{P}_{12} must commute. We can do the calculation for energy eigenstates, $\hat{H}(2, 1)\psi(2, 1) = E\psi(2, 1)$. Then

$$\begin{aligned} \hat{H}(1, 2)\hat{P}_{12}\psi(1, 2) &= \hat{H}(2, 1)\psi(2, 1) = E\psi(2, 1) \\ &= E\hat{P}_{12}\psi(1, 2) = \hat{P}_{12}E\psi(1, 2) \\ &= \hat{P}_{12}\hat{H}(1, 2)\psi(1, 2) \end{aligned} \quad (8.3)$$

or $[\hat{H}, \hat{P}_{12}] = 0$. Therefore, one can find simultaneous eigenstates of \hat{H} and \hat{P} . The eigenvalues of the exchange operator will be ± 1 , since

$$\hat{P}_{12}^2\psi(1, 2) = \psi(1, 2) \quad (8.4)$$

implies $\hat{P}_{12}\psi(1, 2) = \pm\psi(1, 2)$. The eigenstates of the exchange operator are a symmetric state and an antisymmetric state,

$$\begin{aligned} \psi^S(1, 2) &= \frac{1}{\sqrt{2}}(\psi(1, 2) + \psi(2, 1)) \\ \psi^A(1, 2) &= \frac{1}{\sqrt{2}}(\psi(1, 2) - \psi(2, 1)) \end{aligned} \quad (8.5)$$

for which $P = +1$ and -1 , respectively.

This exchange symmetry is called “statistics” for short. It applies to all of the degrees of freedom of the state, both spatial (coordinates or momenta) and internal (spin). It happens that there is a relation between the intrinsic spin of a particle and its statistics, called the “spin-statistics” theorem:

Systems made of particles with integer spin (photons, pions, He^4 atoms, O^{16} nuclei, Rb^{87} atoms) have totally symmetric wave functions and obey Bose-Einstein statistics. They are generically called Bosons.

Systems made of particles with half-integer spin (electrons, neutrinos, He^3 atoms, Li^6 atoms) have totally antisymmetric wave functions and obey Fermi-Dirac statistics. Generically, they are called Fermions.

In these statements, “particle” refers to anything whose internal composition is irrelevant to the situation at hand. A helium-4 atom is composed of two electrons, two protons and two neutrons and therefore is a boson. Any many - Helium-4 system’s wave function is symmetric under exchange of the coordinates of two Helium-4 atoms. The wave function is antisymmetric under the exchange of two protons, for they are fermions. This fact is irrelevant if one is not probing any physics of the protons which is independent of the physics of Helium -4 atoms.

The spin-statistics theorem has its origin in the connection between quantum mechanics (specifically the quantum mechanics of many particle systems, as described by quantum field theory) and special relativity (specifically, causality). Its derivation is outside the scope of this book, but will be found in any good quantum field theory text. It is simply not possible to construct a consistent description of Nature for bosonic spin 1/2 particles or fermionic spin-1 ones. The connection between spin and statistics was presented as an unproven principle by Pauli in the 1920’s, as a way of understanding the spectroscopic properties of atoms. It was proven by him in the context of quantum field theory in 1940. For nonrelativistic quantum mechanical systems, we simply have to take the spin-statistics theorem as a fact and deal with it.

We can build both the symmetric and antisymmetric wave functions for three particle systems by inspection:

$$\psi^S(1, 2, 3) = \frac{1}{\sqrt{6}}(\psi(1, 2, 3) + \psi(2, 3, 1) + \psi(3, 1, 2))$$

$$+\psi(2, 1, 3) + \psi(1, 3, 2) + \psi(3, 2, 1)) \quad (8.6)$$

$$\begin{aligned} \psi^A(1, 2, 3) = & \frac{1}{\sqrt{6}}(\psi(1, 2, 3) + \psi(2, 3, 1) + \psi(3, 1, 2) \\ & - \psi(2, 1, 3) - \psi(1, 3, 2) - \psi(3, 2, 1)). \end{aligned} \quad (8.7)$$

An important special case is that of N noninteracting fermions. Here the Hamiltonian is a sum of identical terms, $\hat{H} = \sum_{i=1}^N \hat{H}_i$. The one particle energy eigenstates obey the equation

$$\hat{H}_i u_{E_k} = E_k u_{E_k} \quad (8.8)$$

while the unsymmetrized N -particle wave functions are simply a product of one particle solutions,

$$\hat{H} u_E(1, 2, 3, \dots, N) = E u_E(1, 2, 3, \dots, N) \quad (8.9)$$

with

$$u_E(1, 2, 3, \dots, N) = u_{E_{k_1}}(1) u_{E_{k_2}}(2) \cdots u_{E_{k_N}}(N) \quad (8.10)$$

and $E = E_1 + E_2 + \cdots + E_N$.

The antisymmetric wave function can be constructed from these solutions by forming the the Slater determinant

$$u^A(1, 2, 3, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} u_{E_1}(1) & u_{E_1}(2) & \cdots \\ u_{E_2}(1) & u_{E_2}(2) & \\ \vdots & & \ddots \end{vmatrix}. \quad (8.11)$$

Determinants change sign if two columns are exchanged. Exchanging two columns exchanges the labels on two particles and the minus sign is just the expected antisymmetry property. If two of the single particle wave functions are identical then two of the rows of the determinant are equal, and the determinant will vanish. This is to be expected, since antisymmetry means that two fermions can never be in the same state.

Statistics apply to the complete wave function, including both space and spin components. If we decompose the wave function for many fermions into spatial and spin wave functions, only one of them need be antisymmetric. The spin wave function could be totally antisymmetric, in which case the spatial wave function must be symmetric. Then the fermions can be in the same spatial state. Equivalently, if the spatial part of the wave function is totally antisymmetric, then the spin wave function is symmetric and the two fermions can be in the same spin state. Explicitly, a (separable) two-body spatial wave function,

which is an exchange eigenstate, is

$$\psi_{AB}(x_1, x_2) = \frac{1}{\sqrt{2}}(u_A(x_1)u_B(x_2) \pm u_A(x_2)u_B(x_2)). \quad (8.12)$$

The symmetric wave function carries the plus sign. To satisfy statistics, the spin wave function in that case must be antisymmetric. The antisymmetric spatial wave function (with its minus sign) requires a symmetric spin wave function.

As an example of this connection, suppose we have a state of two electrons, with the same radial quantum number n , and both with orbital angular momentum $L = 1$. We can write the wave function as written

$$\psi(1, 2) = \psi_{space}(1, 2)\psi_{spin}(1, 2). \quad (8.13)$$

It must be totally antisymmetric. We recall that two spin- $\frac{1}{2}$ particles couple to total $S = 1$ or 0, and that the wave functions are

$$\begin{aligned} |11\rangle &= |\uparrow\rangle |\uparrow\rangle \\ |10\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\rangle |\downarrow\rangle + |\downarrow\rangle |\uparrow\rangle) \\ |1-1\rangle &= |\downarrow\rangle |\downarrow\rangle \\ |00\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\rangle |\downarrow\rangle - |\downarrow\rangle |\uparrow\rangle). \end{aligned} \quad (8.14)$$

The triplet states are symmetric under exchange of the two particles. The singlet state is antisymmetric.

An identical story can be told for the orbital angular momenta. We recall that $1 \otimes 1 = 2 \oplus 1 \oplus 0$. Some of the resulting wave functions are

$$\begin{aligned} |22\rangle &= |11\rangle |11\rangle \\ |11\rangle &= \frac{1}{\sqrt{2}}(|11\rangle |10\rangle - |10\rangle |11\rangle) \\ |00\rangle &= \frac{1}{\sqrt{6}}(|1-1\rangle |11\rangle + |11\rangle |1-1\rangle - 2|10\rangle |10\rangle). \end{aligned} \quad (8.15)$$

We observe that the $L = 2$ and $L = 0$ states are totally symmetric under exchange, and the $L = 1$ states are totally antisymmetric. It is easy to convince one's self that the same

connection will hold for all combinations of $L \otimes L$ – the $L \otimes L = 2L$ states are symmetric, so the $L \otimes L = 2L - 1$ ones are antisymmetric by orthogonality, and so on. To preserve the antisymmetry of the entire wave function, when the electrons are in a spin triplet state they cannot be in $L = 0$ or $L = 2$, and similarly if the electrons are in the spin singlet they cannot have $L = 1$. A two-electron atom with the electrons in the same radial and orbital state will have missing levels. This is in contrast to the situation when the two electrons have different radial quantum numbers. All levels will be present, because the radial part of the wave functions could be made symmetric or antisymmetric to compensate for the symmetry of the space and spin degrees of freedom.

Helium

The helium atom provides an interesting example of the effects of statistics. The two valence electrons couple to either $S = 1$ or $S = 0$. These cases are called “ortho-helium” and “para-helium,” respectively. Because ordinary electromagnetic transitions are electric dipole, the spin of the electron is not affected by emission or absorption of photons. Thus, spectroscopically, helium effectively decouples into two distinct kinds of atom, with its own set of energy levels. Of course, the ground state is para, if both electrons are to sit in the same (lowest) orbital. Helium was a mystery in the pre-quantum days; Bohr-model arguments simply failed when applied to it. Heisenberg was the first to realize that the combination of statistics and interactions could account for the properties of helium.

Again neglecting small spin-dependent terms, its Hamiltonian is

$$\hat{H} = \frac{\hat{p}_1^2}{2m} + \frac{\hat{p}_2^2}{2m} - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{|\vec{r}_1 - \vec{r}_2|}. \quad (8.16)$$

If we discard the Coulomb repulsion between the electrons, the Hamiltonian becomes separable. The wave function is a (symmetrized) product of hydrogenic states

$$\psi(1, 2) = \frac{1}{\sqrt{2}}(\psi_{n_1 l_1 m_1}(r_1)\psi_{n_2 l_2 m_2}(r_2) \pm \psi_{n_1 l_1 m_1}(r_2)\psi_{n_2 l_2 m_2}(r_1))\chi_{s_1 s_2} \quad (8.17)$$

where $\chi_{s_1 s_2}$ represents the spin wave function. The probability to find one of the electrons near r_1 and the other near r_2 is

$$|\psi(1, 2)|^2 d^3 r_1 d^3 r_2 = d^3 r_1 d^3 r_2 [(|\psi_{n_1 l_1 m_1}(r_1)|^2 |\psi_{n_2 l_2 m_2}(r_2)|^2 \pm \text{Re}(\psi_{n_1 l_1 m_1}^*(r_1)\psi_{n_2 l_2 m_2}(r_2)\psi_{n_2 l_2 m_2}^*(r_1)\psi_{n_1 l_1 m_1}(r_2)))]. \quad (8.18)$$

The cross terms in the above expression are often referred to as “exchange terms” or “exchange densities.”

Now let us switch the Coulomb repulsion back on and look at its diagonal matrix elements

$$\begin{aligned} \langle i | \hat{V} | i \rangle = & \int d^3r_1 d^3r_2 \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} (|\psi_{n_1 l_1 m_1}(r_1)|^2 |\psi_{n_2 l_2 m_2}(r_2)|^2 \\ & \pm \text{Re}(\psi_{n_1 l_1 m_1}^*(r_1) \psi_{n_2 l_2 m_2}(r_2) \psi_{n_2 l_2 m_2}^*(r_1) \psi_{n_1 l_1 m_1}(r_2))) \end{aligned} \quad (8.19)$$

In perturbation theory, this quantity is the shift in energy due to the electrons’ Coulomb repulsion. The last set of terms is the exchange interaction. We can write this expression schematically as

$$\langle i | \hat{V} | i \rangle = A \pm B \quad (8.20)$$

The plus sign corresponds to the spin singlet wave function and the minus to the spin triplet wave function. The expression can then be rewritten as expression as

$$\langle i | \hat{V} | i \rangle = A - \frac{1}{2}(1 + \vec{\sigma}_1 \cdot \vec{\sigma}_2)B \quad (8.21)$$

where we recall $\vec{\sigma}_1 \cdot \vec{\sigma}_2 = +1$ for $S = 1$ and -3 for $S = 0$. Even though the Hamiltonian had no explicit spin dependence, statistics induces a spin-dependent energy splitting.

The number B arises from electromagnetic interactions. It can be quite large – electron volts, in typical atoms. This is huge, compared to “intrinsic” spin splittings, arising from, for example, hyperfine interactions. The exchange terms are responsible for the difference in energy between excited states in ortho and para-Helium. It is easy to tell the story, without equations: when the electrons have $S = 1$, their spatial wave function has a node at zero separation, so on average, they are farther away than they would be if they were in the singlet state. Since they are farther away, their energy due to Coulomb repulsion is lower. (This story will be true when the spatial wave functions have no nodes; otherwise it might fail.)

These exchange-related energies are also the origin of spin forces in magnetic materials. Depending on the sign of B , materials can be either ferromagnetic (the lowest state is when the spins are aligned) or antiferromagnetic (the $S = 0$ state is lower). Magnetism is due to the Pauli effect, plus Coulomb repulsion.

Diatomic molecules

In diatomic molecules, the electronic wave function sets up a potential in which the nuclei move. This motion can be described as a mix of rotation and vibration. The lowest energy state of the system are in the ground state vibrational level, and the excitations are rotations, for which the Hamiltonian is

$$\hat{H} = \frac{\hat{L}^2}{2I} \quad (8.22)$$

with I the moment of inertia, equal to μa^2 , with reduced mass μ and internuclear separation a . Therefore, the energy eigenvalues are

$$E_l = \frac{\hbar^2}{2I} l(l+1). \quad (8.23)$$

Of course, l , is an integer, and the coordinate-space wave functions are spherical harmonics, $\langle x|lm\rangle = Y_l^m(\theta, \phi)$. Parity plays the role of the exchange operator for these systems. (We recall, the spherical harmonics are parity eigenstates with eigenvalue $(-)^l$.) For molecules with different nuclei, such as HCl, all l 's are allowed. As in the case of atomic spectroscopy, statistics of identical nuclei will remove states from the expected spectrum.

When the two nuclei are identical bosons, such as $C^{12}C^{12}$, the wave function must be symmetric. The total angular momentum of the C^{12} nucleus is known to be zero, and so statistics only depends on l . Bose statistics demands that only even values of l are allowed.

For molecules with fermionic nuclei, such as H_2 , the situation is similar though slightly more complicated. The proton is a spin- $\frac{1}{2}$ fermion. The spin wave function for the pair will either be a spin singlet or a spin triplet. The corresponding spatial wave functions must be antisymmetric for the triplet and symmetric for the singlet, and only odd l 's are allowed for the triplet and even l 's for the singlet.

The absence of these states was an early warning of new physics in nuclei. At the time of the first analysis along the lines we have made (1929) the neutron was undiscovered and nuclei of charge N and mass number A were thought to be composed of N protons (for charge N) and $N - A$ additional protons plus electrons in lieu of the $N - A$ neutrons. For N^{14} , with $N = 7$, $A = 14$, this would be a system of 21 fermions, with, correspondingly, Fermi statistics. The proton - neutron system has fourteen fermions and Bose statistics. Missing states favored the latter case. The neutron was not discovered until 1932, so this issue persisted for several puzzling years.

Quarks and color

The proton and neutron (and all strongly interacting particles) are not fundamental particles; they are bound states of quarks. Quarks are point-like spin- $\frac{1}{2}$ fermions, like electrons or neutrinos. In addition, they are fractionally charged, carrying a charge of $\frac{2}{3}$ or $-\frac{1}{3}$ (in units of the electronic charge). There are six types of quarks, referred to as “flavors.” The up, charm, and top quarks (just labeled as u , c , and t) have charge $2/3$ while the down, strange and bottom quarks (d , s , b) all have charge $-1/3$. Quarks bind together in two ways (other possibilities may exist, a topic of some current research interest):

Mesons are bound states of a quark and an antiquark. We can imagine naively that their spin must couple as $\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$, so that mesons appear as spin singlets and triplets. Indeed, that is the pattern of low lying states. The lightest mesons have $J = 0$ and the lightest excited states have $J = 1$. Higher states can have higher angular momenta, which can plausibly be inferred to be a combination of spin and orbital angular momentum.

Baryons are bound states of three quarks. Again, naively, we expect the spins of the particles to add as $\frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2}$. The lightest baryons are made of up and down quarks (the lightest quarks). The proton and the neutron are spin- $\frac{1}{2}$ baryons (their quark content is, respectively, uud and udd). Including the strange quark, we would expect to see a set of $J = \frac{3}{2}$ states. They are the decuplet

$$\begin{array}{cccc}
 uuu & uud & udd & ddd \\
 & suu & sud & sdd \\
 & & ssu & ssd \\
 & & & sss
 \end{array} \tag{8.24}$$

and in fact all of these particles are well established resonances,

$$\begin{array}{cccc}
 \Delta^{++} & \Delta^+ & \Delta^0 & \Delta^- \\
 & \Sigma^+ & \Sigma^0 & \Sigma^- \\
 & & \Xi^0 & \Xi^- \\
 & & & \Omega^-
 \end{array} \tag{8.25}$$

But there is a problem with this story: The $J = \frac{3}{2}$ spin wave function is symmetric. For example,

$$|\frac{3}{2}, \frac{3}{2}\rangle = |\uparrow\rangle |\uparrow\rangle |\uparrow\rangle. \tag{8.26}$$

We would expect the space wave function of the lowest energy state to be symmetric. (This can be tested, in the form factors for scattering. No evidence for nodes ever appeared.) We seem to have a violation of the spin-statistics theorem. The resolution (by Greenberg) was to propose that there is another quantum number associated with quarks, called “color,” and that it is the color part of the wave function which carries the antisymmetry. This can be done through the Slater determinant.

$$\begin{vmatrix} r_1 & r_2 & r_3 \\ g_1 & g_2 & g_3 \\ b_1 & b_2 & b_3 \end{vmatrix}. \quad (8.27)$$

This seemed quite contrived at the time, but it turned out to be true. The strong interactions involve the color degree of freedom in the same way that electromagnetism involves charge. It is believed that the strong interactions permanently confine quarks into colorless or “color singlet” objects, for which the color wave function is Eq. 8.27. To pursue this subject would take us too far afield, but here is a suggestion of a check of the idea:

In the proton and neutron, the spin wave functions of the uu or dd pair must be identical. This forces their total spin to be $J = 1$. Now the proton or neutron wave function is $1 \otimes \frac{1}{2} = \frac{1}{2}$. Construct the wave functions and compute the magnetic moments of the two baryons in terms of the magnetic moments of the u and d quarks. Presumably, the masses of the two quarks are nearly equal, because the proton and neutron are nearly mass-degenerate, so the ratio of magnetic moments is the ratio of the charges. How close is your prediction to the observed ratio of magnetic moments, $\mu_n/\mu_p = -0.684$?

Chapter 9

Time independent perturbation theory

General remarks about perturbation theory

A general problem in any practical calculation is that the Hamiltonian cannot be diagonalized exactly. Hamiltonians which describe real world problems are usually infinite dimensional, so writing

$$|\psi\rangle = \sum_j c_j |j\rangle \quad (9.1)$$

and directly attacking

$$\det(\langle i|\hat{H}|j\rangle - \delta_{ij}E) = 0 \quad (9.2)$$

or trying to integrate

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \hat{H} \sum_j c_j(t) |j\rangle \quad (9.3)$$

is not going to get you very far.

However, often H is almost diagonal

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}_1, \quad (9.4)$$

where λ is a small dimensionless parameter and H_0 is time independent and solvable,

$$\hat{H}_0 |\psi_0\rangle = W_0 |\psi_0\rangle. \quad (9.5)$$

Then the idea is to treat the small part of the problem as a perturbation: we assume that the energies and eigenfunctions of the complete system are nearly those of the dominant part of the Hamiltonian. The effect of the small part is treated approximately. Specifically, we write

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}_1. \quad (9.6)$$

and we want to solve

$$\hat{H} |\psi\rangle = W |\psi\rangle. \quad (9.7)$$

The quantity λ is a small expansion parameter. We wish to find $|\psi\rangle$ and W as power series in λ ,

$$|\psi\rangle = |\psi_0\rangle + \lambda |\psi_1\rangle + \lambda^2 |\psi_2\rangle + \dots \quad (9.8)$$

and

$$W = W_0 + \lambda W_1 + \lambda^2 W_2 + \dots \quad (9.9)$$

The time dependent analog is

$$\langle \beta|U(t, t_0)|\alpha\rangle = \delta_{\alpha\beta} e^{-iW_0\alpha(t-t_0)/\hbar} + \lambda U_1(t, t_0) + \lambda^2 U_2(t, t_0) + \dots \quad (9.10)$$

The method of solution of the time independent Schrödinger equation is called Rayleigh - Schrödinger perturbation theory. The time dependent situation is just called “perturbation theory,” but contains as a special case the very useful “Golden Rule” which is the basis for most practical calculations.

Most theoretical calculations begin and end with perturbation theory, so a few more remarks are in order:

To a mathematician, $\lambda \hat{H}_1 \ll H_0$ means that there is a basis where $\langle i | \hat{H} | i \rangle \gg \langle i | \hat{H} | j \rangle$, that the Hamiltonian matrix is diagonal dominant.

Sometimes, one might actually be able to control the parameter λ in a laboratory setting. (For example, the effect of an external magnetic field on an atom can be made small by making the field small.)

Sometimes you can't. An example is

$$\hat{H} = \frac{p^2}{2m_0} + \frac{Ze_0^2}{r} + \left[\frac{p^4}{8m_0c^3} + \dots \vec{L} \cdot \vec{S} \right] \quad (9.11)$$

(where the first two terms are \hat{H}_0 and the terms in brackets are \hat{H}_1) or the interaction of an electron in an electromagnetic field, with the same \hat{H}_0 and $\hat{H}_1 = e/(mc)\vec{A} \cdot \vec{p}$. After the fact, you discover that there is a small expansion parameter, $\alpha = e^2/(\hbar c) = 1/137$ or $\langle v/c \rangle = \alpha$, BUT you can't realize \hat{H}_0 experimentally. In the electromagnetic case, how can you have an electron which doesn't couple to a photon?

This raises an interesting cascade of issues. Just to mention one of them: how are the parameters in \hat{H}_0 (the mass parameter m_0 and the charge e_0 – let's call these objects “coupling constants” in what follows) related to quantities measured in experiment where the Hamiltonian is the full \hat{H} ?

The answer: the measurement of a coupling constant is the results of comparison of a theoretical calculation of an experimental situation, which is fit to experimental data. The coupling constants are determined in the fit. Often, the theoretical calculation is perturbative, and it is done to some finite order of perturbation theory. Then the experimental quantity E is related to theory as

$$E = T_0(e_0, m_0) + \lambda T_1(e_0, m_0) + \lambda^2 T_2(e_0, m_0) + \dots \quad (9.12)$$

Think of E as the experimentally measured charge of the electron. This formula means that e is a function of the “bare parameters” e_0 and m_0 . The quantities e_0 and m_0 have to be

determined from some small set of experiments. Then one can go on to perform additional experiments, and make predictions.

I've never seen this discussed in a quantum mechanics course, but you will see it when you study quantum field theory.

Another thing to notice about Eq. 9.12: the theoretical calculation implicitly assumes that the ... are zero. The theoretical prediction will change when the next order in perturbation theory is computed. This means that all perturbative calculations have an inherent ambiguity, which is usually hard to quantify. Dealing with this is a recurring problem in theoretical physics.

Pure nonperturbative techniques tend not to have universal applicability. There are many special cases. Generally, completeness is lost,

$$|\psi\rangle = \sum_j c_j |j\rangle \rightarrow \sum_j^N c_j |j\rangle \quad (9.13)$$

(this is also the case for many practical perturbative calculations) and then the issue is, how to restore it.

Sometimes there is no good choice for λ . Sometimes, in this situation, a λ is introduced as a bookkeeping device, and set to unity at the end of the calculation.

Rayleigh-Schrödinger perturbation theory

The perturbative expansion is developed simply by substituting the power series expansions for W and $|\psi\rangle$ into the Schrödinger equation

$$(\hat{H}_0 + \lambda \hat{H}_1)(|\psi_0\rangle + \lambda |\psi_1\rangle + \lambda^2 |\psi_2\rangle + \dots) = (W_0 + \lambda W_1 + \lambda^2 W_2 + \dots) \times (|\psi_0\rangle + \lambda |\psi_1\rangle + \lambda^2 |\psi_2\rangle + \dots), \quad (9.14)$$

and equating terms order by order in λ . At zeroth order (of course) is

$$(\hat{H}_0 - W_0) |\psi_0\rangle = 0, \quad (9.15)$$

which we assume is exactly solvable. At first order,

$$(\hat{H}_0 - W_0) |\psi_1\rangle = (W_1 - \hat{H}_1) |\psi_0\rangle. \quad (9.16)$$

In second order we have

$$(\hat{H}_0 - W_0) |\psi_2\rangle = (W_1 - \hat{H}_1) |\psi_1\rangle + W_2 |\psi_0\rangle \quad (9.17)$$

Higher order terms in the series rapidly become unwieldy.

Now let us solve these equations. We assume that $|\psi_0\rangle \equiv |u_m\rangle$ is an eigenfunction of \hat{H}_0 with eigenvalue $W_0 = E_m$. In the formulas we develop, we also assume that the subscript m is a discrete index. It is not really necessarily need to assume this property, just convenient. For a continuous spectrum, sums become integrals. However, it will be necessary to assume that the eigenvalues E_m are non-degenerate.

A glance at the left hand side of Eq. 9.16 tells us that one can add an arbitrary amount of the wave function $|\psi_0\rangle$ to the wave function $|\psi_1\rangle$ without changing anything. We will make use of this fact for our convenience, by defining $|\psi_n\rangle$ to be orthogonal to $|\psi_0\rangle$ (i.e. $\langle\psi_0|\psi_n\rangle = 0$). Taking the inner product of the first order equation with the bra $\langle\psi_0|$ gives

$$\langle\psi_0|(\hat{H}_0 - W_0)|\psi_1\rangle = 0 = \langle\psi_0|(W_1 - \hat{H}_1)|\psi_0\rangle \quad (9.18)$$

or

$$W_1 = \langle\psi_0|\hat{H}_1|\psi_0\rangle \quad (9.19)$$

or (explicitly labeling the state)

$$W_1^m = \langle u_m|\hat{H}_1|u_m\rangle. \quad (9.20)$$

This is a simple, important and useful result. The first order energy shift of the perturbation is just its expectation value in the unperturbed state.

From Eq. 9.17, the second order energy shift is

$$\langle\psi_0|(\hat{H}_0 - W_0)|\psi_2\rangle = 0 = \langle\psi_0|(W_1 - \hat{H}_1)|\psi_1\rangle - W_2 \langle\psi_0|\psi_0\rangle \quad (9.21)$$

or

$$W_2 = \langle\psi_0|(\hat{H}_1 - W_1)|\psi_1\rangle. \quad (9.22)$$

Note that W_1 only depends on $|\psi_0\rangle$ and W_2 only depends on $|\psi_0\rangle$ and $|\psi_1\rangle$. In general to find the energies to order n requires only the wave function to order $n - 1$.

The first order contribution to the perturbed wave function must be a superposition of the unperturbed energy eigenfunctions, excluding $|u_m\rangle$ because the correction is orthogonal to the unperturbed state

$$|\psi_1\rangle = \sum_{n \neq m} c_n |u_n\rangle. \quad (9.23)$$

Rewrite Eq. 9.16 to reflect that we are perturbing about $|u_m\rangle$

$$(\hat{H}_0 - E_m)|\psi_1\rangle = (W_1 - \hat{H}_1)|u_m\rangle \quad (9.24)$$

and substitute for $|\psi_1\rangle$:

$$\sum_{n \neq m} (\hat{H}_0 - E_m)c_n |u_n\rangle = (W_1 - \hat{H}_1)|u_m\rangle \quad (9.25)$$

or

$$\sum_{n \neq m} (E_n - E_m)c_n |u_n\rangle = (W_1 - \hat{H}_1)|u_m\rangle. \quad (9.26)$$

Now take an inner product with by $\langle u_k|$, and, recalling $\langle u_k|u_m\rangle = 0$, we have

$$c_k(E_k - E_m) = W_1 \langle u_k|u_m\rangle - \langle u_k|\hat{H}_1|u_m\rangle \quad (9.27)$$

or

$$c_k = \frac{\langle u_k|\hat{H}_1|u_m\rangle}{E_m - E_k}. \quad (9.28)$$

Thus the complete result for the first-order correction to the wave function is

$$|\psi_1\rangle = \sum_{n \neq m} |u_n\rangle \frac{\langle u_n|\hat{H}_1|u_m\rangle}{E_m - E_n} \quad (9.29)$$

We now see the necessity of our requirement that the eigenvalues be non-degenerate. If they were degenerate, the expression would diverge.

The first order wave function allows us to the second order energy correction. Exploiting the orthogonality of $|\psi_1\rangle$ and $|\psi_0\rangle$, we obtain

$$\begin{aligned} W_2 &= \langle \psi_1|\hat{H}_1|\psi_0\rangle \\ &= \left(\sum_{n \neq m} \frac{\langle u_n|\hat{H}_1|u_m\rangle \langle u_n|}{E_m - E_n} \right) \hat{H}_1 |\psi_0\rangle \\ &= \sum_{n \neq m} \frac{|\langle u_n|\hat{H}_1|u_m\rangle|^2}{E_m - E_n} \end{aligned} \quad (9.30)$$

The complete energy shift through second order is

$$E = E_m + \langle u_m|\hat{H}_1|u_m\rangle + \sum_{n \neq m} \frac{|\langle u_n|\hat{H}_1|u_m\rangle|^2}{E_m - E_n}. \quad (9.31)$$

As a useful mnemonic for remembering the order of terms in the denominator of the second order correction, observe that the second order perturbation always lowers the unperturbed ground state energy. (This is actually a deep statement, to which we will return.)

Two simple examples with oscillators

Let us look at a couple of examples which can actually be solved exactly. They both use the one dimensional simple harmonic oscillator $\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2$. We begin with a perturbation $\hat{H}_1 = \lambda\hat{x}^2$. Of course, we can exactly diagonalize \hat{H} in this case. The full potential is

$$V \rightarrow \frac{1}{2}m\omega^2\hat{x}^2 + \lambda\hat{x}^2 = \frac{1}{2}m\omega'^2\hat{x}^2, \quad (9.32)$$

and so the exact energies are $E_n = \hbar\omega' \left(n + \frac{1}{2}\right)$ with

$$\omega' = \sqrt{\omega^2 + \frac{2\lambda}{m}} = \omega + \frac{\lambda}{m\omega} + \dots \quad (9.33)$$

Perturbation theory should recover the power series. To first order, we expect

$$\Delta E = \frac{\hbar\lambda}{2m\omega}. \quad (9.34)$$

In Rayleigh-Schrödinger perturbation theory, the first order correction is

$$\begin{aligned} E_1 &= \langle 0|\lambda\hat{x}^2|0\rangle \\ &= \frac{\hbar\lambda}{2m\omega} \langle 0|(\hat{a} + \hat{a}^\dagger)^2|0\rangle \\ &= \frac{\hbar\lambda}{2m\omega} \langle 0|\hat{a}^2 + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} + \hat{a}^{\dagger 2}|0\rangle \\ &= \frac{\hbar\lambda}{2m\omega} \end{aligned} \quad (9.35)$$

which agrees (as it should) with the direct expansion of the exact result.

For a slightly less trivial example, consider the perturbation $\hat{H}_1 = \lambda\hat{x}$. This situation also has an exact solution, since

$$\hat{V} = \frac{1}{2}m\omega^2\hat{x}^2 + \lambda\hat{x} = \frac{1}{2}m\omega^2 \left(\hat{x} + \frac{\lambda}{m\omega}\right)^2 - \frac{\lambda^2}{2m\omega^2}. \quad (9.36)$$

The perturbation amounts to a translation in the origin of the oscillator plus an overall constant increment of the energy. The translation does not affect the energy, so

$$\Delta E = -\frac{\lambda^2}{2m\omega^2}. \quad (9.37)$$

Now we approach the problem using Rayleigh-Schrödinger perturbation theory. The first order term vanishes:

$$\lambda \langle 0|\hat{x}|0\rangle = \lambda\sqrt{\frac{\hbar}{2m\omega}} \langle 0|(\hat{a} + \hat{a}^\dagger)|0\rangle = 0. \quad (9.38)$$

For the second order shift, we need to compute

$$\langle \psi_n | \hat{H}_1 | \psi_0 \rangle = \lambda \sqrt{\frac{\hbar}{2m\omega}} \langle \psi_n | (\hat{a} + \hat{a}^\dagger) | \psi_0 \rangle = \lambda \sqrt{\frac{\hbar}{2m\omega}} \delta_{n,1}. \quad (9.39)$$

Therefore, the sum collapses to a single term,

$$\begin{aligned} E_2 &= \sum_{n \neq 0} \frac{|\langle n | \hat{H}_1 | 0 \rangle|^2}{E_0 - E_n} = \frac{|\langle 1 | \hat{H}_1 | 0 \rangle|^2}{E_0 - E_1} \\ &= \frac{\left(\lambda \sqrt{\frac{\hbar}{2m\omega}} \right)^2}{-\hbar\omega} = -\frac{\lambda^2}{2m\omega^2}. \end{aligned} \quad (9.40)$$

The calculation of the energy eigenvalues agrees with the exact result.

The perturbed wave function is

$$\begin{aligned} |\psi\rangle &= |0\rangle + \sum_{n \neq 0} \frac{|n\rangle \langle n | \hat{H}_1 | 0 \rangle}{E_0 - E_n} \\ &= |0\rangle - \frac{\lambda}{\hbar\omega} \sqrt{\frac{\lambda}{2m\omega}} |1\rangle \end{aligned} \quad (9.41)$$

The original wave function for the unshifted ground state is

$$\langle x | 0 \rangle = \left(\frac{\hbar}{\pi m \omega} \right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar} x^2\right). \quad (9.42)$$

The perturbation shifts the origin of the oscillator, so the true wave function has just been translated. Expanding the shifted wave function for small λ ,

$$\left(\frac{\hbar}{\pi m \omega} \right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar} \left(x + \frac{\lambda}{m\omega}\right)^2\right) \simeq \left(\frac{\hbar}{\pi m \omega} \right)^{1/4} \left(1 - \frac{\lambda x}{m\omega} + \dots\right) \exp\left(-\frac{m\omega}{2\hbar} x^2\right). \quad (9.43)$$

The $O(\lambda)$ term is, indeed, the first excited state.

A rearranged series (Brillouin-Wigner perturbation theory)

The initial equation whose perturbative expansion we wish to construct can be rewritten as

$$(\hat{H}_0 - W) |\psi\rangle = -\lambda \hat{H}_1 |\psi\rangle \quad (9.44)$$

and its (formal) solution is

$$|\psi\rangle = -(\hat{H}_0 - W)^{-1} \lambda \hat{H}_1 |\psi\rangle. \quad (9.45)$$

Inserting a complete set of unperturbed states $|n\rangle$ with eigenvalues $\hat{H}_0 |n\rangle = W_n^{(0)} |n\rangle$, Eq. 9.45 becomes

$$|\psi\rangle = - \sum_n \frac{1}{W_n^{(0)} - W} |n\rangle \langle n | \lambda \hat{H}_1 | \psi \rangle \quad (9.46)$$

The sum runs over all states $|n\rangle$. We assume that the exact wave function $|\psi\rangle$ will be close to one of them, which we will label as $|m\rangle$. We separate that term from the series. We also relabel $|\psi\rangle$ as $|u_m\rangle$ guide the eye. We have

$$|u_m\rangle = C |m\rangle + \sum_{n \neq m} \frac{|n\rangle \langle n | \lambda \hat{H}_1 | u_m \rangle}{W_m - W_n^{(0)}} \quad (9.47)$$

where

$$C = \frac{\langle m | \lambda \hat{H}_1 | u_m \rangle}{W_m - W_m^{(0)}}. \quad (9.48)$$

We solve this implicit equation by iteration:

$$\begin{aligned} |u_m\rangle &= C |m\rangle + \sum_{n \neq m} \frac{C |n\rangle \langle n | \lambda \hat{H}_1 | m \rangle}{W_m - W_n^{(0)}} + \sum_{n \neq m, p \neq m} \frac{C |n\rangle \langle n | \lambda \hat{H}_1 | p \rangle \langle p | \lambda \hat{H}_1 | m \rangle}{(W_m - W_n^{(0)})(W_m - W_p^{(0)})} + \dots \\ &= C \left(|m\rangle + \sum_{n \neq m} \frac{Q_{nm} |n\rangle}{W_m - W_n^{(0)}} + \sum_{n \neq m, p \neq m} \frac{Q_{np} Q_{pm} |n\rangle}{(W_m - W_n^{(0)})(W_m - W_p^{(0)})} + \dots \right) \end{aligned} \quad (9.49)$$

where we have defined $Q_{ij} \equiv \langle i | \lambda \hat{H}_1 | j \rangle$. Eq. 9.48 inverts to give $\langle m | \lambda \hat{H}_1 | u_m \rangle = C(W_m - W_m^{(0)})$, so taking the inner product of $|u_m\rangle$ from Eq. 9.49 with $\langle m | \lambda \hat{H}_1$ gives

$$W_m - W_m^{(0)} = Q_{mm} + \sum_{n \neq m} \frac{Q_{mn} Q_{nm}}{(W_m - W_n^{(0)})} + \sum_{n \neq m, p \neq m} \frac{Q_{mn} Q_{np} Q_{pm}}{(W_m - W_n^{(0)})(W_m - W_p^{(0)})} + \dots \quad (9.50)$$

The usefulness of this expansion stems from the fact that one can write down any particular order in the expansion by inspection. For example, the order $N + 1$ contribution is

$$\sum_{i_1 \neq m, \dots, i_N \neq m} \frac{Q_{mi_1} Q_{i_1 i_2} \dots Q_{i_{N-1} i_N} Q_{i_N m}}{(W_m - W_{i_1}^{(0)})(W_m - W_{i_2}^{(0)}) \dots (W_m - W_{i_N}^{(0)})}. \quad (9.51)$$

Of course, it is not completely useful, since the defining equations are implicit. This version of perturbation theory is called “Brillouin-Wigner perturbation theory.” Occasionally, it might be possible to solve the implicit equations iteratively. A more useful reason for describing this method (in the author’s opinion) is this: One can write a series solution for the energy,

$$W_m = W_m^{(0)} + \lambda W_m^{(1)} + \lambda^2 W_m^{(2)} + \dots \quad (9.52)$$

and insert it into the Brillouin-Wigner series. Collecting terms order by order in λ will allow one to generate the much more complicated Rayleigh-Schrödinger expansion in a straightforward way.

Degenerate state perturbation theory

In the presence of degeneracies the expansions in the preceding sections break down due to vanishing energy denominators. Physically, the perturbation expansion fails because its implicit assumption – that the perturbation produces a small effect on the states and energy expectation values – is simply false, when applied to states which are degenerate before the perturbation is applied.

To illustrate this situation, suppose that we have an unperturbed system with two states which are degenerate in zeroth order. Call them $|l\rangle$ and $|m\rangle$. Assume also that there is a small term which can mix them. Call its matrix elements $\lambda \langle i | \hat{H}_1 | j \rangle = \lambda H_{ij}$. The full Hamiltonian is, then,

$$\hat{H} = \begin{pmatrix} E_0 + \lambda H_{ll} & \lambda H_{lm} \\ \lambda H_{ml} & E_0 + \lambda H_{mm} \end{pmatrix}. \quad (9.53)$$

The eigenvalues are

$$W = E_0 + \lambda \left(\frac{H_{ll} + H_{mm}}{2} \right) \pm \lambda \sqrt{\left(\frac{H_{ll} - H_{mm}}{2} \right)^2 + |H_{ml}|^2} \equiv W_{\tilde{l}}, W_{\tilde{m}} \quad (9.54)$$

The energy eigenstates are

$$\begin{aligned} |\tilde{l}\rangle &= \alpha |l\rangle + \beta |m\rangle \\ |\tilde{m}\rangle &= -\beta |l\rangle + \alpha |m\rangle \end{aligned} \quad (9.55)$$

(where of course $|\alpha|^2 + |\beta|^2 = 1$). While the energies are smooth functions of λ , the eigenvectors are not. To see that, call

$$W_{\tilde{l}} = E_0 + \lambda \epsilon. \quad (9.56)$$

Then the eigenvector equation, determining α and β , is

$$(E_0 - (E_0 + \lambda \epsilon))\alpha + \lambda H_{lm}\beta = 0 \quad (9.57)$$

and therefore $\lambda \epsilon \alpha = \lambda H_{lm}\beta$ or

$$\frac{\alpha}{\beta} = \frac{H_{lm}}{\epsilon} \quad (9.58)$$

All λ dependence has disappeared. This illustrates the point, that when a degeneracy is present, it is simply not the case that the eigenfunction of the full Hamiltonian differs from the eigenfunction of \hat{H}_0 by an $O(\lambda)$ correction. Indeed, if the perturbation did not have a diagonal term, $H_{ll} = H_{mm} = 0$, we would find $\alpha/\beta = \pm 1$.

So much for the two state system. Now suppose that we have a zeroth order system with many states, and two (or more) of them are degenerate. We add a perturbation to the system. The way we can deal with it is to first separate the degenerate subspace from the rest of the system, and diagonalize the entire Hamiltonian in it. Then reconsider the entire system, using not the original degenerate states, but the rotated states. Our old Hamiltonian matrix was

$$\hat{H} = \begin{pmatrix} E_0 + \lambda H_{ll} & \lambda H_{lm} & \lambda H_{lk} & \cdots \\ \lambda H_{ml} & E_0 + \lambda H_{mm} & \lambda H_{mk} & \cdots \\ \lambda H_{kl} & \lambda H_{km} & E_k + \lambda H_{kk} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad (9.59)$$

and in the rotated basis it becomes

$$\hat{H} = \begin{pmatrix} W_{\tilde{l}} & 0 & \lambda H_{\tilde{l}k} & \cdots \\ 0 & W_{\tilde{m}} & \lambda H_{\tilde{m}k} & \cdots \\ \lambda H_{k\tilde{l}} & \lambda H_{k\tilde{m}} & E_k + \lambda H_{kk} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (9.60)$$

Now all degeneracies have been lifted, and ordinary Rayleigh-Schrödinger perturbation theory gives the perturbed wave function (to first order) as well as the perturbed energy (to second order). For state $|\tilde{l}\rangle$,

$$|\psi\rangle = |\tilde{l}\rangle + \sum_{j \neq \tilde{l}} \frac{|j\rangle \langle j|\hat{H}_1|\tilde{l}\rangle}{W_{\tilde{l}} - W_j} \quad (9.61)$$

$$E = W_{\tilde{l}} + \lambda^2 \sum_{j \neq \tilde{l}} \frac{|\langle \tilde{l}|\hat{H}_1|j\rangle|^2}{W_{\tilde{l}} - W_j}. \quad (9.62)$$

Notice that there are no $W_{\tilde{l}} - W_{\tilde{m}}$ denominators in the sums because, by design, the perturbation does not connect $|\tilde{l}\rangle$ and $|\tilde{m}\rangle$ – their numerators are exactly zero.

Several comments are in order: First, the author regards the expression “degenerate state perturbation theory” as an oxymoron: the small terms can drastically affect degenerate states. Second, one may encounter situations where states are not degenerate, but their unperturbed energy splitting is small compared to the effect of the perturbation: $E_l - E_m \ll \lambda H_{lm}$, for example. Then it might be advantageous to treat the states as if they were degenerate, diagonalize the subspace exactly, and proceed. Finally, this business can become quite delicate. We have not considered the case where $H_{ll} = H_{mm}$ and $H_{lm} = 0$. In this case the perturbation does not lift the degeneracy, and one must consider higher-order formulas. There is a brief description of this situation in Schiff.

Hydrogen: relativistic correction

We illustrate the uses of perturbation theory in some classic examples drawn from the physics of light atoms.

Recall that the Hamiltonian for a non-relativistic charged particle in a Coulomb potential is $\hat{H}_0 = \hat{p}^2/(2m) - Ze^2/r$. This Hamiltonian leaves out many effects in the physics of the physical electron - proton bound state. The largest of these contribute to what is known as “fine structure,” the relativistic correction to the kinetic energy and the spin-orbit interaction.

The true relation between energy and momentum for a relativistic particle is $E = \sqrt{p^2c^2 + m^2c^4}$. Because the electron’s velocity is small ($v/c \sim \alpha$), the energy can be approximated as

$$E = mc^2 + \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \dots \quad (9.63)$$

and the quartic term can be treated as a perturbation. Its expectation value can be partially evaluated using tricks:

$$\begin{aligned} \langle \hat{H}_{rel} \rangle &= \left\langle -\frac{1}{2m^3c^2} \left(\frac{p^2}{2m} \right)^2 \right\rangle \\ &= \left\langle -\frac{1}{2m^3c^2} \left(\hat{H}_0 + \frac{Ze^2}{r} \right) \left(\hat{H}_0 + \frac{Ze^2}{r} \right) \right\rangle \\ &= -\frac{1}{2m^3c^2} \left((E_n^{(0)})^2 - 2E_n^{(0)}Ze^2 \left\langle \frac{1}{r} \right\rangle + Z^2e^4 \left\langle \frac{1}{r^2} \right\rangle \right). \end{aligned} \quad (9.64)$$

The matrix element $\langle \frac{1}{r} \rangle$ can be computed using the virial theorem,

$$Ze^2 \left\langle \frac{1}{r} \right\rangle = -2E_n^{(0)}. \quad (9.65)$$

The calculation of $\langle \frac{1}{r^2} \rangle$ requires a sum over Laguerre polynomials. It happens to be

$$\left\langle \frac{1}{r^2} \right\rangle = \frac{Z^2}{a_0^2 n^3 (l + \frac{1}{2})}. \quad (9.66)$$

Recalling that the Bohr radius is $a_0 = \frac{\hbar c}{2mc^2}$ and $\frac{e^2}{\hbar c} = \alpha \simeq \frac{1}{137}$, this is

$$\left\langle \frac{1}{r^2} \right\rangle = \frac{Z^2 m^2 c^2 \alpha^2}{\hbar^2 n^3 (l + \frac{1}{2})}. \quad (9.67)$$

The complete expression is

$$E_{rel} = \frac{Z^4 \alpha^2}{n^3} \times \text{Ry} \times \left(\frac{3}{4n} - \frac{2}{2l+1} \right) \quad (9.68)$$

where “Ry” is shorthand for the Rydberg, -13.6 eV. Note that this perturbation lifts the degeneracy between the n and l states. We list a few representative values of the term in parentheses in Table 9. The scale of the correction is about 10^{-5} eV for hydrogen.

state	n, l	value
1s	$n = 1, l = 0$	$\frac{-5}{4}$
2s	$n = 2, l = 0$	$\frac{-39}{24}$
2p	$n = 2, l = 1$	$\frac{-7}{24}$

Hydrogen: spin-orbit coupling

Recall that the electron has a spin magnetic moment

$$\vec{\mu} = -\frac{e}{mc} \vec{S}. \quad (9.69)$$

It moves in the electric field of the proton. The effects of special relativity convert the electric field into a mixture of electric and magnetic field. If the electron were moving with a uniform velocity, this magnetic field would be

$$\vec{B} = -\frac{\vec{v}}{c} \times \vec{E}. \quad (9.70)$$

For a central potential, the electric field is

$$\vec{E} = -\vec{\nabla}V(\vec{r}) = -\vec{r} \frac{1}{r} \frac{\partial V}{\partial r} \quad (9.71)$$

and so the transformed magnetic field would be

$$\vec{B} = -\frac{\vec{v}}{c} \times \left(-\vec{r} \frac{1}{r} \frac{\partial V}{\partial r} \right) = \frac{1}{mc} (\vec{p} \times \vec{r}) \frac{1}{r} \frac{\partial V}{\partial r} = -\frac{1}{mc} \left(\frac{1}{r} \frac{\partial V}{\partial r} \right) \vec{L}. \quad (9.72)$$

However, the electron is not moving with a constant velocity as it orbits the proton. It is accelerating. The correct answer is a factor of two smaller. The extra factor is called Thomas precession term and its rescaling by $1/2$ is the limiting value for non-relativistic velocity and small acceleration. (See Jackson for a more extensive discussion.)

The physical result is unchanged: the electron's magnetic moment interacts with an external magnetic field. The interaction is

$$\hat{H}_{so} = \frac{e}{2m^2c^2} \frac{1}{r} \frac{\partial V}{\partial r} \vec{L} \cdot \vec{S} \quad (9.73)$$

With $V = -Ze^2/r$, this becomes

$$\hat{H}_{so} = \frac{1}{2} \frac{Ze^2}{m^2c^2} \frac{1}{r^3} \vec{L} \cdot \vec{S}. \quad (9.74)$$

As we contemplate a perturbative calculation of the energy shift due to this term, we recall the quantum numbers of the eigenfunctions of that original Hamiltonian. They are n , l , m_l , s , and m_s . We also recall that each level n is $2n^2$ degenerate. The interaction term couples these degenerate states:

$$\vec{L} \cdot \vec{S} = L_z S_z + \frac{1}{2} (L_+ S_- + L_- S_+) \quad (9.75)$$

the first term couples either states with $m'_l = m_l$ and $m'_s = m_s$ and the second two terms couple $m'_l = m_l \pm 1$ and $m'_s = m_s \mp 1$. We are faced with a problem of degenerate state perturbation theory. We can avoid it by finding a basis where the perturbation is diagonal. This is easy to do: recall that the total angular momentum is $\vec{J} = \vec{L} + \vec{S}$ and choose basis states $|j, m_j, l, s\rangle$. Then, because $J^2 = L^2 + S^2 + 2\vec{L} \cdot \vec{S}$, our perturbation is diagonal

$$\vec{L} \cdot \vec{S} = \frac{1}{2} (J^2 - L^2 - S^2). \quad (9.76)$$

The complete spin-orbital energy correction is

$$\Delta E_{so} = \langle n, j, m_j, l, s | \hat{H}_{so} | n, j, m_j, l, s \rangle = \frac{1}{4} \frac{Ze^2 \hbar^2}{m^2 c^2} \left\langle \frac{1}{r^3} \right\rangle (J(J+1) - L(L+1) - S(S+1)). \quad (9.77)$$

Through a tedious calculation, the radial matrix element can be evaluated,

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{Z^3}{a_0^3} \frac{1}{n^2 l (l + \frac{1}{2}) (l + 1)} \quad (9.78)$$

(Note that the numerator of Eq. 9.77 vanishes if $l = 0$.)

All the dimensional factors combine to give a term which is the same size as the relativistic correction,

$$\Delta E_{so} \sim \frac{Z^4 \alpha^2}{n^2} mc^2. \quad (9.79)$$

This is no surprise; both terms come from special relativistic corrections to the nonrelativistic limit. If we sum these two terms, we find

$$\Delta E_{so} + \Delta E_{rel} = -\text{Ry} \frac{Z^4 \alpha^2}{n^3} \left(\frac{1}{j + \frac{1}{2}} - \frac{3}{4n} \right) \quad (9.80)$$

Surprisingly, even though the relativistic and spin-orbital corrections break the degeneracy between states of the same j and different l , in the final analysis, they are degenerate! In hydrogen, the famous special case is the degeneracy of the $2S_{\frac{1}{2}}$ and $2P_{\frac{1}{2}}$ states, while the $2P_{\frac{3}{2}}$ state lies higher. Solving the Dirac equation gives this result immediately.

Hydrogen: hyperfine interaction

Nuclei have the same connection of magnetic moment and spin as electrons. The nuclear magnetic moment is parameterized as

$$\vec{M} = \frac{Zeg_N}{2M_N} \vec{I} \quad (9.81)$$

(where \vec{I} is the nuclear spin, *not* isospin!) and g_N is the gyromagnetic ratio: for the proton it is 2.79 and for the neutron it is $-\frac{2}{3}g_{pro} = -1.91$. The magnetic moment of the nucleus generates a magnetic field which couples to the electron's spin, giving a "hyperfine" splitting. The Hamiltonian is familiar,

$$\hat{H}_{hfs} = \frac{e}{mc} \vec{S} \cdot \vec{B}. \quad (9.82)$$

A short calculation gives the connection between \vec{M} and \vec{B} . It begins with the vector potential of a magnetic dipole,

$$\vec{A} = -\vec{M} \times \vec{\nabla} \left(\frac{1}{r} \right). \quad (9.83)$$

The magnetic field, $\vec{B} = \vec{\nabla} \times \vec{A}$, is

$$\vec{B} = -\vec{\nabla} \times \left(\vec{M} \times \vec{\nabla} \left(\frac{1}{r} \right) \right) = -\vec{M} \nabla^2 \left(\frac{1}{r} \right) + \vec{\nabla} \left(\vec{M} \cdot \vec{\nabla} \left(\frac{1}{r} \right) \right). \quad (9.84)$$

Thus hyperfine Hamiltonian is

$$\hat{H}_{hfs} = \frac{Ze^2 g_N}{2M_N mc^2} \left((\vec{S} \cdot \vec{\nabla})(\vec{I} \cdot \vec{\nabla}) \frac{1}{r} - \vec{S} \cdot \vec{I} \nabla^2 \left(\frac{1}{r} \right) \right). \quad (9.85)$$

We now specialize to the ground state energy of Hydrogen. The wave function is

$$|\psi_{gs}\rangle = \psi_{100}(\vec{r}) |s, m_s, I, m_I\rangle. \quad (9.86)$$

Summing repeated indices, the energy shift is

$$\Delta E_{hfs} = \frac{e^2 g_N}{2M_N m c^2} \langle s, m_s, I, m_I | \int d^3 r |\psi_{100}(\vec{r})|^2 \left(S_i I_j \frac{d}{dx_i} \frac{d}{dx_j} \frac{1}{r} - \vec{S} \cdot \vec{I} \nabla^2 \left(\frac{1}{r} \right) \right) |s, m_s, I, m_I\rangle. \quad (9.87)$$

The first term of this expression as

$$J_{ij} = \int d^3 x |\psi_{100}(\vec{x})|^2 \frac{d}{dx_i} \frac{d}{dx_j} \frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}}. \quad (9.88)$$

If $i \neq j$ the integral vanishes by parity. Thus J_{ij} is proportional to δ_{ij} . Because the ground state wave function of Hydrogen is spherically symmetric, the term is

$$J_{ij} = \delta_{ij} \frac{1}{3} \int d^3 r |\psi_{100}(\vec{r})|^2 \sum_{i=1}^3 \frac{d^2}{dx^2} \left(\frac{1}{r} \right) = \delta_{ij} \frac{1}{3} \int d^3 r |\psi_{100}(\vec{r})|^2 \nabla^2 \left(\frac{1}{r} \right). \quad (9.89)$$

Recalling that $\nabla^2 \frac{1}{r} = -4\pi \delta(\vec{r})$, the integral over the Dirac delta function yields

$$J = -\frac{4\pi}{3} |\psi_{100}(0)|^2, \quad (9.90)$$

which combines with the second term to give

$$\Delta E_{hfs} = \frac{e^2 g_N}{2M_N m c^2} \frac{8\pi}{3} |\psi_{100}(0)|^2 \langle \vec{S} \cdot \vec{I} \rangle. \quad (9.91)$$

Since $|\psi_{100}|^2 = 1/(\pi a_0^3) = (m^3 c^3 \alpha^3)/(\pi \hbar^2)$, we have

$$\Delta E_{hfs} = g_N \frac{m}{M_N} \alpha^4 m c^2 \frac{4}{3} \langle \vec{S} \cdot \vec{I} \rangle. \quad (9.92)$$

This correction is on the order of $\frac{m}{M_N} \alpha^2 \text{Ry}$, about 1/2000 the size of the fine structure. In hydrogen both the proton and the electron are spin- $\frac{1}{2}$ particles, so the total angular momentum is

$$\vec{F} = \vec{I} + \vec{S} \quad (9.93)$$

and

$$\vec{S} \cdot \vec{I} = \frac{\hbar^2}{2} \left(F(F+1) - \frac{3}{2} \right). \quad (9.94)$$

F is either 0 or 1, so

$$\vec{S} \cdot \vec{I} = \hbar^2 \times \begin{cases} +\frac{1}{4} & F = 1 \\ -\frac{3}{4} & F = 0 \end{cases} \quad (9.95)$$

This splits the ground state of hydrogen. A calculation of the numerical factor shows that the frequency of the splitting is 1420 MHz. Transitions between the two hyperfine states generate this famous spectral line, used by astronomers to determine the density of hydrogen in interstellar clouds.

Before we turn to the next set of examples, let's pause to think about something which might seem obvious, but may not be – parameter counting. We began discussing hydrogen with the Hamiltonian for a particle (the electron) in a Coulomb potential,

$$H = \frac{p^2}{2m_e} - \frac{Ze^2}{r}. \quad (9.96)$$

All observables (energies, expectation values of functions of p and r) depend on only two parameters, m_e and the combination Ze^2 – let's say $Z\alpha$ where $\alpha = e^2/(\hbar c)$ is dimensionless. We could determine these parameters from two experimental inputs (compare them to two theoretical calculations) and then everything else we would calculate would be a prediction.

Perhaps we would like to be more accurate: a single electron atom is not a particle in a Coulomb potential, it is a two-body system. We know how to account for that; we replace the electron mass by the reduced mass $m = m_e m_N / (m_e + m_N)$. This gives more accuracy, but at the cost of adding another parameter, the mass of the nucleus m_N . We need an additional experiment to fix it.

Now we want to be more accurate and we think about fine structure. New dynamics must come in: the spin of the electron. We need to know that the electron has spin 1/2. And there is a new parameter: the magnetic moment of the electron, or maybe better to say, the electron's g -factor, $g = 2$ so $\vec{\mu} = g e / (2mc) \vec{S}$. We are up to four experimental parameters (maybe five, counting the electron's spin). For the final topic in this section there are the analogs for the nucleus: the nuclear spin (and different nuclei have different values for their spin) and the nuclear g -factor g_N . As we demand more and more accuracy in our theoretical calculations (or perhaps better to say, more and more fidelity between theory and experiment) we are often forced to add more and more inputs from experiment. (But not always: the electron's g -factor is not quite equal to 2; using the quantum theory of the electromagnetic field we can compute it more accurately – to a point – before we need new experimental inputs.) This might not sound like a good thing. Of course, we can now predict more quantities with higher accuracy, so it may be a good thing, after all. Either way, that's the way life is!

Applied magnetic fields and spectroscopy

When a single-electron atom is placed in an external magnetic field, the electron feels an interaction

$$\hat{H}_I = -\vec{\mu}_{tot} \cdot \vec{B} \quad (9.97)$$

where the total magnetic moment is a sum of an orbital and a spin part, $\vec{\mu}_{tot} = \vec{\mu}_{orb} + \vec{\mu}_{spin}$ and

$$\vec{\mu}_{orb} = -\frac{e}{2mc} \vec{L} \quad \vec{\mu}_{spin} = -\frac{e}{mc} \vec{S}. \quad (9.98)$$

Thus

$$\hat{H}_I = -\frac{e}{2mc} (\vec{L} + 2\vec{S}) \cdot \vec{B}. \quad (9.99)$$

Recall that the atomic Hamiltonian, neglecting relativistic energy, but keeping the spin-orbital interaction, is

$$\hat{H} = \frac{p^2}{2m} - \frac{Ze^2}{r} + \frac{1}{2m^2c^2} \frac{Ze^2}{r^2} \vec{L} \cdot \vec{S} \quad (9.100)$$

We should, in principle, treat the magnetic term and the spin-orbital term on an equal footing. This is difficult, because the preferred direction of the magnetic field breaks rotational invariance. The two easy limits to analyze are the cases $\hat{H}_I \ll \hat{H}_{s-o}$ (the Zeeman effect) or $\hat{H}_I \gg \hat{H}_{s-o}$ (the Paschen-Back effect).

For the Zeeman effect the optimum basis is states that are diagonal in j , m_j , l , and s . Let $\vec{B} = |\vec{B}|\hat{z}$. Then the perturbing Hamiltonian is

$$\hat{H}_I = \frac{eB}{2mc} (L_z + S_z + S_z) = \frac{eB}{2mc} (J_z + S_z) \quad (9.101)$$

The first order energy shift is

$$\Delta E = \frac{eB}{2mc} \langle j, m_j, l, s | J_z + S_z | j, m_j, l, s \rangle. \quad (9.102)$$

Matrix elements of the operator $(J_z + S_z)$ are readily evaluated using the projection theorem

$$\langle j, m_j, l, s | S_z | j, m_j, l, s \rangle = \frac{\langle j, m_j, l, s | \vec{J} \cdot (\vec{J} + \vec{S}) | j, m_j, l, s \rangle}{\hbar^2 j(j+1)} \langle j, m_j, l, s | J_z | j, m_j, l, s \rangle \quad (9.103)$$

The identity

$$(\vec{J} - \vec{S})^2 = J^2 + S^2 - 2\vec{J} \cdot \vec{S} = L^2 \quad (9.104)$$

yields

$$\vec{J} \cdot (\vec{J} + \vec{S}) = J^2 + \frac{1}{2}(J^2 - L^2 + S^2) \quad (9.105)$$

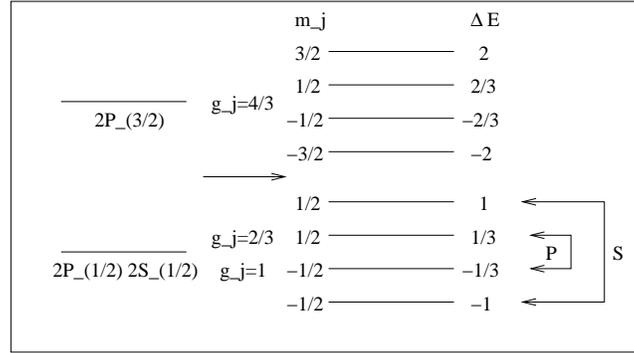


Figure 9.1: Zeeman splitting in the $n = 2$ levels of hydrogen.

and the energy shift is

$$\Delta E = \frac{eB\hbar}{2mc} \left(1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)} \right) m_j \tag{9.106}$$

The factor in the brackets is known as the Lande g-factor g_j , and the energy shift is given compactly as

$$\Delta E = \frac{eB\hbar}{2mc} g_j m_j. \tag{9.107}$$

For a single electron atom, the total angular momentum will be either $j = l + \frac{1}{2}$ or $j = l - \frac{1}{2}$. For these special cases the Lande g-factor is

$$l + \frac{1}{2} \rightarrow g_j = 1 + \frac{(l + \frac{1}{2})(l + \frac{3}{2}) - l(l+1) + \frac{3}{4}}{2(l + \frac{1}{2})(l + \frac{3}{2})} = 1 + \frac{1}{2l+1} \tag{9.108}$$

$$l - \frac{1}{2} \rightarrow g_j = 1 + \frac{(l - \frac{1}{2})(l + \frac{1}{2}) - l(l+1) + \frac{3}{4}}{2(l - \frac{1}{2})(l + \frac{1}{2})} = 1 - \frac{1}{2l+1} \tag{9.109}$$

As an example, consider the $n = 2$ levels of Hydrogen. Fig. 9.1 shows the splitting in units of $eB\hbar/(2mc)$. The magnetic field breaks degeneracy between $2P_{1/2}$ and $2S_{1/2}$ levels.

The other limit is one of strong magnetic field. In this case we can either neglect the spin-orbit coupling term, or treat it as a perturbation to the Hamiltonian after diagonalizing the magnetic moment term. Again we assume the magnetic field is in the z-direction. The interaction Hamiltonian is

$$\hat{H}_I = \frac{e\hbar B}{2mc} (L_z + 2S_z) \tag{9.110}$$

where χ is the electric susceptibility. The change in energy is

$$\Delta E = \frac{1}{2} \vec{P} \cdot \vec{E} = \frac{1}{2} \chi E^2 \quad (9.115)$$

This energy shift should arise as a second-order perturbative result. It is

$$E_{100}^{(2)} = e^2 E^2 \sum_{nlm \neq 100} \frac{|\langle n, l, m | z | 100 \rangle|^2}{E_{100} - E_{nlm}}. \quad (9.116)$$

The matrix element is

$$\langle n, l, m | z | 100 \rangle = \int r^2 d^3 r d\Omega \phi_{nlm} r \cos(\theta) \phi_{100}. \quad (9.117)$$

Its angular part is

$$\int d\Omega Y_{lm}^* \cos(\theta) Y_{00}. \quad (9.118)$$

Because $\frac{1}{\sqrt{3}} Y_{10} = \cos(\theta) Y_{00}$, the angular integral is

$$\int d\Omega Y_{lm}^* \cos(\theta) Y_{00} = \frac{1}{\sqrt{3}} \delta_{l,1} \delta_{m,0} \quad (9.119)$$

Thus, the matrix element from the ground state is non vanishing only to the $m = 0$ state. Physically this happens because the perturbation conserves L_z .

The evaluation of the sum of terms involving $|\langle n00 | z | 100 \rangle|^2$ is nontrivial. Parameterizing the result as $|\langle n00 | z | 100 \rangle|^2 = a_0^2 f(n)$, where $f(n)$ is some dimensionless function of n , the series becomes

$$\Delta E_{100}^{(2)} = -e^2 E^2 a_0^2 \sum_{n=2}^{\infty} \frac{f(n)}{\frac{1}{2} \alpha^2 m c^2 \left(1 - \frac{1}{n^2}\right)} = -2a_0^3 E^2 \sum_{n=2}^{\infty} \frac{f(n) n^2}{n^2 - 1} \quad (9.120)$$

The induced dipole moment is related to the energy shift by $\Delta E = -\vec{p} \cdot \vec{E}$, so

$$p_z = -4E a_0^3 \sum_n \frac{f(n) n^2}{n^2 - 1}. \quad (9.121)$$

The electric susceptibility (per atom) is then

$$\chi = \frac{\partial \vec{p}}{\partial \vec{E}} = a_0^3 \sum_n \frac{f(n) n^2}{n^2 - 1}. \quad (9.122)$$

It happens that the dimensionless series in Eq. 9.122 sums to a value of 1.125. The contribution of particular terms is shown in Table 9.

n	$\sum_n \frac{f_n}{1 - \frac{1}{n^2}}$
2	0.74
3	0.10
\vdots	\vdots
	1.125

While the precise value is not easy to obtain, it is possible to find upper and lower bounds on the sum using simple arguments.

The $n = 2$ term by itself gives a lower bound on the sum. For an upper bound, consider the inequality

$$\sum_{n=2}^{\infty} \frac{|\langle 100 | \hat{H}_I | nlm \rangle|^2}{E_1 - E_n} \leq \frac{1}{E_1 - E_2} \sum_n |\langle 100 | \hat{H}_I | nlm \rangle|^2. \quad (9.123)$$

Completeness collapses the sum to

$$\frac{1}{E_1 - E_2} \sum_{nlm} \langle 100 | \hat{H}_I | nlm \rangle \langle nlm | \hat{H}_I | 100 \rangle = \frac{1}{E_1 - E_2} \langle 100 | \hat{H}_I^2 | 100 \rangle. \quad (9.124)$$

Because of rotational invariance,

$$\langle 100 | z^2 | 100 \rangle = \langle 100 | x^2 | 100 \rangle = \langle 100 | y^2 | 100 \rangle = \frac{1}{3} \langle 100 | r^2 | 100 \rangle = a_0^2. \quad (9.125)$$

Our inequality is

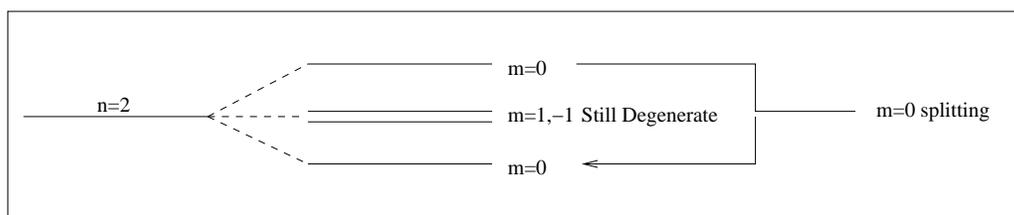
$$\begin{aligned} \Delta E &> \frac{e^2 \vec{E}^2}{E_{100} - E_{210}} \sum_{nlm} |\langle 100 | z | nlm \rangle|^2 \\ &= \frac{e^2 \vec{E}^2}{E_{100} - E_{210}} \langle 100 | z^2 | 100 \rangle \\ &= -2\vec{E}^2 a_0^3 \frac{1}{1 - \frac{1}{4}}. \end{aligned} \quad (9.126)$$

$$(9.127)$$

In summary, we have found

$$0.74 < \sum_n \frac{f_n}{1 - \frac{1}{n^2}} < 1.333. \quad (9.128)$$

So much for the ground state. Now consider the Stark effect in the $n = 2$ sector. Neglecting spin, this level is fourfold degenerate (one S-wave and 3 P-wave states). The

Figure 9.3: Stark splitting in the $n = 2$ levels of hydrogen.

Hamiltonian for the interaction is

$$H = \begin{bmatrix} & |211\rangle & |21-1\rangle & |200\rangle & & |210\rangle \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \langle 210|z|200\rangle \\ 0 & 0 & \langle 200|z|210\rangle & 0 \end{bmatrix} & |211\rangle \\ & |21-1\rangle \\ & |200\rangle \\ & |210\rangle \end{bmatrix} \quad (9.129)$$

We have a problem in degenerate state perturbation theory, whose resolution involves diagonalizing the degenerate subspace. This is straightforward. The eigenkets and eigenenergies are

$$|\psi_{\pm}\rangle = \frac{1}{\sqrt{2}}(|200\rangle \pm |210\rangle) \quad (9.130)$$

$$E_{\pm} = \mp 3a_0. \quad (9.131)$$

Because the matrix element $\langle 2lm|z|2l'm'\rangle$ is only non-zero for $m' = m$ and $l' = 1, l = 0$ or $l' = 0, l = 1$, the other two eigenstates remain degenerate, unaffected by the perturbation. The pattern of splitting is illustrated in Fig. 9.3. The spin-orbit term would lift the remaining degeneracy.

Helium

Keeping only Coulomb terms, the Hamiltonian for the two electrons of Helium is

$$H = \frac{p_1^2}{2m} - \frac{Ze^2}{r_1} + \frac{p_2^2}{2m} - \frac{Ze^2}{r_2} + \frac{e^2}{|\vec{r}_2 - \vec{r}_1|} = H_1 + H_2 + V. \quad (9.132)$$

Neglecting spin, statistics, and the inter-electron repulsion, we take as the unperturbed Hamiltonian $H_0 = H_1 + H_2$ and use its eigenfunctions as our zeroth-order states. The

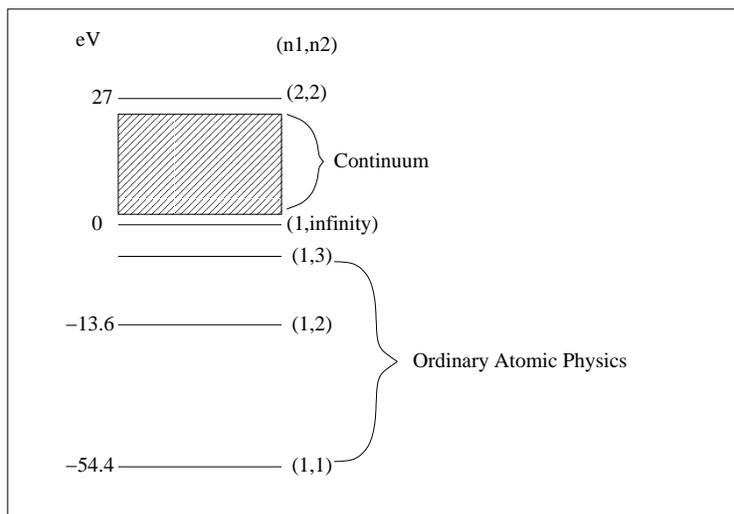


Figure 9.4: Spectrum of helium.

potential term is treated perturbatively. The unperturbed wave function is a product of hydrogenic wave functions

$$\psi(r_1, r_2) = u_{n_1 l_1 m_1}(r_1) u_{n_2 l_2 m_2}(r_2) \quad (9.133)$$

and the zeroth-order energies are just

$$E = E_{n_1} + E_{n_2} = \frac{-1}{2} mc^2 (Z\alpha)^2 \left(\frac{1}{n_1^2} + \frac{1}{n_2^2} \right). \quad (9.134)$$

This gives a ground state energy for He of approximately -108.8 eV. Note that ionization energy (the cost to completely remove one electron while leaving the other in the $n = 1$ state) is exactly half this, -54.4 eV. Taking this value as the zero point energy of the system, we plot the excited states spectrum in Fig. 9.4. Notice that if both electrons are raised above the $n = 1$ state, the atom can “auto-ionize” – the lowest energy state is in the ionization continuum. Therefore, these states are intrinsically unstable. To describe an ordinary atom, we must leave at least one electron in its ground state.

In Chapter 8, we described the interplay of spin and statistics on the spectrum of helium. Recall that electrons are spin- $\frac{1}{2}$ fermions. The total spin is therefore $S = 1$ or $S = 0$.

For ortho-Helium, the electrons have $S = 1$. These spin states are spin symmetric, so the spatial wave function must be totally anti-symmetric. This forbids double occupancy in the ground state. Para-Helium states have $S = 0$. This state is antisymmetric, so the spatial wave function must be symmetric. The ground state of Helium must be para-He.

We now consider the effect of Coulomb repulsion of the electrons, treated as a perturbation, on the energy levels. Take the ground state, first. The unperturbed ground state wave function is

$$\psi(\vec{r}_1, \vec{r}_2) = U_{100}(r_1)U_{100}(r_2) \frac{|\uparrow\rangle |\downarrow\rangle - |\downarrow\rangle |\uparrow\rangle}{\sqrt{2}} \quad (9.135)$$

and

$$U_{100}(r) = \frac{2}{\sqrt{4\pi}} \left(\frac{Z}{a_0}\right)^{3/2} e^{-\frac{Zr}{a_0}}. \quad (9.136)$$

The first order energy shift is

$$\begin{aligned} \Delta E &= e^2 \int d^3r_1 d^3r_2 \frac{|U_{100}(r_1)|^2 |U_{100}(r_2)|^2}{|\vec{r}_1 - \vec{r}_2|} \\ &= \left(\frac{Z}{a_0}\right)^6 \frac{e^2}{\pi^2} \int r_1^2 dr_1 d\Omega_1 r_2^2 dr_2 d\Omega_2 \frac{\exp[-2Z/a_0(r_1 + r_2)]}{|\vec{r}_1 - \vec{r}_2|}. \end{aligned} \quad (9.137)$$

To evaluate this expression, expand $1/|\vec{r}_1 - \vec{r}_2|$ in Legendre polynomials. Pick the z -axis along either r_1 or r_2 . we have

$$\Delta E = \left(\frac{Z}{a_0}\right)^6 \frac{e^2}{\pi^2} \sum_{l=0}^{\infty} \int r_1^2 dr_1 d\Omega_1 r_2^2 dr_2 d\Omega_2 \exp[-2Z/a_0(r_1 + r_2)] \frac{r_1^l}{r_1^{l+1}} P_l(\cos\theta) \quad (9.138)$$

Only the $l = 0$ term will survive the sum. The integral is straightforward and yields

$$\Delta E = \frac{5Ze^2}{8a_0} = \frac{5}{4}Z \times Ry. \quad (9.139)$$

(We could have anticipated the overall scale of e^2/a_0 on dimensional grounds.) Then the first order shift in ground state energy is

$$E_{ground} = E_0 + \Delta E = -108.8 \text{ eV} + 34 \text{ eV} = -74.8 \text{ eV} \quad (9.140)$$

(with respect to the energy with both electrons removed, or $-74.8 + 54.4 = -20.4$ eV compared to the 1SHe^+ ion. This is not quite the experimental result of $E_{ground} = -78.975$ eV. We will revisit this result in Chapter 10 and improve on it.

Now we turn to the excited states. The para-wave function is

$$U(1,2) = \frac{U_a(1)U_b(2) + U_b(1)U_a(2)}{\sqrt{2}} \quad (9.141)$$

while the ortho-wave function is

$$U(1,2) = \frac{U_a(1)U_b(2) - U_b(1)U_a(2)}{\sqrt{2}}. \quad (9.142)$$

The labels a and b are the set of quantum numbers associated with those states.

The first order energy shift is

$$\begin{aligned}
 \Delta E_1^{(\pm)} &= \frac{e^2}{2} \int d^3r_1 d^3r_2 \frac{1}{|\vec{r}_1 - \vec{r}_2|} (U_a(1)U_b(2) \pm U_b(1)U_a(2))^* (U_a(1)U_b(2) \pm U_b(1)U_a(2)) \\
 &= e^2 \int d^3r_1 d^3r_2 \frac{1}{|\vec{r}_1 - \vec{r}_2|} [|U_a(1)|^2 |U_b(2)|^2 \pm (U_a^*(1)U_b(1)) (U_a(2)U_b^*(2))] \\
 &= J \pm K
 \end{aligned} \tag{9.143}$$

where the (+) is for the ortho states, and the (-) is for the para states. J is called the “direct term” and has a classical interpretation as the Coulomb interaction of two charge densities $\rho_a(r) \equiv -e|U_a(r)|^2$.

$$J = \int d^3r_1 \int d^3r_2 \frac{\rho_a(r_1)\rho_b(r_2)}{|\vec{r}_1 - \vec{r}_2|}. \tag{9.144}$$

K is called the “exchange energy.” It has no classical analog.

Although spin had nothing to do with the energy difference between para and ortho states, we recall that we can write the energy shift as an effective spin-spin interaction

$$\Delta E_{nl} = J_{nl} - \frac{1}{2}(1 + \sigma_1 \cdot \sigma_2)K_{nl}. \tag{9.145}$$

$\sigma_1 \cdot \sigma_2 = 1$ or -3 for $S = 1$ or 0 , respectively. K_{nl} will be a number of the order of electron volts. It is a much larger number than an intrinsic dipole-dipole interaction, which is the size of a fine structure splitting. The large spin-dependent energies seen in magnetic materials arise due to the effects of statistics.

J is certainly positive. but what about K ? If $l = n - 1$ then the U_{nlm} 's have no nodes, implying that K is also positive. Physically, we would also expect $K > 0$ because the $S=1$ states would tend to be spread out more than the $S = 0$ ones. This physical argument forms the basis for Hund's rule for the spectroscopy of multi-electron atoms: “Higher spin states are shifted lower in energy,” Like all good rules, it has exceptions.

Chapter 10

Variational methods

Consider replacing the true ground state wave function $|\psi_0\rangle$ associated with a Hamiltonian H by any other trial wave function $|\psi\rangle$, which obeys the same boundary conditions as $|\psi_0\rangle$. It is a remarkable fact that the expectation value of the Hamiltonian in the trial state $|\psi\rangle$ is an upper bound on the exact ground state energy of H :

$$E_{trial} = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} \geq E_{true} \quad (10.1)$$

Furthermore, $E_{trial} = E_{true}$ if and only if the trial wave function happens to be equal to the true ground state wave function. To prove this result, expand the normalized trial state in a superposition of energy eigenstates

$$|\psi\rangle = \sum c_n |\psi_n\rangle \quad (10.2)$$

and compute

$$\langle\psi|H|\psi\rangle = \sum_{m,n} c_m^* c_n \langle\psi_m|H|\psi_n\rangle. \quad (10.3)$$

Orthogonality collapses the sum to

$$\langle\psi|H|\psi\rangle = \sum_n |c_n|^2 E_n \quad (10.4)$$

and since all the E_n 's are greater than E_0 ,

$$\langle\psi|H|\psi\rangle \geq E_0 \sum_n |c_n|^2 \quad (10.5)$$

or

$$\langle\psi|H|\psi\rangle \geq E_0 \quad (10.6)$$

since the sum is unity.

This remarkable fact has practical consequences: Variational methods form a vast and important part of the literature of approximate solutions of the Schrödinger equation. In many cases the bounds are coarse and only of qualitative use. However, one can use the variational principle to construct better and better approximations to the ground state wave function. The energy tells whether one wave function is “better” than another by giving the winner’s energy a lower value. Sometimes these calculations can be made extraordinarily precise.

There are various ways to exploit the variational principle: In the most commonly seen case, the the variational wave function is characterized by a set of parameters, α_i . The energy is extremized by the constraint that

$$\frac{\partial E(\alpha)}{\partial \alpha_i} = 0 \quad (10.7)$$

The solution of these equations give the “best” (minimizing) α_i 's (and hence the best trial wave function accessible to the parameterization). Generally, the normalization condition of the state is α dependent. We can work directly with Eq. 10.7, or we can define

$$E \equiv \frac{\langle \psi(\alpha) | \hat{H} | \psi(\alpha) \rangle}{\langle \psi(\alpha) | \psi(\alpha) \rangle}. \quad (10.8)$$

The minimization condition becomes

$$\frac{\partial E}{\partial \alpha_i} = \frac{1}{\langle \psi | \psi \rangle} \frac{\partial}{\partial \alpha_i} \langle \psi | \hat{H} | \psi \rangle - \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle^2} \frac{\partial}{\partial \alpha_i} \langle \psi | \psi \rangle = 0 \quad (10.9)$$

or

$$\frac{\partial}{\partial \alpha_i} \langle \psi | \hat{H} | \psi \rangle - E \frac{\partial}{\partial \alpha_i} \langle \psi | \psi \rangle = 0. \quad (10.10)$$

The reader might recognize this as a constrained minimization problem, where E plays the role of a Lagrange multiplier.

Variational calculations can be used for excited states, but in that case the trial wave function must be orthogonal to the exact ground state, and to all states with lower energy than the one desired, for all values of the variational parameters. If that is the case, then one can replace the E_0 of Eq. 10.5 by

$$\langle \psi | \hat{H} | \psi \rangle = \sum_j E_j |c_j|^2 \geq E_1 \sum_j |c_j|^2 \quad (10.11)$$

and one has a variational bound. Otherwise, the bound on the energy is only that it is greater than the ground state. If the variational wave function can do so, it will prefer to increase the amplitudes of lower energy states: the variational wave function always finds the lowest energy that it can. Sometimes, it is easy to orthogonalize the trial wave function against other states. For example, in a central potential problem, an $l = 1$ trial wave function will bound the energy of the lowest $l = 1$ state, regardless of whether there is a lower $l = 0$ state. Sometimes, it is hard to do this. An example would be to find a bound on the second-lowest even-parity state.

Typically, variational results for energies are much better than variational results for matrix elements. To see this, suppose that the trial wave function consists of the ground state plus a small admixture of some orthogonal excited state, $|\psi\rangle = |\psi_0\rangle + \epsilon |\psi_1\rangle$ Then

$$\frac{\langle \psi(\alpha) | \hat{H} | \psi(\alpha) \rangle}{\langle \psi(\alpha) | \psi(\alpha) \rangle} = \frac{E_0 + \epsilon^2 E_1}{1 + \epsilon^2} = E_0 + O(\epsilon^2) \quad (10.12)$$

If the ground state wave function is known to ten per cent then the true ground state energy can be computed to within one per cent. Expectation values for operators are less forgiving:

$$\frac{\langle \psi | \hat{O} | \psi \rangle}{\langle \psi | \psi \rangle} = \langle \psi_0 | \hat{O} | \psi_0 \rangle + \epsilon \left(\langle \psi_0 | \hat{O} | \psi_1 \rangle + \langle \psi_1 | \hat{O} | \psi_0 \rangle \right) + \dots \quad (10.13)$$

The expectation value of the operator can be determined to within ten per cent, and (unlike the for energy) the sign of the extra piece is unknown. Exceptions to this rule (usually only of academic interest) occur if the operator is anti-Hermitian, or if the operator is diagonal in the trial (and true) bases.

A simple harmonic oscillator example

We assume the standard Hamiltonian $\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega\hat{x}^2$, and attempt to variationally bound the ground state energy using the trial wave function $|\psi\rangle = Ne^{-\beta x^2/2}$. The expectation value of the potential term of the Hamiltonian is

$$\langle x^2 \rangle = \frac{\int dx x^2 e^{-\beta x^2}}{\int dx e^{-\beta x^2}}. \quad (10.14)$$

Differentiating under the integral, we write

$$\int dx x^2 e^{-\beta x^2} = -\frac{\partial}{\partial \beta} \int dx e^{-\beta x^2} \quad (10.15)$$

so

$$\langle x^2 \rangle = -\sqrt{\beta} \frac{\partial}{\partial \beta} \frac{1}{\sqrt{\beta}} = \frac{1}{2\beta}. \quad (10.16)$$

Similarly,

$$\frac{\partial^2 \psi}{\partial x^2} = (-\beta + \beta^2 x^2) e^{-\beta x^2} \quad (10.17)$$

and the expectation value of the kinetic energy is

$$\langle \text{KE} \rangle = -\frac{\hbar^2}{2m} (-\beta + \beta^2 \langle x^2 \rangle) = \frac{\hbar^2 \beta}{4m}. \quad (10.18)$$

The total energy is $E(\beta) = \frac{\hbar^2 \beta}{4m} + \frac{m\omega^2}{4\beta}$. The β -dependence of the energy is shown in Fig. 10.1. Obviously, there is a minimizing value of $\beta = m\omega/\hbar$, for which $E(\beta) = \hbar\omega/2$. Of course, this happens to be the exact result, because one of the members of the family of trial solutions happened to be the exact ground state.

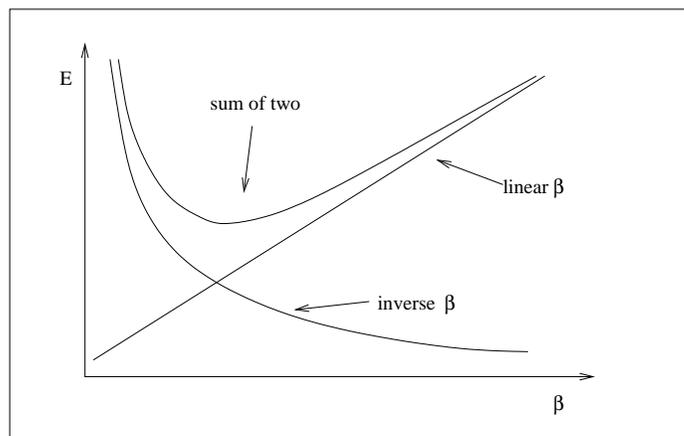


Figure 10.1: Dependence of the oscillator energy on the variational parameter.

Helium ground state energy

Recall that the calculation of the ground state energy of helium, which treated the inter-electron repulsion perturbatively, did not agree very well with experiment. A variational calculation should be able to improve this result. The Hamiltonian for helium was

$$\hat{H} = \frac{\hat{p}_1^2}{2m} - \frac{Ze^2}{r_1} + \frac{\hat{p}_2^2}{2m} - \frac{Ze^2}{r_2} + \frac{e^2}{|\vec{r}_1 - \vec{r}_2|}. \quad (10.19)$$

Let us assume that the trial wave function is separable, $\psi(r_1, r_2) = \psi(r_1)\psi(r_2)$ with

$$\psi(r) = \frac{1}{\sqrt{4\pi}} \left(\frac{Z^*}{a_0} \right)^{3/2} \exp(-Z^*r/a_0). \quad (10.20)$$

This is a product of hydrogen-like wave functions, and the variational parameter is Z^* . As a trick to do this calculation, notice that if

$$\left(\frac{\hat{p}^2}{2m} - \frac{Z^*e^2}{r} \right) \psi(r) = \epsilon \psi(r) \quad (10.21)$$

then

$$\epsilon = \frac{1}{2} Z^* \alpha^2 m_e c^2. \quad (10.22)$$

The expectation value of the energy in the trial state is

$$E(Z^*) = \int d^3r_1 d^3r_2 \psi^*(r_1) \psi^*(r_2) \left(\frac{\hat{p}_1^2}{2m} - \frac{Ze^2}{r_1} + \frac{\hat{p}_2^2}{2m} - \frac{Ze^2}{r_2} + \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \right) \psi(r_1) \psi(r_2) \quad (10.23)$$

Rewriting

$$\frac{p^2}{2m} - \frac{Ze^2}{r} = \frac{p^2}{2m} - \frac{Z^*e^2}{r} - \frac{(Z - Z^*)e^2}{r}, \quad (10.24)$$

this energy is

$$E(Z^*) = 2\epsilon - 2(Z - Z^*)e^2 \left\langle \frac{1}{r} \right\rangle_{Z^*} + e^2 \left\langle \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right\rangle_{Z^*}. \quad (10.25)$$

The virial theorem tells us that

$$e^2 \left\langle \frac{1}{r} \right\rangle_{Z^*} = \frac{Z^*e^2}{a_0} = Z^*\alpha^2 m_e c^2. \quad (10.26)$$

We previously found (see Eq. 9.139) that

$$e^2 \left\langle \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right\rangle_{Z^*} = \frac{5}{4} Z^* \frac{1}{2} \alpha^2 m_e c^2 \quad (10.27)$$

Thus the total energy is

$$E(Z^*) = -\frac{1}{2} \alpha^2 m_e c^2 \left(-2Z^{*2} + 4ZZ^* - \frac{5}{4}Z^* \right). \quad (10.28)$$

Minimizing with respect to Z^* gives $Z^* = Z - 5/16$ and

$$E = -\frac{1}{2} \alpha^2 m_e c^2 \left(2\left(Z - \frac{5}{16}\right) \right)^2 \quad (10.29)$$

There is an obvious physical interpretation of this result: Z^* can be thought of as an effective nuclear charge seen by each electron, which is suppressed because of screening by the other electron. The variational ground state energy of Helium is $E_0 = -77.38$ eV, as compared to the true ground state energy of -78.975 eV and the result from perturbation theory of -74.8 eV. A better trial wave function could be used to lower the variational bound still more: Bethe and Jackiw offer some examples.

Ritz variational method

The Ritz variational method is often used in research, but the author has not found it to be well documented in elementary texts. The idea is to take a restricted (or finite) basis of states, which we assume to be orthonormal, and write our trial state as a superposition of them

$$|\psi\rangle = \sum_{i=1}^N \alpha_i |j\rangle. \quad (10.30)$$

We diagonalize the Hamiltonian in the basis. This gives us a set of N energies $\{E\} = E_0, E_1, \dots, E_{N-1}$. The lowest energy is a variational upper bound on the true ground state energy.

The proof of this statement uses Eq. 10.9. The α_j 's are variational parameters (note that we regard α_j^* and α_j as different). We seek an extremum while regarding E as the Lagrange multiplier which preserves normalization

$$\frac{\partial}{\partial \alpha_i^*} \left\{ \langle \psi | \hat{H} | \psi \rangle - E \langle \psi | \psi \rangle \right\} = 0 \quad (10.31)$$

Noting that

$$\langle \psi | \psi \rangle = \sum_{m=1}^N \alpha_m^* \alpha_m \quad (10.32)$$

and writing

$$\langle \psi | \hat{H} | \psi \rangle = \sum_{m,n} \alpha_m^* \alpha_n h_{mn} \quad (10.33)$$

with $h_{mn} = \langle m | \hat{H} | n \rangle$, Eq. 10.31 becomes

$$\frac{\partial}{\partial \alpha_i^*} \left\{ \sum_{mn} \alpha_m^* h_{mn} \alpha_n - \sum_m \alpha_m^* \alpha_m E \right\} = 0 \quad (10.34)$$

whose solution involves the set of N linear equations

$$h_{in} \alpha_n - \delta_{in} E \alpha_n = 0. \quad (10.35)$$

There is only a solution when

$$\det(h_{nm} - \delta_{nm} E) = 0. \quad (10.36)$$

This is, of course, the eigenvalue equation in the restricted basis. Its lowest energy solution is thus our variational upper bound.

As a particular case of this equation, recall the formula for the correction to the ground state energy from second-order perturbation theory.

$$E_n = E_n^0 + \sum_{l \neq n} \frac{|\langle n | \hat{H}_1 | l \rangle|^2}{E_n^0 - E_l^0}. \quad (10.37)$$

Note that the contribution of every term to the lowest state's energy is negative. Why is this? Perturbation theory represents an approximate diagonalization of the Hamiltonian

matrix. Suppose we restrict the sum to a finite basis. The approximate ground state wave function will be given by the familiar sum

$$|\psi_0\rangle = |\psi_0^0\rangle + \sum_{i \neq 0}^N c_i |\psi_i^0\rangle. \quad (10.38)$$

This sum is equivalent to a variational choice for the wave function. As we add more states (increasing N) the wave function must improve, and the variational bound on the energy must become lower and lower.

A research example

I spend a lot of my time (and a computer's time) finding eigenmodes of complicated Hamiltonians. In my problems, space is discretized on a lattice, the wave functions are column vectors whose length is the number of discretization points, and the Hamiltonians are very large sparse matrices acting on the vectors. Here is an algorithm (see T. Kalkreuter and H. Simma, *Comp. Phys. Comm.* 93, 33 (1996)) which I find useful:

Suppose that we are given a Hamiltonian and we want to find the N lowest energies wave functions corresponding to those energies. We start off with an initial set of trial wave functions $\{\psi_0^J\}$. We then perform the following sequence of operations to refine them:

- 1) Diagonalize the Hamiltonian in the subspace of the set of wave functions $\{\psi_0^J\}$.
- 2) Take the lowest energy wave function (call it $|\phi_0\rangle$) from the set. Construct an orthogonal wave function by the application of H followed by projection

$$|\phi_1\rangle = \hat{H} |\phi_0\rangle - |\phi_0\rangle \langle \phi_0 | \hat{H} | \phi_0 \rangle \quad (10.39)$$

and normalize it. Then diagonalize the Hamiltonian in the 2×2 subspace of the states $|\phi_0\rangle$ and $|\phi_1\rangle$. Find the lowest energy state (call it $|\phi'_0\rangle$).

- 3) Replace $|\phi_0\rangle$ by $|\phi'_0\rangle$ and repeat Step 2 some number of times, until the change in energy of the state is small.
- 4) After we are confident that our lowest energy state is reasonably converged, take the first excited state from step 1 and orthogonalize it (using Gram-Schmidt or some other finite vector orthogonalization process) with respect to the lowest energy state. Then

go to step 2 and repeat the entire process for the excited state. Orthogonalize with respect to the ground state, as well. When the first state seems stable, repeat for the third state, and so on. During the construction of states for the N th level, all trial states must be orthogonalized against all the lower energy states.

- 5) When all N states have been “improved,” return to step 1 with them, and repeat the whole procedure.

Typically, the quality of the variational state degrades as its energy rises, so if N states are needed, it might be a good idea to work in a basis which is larger than N , and throw away the extra states.

In contrast to low eigenmodes, finding the eigenmode of a matrix H with the highest eigenvalue, when the spectrum of H is bounded from above, is simple. We start by taking any trial vector $|\psi\rangle$. It has the usual expansion in energy eigenfunctions of H , $|\psi\rangle = \sum_j \alpha_j |j\rangle$. Now consider the state constructed by repeated application of H . It is

$$|\psi_N\rangle = \hat{H}^N |\psi\rangle = \sum_j \alpha_j E_j^N |j\rangle \quad (10.40)$$

The state with the largest coefficient in the sum is the one with the largest eigenvalue, so $|\psi_N\rangle$ is merely this state, rescaled, plus small corrections.

Chapter 11

Time dependent perturbation theory

Just as we did for energy levels, approximate solutions may be constructed for the time dependent Schrödinger equation. The story begins similarly: We assume that the Hamiltonian can be written as $\hat{H} = \hat{H}_0 + \hat{V}$, where H_0 is “large” in some sense, time independent, and can be diagonalized. We further assume that V may be regarded as a perturbation. The time evolution will be developed in terms of the eigenfunctions of H_0 .

It is very useful to replace the time dependent wave function by a time evolution operator, and to find matrix elements of that operator. It is defined through

$$|\psi(t)\rangle = \hat{U}(t, t_0) |\psi(t_0)\rangle \quad (11.1)$$

and it obeys an equation of motion

$$i\hbar \frac{\partial \hat{U}(t, t_0)}{\partial t} = \hat{H} \hat{U}(t, t_0). \quad (11.2)$$

Matrix elements of U connect initial states at some early time with final states at some later time. What they encode is called the “transition amplitude” from the initial state to the final state. It is

$$\begin{aligned} A &= \langle \beta(t) | \hat{U}(t, t_0) | \alpha(t_0) \rangle \\ &= U_{\beta\alpha}(t, t_0) \end{aligned} \quad (11.3)$$

Often, the early time is chosen to be far in the past ($t_0 \rightarrow -\infty$) and the late time is far in the future ($t \rightarrow \infty$) – “far” in the sense of some natural time scale in the problem. In order to specify the transition amplitude, we must be able to specify the initial and final states $|\alpha(t_0)\rangle$ and $|\beta(t)\rangle$. However, we can only construct the wave function if we can diagonalize the Hamiltonian. This is a problem, if $\hat{H} = \hat{H}_0 + \hat{V}$ and only H_0 can be diagonalized. The way to evade this problem is to define $V(t)$ so that it vanishes at the early or late time where we specify the wave functions. Sometimes this is easy, when V is explicitly switched on and off at specific times, for example $V(t') \propto \theta(t' - t_0) - \theta(t - t')$. Sometimes the amplitude of the perturbation is constant in time. To make sense of the calculation, in that case we must assume that the perturbation is switched on and off adiabatically at early and late times. One way to do this is to replace $V(t)$ by $V(t) \exp(-\epsilon|t|)$, and then to take ϵ to zero, cautiously.

Path integral formulation of time-dependent perturbation theory

Constructing the formulas for time-dependent perturbation theory beginning with the path integral is a bit more transparent than working directly with the equation of motion. We will

begin with this approach. We write the Hamiltonian as $\hat{H} = \hat{H}_0 + \hat{V}$. We assume that the initial and final states are eigenfunctions of \hat{H}_0 . The matrix element of the time evolution operator is

$$U_{\beta\alpha}(t) = \langle \beta | e^{-i\hat{H}t/\hbar} | \alpha \rangle. \quad (11.4)$$

We slice up the exponential into a set of N infinitesimal intervals where $N\Delta t = t$.

$$U_{\beta\alpha}(t) = \lim_{N \rightarrow \infty} \langle \beta | \left(e^{-i(\hat{H}_0 + \hat{V})\Delta t/\hbar} \right)^N | \alpha \rangle \quad (11.5)$$

Inserting complete states of \hat{H}_0 between the exponential factors gives

$$U_{\beta\alpha}(t) = \lim_{N \rightarrow \infty} \sum_{i_1} \sum_{i_2} \cdots \sum_{i_{N-1}} \langle \beta | e^{-i(\hat{H}_0 + \hat{V})\Delta t/\hbar} | i_1 \rangle \langle i_1 | e^{-i(\hat{H}_0 + \hat{V})\Delta t/\hbar} | i_2 \rangle \langle i_2 | \cdots | \alpha \rangle. \quad (11.6)$$

So far, we have not used the fact that the size of V is small compared to H_0 . Now we do so. We expand each exponential in a power series in V

$$e^{-i(\hat{H}_0 + \hat{V})\Delta t/\hbar} = e^{-i\hat{H}_0\Delta t/\hbar} \left[1 - \frac{i\Delta t}{\hbar} V + O(\Delta t^2) \right]. \quad (11.7)$$

The transition amplitude becomes

$$\begin{aligned} U_{\beta\alpha}(t) &= \lim_{N \rightarrow \infty} \sum_{i_1} \sum_{i_2} \cdots \sum_{i_{N-1}} \langle \beta | e^{-i\hat{H}_0\Delta t/\hbar} \left(1 - \frac{i\Delta t}{\hbar} V(t_n) \right) | i_1 \rangle \times \\ &\quad \langle i_1 | e^{-i\hat{H}_0\Delta t/\hbar} \left(1 - \frac{i\Delta t}{\hbar} V(t_n) \right) | i_2 \rangle \langle i_2 | \cdots | \alpha \rangle. \end{aligned} \quad (11.8)$$

Let us take this long formula and regroup it into a sum of terms order by order in powers of V . At zeroth order we have

$$U_{\beta\alpha}^{(0)}(t) = \lim_{N \rightarrow \infty} \sum_{i_1} \sum_{i_2} \cdots \sum_{i_{N-1}} \langle \beta | e^{-i\hat{H}_0\Delta t/\hbar} | i_1 \rangle \langle i_1 | e^{-i\hat{H}_0\Delta t/\hbar} | i_2 \rangle \langle i_2 | \cdots | \alpha \rangle \quad (11.9)$$

which collapses to

$$U_{\beta\alpha}^{(0)}(t) = \langle \beta | e^{-i\hat{H}_0 t/\hbar} | \alpha \rangle = \delta_{\beta\alpha} e^{-iE_\alpha t/\hbar} \quad (11.10)$$

because all the states are eigenstates of H_0 .

In first order we need one V out of the product of terms in the big sum. In each time slice there is a $1 - \frac{i\Delta t}{\hbar} V$. The result will be a sum of terms, one for each time slice. To see what we get, let us imagine that there are three time slices. Then the evolution operator is

$$U_{\beta\alpha}(t) = \sum_i \sum_j \langle \beta | e^{-i(\hat{H}_0 + V)\Delta t/\hbar} | i \rangle \langle i | e^{-i(\hat{H}_0 + V)\Delta t/\hbar} | j \rangle \langle j | e^{-i(\hat{H}_0 + V)\Delta t/\hbar} | \alpha \rangle. \quad (11.11)$$

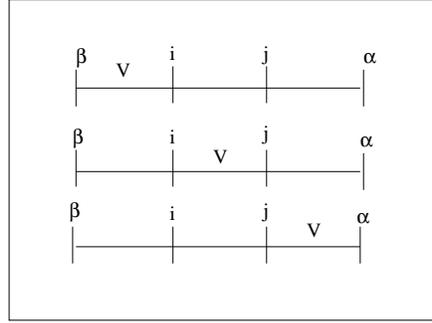


Figure 11.1: Representation of Eq. 11.12.

The order V^1 contribution is

$$\begin{aligned}
 U_{\beta\alpha}(t) = & \langle\beta|e^{-i\hat{H}_0\Delta t/\hbar} \left(\frac{i\Delta t}{\hbar}V\right) |i\rangle \langle i|e^{-i\hat{H}_0\Delta t/\hbar}|j\rangle \langle j|e^{-i\hat{H}_0\Delta t/\hbar}|\alpha\rangle + \\
 & \langle\beta|e^{-i\hat{H}_0\Delta t/\hbar}|i\rangle \langle i|e^{-i\hat{H}_0\Delta t/\hbar} \left(\frac{i\Delta t}{\hbar}V\right) |j\rangle \langle j|e^{-i\hat{H}_0\Delta t/\hbar}|\alpha\rangle + \\
 & \langle\beta|e^{-i\hat{H}_0\Delta t/\hbar}|i\rangle \langle i|e^{-i\hat{H}_0\Delta t/\hbar}|j\rangle \langle j|e^{-i\hat{H}_0\Delta t/\hbar} \left(\frac{i\Delta t}{\hbar}V\right) |\alpha\rangle. \quad (11.12)
 \end{aligned}$$

This long formula is represented pictorially by Fig. 11.1. The matrix elements in which there is no interaction (i.e. no V terms) give

$$\langle i|e^{-i\hat{H}_0\Delta t/\hbar}|j\rangle = \delta_{ij}e^{-iE_i\Delta t/\hbar}. \quad (11.13)$$

Since the initial and final states, as well as the intermediate states are eigenstates of H_0 , Fig. 11.1 has the following interpretation: In the earliest time the state begins in eigenstate $|\alpha\rangle$ and then propagates through time into state $|i\rangle$, which must be identical to it. The propagation through time without any interaction does not change the physical content of the wave function, it just introduces an overall multiplication by a phase. The state then propagates for another time Δt into state $|j\rangle$ which again must be equal to state $|i\rangle$. In the third time slice, where the transition is from $|j\rangle$ to $|\beta\rangle$ there is something new – the perturbing potential presents itself to transform the state. One can interpret the other lines in the picture similarly.

Summing over the three first-order contributions gives

$$U_{\beta\alpha}^{(1)}(t_f, t_i) = -\frac{i\Delta t}{\hbar} \sum_{\text{all slices } t_j} \langle\beta|e^{-iE_\beta(t_f-t_j)/\hbar}V(t_j)e^{-iE_\alpha(t_j-t_i)/\hbar}|\alpha\rangle. \quad (11.14)$$

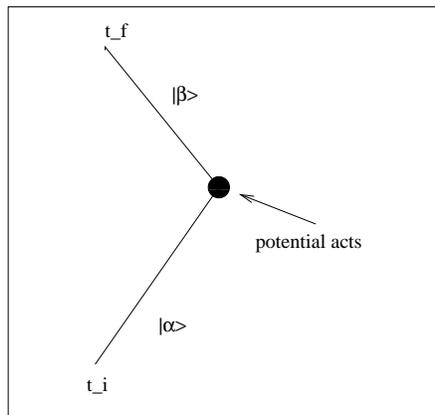


Figure 11.2: Space-time representation of the first order transition amplitude.

Converting the sum over time slices into an integral yields the first order expression for the transition amplitude

$$U_{\beta\alpha}^{(1)}(t_f, t_i) = -\frac{i}{\hbar} \int_{t_i}^{t_f} dt e^{-iE_\beta(t_f-t)/\hbar} \langle \beta | V(t) | \alpha \rangle e^{-iE_\alpha(t-t_i)/\hbar}. \quad (11.15)$$

We can represent this process by a cartoon, Fig. 11.2. The lines represent the propagation of the system in its initial and final state. The vertex represents the action of the perturbation, which transforms the system from its initial state to its final state. The interaction only acts at one point in time, but which time this is will be unconstrained. We must integrate over all times where it acts to get the full amplitude for the process.

Note that in the formula for the first order amplitude, there is no trace of any intermediate states. All we have are the initial and final state, which we specified. The system evolves from the initial time until the perturbation acts, simply by rotating its phase (the $e^{-iE_\alpha(t-t_i)}$ factor) and after the perturbation has acted and the system has gone into the final state, its phase continues to rotate, but now controlled by the energy E_β .

Next, let us go to second order. Again we illustrate the expansion using only three time slices. The representation of the amplitude is shown in Fig. 11.3. We have not shown terms which are order V^2 at a single time slice. In the $N \rightarrow \infty$ limit they represent a vanishingly small component compared to the terms with a single V on each time slice. Compare Fig. 11.4. The second order contribution to the transition amplitude is

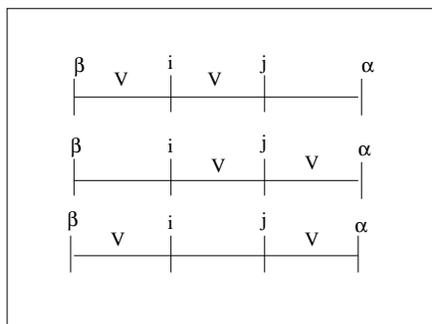


Figure 11.3: Pictorial representation of the second order transition amplitude, with three time slices.

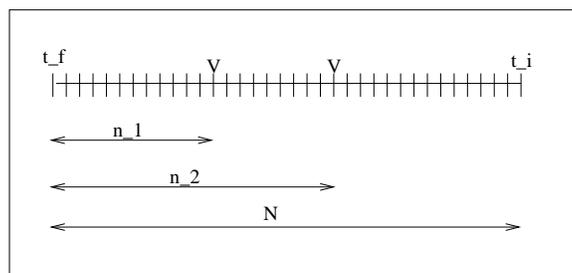


Figure 11.4: The second order amplitude with many time slices.

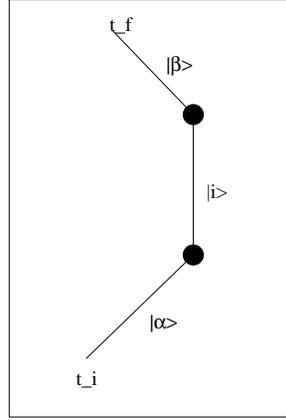


Figure 11.5: Space-time picture of the second order transition amplitude.

$$U_{\beta\alpha}^{(2)} = \left(-\frac{i\Delta t}{\hbar}\right)^2 \sum_i e^{-iE_\beta n_1 \Delta t/\hbar} \langle \beta | V(t_{n_1}) | i \rangle e^{-iE_i(n_2 - n_1) \Delta t/\hbar} \langle i | V(t_{n_2}) | \alpha \rangle e^{-iE_\alpha(N - n_2) \Delta t/\hbar}. \quad (11.16)$$

Again promoting time to a continuous variable, we have

$$U_{\beta\alpha} = \left(-\frac{i}{\hbar}\right)^2 \sum_i \int_{t_i}^{t_f} dt_2 \int_{t_i}^{t_2} dt_1 e^{-iE_\beta(t_f - t_i)/\hbar} \langle \beta | V(t_2) | i \rangle e^{-iE_i(t_2 - t_1)/\hbar} \langle i | V(t_1) | \alpha \rangle e^{-iE_\alpha(t_1 - t_i)/\hbar}. \quad (11.17)$$

We can represent this as another cartoon in Fig. 11.5. Now the perturbation acts twice, at a pair of time points. Between these points, the particle propagates as an eigenstate of H_0 .

Note the sum over intermediate states. In principle, the initial application of V could take the initial state into any intermediate state, as long as the final application of V can connect it to the final state. To groom the formula, let us define a potential (the “interaction representation” potential)

$$\hat{V}(t) = e^{i\hat{H}_0 t/\hbar} V(t) e^{-i\hat{H}_0 t/\hbar} \quad (11.18)$$

and the “interaction representation” evolution operator

$$\hat{U}(t_f, t_i) = e^{-i\hat{H}_0 t_f/\hbar} U(t_f, t_i) e^{i\hat{H}_0 t_i/\hbar}. \quad (11.19)$$

With these definitions, the entire series can be written more compactly as

$$\begin{aligned} \hat{U}(t_f, t_i) = & 1 - \frac{i}{\hbar} \int_{t_i}^{t_f} dt_1 \hat{V}(t_1) + \left(-\frac{i}{\hbar}\right)^2 \int_{t_i}^{t_f} dt_2 \int_{t_i}^{t_2} dt_1 \hat{V}(t_2) \hat{V}(t_1) \\ & + \left(-\frac{i}{\hbar}\right)^3 \int_{t_i}^{t_f} dt_3 \int_{t_i}^{t_3} dt_2 \int_{t_i}^{t_2} dt_1 \hat{V}(t_3) \hat{V}(t_2) \hat{V}(t_1) + \dots \end{aligned} \quad (11.20)$$

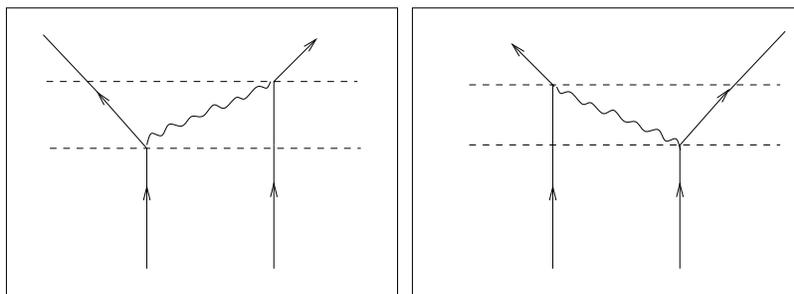


Figure 11.6: Electron-electron scattering: two intermediate states.

This is called the “Dyson series” for the evolution operator. Note the nested time integrals, preserving the ordering in the multiple time integrals of the temporal points where the potential acts.

As an example of the sum over intermediate states, imagine (rather poetically) computing the transition amplitude for electron - electron scattering. The interaction term involves the emission and absorption of a photon from the electron. The initial state $|\alpha\rangle$ is a state in which the two electrons have momenta \vec{p}_1 and \vec{p}_2 . The final state $|\beta\rangle$ is also a two-electron state where the particles carry momenta \vec{k}_1 and \vec{k}_2 . All we know about this process is that two electrons came in, scattered off one another, and two electrons came out. We do not know which electron emitted a photon, and which one absorbed it. We must sum over both possibilities. The situation is shown in Fig. 11.6. In the left panel, the intermediate state consists of an electron with momentum p_2 , another electron with momentum k_1 , and a photon, while in the right panel the intermediate state’s electrons carry momentum p_1 and k_2 . These intermediate states are different.

Hamiltonian formulation of time-dependent perturbation theory

We repeat the derivation of the Dyson series, by a direct attack on the equation of motion of the evolution operator. Again writing $H = H_0 + V$, it is

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) = (\hat{H}_0 + V) U(t, t_0). \quad (11.21)$$

Our calculation simplifies if we work with the interaction representation perturbation and evolution operator, Eqs. 11.18 and 11.19. With these replacements, Eq. 11.21 becomes

$$i\hbar \frac{\partial}{\partial t} \left[e^{-i\hat{H}_0 t/\hbar} \hat{U} e^{i\hat{H}_0 t/\hbar} \right] = (\hat{H}_0 + V) e^{-i\hat{H}_0 t/\hbar} \hat{U} e^{i\hat{H}_0 t_0/\hbar}. \quad (11.22)$$

The left hand side is

$$i\hbar \frac{\partial}{\partial t} \left[e^{-i\hat{H}_0 t/\hbar} \hat{U} e^{i\hat{H}_0 t_0/\hbar} \right] = \left[\hat{H}_0 e^{-i\hat{H}_0 t/\hbar} \hat{U} + e^{-i\hat{H}_0 t/\hbar} i\hbar \frac{\partial \hat{U}}{\partial t} \right] e^{i\hat{H}_0 t_0/\hbar}, \quad (11.23)$$

so

$$\left[\hat{H}_0 e^{-i\hat{H}_0 t/\hbar} \hat{U} + e^{-i\hat{H}_0 t/\hbar} i\hbar \frac{\partial \hat{U}}{\partial t} \right] e^{i\hat{H}_0 t_0/\hbar} = (H_0 + V) e^{-i\hat{H}_0 t/\hbar} \hat{U} e^{i\hat{H}_0 t_0/\hbar} \quad (11.24)$$

or

$$i\hbar \frac{\partial \hat{U}}{\partial t} = \left(e^{i\hat{H}_0 t/\hbar} V e^{-i\hat{H}_0 t/\hbar} \right) \hat{U} = \hat{V} \hat{U}. \quad (11.25)$$

This is just a first order differential equation. Integrating it, we find

$$i\hbar \left[\hat{U}(t, t_0) - \hat{U}(t_0, t_0) \right] = \int_{t_0}^t dt' \hat{V}(t') \hat{U}(t', t_0). \quad (11.26)$$

Because $\hat{U}(t_0, t_0) = 1$, the evolution operator satisfies the integral equation

$$\hat{U}(t, t_0) = 1 - \frac{i}{\hbar} \int_{t_0}^t dt' \hat{V}(t') \hat{U}(t', t_0). \quad (11.27)$$

Like all Volterra integral equations, it can be solved by iteration:

$$\hat{U}(t, t_0) = 1 - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{V}(t_1) + \left(-\frac{i}{\hbar} \right)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \hat{V}(t_1) \hat{V}(t_2) + \dots \quad (11.28)$$

The operators $V(t)$ may not commute at different times, and so it is important to preserve their time ordering. This is Eq. 11.20, derived more succinctly, although perhaps less transparently.

The two-state system as an example

Let us check our perturbative formalism against an exactly - solvable example, the two-state system

$$\hat{H} = \hbar \begin{pmatrix} \lambda & \epsilon \\ \epsilon & -\lambda \end{pmatrix} \quad (11.29)$$

Suppose that at time $t = 0$ the system is in state

$$\psi = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (11.30)$$

and that our object is the time-dependent transition amplitude into the state

$$\psi = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (11.31)$$

at a later time. This problem can, of course, be solved exactly. Recall that we can construct the solution as an expansion in energy eigenfunctions

$$\psi(t) = A_+ \psi_+ e^{-i\omega t} + A_- \psi_- e^{i\omega t} \quad (11.32)$$

The eigenvalues of H are $\pm\omega = \pm\sqrt{\lambda^2 + \epsilon^2}$. The un-normalized solutions can be written as

$$\psi_+ = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (11.33)$$

with

$$(\lambda - \omega)\alpha + \epsilon\beta = 0, \quad (11.34)$$

or

$$\psi_+ = C \begin{pmatrix} \epsilon \\ \omega - \lambda \end{pmatrix}. \quad (11.35)$$

The orthogonal eigenstate is

$$\psi_- = C \begin{pmatrix} \omega - \lambda \\ -\epsilon \end{pmatrix}. \quad (11.36)$$

Then Eq. 11.32 becomes

$$\psi(t) = A_+ \begin{pmatrix} \epsilon \\ \omega - \lambda \end{pmatrix} e^{-i\omega t} + A_- \begin{pmatrix} \omega - \lambda \\ -\epsilon \end{pmatrix} e^{i\omega t} \quad (11.37)$$

The C 's have been absorbed into other constants without loss of generality. The $t = 0$ boundary condition requires that

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = A_+ \begin{pmatrix} \epsilon \\ \omega - \lambda \end{pmatrix} + A_- \begin{pmatrix} \omega - \lambda \\ -\epsilon \end{pmatrix} \quad (11.38)$$

and so

$$A_+ = \frac{\epsilon}{\epsilon^2 + (\omega - \lambda)^2} \quad A_- = \frac{\omega - \lambda}{\epsilon^2 + (\omega - \lambda)^2}. \quad (11.39)$$

The amplitude into the the "down" state at a time t is

$$\begin{pmatrix} 0 & 1 \end{pmatrix} \psi(t) = (\omega - \lambda)A_+ e^{-i\omega t} - \epsilon A_- e^{i\omega t} = \frac{\epsilon(\omega - \lambda)}{\epsilon^2 + (\omega - \lambda)^2} (e^{-i\omega t} - e^{i\omega t}). \quad (11.40)$$

This is exact. To compare it to the perturbative solution, it is necessary to expand it in a series in ϵ . Recall that $\omega = \sqrt{\lambda^2 + \epsilon^2} = \lambda + \epsilon^2/(2\lambda)$. With this result, the coefficient in front of the harmonic time dependence becomes

$$\frac{\epsilon(\omega - \lambda)}{\epsilon^2 + (\omega - \lambda)^2} \rightarrow \frac{\epsilon \times \frac{\epsilon^2}{2\lambda}}{\epsilon^2 + \left(\frac{\epsilon^2}{2\lambda}\right)^2} = \frac{1}{2} \frac{\epsilon}{\lambda}. \quad (11.41)$$

Next we compute the matrix of the evolution operator perturbatively. We regard $\lambda \gg \epsilon$ and identify H_0 and V accordingly. In lowest order, the transition amplitude is

$$U_{\downarrow\uparrow}(t) = \frac{i}{\hbar} \int_0^t dt' e^{i\lambda(t-t')} \langle \downarrow | V | \uparrow \rangle e^{-i\lambda(t'-0)}. \quad (11.42)$$

The signs on the exponentials are due to the signs on the energies. Evaluating the matrix element gives

$$\begin{aligned} U_{\downarrow\uparrow}(t) &= -\frac{i}{\hbar} e^{-\lambda t} \hbar \epsilon \int_0^t dt' e^{-2i\lambda t'} \\ &= \frac{\epsilon}{2\lambda} [e^{-i\lambda t} - e^{i\lambda t}]. \end{aligned} \quad (11.43)$$

With the approximation $\omega = \lambda$ in the exponentials, the perturbative result agrees with the expansion of the exact solution.

Special cases of time-dependent perturbation theory

A number of special cases of time dependence are often encountered in the literature. To begin, $V(t)$ could simply be some arbitrary function of time. Then the transition amplitude just has to be computed in a case-by-case basis. The only reason to list this possibility (which is a staple of examination questions) is to remark: in general, the transition amplitude can connect initial and final states of different energy. This should come as no surprise: if the Hamiltonian is time dependent, energy is not conserved.

Next, we have the so-called ‘‘Sudden Approximation.’’ The potential is a simple impulse in time. This can be represented as

$$V = V_0 \delta(t - t_0) \quad (11.44)$$

Of course V_0 itself might be complicated. The lowest order time evolution operator is (rather trivially)

$$U_{\beta\alpha} = -\frac{i}{\hbar} e^{-E_\beta t_2/\hbar} e^{iE_\alpha t_1/\hbar} \langle \beta | V | \alpha \rangle. \quad (11.45)$$

The probability that the state is in $|\beta\rangle$ at a time t is then

$$|U_{\beta\alpha}|^2 = \frac{|\langle\beta|V|\alpha\rangle|^2}{\hbar^2}. \quad (11.46)$$

Potentials which are harmonic in time

$$V(t) = V_0 e^{-i\omega t} \quad \text{or} \quad V(t) = V_0 \cos(\omega t) \quad (11.47)$$

are often encountered. Since the expression for the evolution operator can almost be written down by inspection after going through the case of a constant potential, the evaluation of the evolution operator in this case is left as an exercise for the reader. (The answer will be given below).

Finally, the potential may be constant in time. This is something of a misnomer, because we are formally required to switch the potential on and off adiabatically. Neglecting this formal requirement, the transition amplitude becomes

$$U_{\beta\alpha} = -\frac{i}{\hbar} \langle\beta|V|\alpha\rangle e^{-iE_\beta t_2/\hbar} e^{iE_\alpha t_1/\hbar} \int_{t_1}^{t_2} e^{-i(E_\alpha - E_\beta)t/\hbar} dt. \quad (11.48)$$

The integral is straightforward to evaluate: it is

$$\int_{t_1}^{t_2} e^{-i(E_\alpha - E_\beta)t/\hbar} dt = \frac{e^{-i(E_\alpha - E_\beta)t_2/\hbar} - e^{-i(E_\alpha - E_\beta)t_1/\hbar}}{-i(E_\alpha - E_\beta)/\hbar}. \quad (11.49)$$

Just for simplicity, let us set $t_1 \rightarrow 0$. We also define $\hbar\omega_{\alpha\beta} = E_\alpha - E_\beta$. The transition probability to be found in state $|\beta\rangle$ at time t is

$$|U(t)_{\beta\alpha}|^2 = \frac{1}{\hbar^2} |\langle\beta|V|\alpha\rangle|^2 \left\{ \frac{2 - 2\cos(\omega_{\alpha\beta}t)}{\omega_{\alpha\beta}^2} \right\} \quad (11.50)$$

In this expression, there is no reason for a transition not to connect states of different energy; there is a transition probability when E_α is different from E_β . Energy is not conserved. This is no surprise, the potential was nonzero for a finite interval t – the Hamiltonian is time dependent.

For a harmonic perturbation, with $V \sim \exp(-i\omega t)$, this becomes

$$|U(t)_{\beta\alpha}|^2 = \frac{1}{\hbar^2} |\langle\beta|V|\alpha\rangle|^2 \left\{ \frac{2 - 2\cos((\omega_{\alpha\beta} - \omega)t)}{(\omega_{\alpha\beta} - \omega)^2} \right\}. \quad (11.51)$$

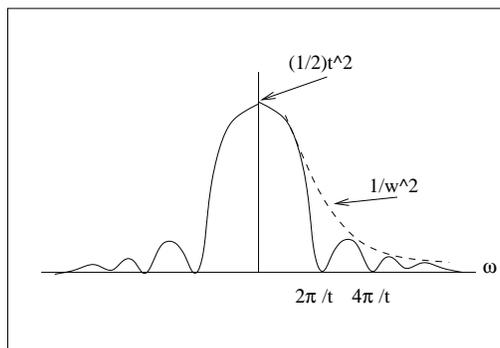


Figure 11.7: Frequency dependence vs. $\omega_{\alpha\beta}$ of the transition amplitude of Eq. 11.50.

When our spectrum is discrete, this is the end of the story. If we set $\omega = \omega_{\alpha\beta}$, $|U_{\beta\alpha}(t)|^2$ diverges. Perturbation theory has broken down. This is actually no surprise; recall the discussion of the exact time dependence of the two state system in Ch. 3, especially Eq. 3.50. At resonance, when $\omega = \omega_{\alpha\beta}$, $|U_{\beta\alpha}(t)|^2$ oscillates between zero and unity. Of course, $|U_{\beta\alpha}(t)|^2 = 1$ is not a small number; there is no reason for perturbation theory to still be applicable in this case. We would only expect it would be valid when the transition probability was small. The mechanical way perturbation theory fails is that it does not include the width of the Lorentzian which regulates the expression.

Fermi's Golden Rule

Now we consider a situation in which we have many final states $|\beta\rangle$ whose energies E_β are all nearly equal. Our perturbing Hamiltonian V will be assumed to be a constant in time. In this case, two things happen: first, the transition probability $|U_{\beta\alpha}(t_f)|^2$ is proportional to t_f – the rate of the transition $\alpha \rightarrow \beta$ is a constant in time. And second, transitions only occur when E_β is very close to E_α , and at long times only when E_β is equal to E_α : energy is conserved in the transition.

These results follow from the behavior of Eq. 11.50. We can think of the object in the curly brackets as a function of $\omega_{\alpha\beta}$. We plot it in Fig. 11.7. At the origin where $\omega_{\alpha\beta}t$ is small, the function approaches $\frac{1}{2}t^2$. The envelope of the function falls away as $1/\omega_{\alpha\beta}^2$. Nodes of the function are found at integer multiples of $2\pi/t$. They become ever closer to the origin as t grows. The width of the function collapses to zero like $1/t$. This all means that as t grows, the function is dominated more and more by the peak around $\omega_{\alpha\beta} = 0$.

We interpret this result as meaning that at long times only transitions where $\omega_{\alpha\beta} = 0$ are allowed: the final and initial states have the same energy and energy is a conserved quantity. Furthermore, consider transitions into any of the related states β . Then the total transition probability out of state $|\alpha\rangle$ is the integral over the curve. In the limit of tiny $\omega_{\alpha\beta}$, the area under the curve, the product of its height times its width, is proportional to t , $P(t_f) = Wt_f$. The constant of proportionality W is the rate of transitions per unit time – it is called the “transition probability per unit time.” To flesh this out a bit more, suppose that there are $\rho(E_\beta)dE_\beta$ states available in the energy interval dE_β . The total transition rate (transitions per unit time) will be

$$W = \lim_{t_f \rightarrow \infty} \frac{1}{t_f} \int P(\alpha \rightarrow \beta) \rho(E_\beta) dE_\beta. \quad (11.52)$$

This is the rate at which states of energy E_β are populated.

A physics application of this idea is the lifetime of an excited state. The number of decays of N parent states at time t is

$$\frac{dN(t)}{dt} = -WN(t). \quad (11.53)$$

The solution of this equation is the usual formula of radioactive decay,

$$N(t) = N_0 \exp(-Wt). \quad (11.54)$$

so that $1/W$, the “inverse transition probability per unit time,” is also the lifetime of the state.

To get a practical formula, proceed in stages. Inserting the explicit time-dependence of Eq. 11.50 into Eq. 11.52 yields

$$W = \lim_{t_f \rightarrow \infty} \int \frac{1}{\hbar^2} |\langle \beta | V | \alpha \rangle|^2 \left\{ \frac{2 - 2 \cos(\omega_{\alpha\beta} t_f)}{\omega_{\alpha\beta}^2 t_f} \right\} \rho(E_\beta) dE_\beta. \quad (11.55)$$

As we argued, at large times the object in curly brackets becomes strongly peaked about $\omega_{\alpha\beta} = 0$. We have also argued that in that limit, its integral (over ω or energy) is a constant. To find the constant, pause to define the function $I(t)$ as

$$I(t) = \frac{1}{\hbar^2 t} \left| \int_0^t e^{i\omega_{\alpha\beta} t} dt \right|^2 = \frac{2(1 - \cos(\omega_{\alpha\beta} t))}{(\hbar\omega_{\alpha\beta})^2 t} \quad (11.56)$$

To find the constant of proportionality, write it as

$$I(t) = A\delta(\hbar\omega_{\alpha\beta}). \quad (11.57)$$

and integrate the expression over an energy interval. If it is proportional to a delta function, the integral will be equal to a constant and the value of the constant will normalize the delta function. So

$$A = \int dEI(t) = \int_{-\infty}^{\infty} d(\hbar\omega) \frac{2(1 - \cos(\omega t))}{(\hbar\omega)^2 t} \quad (11.58)$$

or

$$A = \frac{2}{\hbar} \int_{-\infty}^{\infty} dx \frac{1 - \cos(x)}{x^2} = \frac{2\pi}{\hbar} \quad (11.59)$$

Thus the transition probability per unit time is

$$W_{\beta\alpha} = \int \frac{2\pi}{\hbar} |\langle\beta|V|\alpha\rangle|^2 \delta(E_\beta - E_\alpha) \rho(E_\beta) dE_\beta. \quad (11.60)$$

This expression forms the basis for many practical calculation of time evolution. It is such an important result, that Fermi called it the “Golden Rule.” We will make extensive use of it. I like to build the variations on what I call the “useless form of the Golden Rule,”

$$W_{\beta\alpha} = \frac{2\pi}{\hbar} |\langle\beta|V|\alpha\rangle|^2 \delta(E_\beta - E_\alpha). \quad (11.61)$$

It is useless because the presence of the delta function means that the expression is either zero or infinity! But it is a useful building block. Many physical systems have a continuum of energy eigenstates, which are highly degenerate. The properties of the states, the so-called “density of states,” generally has nothing to do with the physics of the process which populates them ($|\langle\beta|V|\alpha\rangle$). When the transition probability per unit time is summed over the available “phase space” of these possible states the delta-function will select a subspace of them which can be populated by the transition process.

For example consider the decay of a 2P state of hydrogen into the 1S state plus a photon. There are many potential final states: in principle, while the photon must have energy $E_\gamma = E_{2P} - E_{1S}$, it could be emitted in any arbitrary direction. The direction of its wave number \vec{k} is arbitrary, even though $E_\gamma = \hbar c |\vec{k}|$ is fixed.

We could equally well have considered the case of many initial states and one final state. Imagine illuminating an atom in its ground state with broad-spectrum radiation of intensity $I(\omega)$. Our transition is $1S + \gamma \rightarrow 2P$. This is a harmonic perturbation, $V \sim \exp(-i\omega t)$, and the $I(t)$ expression in W is modified to

$$I(t) = \frac{2(1 - \cos((\omega_{\beta\alpha} - \omega)t))}{(\hbar(\omega_{\beta\alpha} - \omega))^2 t} \rightarrow \frac{2\pi}{\hbar} \delta(E_\beta - E_\alpha - \hbar\omega). \quad (11.62)$$

The density of final states $\rho(E_\beta)dE_\beta$ is replaced by $I(\omega)d\omega$ in the summation over what is absorbed. Then the total absorption rate is

$$W_{\beta\alpha} = \int I(\omega)d\omega \frac{2\pi}{\hbar} |\langle\beta|V|\alpha\rangle|^2 \delta(E_\beta - E_\alpha - \hbar\omega) = I(\omega = \frac{E_\beta - E_\alpha}{\hbar}) \frac{2\pi}{\hbar^2} |\langle\beta|V|\alpha\rangle|^2. \quad (11.63)$$

Now return to Eq. 11.60. Physicists like to write the differential transition probability as

$$dW_{\beta\alpha} = \frac{2\pi}{\hbar} |\langle\beta|V|\alpha\rangle|^2 \delta(E_\beta - E_\alpha) \rho(E_\beta) dE_\beta. \quad (11.64)$$

For a harmonic potential, $V(t) \propto \cos(\omega t)$, the single delta-function becomes

$$\frac{1}{4} [\delta(E_\beta - E_\alpha - \hbar\omega) + \delta(E_\beta - E_\alpha + \hbar\omega)]. \quad (11.65)$$

The two terms have simple interpretations: the two $\exp(\pm i\omega t)$ parts of the cosine have the correct frequency dependence to add or subtract energy from the initial and final state, and cause transitions up or down in energy.

All this discussion was in the context of first order perturbation theory. Let us see what happens in second order. The evolution operator is

$$U_{\beta\alpha}^{(2)} = \sum_{\gamma} \left(-\frac{i}{\hbar}\right)^2 \int_{t_i}^{t_f} dt_1 \int_{t_i}^{t_1} dt_2 e^{-i\omega_\beta t_f} e^{-i(\omega_\gamma - \omega_\beta)t_1} e^{-i(\omega_\alpha - \omega_\gamma)t_2} e^{i\omega_\alpha t_i} \langle\beta|V|\gamma\rangle \langle\gamma|V|\alpha\rangle. \quad (11.66)$$

We evaluate the integral over dt_2 to give

$$U_{\beta\alpha} = \sum_{\gamma} \left(-\frac{i}{\hbar}\right)^2 \langle\beta|V|\gamma\rangle \langle\gamma|V|\alpha\rangle e^{-i\omega_\beta t_f} e^{i\omega_\alpha t_i} \int_{t_i}^{t_f} dt_1 \left(\frac{e^{-i(\omega_\alpha - \omega_\gamma)t_1} - e^{-i(\omega_\alpha - \omega_\gamma)t_i}}{-i(\omega_\alpha - \omega_\gamma)} \right) e^{-i(\omega_\gamma - \omega_\beta)t_1}. \quad (11.67)$$

$$(11.68)$$

The first term in the long parentheses in this expression is

$$U_{\beta\alpha}^{(2)} = -\frac{i}{\hbar} \sum_{\gamma} \frac{\langle\beta|V|\gamma\rangle \langle\gamma|V|\alpha\rangle}{E_\alpha - E_\gamma} e^{-i\omega_\beta t_f} e^{i\omega_\alpha t_i} \int_{t_i}^{t_f} dt e^{-i(\omega_\beta - \omega_\alpha)t}. \quad (11.69)$$

In the limit $t_i - t_f \rightarrow \infty$ the integral forces $\hbar\omega_\beta = \hbar\omega_\alpha$, just as in the equation for first order perturbation theory. At long times the integral over the second term in the long parentheses becomes small compared to the integral over the first term. Dropping the second term and

keeping the overall time integral from the first term gives us the second order expression for the transition amplitude

$$U_{\beta\alpha}(t) = \left[\delta_{\alpha\beta} - \frac{i}{\hbar} \{ \langle \beta | V | \alpha \rangle + \sum_{\gamma} \frac{\langle \beta | V | \gamma \rangle \langle \gamma | V | \alpha \rangle}{E_{\alpha} - E_{\gamma}} + \dots \} \right] \times \int_{t_i}^{t_f} dt e^{-i(\omega_{\beta} - \omega_{\alpha})t} \times e^{-i(\omega_{\beta}t_f - \omega_{\alpha}t_i)}. \quad (11.70)$$

The last factor, $\exp(-i(\omega_{\beta}t_f - \omega_{\alpha}t_i))$, will disappear when we take the absolute square. The integral from t_i to t_f is the same expression that we have already found in lowest order. We treat it exactly as we did when we derived the first order Golden Rule. It will (again) enforce energy conservation in the sense that the transition probability will vanish unless $\omega_{\beta} = \omega_{\alpha}$, and the transition probability per unit time will be a constant. The same thing happens when we go to higher and higher order in the perturbative expansion for $U_{\beta\alpha}(t)$.

The term in braces in Eq. 11.70 is called the “T-matrix.” At second order it is

$$\langle \beta | T | \alpha \rangle = \langle \beta | V | \alpha \rangle + \sum_{\gamma} \frac{\langle \beta | V | \gamma \rangle \langle \gamma | V | \alpha \rangle}{E_{\alpha} - E_{\gamma}} + \dots \quad (11.71)$$

The upshot of this calculation is that we can write an expression for the Golden Rule which goes beyond first order perturbation theory, simply by replacing V by T :

$$W_{\beta\alpha} = \frac{2\pi}{\hbar} |\langle \beta | T | \alpha \rangle|^2 \delta(E_{\beta} - E_{\alpha}). \quad (11.72)$$

Notice that the sum in Eq. 11.71 runs over all intermediate states, regardless of energy. Referring to the sketch of the process in Fig. 11.5, we interpret this result by saying that because the system only remains in state $|\gamma\rangle$ for a finite time interval, energy need not be conserved in the transition from $|\alpha\rangle$ to $|\gamma\rangle$ or from $|\gamma\rangle$ to $|\beta\rangle$. Alternatively, we can say that only on small time scales, where $t \sim 1/(\omega_{\alpha} - \omega_{\gamma})$, are transitions possible, which do not conserve energy.

In casual introductions to quantum mechanics, one often sees discussions in which the “position-momentum uncertainty principle” $\Delta x \Delta p > \hbar$ is introduced along with an “energy-time uncertainty principle,” $\Delta E \Delta t > \hbar$. They really have very different meanings: the usual uncertainty principle is a statement about states and observables, which at heart comes from the non-commutativity of operators. The energy-time uncertainty principle is quite different. It is a “principle” only in the sense that we have just described, that processes need not conserve energy over short time intervals.

Chapter 12

Electromagnetic interactions in semiclassical approximation

In this chapter we use time dependent perturbation theory to analyze the physics of the absorption and emission of electromagnetic radiation by quantum mechanical systems. This subject is of great practical importance, since electromagnetism is the most common experimental probe used to study matter.

We will analyze radiative transitions using what is called the “semi-classical approximation.” In this approximation, the “matter” part of the system, the charged particles which emit and absorb radiation, will be described using the Schrödinger equation. The electromagnetic field, however, will be described as if it were classical. This approximation is (of course) inconsistent. There is no honest way to describe radiation processes as involving the emission and absorption of photons when radiation is classical. However, the reader who is willing to live temporarily with this inconsistency will find that the formalism captures essentially all the practical parts of the subject of radiation: selection rules, lifetimes of states, probabilities for emission of radiation into particular directions or polarizations. The price to be paid is the necessity to adopt some ad-hoc rules for how to initiate calculations and how to interpret their results. We will return later to construct a fully quantum mechanical description of the electromagnetic field.

Einstein relations

In 1917 Einstein set the stage for the interaction of radiation with matter. In this paper (Physikalische Zeitschrift **18**, 121 (1917), reprinted and translated in D. ter Haar, “The old quantum theory,” (Pergamon, 1967)) he described the three kinds of processes involving radiation and matter, and relations between them. 1917 was of course before the discovery of quantum mechanics, and one of the successes of quantum mechanics was to verify and quantify his work. Einstein envisioned three kinds of processes involving electromagnetism and matter:

- 1) Spontaneous emission: While this process has no classical analogue, it comes first on Einstein’s list. A system in state $|n\rangle$, with energy E_n , spontaneously decays into another state $|m\rangle$ of energy E_m by the emission of a photon of energy $\hbar\omega = E_n - E_m$. Einstein parameterized the probability of the decay dW in a time dt by

$$dW = A_{mn}dt \quad (12.1)$$

where A_{mn} is called an “Einstein coefficient.” We recognize it as the transition probability per unit time for the initial state to decay to the final state, with the emission

of a photon.

- 2) Absorption: A system in state $|m\rangle$ may absorb a photon from a source of energy $\hbar\omega$ and is then promoted to state $|n\rangle$. In modern language, the rate for this process is parameterized with an absorption cross section, defined as

$$\sigma \equiv \frac{\text{energy absorbed by atom}}{\text{unit time} \times \text{energy flux}} \quad (12.2)$$

Note the units of σ – length squared (hence the name). The cross section represents an inherent strength of a reaction. Equivalently, one can consider the absorption rate itself. This quantity should be proportional to the intensity of the radiation of the appropriate frequency, $I(\omega)$, illuminating the system. The absorption rate is parameterized through a second Einstein coefficient B_{nm} ,

$$\text{Rate of Absorption} = I(\omega) \times B_{nm}. \quad (12.3)$$

We have flipped the subscript to label the states as we did for spontaneous emission, but to represent the transition as going from the lower energy state to the higher one.

- 3) Stimulated or induced emission: Finally, a system illuminated by radiation might emit radiation as a result of the presence of the electromagnetic field. This rate can also be parameterized with an Einstein coefficient for emission, defined through

$$\text{Rate of Emission} = I(\omega) \times B_{mn}. \quad (12.4)$$

The three Einstein coefficients can be related by imagining that they take place in a cavity filled with black body radiation. The intensity of radiation (energy density per frequency interval) is given by Planck's formula

$$I(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\hbar\omega/k_B T} - 1}. \quad (12.5)$$

Suppose that the cavity also contains atoms which are in thermal equilibrium with the radiation. Also suppose, for simplicity, that these atoms have two levels $|m\rangle$ and $|n\rangle$ with energies $E_n > E_m$, and that $E_n - E_m = \hbar\omega$. The three kinds of radiative processes (spontaneous and stimulated emission, and absorption) can all occur in the cavity.

In equilibrium the rate at which excited state atoms are converted to the ground state must be equal to the rate at which atoms are removed from the ground state. Calling N_n and N_m the number of atoms in each state, the equilibrium condition is

$$\text{Rate}(n \rightarrow m) = \text{Rate}(m \rightarrow n) \quad (12.6)$$

or

$$N_n(A_{mn} + B_{mn}I(\omega)) = N_m(B_{nm}I(\omega)). \quad (12.7)$$

Solving this equation for the intensity gives

$$I(\omega) = \frac{A_{mn}}{\frac{N_m}{N_n}B_{nm} - B_{mn}}. \quad (12.8)$$

If we assume that the relative numbers of particles in each state will be given by Boltzmann statistics (i.e. $N_m/N_n = \exp(-(-E_m - E_n)/kT)$), Eq. 12.8 becomes

$$I(\omega) = \frac{A_{mn}}{e^{\hbar\omega/kT}B_{nm} - B_{mn}}. \quad (12.9)$$

But the intensity is given by the Planck formula. Therefore the rate of absorption (per atom, per frequency interval) and rate of stimulated emission must be equal: $B_{mn} = B_{nm}$. Additionally, the rate for spontaneous emission must be related to the rate of absorption:

$$A_{mn} = \frac{\hbar\omega^3}{\pi^2c^3}B_{nm}. \quad (12.10)$$

This was as far as Einstein could proceed in 1917. Along the way in our development of electromagnetic interactions, we will be able to compute the values of the A and B coefficients and verify Eq. 12.10.

Necessary electrodynamic preliminaries

No professional calculation involving electromagnetism in quantum mechanics uses MKS units. Most quantum field theory books use Lorentz-Heaviside units, a variation on CGS with some rearranged 4π 's. We will work with the slightly more familiar CGS convention, where the four Maxwell's equations are

$$\begin{aligned} \vec{\nabla} \times \vec{E} + \frac{1}{c} \frac{\partial \vec{H}}{\partial t} &= 0 & \vec{\nabla} \times \vec{H} - \frac{1}{c} \frac{\partial \vec{E}}{\partial t} &= \frac{4\pi}{c} \vec{J} \\ \vec{\nabla} \cdot \vec{B} &= 0 & \vec{\nabla} \cdot \vec{E} &= 4\pi\rho \end{aligned} \quad (12.11)$$

In Gaussian units, $\vec{D} = \vec{E}$ and $\vec{H} = \vec{B}$ in free space, and I seem to have been casual about writing \vec{B} or \vec{H} in my expressions. Introducing a scalar potential ϕ and a vector potential \vec{A} allows us to write the fields as

$$\vec{E} = -\vec{\nabla}\phi - \frac{1}{c} \frac{\partial \vec{A}}{\partial t}; \quad \vec{H} = \vec{\nabla} \times \vec{A} \quad (12.12)$$

Recall that the scalar potential and the vector potential are not unique quantities. We can perform a gauge transformation of the fields without the loss of any physical content in the solutions.

$$\vec{A} \rightarrow \vec{A} + \vec{\nabla}\chi \quad \phi \rightarrow \phi - \frac{1}{c} \frac{\partial\chi}{\partial t} \quad (12.13)$$

We will treat the interaction of radiation and matter perturbatively. The analog of zeroth-order solutions of the Schrödinger equation for the radiation field are the solutions of the Maxwell equations in free space, $\vec{J} = 0$ and $\rho = 0$. A convenient gauge choice is Coulomb gauge: $\phi(x, t) = 0$ and $\vec{\nabla} \cdot \vec{A}(x, t) = 0$. Then the vector potential is a solution of the wave equation

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} = 0 \quad (12.14)$$

for which

$$\vec{A}(\vec{x}, t) = \vec{A}_0 e^{i(\vec{k} \cdot \vec{x} - \omega t)} + \vec{A}_0^* e^{-i(\vec{k} \cdot \vec{x} - \omega t)} \quad (12.15)$$

We can also write this solution as

$$\vec{A}(\vec{x}, t) = \vec{\epsilon} a \cos(\vec{k} \cdot \vec{x} - \omega t) \quad (12.16)$$

where $\vec{\epsilon}$ is a unit polarization vector. The gauge condition, $\vec{\nabla} \cdot \vec{A} = 0$ imposes the so-called transversality condition on the vector potential, $\vec{A}_0 \cdot \vec{k} = 0$. Thus $\vec{k} \cdot \vec{\epsilon} = 0$. In Coulomb gauge the two independent polarizations signal the two polarization states of light. (We will always assume linear polarization in our formulas.)

The amplitude of the vector potential can be expressed in terms of the intensity of the radiation. In this book we adopt the convention that “intensity” means energy density. Depending on the physical process under consideration there are two choices for the amplitude.

The first choice is appropriate when we consider an external source of radiation, appropriate to absorption or stimulated emission. We define the intensity of the radiation in a frequency interval $\Delta\omega$ as the time average of

$$I(\omega)\Delta\omega = \frac{1}{8\pi} (E^2 + H^2). \quad (12.17)$$

The fields are given by appropriate derivatives of Eq. 12.16:

$$\vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} = -\frac{\omega}{c} \vec{\epsilon} a \sin(\vec{k} \cdot \vec{x} - \omega t) \quad (12.18)$$

and

$$\vec{H} = \vec{\nabla} \times \vec{A} = -\vec{k} \times \vec{\epsilon} a \sin(\vec{k} \cdot \vec{x} - \omega t). \quad (12.19)$$

. The instantaneous intensity is

$$\begin{aligned} I(\omega)\Delta\omega &= \frac{1}{8\pi}|a|^2 \sin^2(\vec{k} \cdot \vec{x} - \omega t) \left\{ \frac{\omega^2}{c^2} \vec{\epsilon} \cdot \vec{\epsilon} + |\vec{\epsilon} \times \hat{k}|^2 |k|^2 \right\} \\ &= \frac{1}{4\pi} \frac{\omega^2}{c^2} \sin^2(\vec{k} \cdot \vec{x} - \omega t) \end{aligned} \quad (12.20)$$

The time average converts the $\sin^2()$ into a $\frac{1}{2}$, so

$$I(\omega)\Delta\omega = \frac{\omega^2}{8\pi c^2} |a|^2 \quad (12.21)$$

and the squared normalization of the vector potential is

$$|a|^2 = \frac{8\pi c^2}{\omega^2} I(\omega)\Delta\omega. \quad (12.22)$$

The second choice for a normalization factor comes when we want to represent the radiation field associated with a single photon. (This is the first of several stories we must tell when describing the quantum electromagnetic field semi-classically.) A single photon with frequency ω has an energy $\hbar\omega$ or a time averaged energy density (energy per unit volume) of $\hbar\omega/V$. To write a solution for Maxwell's equations which encodes this constraint, we simply express the energy in the field in a volume V as

$$E = V \frac{E^2 + H^2}{8\pi} = \frac{V}{8\pi} |a|^2 \left\{ \frac{\omega^2}{c^2} + k^2 \right\} \sin^2(\vec{k} \cdot \vec{x} - \omega t). \quad (12.23)$$

Again time averaging the $\sin^2()$ term gives

$$E = V \frac{1}{8\pi} |a|^2 \frac{\omega^2}{c^2} \quad (12.24)$$

and because we are dealing with photons we force this to be equal to $\hbar\omega$. Thus

$$|a|^2 = \frac{8\pi\hbar c^2}{\omega V}. \quad (12.25)$$

For future reference, the complete expression for the electromagnetic field with "photon normalization" is

$$\begin{aligned} \vec{A}(\vec{x}, t) &= \left(\frac{8\pi\hbar c}{\omega V} \right)^{1/2} \vec{\epsilon} \cos(\vec{k} \cdot \vec{x} - \omega t) \\ &= \left(\frac{2\pi\hbar c}{\omega V} \right)^{1/2} \vec{\epsilon} \left(e^{i(\vec{k} \cdot \vec{x} - \omega t)} + e^{-i(\vec{k} \cdot \vec{x} - \omega t)} \right) \end{aligned} \quad (12.26)$$

Notice the factor of volume. It will disappear in final answers.

Electromagnetic transition amplitudes

The Hamiltonian for a particle in an electromagnetic field is

$$\hat{H} = \frac{1}{2m} \left(\hat{p} - \frac{e\vec{A}}{c} \right)^2 + V(x). \quad (12.27)$$

We decompose this into a zeroth order Hamiltonian

$$\hat{H}_0 = \frac{p^2}{2m} + V(x) \quad (12.28)$$

and a perturbation

$$V = -\frac{e}{2mc} (\vec{p} \cdot \vec{A} + \vec{A} \cdot \vec{p}) + \frac{e^2 A^2}{2mc^2}. \quad (12.29)$$

We groom the momentum dependence, writing

$$(\vec{p} \cdot \vec{A} + \vec{A} \cdot \vec{p}) \psi = (\vec{p} \cdot \vec{A}) \psi + \vec{A} \cdot (\vec{p}\psi) + \vec{A} \cdot (\vec{p}\psi) \quad (12.30)$$

where

$$\vec{p} \cdot \vec{A} = \frac{\hbar}{i} \vec{\nabla} \cdot \vec{A}. \quad (12.31)$$

In Coulomb gauge, $\vec{\nabla} \cdot \vec{A} = 0$, so

$$(\vec{p} \cdot \vec{A} + \vec{A} \cdot \vec{p}) = 2\vec{A} \cdot \vec{p} \quad (12.32)$$

and the interaction term becomes

$$V = -\frac{e}{mc} \vec{A} \cdot \vec{p} + \frac{e^2 A^2}{2mc^2}. \quad (12.33)$$

In the remainder of this chapter we will only consider lowest-order processes, and so we will ignore the term proportional to A^2 .

Assume that at $t = 0$ the vector potential is switched on. The lowest-order transition amplitude from an initial state $|n\rangle$ to a final state $|k\rangle$ is

$$U_{kn}(t) = -\frac{i}{\hbar} \int_0^t dt' \langle k|V(t')|n\rangle e^{i\omega_{kn}t'} \quad (12.34)$$

The time dependent potential is

$$V(t') = -\frac{ea}{mc} \left(e^{i(\vec{k} \cdot \vec{x} - \omega t')} + e^{-i(\vec{k} \cdot \vec{x} - \omega t')} \right) \vec{\epsilon} \cdot \vec{p}. \quad (12.35)$$

This gives the transition amplitude as a sum of two terms,

$$U_{kn}(t) = -\frac{\langle k|V_1|n\rangle}{\hbar} \frac{e^{i(\omega_{kn}-\omega)t} - 1}{\omega_{kn} - \omega} - \frac{\langle k|V_2|n\rangle}{\hbar} \frac{e^{i(\omega_{kn}+\omega)t} - 1}{\omega_{kn} + \omega}. \quad (12.36)$$

The coordinate-space matrix elements are given in terms of the final state and initial space wave functions ψ_k and ψ_n , as

$$\langle k|V_1|n\rangle = \frac{ie\hbar a}{mc} \int d^3r \psi_k^* e^{i\vec{k}\cdot\vec{r}} \vec{\epsilon} \cdot \vec{\nabla} \psi_n \quad (12.37)$$

and

$$\langle k|V_2|n\rangle = \frac{ie\hbar a}{mc} \int d^3r \psi_k^* e^{-i\vec{k}\cdot\vec{r}} \vec{\epsilon} \cdot \vec{\nabla} \psi_n. \quad (12.38)$$

The two terms are responsible for different phenomena. The first term will give a large transition amplitude only when $\omega_{kn} = \omega$. This represents absorption of radiation. Similarly, the second term is responsible for emission.

Golden Rule expression for absorption

We now have all the pieces in hand to compute the transition probability per unit time for absorption

$$\begin{aligned} W &= \frac{2\pi}{\hbar} |\langle k|V_1|n\rangle|^2 \delta(E_k - E_n - \hbar\omega) \\ &= \frac{2\pi\hbar e^2 |a|^2}{m^2 c^2} \left| \int d^3r \psi_k^* e^{i\vec{k}\cdot\vec{r}} \vec{\epsilon} \cdot \vec{\nabla} \psi_n \right|^2 \delta(E_k - E_n - \hbar\omega). \end{aligned} \quad (12.39)$$

We must deal, somehow, with the delta-function. Let us do this by considering absorption from a beam of radiation. We know that the normalization of the vector potential is given by Eq. 12.22. Using this expression for $|a|^2$ gives the differential transition probability per unit second from a frequency interval $\Delta\omega$ of

$$\Delta W = \frac{4\pi^2 c^2}{\omega^2} I(\omega) \frac{e^2}{m^2 c^2} \hbar \Delta\omega \delta(\Delta E - \hbar\omega) |M|^2 \quad (12.40)$$

where M is the long expression for the coordinate space matrix element. We compute the total absorption rate by replacing the finite widths by infinitesimals and integrating over the intensity profile:

$$\begin{aligned} \text{Rate} = W &= \int d\omega \delta(\Delta E - \hbar\omega) I(\omega) \times \frac{4\pi^2 \hbar c^2 e^2}{\omega^2 m^2 c^2} |M|^2 \\ &= I \left(\omega_{fi} = \frac{\Delta E}{\hbar} \right) B_{fi}. \end{aligned} \quad (12.41)$$

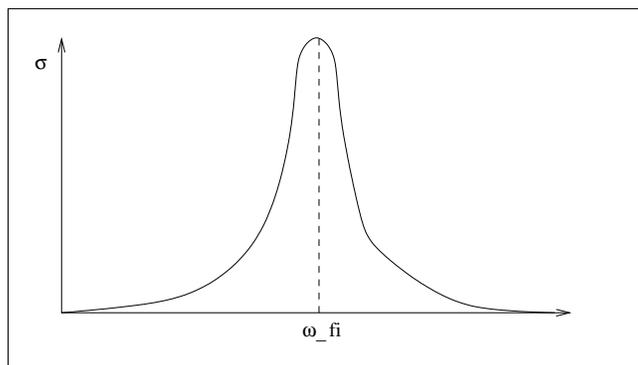


Figure 12.1: A schematic absorption cross section as a function of frequency.

B_{fi} is the Einstein coefficient for absorption.

One can also interpret the Golden Rule result a little differently. The delta-function can be regarded as the infinitesimal limit of a sharp, narrow function such as a Lorentzian.

$$\delta(\omega - \omega_{fi}) = \lim_{\gamma \rightarrow 0} \frac{\gamma}{2\pi} \frac{1}{(\omega - \omega_{fi})^2 + \frac{\gamma^2}{4}} \quad (12.42)$$

Then the transition probability per unit time can be expressed in terms of an absorption cross section, defined as

$$\sigma = \frac{\text{energy absorbed per unit time}}{\text{energy flux}}. \quad (12.43)$$

Thus

$$\begin{aligned} \sigma &= \frac{\hbar\omega \times W}{c \left(\frac{1}{2\pi} \frac{\omega^2}{c^2} |a|^2 \right)} \\ &= \frac{4\pi^2 \hbar^2 e^2}{m\omega^2 c} |M_{fi}|^2 \delta(\Delta E - \hbar\omega) \end{aligned} \quad (12.44)$$

or

$$\sigma = \frac{4\pi^2 \hbar^2 e^2}{m\omega^2 c} |M_{fi}|^2 \frac{\gamma}{2\pi} \frac{1}{(\omega - \omega_{fi})^2 + \frac{\gamma^2}{4}}. \quad (12.45)$$

The frequency dependence of the absorption cross section is shown in Fig. 12.1, a sharp absorption line.

Stimulated emission

The calculation of the rate for stimulated emission is almost identical to the one for absorption. The Golden Rule tells us that it is

$$W_{k \rightarrow n} = \frac{2\pi}{\hbar} |\langle n|V_2|k \rangle|^2 \delta(E_n + \hbar\omega - E_k) \quad (12.46)$$

where the matrix element is

$$\langle n|V_2|k \rangle = \int d^3r \psi_n^*(r) e^{-i\vec{k}\cdot\vec{r}} \vec{\epsilon} \cdot \vec{\nabla} \psi_k(r). \quad (12.47)$$

Contrast this expression with the corresponding one for absorption,

$$\langle k|V_1|n \rangle = \int d^3r U_k^*(r) e^{i\vec{k}\cdot\vec{r}} \vec{\epsilon} \cdot \vec{\nabla} U_n(r). \quad (12.48)$$

Only the squared modulus of the matrix element goes into the transition probability. Because of transversality, we can integrate by parts to show that

$$\int d^3r \psi_k^*(r) e^{i\vec{k}\cdot\vec{r}} \vec{\epsilon} \cdot \vec{\nabla} \psi_n(r) = - \left[\int d^3r \psi_n^*(r) e^{-i\vec{k}\cdot\vec{r}} \vec{\epsilon} \cdot \vec{\nabla} \psi_k(r) \right]^*. \quad (12.49)$$

or

$$\langle n|V_2|k \rangle = - \langle k|V_1|n \rangle^*. \quad (12.50)$$

This means that the two Einstein coefficients are equal, $B_{kn} = B_{nk}$. Physically, this is a consequence of the time reversal invariance of electromagnetism.

One body phase space

The transition probability per unit time for spontaneous emission is

$$W_{fi} = \frac{2\pi}{\hbar} |\langle f|V|i \rangle|^2 \delta(E_f + \hbar\omega - E_i) \quad (12.51)$$

where the perturbation is

$$V = \frac{e\hbar}{2mc} \left(\frac{2\pi\hbar c^2}{\omega V} \right)^{1/2} e^{i\vec{k}\cdot\vec{r}} \vec{\epsilon} \cdot \vec{\nabla}. \quad (12.52)$$

We must now confront the delta function. We do this by realizing that there are many possible states for the photon which can satisfy the delta function. Its energy $\hbar\omega = E_i - E_f$, and hence the magnitude of the photon's wave number, $\hbar|\vec{k}| = \hbar\omega/c$ is fixed, but the

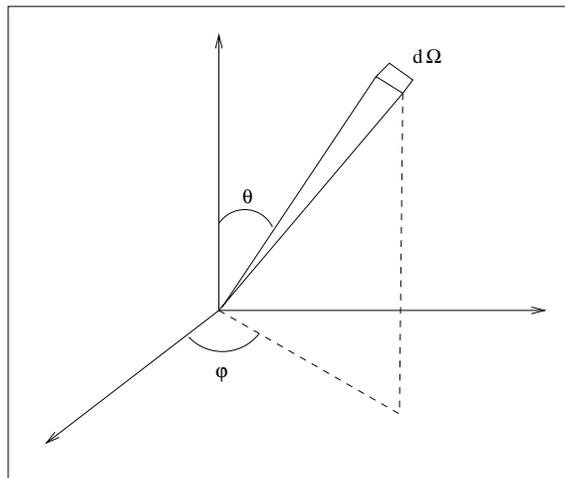


Figure 12.2: Geometry for the calculation of one body phase space: the photon is emitted into a region of solid angle $d\Omega$.

orientation of \vec{k} – the direction of emission of the photon – can be arbitrary. (For the careful reader, we are outside classical radiation theory when we do this.)

Place the decaying system at the origin of a coordinate system. The photon will be emitted into some solid angle $d\Omega$, as shown in Fig. 12.2. We consider the (differential) transition probability per unit time for the decay to occur while the photon is observed in an element of solid angle $d\Omega$. This rate is proportional to the number of available quantum states into which the photon can decay. This number is called a “phase space factor.” The differential transition probability becomes

$$\Delta W = \left[\frac{d\rho}{d\Omega} \right] d\Omega \frac{2\pi}{\hbar} |\langle f|V|i\rangle|^2 \delta(\Delta E - E_\gamma). \quad (12.53)$$

Any less-differential transition probability will involve an integral which will include the delta function. Rather than work directly with Eq. 12.53, let us proceed more generally.

Begin by thinking of the photon as being confined in a cubic box whose sides have length L . For convenience, we impose periodic boundary conditions on the solution of the wave equation. The components wave numbers of the modes will be quantized in units of $2\pi/L$ (for example, $k_x = \frac{2\pi}{L}n$, $n = 0, 1, \dots$). Therefore in a range of wave numbers between k_x and $k_x + dk_x$ there are $L/(2\pi)dk_x$ modes. Recalling that $p = \hbar k$, the number of modes between \vec{p} and $\vec{p} + d\vec{p}$ is

$$d\rho = \left(\frac{L}{2\pi\hbar} \right)^3 dp_x dp_y dp_z. \quad (12.54)$$

Since L^3 is the volume of the box, we can write this simply as

$$d\rho = \frac{V}{(2\pi\hbar)^3} d^3p. \quad (12.55)$$

This is the differential number of free particle states per momentum bin in a box of volume V , and hence this is our differential phase space factor $d\rho(E)$. (Often in the literature, the “differential” is inconsistently omitted, and the right side of Eq. 12.55 is written as $\rho(E)$.)

With this result, the transition probability for emission of a photon into an infinitesimal momentum (or wave number) bin d^3p is

$$dW = \frac{2\pi}{\hbar} |\langle f|V|i\rangle|^2 \delta(E_f - E_i + \hbar\omega) \frac{d^3p}{(2\pi\hbar)^3} V. \quad (12.56)$$

The overall factor of V seems peculiar. However, recall the normalization convention for a “single photon” electromagnetic wave:

$$\vec{A} = \left(\frac{2\pi\hbar c}{\omega V} \right)^{1/2} \vec{\epsilon} \left(e^{i(\vec{k}\cdot\vec{x} - \omega t)} + e^{-i(\vec{k}\cdot\vec{x} - \omega t)} \right) \quad (12.57)$$

Its factor of $1/\sqrt{V}$ will, when squared, cancel the factor of V from the phase space.

A final rewriting of the Golden Rule expression uses the relativistic relation between energy and momentum, $E = |p|c$, to convert the momentum integral to an integral over energy. In spherical coordinates

$$d^3p = \frac{d^3E_\gamma}{c^3} = \frac{E_\gamma^2 dE_\gamma d\Omega}{c^3}. \quad (12.58)$$

The transition probability per unit time for emission of a photon of energy E_γ into a unit of solid angle $d\Omega$ is, therefore,

$$dW = \frac{V}{(2\pi\hbar)^3} \frac{E_\gamma^2 dE_\gamma d\Omega}{c^3} |\langle f|V|i\rangle|^2 \frac{2\pi}{\hbar} \delta(E_\gamma - (E_i - E_f)) \quad (12.59)$$

This expression is readily integrated over the energy of the photon. The energy-conserving delta function gives sharp spectral lines to the emission spectrum. The differential decay rate $\frac{dW}{d\Omega}$ may have nontrivial angular dependence, which can be used to diagnose the quantum numbers of the initial and final states. The total decay rate summed over all final states $|f\rangle$ is the inverse lifetime of the state, $\sum_f W_{fi} = \Gamma = \frac{1}{\tau_i}$.

Multipole expansions

The matrix element in the Golden Rule expression for spontaneous emission is

$$\langle f|V|i\rangle = -\frac{ie\hbar}{mc} \left(\frac{2\pi\hbar c^2}{\omega V} \right)^{\frac{1}{2}} \int d^3r \psi_f^*(r) e^{i\vec{k}\cdot\vec{r}} \vec{\epsilon} \cdot \vec{\nabla} \psi_i(r). \quad (12.60)$$

It is obviously nontrivial to calculate. Fortunately, in most cases it is not necessary to exactly evaluate it. In typical atomic transitions the wavelengths of the photons are thousands of Angstroms while the size of the atom is only about an Angstrom. Thus $e^{i\vec{k}\cdot\vec{r}} \simeq 1$ everywhere that the atomic wave functions are nonzero. Setting $e^{i\vec{k}\cdot\vec{r}} = 1$ is called the dipole approximation. To see where this name comes from, write

$$\int d^3r \psi_f^*(r) \vec{\nabla} \psi_i(r) = \frac{i}{\hbar} \langle f|\vec{p}|i\rangle. \quad (12.61)$$

Use $\vec{p} = m\frac{d\vec{r}}{dt}$, the Heisenberg equation of motion, and the fact that $|f\rangle$ and $|i\rangle$ are eigenstates of the Hamiltonian to write

$$\langle f|\frac{d\vec{r}}{dt}|i\rangle = \frac{d}{dt} \langle f|\vec{r}|i\rangle = \frac{i}{\hbar} \langle f|[\hat{H}, \vec{r}]|i\rangle = i\omega_{fi} \langle f|\vec{r}|i\rangle \quad (12.62)$$

The matrix element in Eq. 12.61 becomes the matrix element of the dipole operator, $-e\vec{r}$,

$$\int d^3r \psi_f^*(r) \vec{\nabla} \psi_i(r) = -\frac{m\omega_{fi}}{\hbar} \int d^3r \psi_f^*(r) \vec{r} \psi_i(r), \quad (12.63)$$

so we see how the dipole approximation got its name.

In dipole approximation the Einstein coefficient is

$$B_{kn} = \frac{4\pi^2 e^2}{\hbar^2} |\langle k|\vec{\epsilon} \cdot \vec{r}|n\rangle|^2. \quad (12.64)$$

$\langle k|\vec{r}|n\rangle = 0$ defines a “forbidden transition.” It may only be forbidden in dipole approximation, not to all orders, and so the actual transition rate will be computed by expanding

$$e^{i\vec{k}\cdot\vec{r}} = 1 + i\vec{k} \cdot \vec{r} - \frac{1}{2}(\vec{k} \cdot \vec{r})^2 + \dots \quad (12.65)$$

and looking for non-vanishing terms. If the first $(n-1)$ terms in this series vanish the Γ is proportional to the n th term in the series. In atomic transitions, kr is on the order of 10^{-3} to 10^{-4} and “forbidden” in the dipole approximation means truly forbidden for all practical purposes.. In nuclear decays, kr may be on the order of 0.1 and so higher order transitions are much more common.

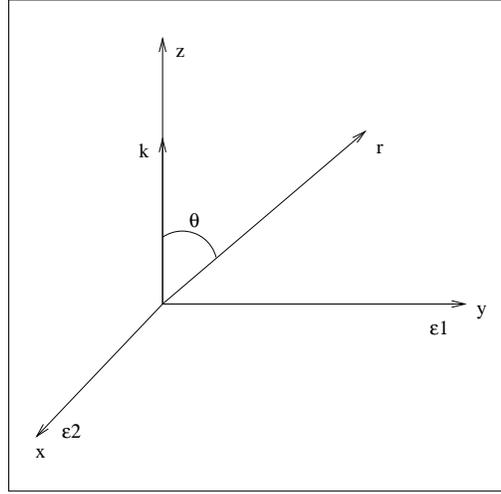


Figure 12.3: Geometry for evaluating the polarization sum.

Return to spontaneous emission. In dipole approximation, the transition probability per unit time for emission of a photon into an element of solid angle $d\Omega$ is

$$\frac{d\Gamma}{d\Omega} = \frac{V}{(2\pi\hbar)^3} \times \frac{E_\gamma^2 dE_\gamma}{c^3} \delta(E_{fi} - E_\gamma) \frac{2\pi}{\hbar} \left(\frac{e\hbar}{mc} \left(\frac{2\pi\hbar c^2}{\omega V} \right)^{\frac{1}{2}} \right)^2 \times |\vec{\epsilon} \cdot \langle f | \vec{\nabla} | i \rangle|^2. \quad (12.66)$$

Combining terms, exploiting Eq. 12.63. and integrating over the delta function gives

$$\frac{d\Gamma}{d\Omega} = \frac{\alpha c k_{fi}^3}{2\pi} |\vec{\epsilon} \cdot \langle f | \vec{r} | i \rangle|^2. \quad (12.67)$$

Commonly, the polarization of the emitted photon is not detected. In that case we must sum over both of the projected polarizations. Picking the \hat{z} direction to lie along the direction of the outgoing photon, as shown in Fig. 12.3, we have

$$\begin{aligned} \sum_{pol} |\vec{\epsilon} \cdot \vec{M}|^2 &= |\vec{\epsilon}_1 \cdot \vec{M}|^2 + |\vec{\epsilon}_2 \cdot \vec{M}|^2 \equiv |M_x|^2 + |M_y|^2 \\ &= |M_x|^2 + |M_y|^2 + |M_z|^2 - |M_z|^2 \\ &= |\vec{M}|^2 - \frac{|\vec{k} \cdot \vec{M}|^2}{k^2} \\ &= |M|^2 (1 - \cos^2 \theta) \end{aligned} \quad (12.68)$$

and thus

$$|\langle f | \vec{\epsilon} \cdot \vec{r} | i \rangle|^2 = |\langle f | \vec{r} | i \rangle|^2 \sin^2 \theta. \quad (12.69)$$

Eq. 12.67 becomes

$$\frac{d\Gamma}{d\Omega} = \frac{\alpha c k_{fi}^3}{2\pi} \sin^2 \theta |\langle f | \vec{r} | i \rangle|^2. \quad (12.70)$$

Performing the angular integral gives

$$\begin{aligned} \Gamma &= \frac{\alpha c k_{fi}^3}{2\pi} |\langle f | \vec{r} | i \rangle|^2 \int d\Omega \sin^2 \theta \\ &= \frac{\alpha c k_{fi}^3}{2\pi} |\langle f | \vec{r} | i \rangle|^2 \int_0^{2\pi} d\phi \int_{-1}^1 d(\cos \theta) (1 - \cos^2 \theta) \\ &= \frac{4\alpha c k_{fi}^3}{3} |\langle f | \vec{r} | i \rangle|^2 \end{aligned} \quad (12.71)$$

This is also the Einstein coefficient for spontaneous emission, A_{fi} , and $1/\Gamma$ is the lifetime of the state τ (assuming, of course, that this is the only allowed decay).

Note that there may be a nontrivial angular dependence in $\langle f | \vec{r} | i \rangle$ but the polarization sum followed by the integration over all space washes it out.

Note also that Γ scales like k_{fi}^3 . Since the energy difference between the two levels is equal to $\hbar c k_{fi}$, the larger the energy gap, the quicker the decay.

“Selection rules” tell us when $\Gamma = 0$. We can develop a sense for them by a looking at a series of examples. First, suppose the system is a one-dimensional simple harmonic oscillator oriented along the z-axis. The dipole matrix element is

$$\langle f | \vec{r} | i \rangle = \hat{z} \langle f | z | i \rangle = \int dz \psi_{n_f}^* z \psi_{n_i} \quad (12.72)$$

The operator, z can be expressed in terms of raising and lowering operators, yielding

$$\langle f | \vec{r} | i \rangle = \int dz \psi_{n_f}^* \sqrt{\frac{\hbar}{2m\omega}} (\sqrt{n_i} \delta_{n_f, n_i-1} + \sqrt{n_i+1} \delta_{n_f, n_i+1}) \psi_{n_i} \quad (12.73)$$

Therefore, the only allowed transitions are those with $\Delta n = 1$.

A physical example of this selection rule is in the spectroscopy of diatomic molecules. The nuclei vibrate and rotate in the minimum of a potential due to the electrons. Transitions between these states have selection rules: for the vibrational levels, which are harmonic oscillator states, the $\Delta v = 1$ selection rule is what we have just derived. Angular momentum controls the selection rules for changes in the rotational levels, and we consider that case next.

The inverse lifetime for the N th state (from the decay $N \rightarrow N - 1$ is

$$\frac{1}{\tau} = \frac{4e^2\omega^2}{3\hbar c^4} \left(\frac{\hbar N}{2m\omega} \right). \quad (12.74)$$

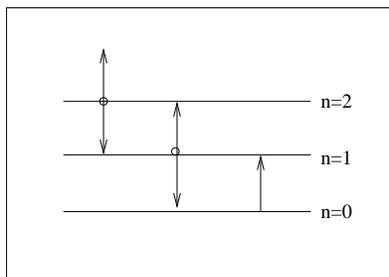


Figure 12.4: Diagram of allowed transitions between harmonic oscillator states.

The higher energy states decay faster. Notice that there is only a single spectral line!

In single electron atoms, or more generally, in transitions between states in a central potential, angular momentum considerations control selection rules. Neglect spin for the moment. The wave functions separate into radial and angular parts. $\psi(\vec{r}) = R(\vec{r})Y_{lm}(\theta, \phi)$. Write \vec{r} in a spherical tensor basis,

$$\vec{r} = \left(-\frac{\hat{x} - i\hat{y}}{\sqrt{2}}, \hat{z}, \frac{\hat{x} + i\hat{y}}{\sqrt{2}} \right) = r \left(-\frac{\hat{x} - i\hat{y}}{\sqrt{2}}, \hat{z}, \frac{\hat{x} + i\hat{y}}{\sqrt{2}} \right). \quad (12.75)$$

The required matrix element is

$$\begin{aligned} \langle f | \vec{r}_m | i \rangle &= \int r^2 dr R_f^* r R_i \int d\Omega Y_{l_f}^{m_f*} Y_1^m Y_{l_i}^{m_i} \\ &= \int r^3 dr R_f^* R_i \times I \end{aligned} \quad (12.76)$$

where I can be evaluated through the three spherical harmonic formula, Eq. 7.116,

$$I = \sqrt{\frac{2l_i + 1}{2l_f + 1}} \langle l_i, 1; 0, 0 | l_i, 1; l_f, 0 \rangle \langle l_i, 1; m_i, m | l_i, 1 | l_f, m_f \rangle. \quad (12.77)$$

For the matrix element of a particular component of \vec{r} to be non-vanishing, we need $m_f = m + m_i$. The Clebsch-Gordon coefficients allow $l_f = l_i + 1$ or $l_f = l_i - 1$, but $l_f = l_i$ is forbidden by parity. Thus we have the selection rules

$$\begin{aligned} \Delta m &= \pm 1, 0 \\ \Delta l &= \pm 1 \end{aligned} \quad (12.78)$$

As an example, let us calculate the rate of the $2p \rightarrow 1s$ electric dipole transition in hydrogen, neglecting spin. The $2p$ states are

$$|1m\rangle = Y_1^m(\Omega)R_{21}(r) \quad (12.79)$$

and the $1s$ state is

$$|00\rangle = Y_0^0(\Omega)R_{10}(r) \quad (12.80)$$

As $\hat{r}Y_0^{0*} = \sqrt{\frac{1}{3}}(Y_1^{-1}, Y_1^0, Y_1^1)$, it is easy to see that

$$\begin{aligned} \langle 00|\vec{r}|11\rangle &= -\sqrt{\frac{1}{3}}R\frac{\hat{x} + i\hat{y}}{\sqrt{2}} \\ \langle 00|\vec{r}|10\rangle &= \sqrt{\frac{1}{3}}R\hat{z} \\ \langle 00|\vec{r}|1-1\rangle &= \sqrt{\frac{1}{3}}R\frac{\hat{x} - i\hat{y}}{\sqrt{2}} \end{aligned} \quad (12.81)$$

where R is the radial integral (which is the same in all three cases). The squared modulus of all three of these matrix elements is equal, so the inverse lifetime of all of the $2p$ states is the same,

$$\frac{1}{\tau} = \frac{4e^2}{3\hbar c} \frac{1}{c^2} \left(\frac{\hbar\omega}{\hbar}\right)^2 \frac{R^2}{3} = \frac{4}{9}\alpha \frac{(E_{2p} - E_{1s})^3}{(\hbar c)^3} R^2 c. \quad (12.82)$$

Inserting numbers, the lifetime is $\tau \simeq 1.6 \times 10^{-9} s$.

This treatment did not include spin. Because the operator driving the transition, \vec{r} , is spin-independent, the transition cannot change the spin. The selection rule is $\Delta s = 0$, $\Delta m_s = 0$.

Finally, we can have states which are simply labeled by quantum numbers for total angular momentum, j and m . Because \vec{r} is a rank-1 tensor (a vector) the Wigner-Eckhart theorem tells us

$$\langle f j_f m_f | \vec{r} | i j_i m_i \rangle = -\frac{4\pi}{3} \langle |V_q^{(1)}| \rangle = \langle j_i, 1; m_i, q | j_i, 1; j_f, m_f \rangle \frac{\langle f j_f || r || i j_i \rangle}{\sqrt{2j_i + 1}}. \quad (12.83)$$

We are now equipped to calculate the intensities of spectral lines. Two kinds of problems are commonly encountered: First, we are given particular initial and final states, all of whose angular momentum quantum numbers are specified. An example of such a problem would be to imagine an atom in an external magnetic field. The initial and final states are Zeeman-split. What are the intensities of all the spectral lines between the two multiplets?

In the second example, we cannot tell which m_i goes to which m_f . We want to compute the absolute intensity of the spectral line, or the relative intensities of two lines from states with the same radial wave functions. In this case we must average over the initial states and

sum over the final states,

$$|\langle f|\vec{r}|i\rangle|^2 \rightarrow \sum_{m_f} \sum_{m_i} \frac{|\langle f j_f m_f | \vec{r} | i j_i m_i \rangle|^2}{2j_i + 1} \quad (12.84)$$

Again writing \vec{r} in a spherical tensor basis gives

$$\langle f j_f m_f | r_q | i j_i m_i \rangle = \langle j_i, 1; m_i, q | j_i, 1; j_f, m_f \rangle \frac{\langle f j_f || r || i j_i \rangle}{\sqrt{2j_i + 1}}. \quad (12.85)$$

Up to proportionality factors,

$$\Gamma = \sum_{m_i} \sum_{m_f} \sum_{q=-1}^1 \frac{|\langle f j_f || r || i j_i \rangle|^2}{(2j_i + 1)^2} |\langle j_i, 1; m_i, q | j_i, 1; j_f, m_f \rangle|^2. \quad (12.86)$$

The reduced matrix element has no m_i, m_f or q dependence, so

$$\Gamma = \frac{|\langle f j_f || r || i j_i \rangle|^2}{(2j_i + 1)^2} \sum_{m_i} \sum_{m_f} \sum_{q=-1}^1 \langle j_i, 1; j_f, m_f | j_i, 1; m_i, q \rangle \langle j_i, 1; m_i, q | j_i, 1; j_f, m_f \rangle \quad (12.87)$$

The sum collapses to unity from completeness, leaving

$$\Gamma \propto \frac{(2j_f + 1)}{(2j_i + 1)^2} |\langle f j_f || r || i j_i \rangle|^2. \quad (12.88)$$

To compute the reduced matrix element, one would typically calculate the full transition probability for some simple choice of m_i and m_f and then compare it to a calculation using the Wigner-Eckhart theorem.

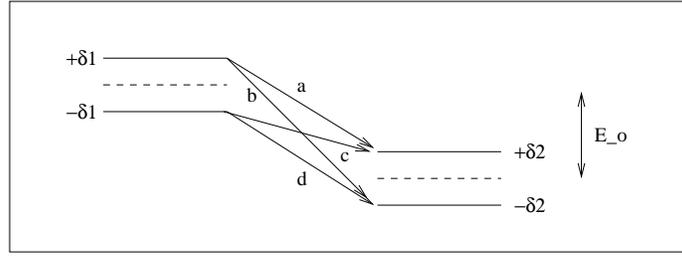
Example: $S \rightarrow P$ transitions in a single electron atom

We illustrate all this machinery in the simple case of an electron in an S state of a single-electron atom, which decays to a lower energy P state. First we consider the transition $S_{1/2} \rightarrow P_{1/2}$. We decompose the states into an $L - S$ basis:

$$|P_{1/2}\rangle = \begin{cases} |\frac{1}{2}, \frac{1}{2}\rangle = \sqrt{\frac{2}{3}} |11\rangle |\downarrow\rangle - \sqrt{\frac{1}{3}} |10\rangle |\uparrow\rangle \\ |\frac{1}{2}, -\frac{1}{2}\rangle = \sqrt{\frac{2}{3}} |1-1\rangle |\uparrow\rangle - \sqrt{\frac{1}{3}} |10\rangle |\downarrow\rangle \end{cases} \quad (12.89)$$

and

$$|S_{1/2}\rangle = \begin{cases} |\frac{1}{2}, \frac{1}{2}\rangle = |00\rangle |\uparrow\rangle \\ |\frac{1}{2}, -\frac{1}{2}\rangle = |00\rangle |\downarrow\rangle \end{cases} \quad (12.90)$$


 Figure 12.5: $S_{1/2} \rightarrow P_{1/2}$ transitions in an external magnetic field.

The following matrix elements are nonzero:

$$\begin{aligned}
 \left\langle 3\frac{1}{2}\frac{1}{2} \left| \vec{r} \right| 2\frac{1}{2}\frac{1}{2} \right\rangle &= -\sqrt{\frac{1}{3}} \frac{R}{\sqrt{3}} \hat{z} & (12.91) \\
 \left\langle 3\frac{1}{2}\frac{1}{2} \left| \vec{r} \right| 2\frac{1}{2}\frac{1}{2} \right\rangle &= \sqrt{\frac{2}{3}} \frac{R}{\sqrt{3}} \left(\frac{\hat{x} - i\hat{y}}{\sqrt{2}} \right) \\
 \left\langle 3\frac{1}{2}-\frac{1}{2} \left| \vec{r} \right| 2\frac{1}{2}\frac{1}{2} \right\rangle &= \sqrt{\frac{2}{3}} \langle 00 | \vec{r} | 11 \rangle \\
 \left\langle 3\frac{1}{2}\frac{1}{2} \left| \vec{r} \right| 2\frac{1}{2}\frac{1}{2} \right\rangle &= -\sqrt{\frac{1}{3}} \langle 00 | \vec{r} | 11 \rangle
 \end{aligned}$$

The relative intensities of the transitions will just be given by the ratios of the squares of the Clebsch-Gordon coefficients. In an external magnetic field, the splittings in the $S_{1/2}$ and $P_{1/2}$ multiplets are different (the Landé g -factors are different). This gives the following spectrum and relative intensity.

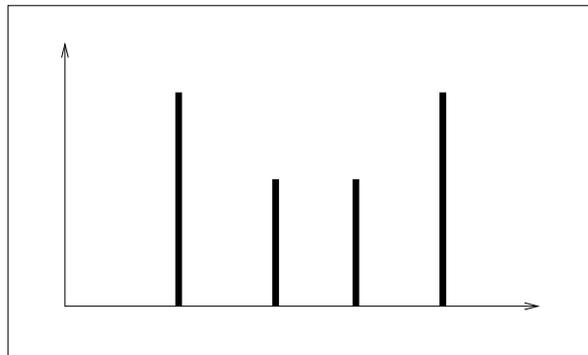
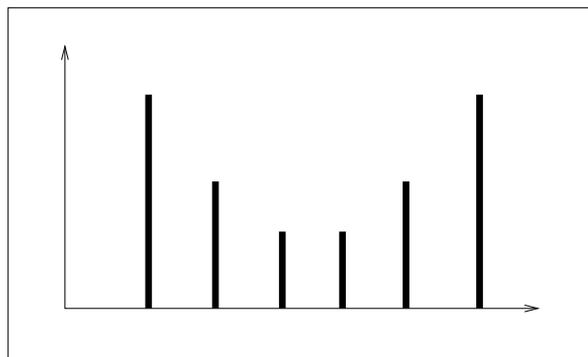
	Strength
$E_a = E_0 + \delta_1 - \delta_2$	1
$E_b = E_0 + \delta_1 + \delta_2$	2
$E_c = E_0 - \delta_1 - \delta_2$	2
$E_d = E_0 - \delta_1 + \delta_2$	1

(12.92)

A photograph of the multiplet of lines would resemble Fig. 12.6. The total transition rate will be $\frac{R^2}{3} \left(\frac{1}{3} + \frac{2}{3} + \frac{2}{3} + \frac{1}{3} \right) = (2/3)R^2$.

The $S_{1/2} \rightarrow P_{3/2}$ family of transitions is similar. The $2P_{3/2}$ states as

$$2P_{3/2} = \begin{cases} \left| \frac{3}{2}\frac{3}{2} \right\rangle = |11\rangle |\uparrow\rangle \\ \left| \frac{3}{2}\frac{3}{2} \right\rangle = \sqrt{\frac{1}{3}} |11\rangle |\downarrow\rangle + \sqrt{\frac{2}{3}} |10\rangle |\uparrow\rangle \\ \left| \frac{3}{2}-\frac{1}{2} \right\rangle = \sqrt{\frac{1}{3}} |1-1\rangle |\uparrow\rangle + \sqrt{\frac{2}{3}} |10\rangle |\downarrow\rangle \\ \left| \frac{3}{2}-\frac{3}{2} \right\rangle = |1-1\rangle |\downarrow\rangle \end{cases} \quad (12.93)$$

Figure 12.6: Spectral lines in a transition $S_{1/2} \rightarrow P_{1/2}$.Figure 12.7: Spectral lines in a transition $S_{1/2} \rightarrow P_{3/2}$.

Again we have the spin selection rule $\Delta m_s = 0$. This leaves six spectral lines. The relative intensities will be

	rel. intensity	
$ \langle 3\frac{1}{2}\frac{1}{2} \vec{r} 2\frac{3}{2}\frac{3}{2} \rangle ^2 = \frac{R^2}{3}$	3	(12.94)
$ \langle 3\frac{1}{2}\frac{1}{2} \vec{r} 2\frac{3}{2}\frac{1}{2} \rangle ^2 = \frac{R^2}{3} \times \frac{2}{3}$	2	
$ \langle 3\frac{1}{2}\frac{1}{2} \vec{r} 2\frac{3}{2}-\frac{1}{2} \rangle ^2 = \frac{R^2}{3} \times \frac{1}{3}$	1	

The spectrum is shown in Fig. 12.7. The complete m -averaged rates are then

$$\begin{aligned} \frac{1}{2} \sum_{s_i} \sum_{m_f} \Gamma(3S_{1/2} \rightarrow 2P_{1/2}) &= \frac{1}{2} \times 2 \times \frac{R^2}{3} \\ \frac{1}{2} \sum_{s_i} \sum_{m_f} \Gamma(3S_{1/2} \rightarrow 2P_{3/2}) &= \frac{1}{2} \times 4 \times \frac{R^2}{3} \end{aligned} \quad (12.95)$$

If there were no Zeeman field splitting these lines, we would see two spectral lines with a relative intensity of two to one. This ratio, plus the relative intensities of the Zeeman-split lines are diagnostics which, in principle should allow us to assign quantum numbers to the

states. (That's how it was done in the old days!)

We have only scratched the surface of the vast, technical subject of radiative processes, but perhaps we should leave it, to consider some applications which take us slightly farther afield.

Completing the story of blackbody radiation and Einstein coefficients

At the start of this chapter we recalled the Einstein relation between the A and B coefficients for emission and absorption by atoms in a black body, Eq. 12.10. The dipole A and B coefficients were also found to be

$$A_{fi} = W_{fi} = \frac{4}{3} \frac{e^2}{\hbar c^3} \omega^3 |\langle f | \vec{r} | i \rangle|^2; \quad (12.96)$$

$$B_{fi} = \frac{4\pi^2 e^2}{\hbar^2} |\langle f | \vec{\epsilon} \cdot \vec{r} | i \rangle|^2. \quad (12.97)$$

The B coefficient is the coefficient for absorption of radiation with a particular polarization, emanating from a particular direction. To reproduce the Einstein relation, we need to average B_{fi} over all incident polarizations and over all directions of the incident radiation. There are two polarizations so we need to calculate

$$\frac{1}{2} \sum_{\epsilon} B_{fi} = \frac{2\pi^2 e^2}{\hbar^2} \sum_{\epsilon} |\langle f | \vec{\epsilon} \cdot \vec{r} | i \rangle|^2 \quad (12.98)$$

We set up our coordinate system as in Fig. 12.8, where $\vec{\epsilon}_1$, $\vec{\epsilon}_2$ and \vec{k} make up a right-handed coordinate system. The polarization average is

$$\frac{1}{2} \sum_{\epsilon} B_{fi} = \frac{2\pi^2 e^2}{\hbar^2} \sin^2 \theta |\langle f | \vec{r} | i \rangle|^2. \quad (12.99)$$

The black body presents an isotropic distribution of radiation to be absorbed, and so we must also average over all angles.

$$\begin{aligned} \left\langle \frac{1}{2} \sum_{\epsilon} B_{fi} \right\rangle &= \frac{2\pi^2 e^2}{\hbar^2} |\langle f | \vec{r} | i \rangle|^2 \frac{\int d\phi \int d\theta \sin^2 \theta}{4\pi} \\ &= \frac{2\pi^2 e^2}{\hbar^2} |\langle f | \vec{r} | i \rangle|^2 \times \frac{1}{2} \times \frac{4}{3} = \frac{4\pi^2 e^2}{3\hbar^2} |\langle f | \vec{r} | i \rangle|^2 \end{aligned} \quad (12.100)$$

This gives

$$\frac{A_{fi}}{B_{fi}} = \frac{\frac{4e^2\omega^3}{4\hbar c^3}}{\frac{4\pi^2 e^2}{3\hbar^2}} = \frac{\hbar\omega^3}{\pi^2 c^3} \quad (12.101)$$

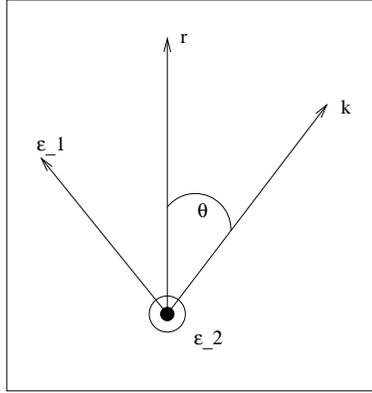


Figure 12.8: Coordinate system for polarizations and wave numbers, for averaging the B coefficient.

or

$$B_{fi} = \frac{\pi^2 c^3}{\hbar \omega^3} A_{fi} \quad (12.102)$$

which is the Einstein relation.

Spin-flip transitions

Both the proton and the electron are spin- $\frac{1}{2}$ fermions, so in the $1S$ state of hydrogen, the total spin of the two particles is $\vec{F} = \vec{S}_p + \vec{S}_e = 0, 1$. The $F = 1$ and $F = 0$ states are split by the hyperfine interaction, with a separation of 1420 MHz. Recall from Ch. 9 that the higher energy state is the $F = 1$ state. This state can decay to the $F = 0$ state if the electron flips its spin. The perturbing Hamiltonian is the usual one for a magnetic moment in a magnetic field,

$$\hat{H} = -\vec{\mu} \cdot \vec{B} \quad \vec{\mu} = \frac{e\hbar}{2mc} \vec{\sigma}. \quad (12.103)$$

where this time the field is that of the emitted photon. (We ignore the contribution of the proton because its magnetic moment is tiny compared to the electron's. In positronium there would be a contribution from both particles.) The photon's vector potential is

$$\vec{A}_{\text{photon}} = \left(\frac{2\pi\hbar c^2}{\omega V} \right)^{\frac{1}{2}} \vec{\epsilon} \left[e^{i(\vec{k} \cdot \vec{x} - \omega t)} + e^{-i(\vec{k} \cdot \vec{x} - \omega t)} \right] \quad (12.104)$$

and so the magnetic field is

$$\vec{B} = \vec{\nabla} \times \vec{A}_{\text{photon}} = i\vec{k} \times \vec{\epsilon} \left(\frac{2\pi\hbar c^2}{\omega V} \right)^{\frac{1}{2}} \vec{\epsilon} \left[e^{i(\vec{k} \cdot \vec{x} - \omega t)} - e^{-i(\vec{k} \cdot \vec{x} - \omega t)} \right]. \quad (12.105)$$

Thus the Hamiltonian is

$$\hat{H} = -\frac{ie\hbar}{2mc} \left(\frac{2\pi\hbar c^2}{\omega V} \right)^{\frac{1}{2}} \sigma \cdot (\vec{k} \times \vec{\epsilon}) \left[e^{i(\vec{k} \cdot \vec{x} - \omega t)} - e^{-i(\vec{k} \cdot \vec{x} - \omega t)} \right] \quad (12.106)$$

The two exponential terms will give rise to transitions up and down in energy between the two states. The 1420 MHz line is in the microwave band and the wavelength is approximately 21 cm, which means

$$kr \approx \frac{2\pi}{21 \text{ cm}} \times \frac{1}{2} \text{Å} \approx 1.2 \times 10^{-7}. \quad (12.107)$$

The dipole approximation should work well. The final polarization summed differential transition probability per unit time is then

$$d\Gamma = \frac{1}{8\pi} \alpha \frac{E_\gamma}{(mc^2)^2} \hbar c^2 dE_\gamma \delta(E_\gamma - \Delta E) \sum_{\epsilon} |\langle f | \vec{\sigma} \cdot (\vec{k} \times \vec{\epsilon}) | i \rangle|^2. \quad (12.108)$$

We evaluate the matrix element between the states

$$\begin{aligned} |i\rangle &= |10\rangle = \frac{1}{\sqrt{2}} (|\uparrow_e \downarrow_p\rangle + |\downarrow_e \uparrow_p\rangle) \\ |f\rangle &= |00\rangle = \frac{1}{\sqrt{2}} (|\uparrow_e \downarrow_p\rangle - |\downarrow_e \uparrow_p\rangle). \end{aligned} \quad (12.109)$$

Only the matrix element of σ_z will be non-vanishing, and it is

$$\frac{1}{2} (\langle \uparrow \downarrow | - \langle \downarrow \uparrow |) \vec{\sigma}_{e,z} (|\uparrow \downarrow\rangle + |\downarrow \uparrow\rangle) = 1. \quad (12.110)$$

To deal with the polarizations, we define a coordinate system where $\hat{e}_1 = \hat{x}$ and $\hat{e}_2 = \hat{k} \times \hat{e}_1$, as shown in Fig. 12.9. With this choice, $(\vec{k} \times \hat{e}_2)_z = 0$ and $(\vec{k} \times \hat{e}_1)_z = -k \sin \theta$. The inverse lifetime is thus

$$\Gamma = \frac{1}{\tau} = \frac{\alpha}{8\pi} \frac{E_\gamma^3}{(mc^2)^2} \frac{c}{\hbar c} \int d\phi d\cos\theta \sin^2\theta = \frac{\alpha}{3} \frac{E_\gamma^3}{(mc^2)^2} \frac{c}{\hbar c}. \quad (12.111)$$

We have used the fact that $|\vec{k}| = k = E_\gamma/(\hbar c)$. The numerical value of the lifetime is about 2×10^{14} sec. In the real world, this is too slow to compete with more likely processes, like collisional de-excitation. However, this is an interesting decay mechanism in the physics of heavy quark systems.

Note that the perturbation does not couple to the spatial part of the wave function, and so the matrix element could be calculated without knowledge of it. If the initial and final states had different spatial wave functions (for example, in the transition $2S \rightarrow 1S$), the transition would vanish in dipole approximation. One would have to go to higher order in the multipole expansion to get a nonzero rate.

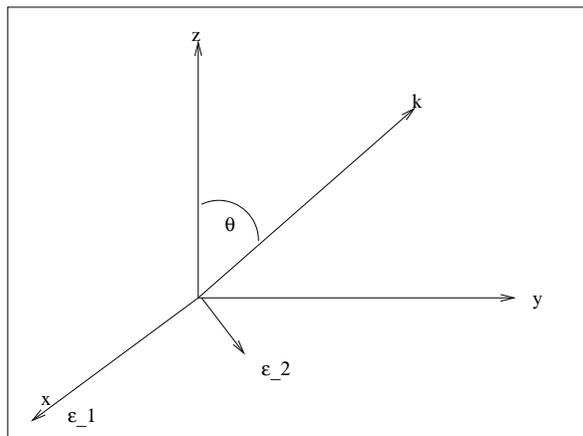


Figure 12.9: Coordinate system for the spin-flip matrix element'

Line breadth

In Ch. 11 we introduced the “energy-time uncertainty relation” as the connection between the time over which a transition probability was observed and the degree of energy non-conservation which could occur. We now introduce a second version of this relation, in a connection between the lifetime of a state and the width of the spectral line associated with transitions as the state decays. We make this connection using perturbation theory, although it is more general.

Suppose that long ago the system was in a particular state. As its wave function evolves with time, one may ask for the probability that the system remains in the same state far in the future. When we answer this question, we must be careful to switch the perturbing potential on and off adiabatically. The physical reason for doing this is that if the perturbation was switched on, we could not construct eigenstates of the zeroth order Hamiltonian. The mathematical reason for doing this is that the calculation is delicate. Otherwise, we assume that the perturbation is time-independent. Thus, we write

$$V \rightarrow V e^{\eta t} \quad (12.112)$$

and take the limit $\eta \rightarrow 0$ at the end of the calculation. We work in the interaction representation. The $t = -\infty$ state is $|\psi(t)\rangle = |i\rangle$.

Consider first the transition amplitude to a state $|n\rangle$, which is not the same as the initial

state. The first order transition amplitude is

$$U_{ni}(t) = 0 - \frac{i}{\hbar} \langle n|V|i\rangle \int_{t_0}^t dt' e^{\eta t'} e^{i\omega_n t'} e^{-i\omega_i t'} \quad (12.113)$$

Performing the time integral and then and (setting $t_0 \rightarrow -\infty$) gives

$$U_{ni}(t) = -\frac{i}{\hbar} \frac{\langle n|V|i\rangle}{i(\omega_n - \omega_i) + \eta} e^{(\eta+i(\omega_n-\omega_i))t} \quad (12.114)$$

Let us check that this gives the Golden Rule result, setting $\eta \rightarrow 0$ at the end of the calculation). The transition probability is

$$|U_{ni}(t)|^2 = \frac{1}{\hbar^2} \frac{|\langle n|V|i\rangle|^2}{(\omega_n - \omega_i)^2 + \eta^2} e^{2\eta t} \quad (12.115)$$

If the transition probability $|U_{ni}(t)|^2$ is equal to the time interval times the (time independent) transition probability per unit time, $|U_{ni}(t)|^2 = t \times W_{ni}$, then we can write $W_{ni} = d|U_{ni}(t)|^2/dt$ because W_{ni} should by definition have no explicit time dependence. Performing this calculation gives

$$W_{ni} = \frac{2\eta}{\hbar^2} \frac{|\langle n|V|i\rangle|^2}{(\omega_n - \omega_i)^2 + \eta^2} e^{2\eta t} \quad (12.116)$$

As $\eta \rightarrow 0$ $e^{2\eta t} \rightarrow 1$, but evaluating $\frac{\eta}{\eta^2 + \omega^2}$ requires care. The limiting form which is appropriate in the sense of contour integration is

$$\lim_{\eta \rightarrow 0} \frac{\eta}{\eta^2 + \omega^2} = \pi \delta(\omega). \quad (12.117)$$

The transition probability per unit time takes its familiar form

$$W_{ni} = \frac{2\pi}{\hbar} \delta(E_i - E_n) |\langle n|V|i\rangle|^2 \quad (12.118)$$

after rescaling the argument of the delta function.

Next we calculate the amplitude for the state to remain itself, when probed at very long times in the future. The transition amplitude is

$$\begin{aligned} U_{ii}(t) &= 1 - \frac{i}{\hbar} \langle i|V|i\rangle \int_{t_0}^t dt' e^{\eta t'} \\ &+ \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 e^{\eta t_1} e^{\eta t_2} \sum_m |\langle i|V|m\rangle|^2 e^{i\omega_i t_1} e^{-i\omega_m(t_1-t_2)} e^{-i\omega_i t_2} \end{aligned} \quad (12.119)$$

The first order term integrates to

$$-\frac{i}{\hbar} \langle i|V|i\rangle \int_{t_0}^t dt' e^{\eta t'} = -\frac{i}{\hbar} \frac{\langle i|V|i\rangle}{\hbar\eta} e^{\eta t}. \quad (12.120)$$

The second order term is

$$\begin{aligned} U_{ii}^{(2)} &= \left(-\frac{i}{\hbar}\right)^2 \sum_m |\langle i|V|m\rangle|^2 \frac{1}{-i(\omega_m - \omega_i) + \eta} \int_{-\infty}^t dt_1 e^{(\eta+i(\omega_i-\omega_m))t_1} e^{-i(\omega_i-\omega_m)t_1} e^{\eta t_1} \\ &= \left(-\frac{i}{\hbar}\right)^2 c \sum_m \frac{|\langle i|V|m\rangle|^2}{-i(\omega_i - \omega_m) + \eta} \int_{-\infty}^t dt_1 e^{2\eta t_1}. \end{aligned}$$

We separate the $m = i$ from the sum, so

$$U_{ii}^{(2)} = \left(-\frac{i}{\hbar}\right)^2 \frac{|\langle i|V|i\rangle|^2}{2\eta^2} e^{2\eta t} + \left(-\frac{i}{\hbar}\right) \sum_{m \neq i} \frac{|\langle i|V|m\rangle|^2}{2\eta^2 \hbar (E_i - E_m + i\hbar\eta)}. \quad (12.121)$$

We now want to show that

$$i\hbar \frac{dU_{ii}}{dt} = \Delta_i U_{ii} \quad (12.122)$$

or equivalently

$$U_{ii}(t) = e^{-i\Delta_i t/\hbar}, \quad (12.123)$$

where Δ_i is a constant. To show that

$$\frac{dU_{ii}/dt}{U_{ii}} = -\frac{i}{\hbar} \Delta_i \quad (12.124)$$

for constant Δ_i is equivalent. We do this by beginning with the explicit ratio

$$\frac{dU_{ii}/dt}{U_{ii}} = \frac{-\frac{i}{\hbar} \langle i|V|i\rangle e^{2\eta t} + \left(-\frac{i}{\hbar}\right)^2 \frac{\langle i|V|i\rangle}{\eta} e^{2\eta t} + \left(-\frac{i}{\hbar}\right) \sum_{m \neq i} \frac{|\langle i|V|m\rangle|^2}{E_i - E_m + i\hbar\eta}}{1 - \frac{i}{\hbar} \frac{\langle i|V|i\rangle}{\eta} e^{2\eta t} + \dots} \quad (12.125)$$

We are doing perturbation theory, so we simply expand the ratio as a power series in V ,

$$\begin{aligned} \frac{dU_{ii}/dt}{U_{ii}} &= -\frac{i}{\hbar} \langle i|V|i\rangle \left(1 + \frac{i}{\hbar} \frac{\langle i|V|i\rangle}{\eta}\right) + \left(-\frac{i}{\hbar}\right)^2 \frac{|\langle i|V|i\rangle|^2}{\eta} + \left(-\frac{i}{\hbar}\right) \sum_{m \neq i} \frac{|\langle i|V|m\rangle|^2}{E_i - E_m + i\hbar\eta} \\ &= -\frac{i}{\hbar} \left(\langle i|V|i\rangle + \sum_{m \neq i} \frac{|\langle i|V|m\rangle|^2}{E_i - E_m + i\hbar\eta} \right) \end{aligned} \quad (12.126)$$

$$(12.127)$$

or

$$\Delta_i = \langle i|V|i\rangle + \sum_{m \neq i} \frac{|\langle i|V|m\rangle|^2}{E_i - E_m + i\hbar\eta}. \quad (12.128)$$

To make sense of this relation, recall that we have been working in interaction representation. We convert it to Schrödinger representation by re-introducing the zeroth-order energy into the exponential, so that

$$U_{ii}(t) = \exp(-i(E_i + \Delta_i)t/\hbar). \quad (12.129)$$

Δ_i is complex. Its real part corresponds to an energy shift. To first order $\Delta_i = \langle i|V|i\rangle$ which is the usual result from time-independent perturbation theory. The second order expression looks familiar, but what about the $i\hbar\eta$ in the denominator? We recall the identity (only defined under an integral)

$$\frac{1}{x+i\epsilon} = \mathcal{P}\frac{1}{x} + i\pi\delta(x) \quad (12.130)$$

where \mathcal{P} stands for the principal value. With this relation we can decompose the second order piece of Δ_i into a real part,

$$\text{Re}(\Delta_i^{(2)}) = \mathcal{P} \sum_{m \neq i} \frac{|\langle i|V|m\rangle|^2}{E_i - E_m} \quad (12.131)$$

which is just the second order energy shift, and an imaginary part,

$$\text{Im}(\Delta_i^{(2)}) = \pi \sum_{m \neq i} |\langle i|V|m\rangle|^2 \delta(E_i - E_m) = \frac{\hbar}{2} \sum_{m \neq i} W_{mi} \quad (12.132)$$

which is just the sum of W_{mi} transition probabilities per unit time, scaled by \hbar to become energies. The amplitude for the state to remain itself is thus

$$U_{ii}(t) = \exp^{-i\text{Re}(\Delta_i)t/\hbar} \exp^{-\sum_{m \neq i} W_{mi}t/2}. \quad (12.133)$$

The squared modulus of this quantity gives probability that the state persists. It is

$$|U_{ii}(t)|^2 = \exp -t/\tau \quad (12.134)$$

where

$$\frac{1}{\tau} = \sum_{m \neq i} W_{mi}. \quad (12.135)$$

This makes complete sense: the inverse lifetime of the state is the sum of transition probabilities per unit time for decays out of the initial state.

Notice that the state itself no longer evolves in time as a pure phase. It is still the case that

$$|i(t)\rangle = U_{ii}(t) |i\rangle, \quad (12.136)$$

but now

$$U_{ii}(t) = \exp\left(-\frac{i}{\hbar}\epsilon t - \frac{\Gamma t}{2\hbar}\right) \quad (12.137)$$

Carry this function into the energy domain by Fourier transform:

$$\begin{aligned} f(E) &= \int dt e^{iEt/\hbar} U_{ii}(t) \\ &= \int dt e^{it(E-\epsilon)/\hbar} e^{-\Gamma t/2\hbar} \\ &= \frac{1}{E - \epsilon - \frac{i\Gamma}{2}} \end{aligned} \quad (12.138)$$

The absolute square of this expression is

$$|f(E)|^2 = \frac{1}{(E - \epsilon)^2 + \frac{\Gamma^2}{4}}, \quad (12.139)$$

the line shape for a “Lorentzian.” We encountered this function before, in Eq. 12.42. A sharp energy state, whose time evolution is given by a simple $\exp(-i\omega t)$, would be a delta-function in the energy domain. The fact that the state can decay broadens the delta function. The width of the state – its uncertainty in energy – is equal to Γ , which is \hbar/τ . Thus $\Gamma\tau = \Delta E\tau = \hbar$. This is the careful expression of the “energy-time uncertainty relation.”

The quantity Γ can be observed in spectral lines from transitions involving the state. Because the state is not sharp, the spectral line will also have a Lorentzian line shape. We are describing what is called the “natural width” of the spectral line. In practice, many other effects can contribute to nonzero line breadth, including Doppler broadening, the shift of the photon’s frequency because the source is moving. We will see Lorentzians again, in Ch. 13.

Coupled channels

An important ingredient in the study of condensed atomic gases is a so-called “Feshbach resonance.” The idea is that perturbing the system with an external electric or magnetic field, one can drive an energy level very close to zero energy. Then the system will become strongly interacting, in the sense that the scattering length will diverge. (This point will be discussed in Ch. 13.) We can give a schematic illustration of this phenomenon through the following solvable model.

Suppose we have a system with a set of “ordinary” energy levels ϵ_i where $i = 1, 2, \dots, N$. Also let us have one “special” energy level which we denote by ϵ_0 . We assume that we

have control over the “special” energy level ϵ_0 through some laboratory apparatus. The “ordinary” energy levels are also assumed to be decoupled from one another, but not from the “special” energy level. We can then write down the Hamiltonian in matrix form

$$\hat{H} = \begin{bmatrix} \epsilon_0 & V_{01} & V_{02} & V_{03} & \cdots \\ V_{10} & \epsilon_1 & 0 & 0 & \cdots \\ V_{20} & 0 & \epsilon_2 & 0 & \cdots \\ V_{30} & 0 & 0 & \epsilon_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (12.140)$$

One can, of course, find the energy eigenvalues of this system by solving the secular equation

$$\det[\hat{H} - \delta_{ij}E] = 0. \quad (12.141)$$

Write out the determinant as

$$\begin{aligned} 0 = (\epsilon_0 - E) \det \begin{bmatrix} \epsilon_1 - E & 0 & \cdots \\ 0 & \epsilon_2 - E & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} - V_{01} \det \begin{bmatrix} V_{10} & 0 & \cdots \\ V_{20} & \epsilon_2 - E & \cdots \\ V_{30} & 0 & \epsilon_3 - E \end{bmatrix} \\ + V_{02} \begin{bmatrix} V_{10} & \epsilon_1 - E & \cdots \\ V_{20} & 0 & \cdots \\ V_{30} & 0 & \epsilon_3 - E \end{bmatrix} + \cdots \end{aligned} \quad (12.142)$$

This is

$$0 = (\epsilon_0 - E) \prod_i (\epsilon_i - E) - |V_{01}|^2 (\epsilon_2 - E)(\epsilon_3 - E) \cdots - |V_{01}|^2 (\epsilon_1 - E)(\epsilon_3 - E) \cdots + \cdots \quad (12.143)$$

or

$$0 = (\epsilon_0 - E) \prod_i (\epsilon_i - E) - \sum_i |V_{0i}|^2 \prod_{j \neq i} (\epsilon_j - E). \quad (12.144)$$

There are $N + 1$ roots to this characteristic equation. Let's focus on the case where one of them, E is assumed to be close to ϵ_0 . In that case E is given by the implicit equation

$$(\epsilon_0 - E) - \sum_i \frac{|V_{0i}|^2}{(\epsilon_i - E)} = 0 \quad (12.145)$$

or

$$E = \epsilon_0 + \sum_i \frac{|V_{0i}|^2}{(E - \epsilon_i)}. \quad (12.146)$$

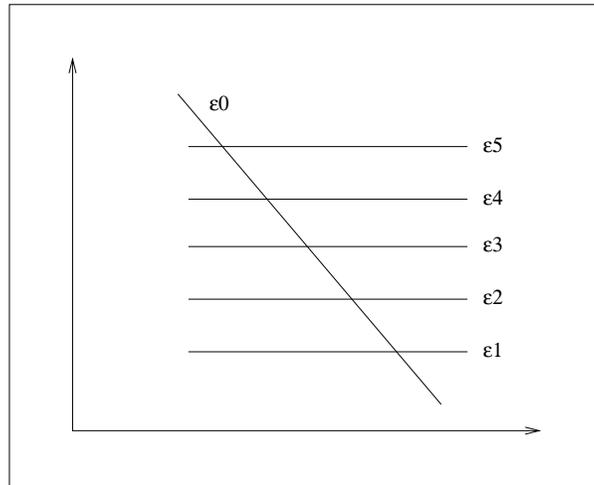


Figure 12.10: Naively tuning ϵ_0 across the other e_i 's.

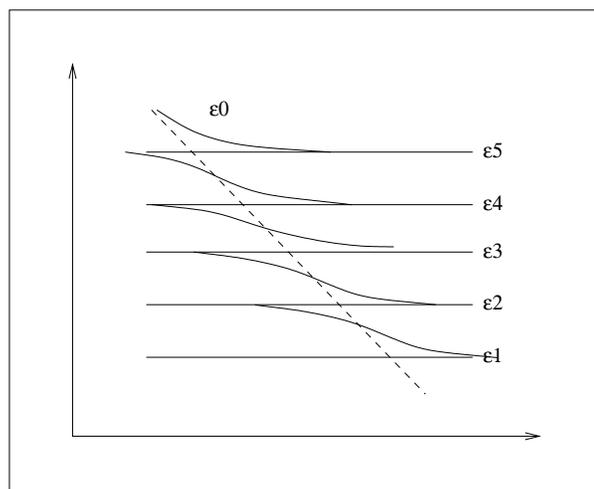


Figure 12.11: As the energy ϵ_0 is tuned across the other e_i 's, the true energy levels mix and repel.

Imagine varying our laboratory apparatus so that we vary the energy ϵ_0 . The naive energy spectrum for the system is shown in Fig. 12.10. But we know that in quantum mechanics levels never cross. A more accurate picture would be shown in Fig. 12.11.

The secular equation knows about level repulsion and mixing. Single out another level in the sum, ϵ_1 . The secular equation is

$$\epsilon_1 - E = \frac{|V_{01}|^2}{(\epsilon_0 - E)} + \sum_{i \neq 1} \frac{|V_{0i}|^2}{(\epsilon_i - E)} \frac{(\epsilon_1 - E)}{(\epsilon_0 - E)} \quad (12.147)$$

When $\epsilon_0 \approx \epsilon_1$ this simplifies to

$$\epsilon_1 - E = \frac{|V_{01}|^2}{\epsilon_0 - E} + \text{small corrections} \quad (12.148)$$

This is the secular equation for the Hamiltonian

$$\hat{H} = \begin{bmatrix} \epsilon_1 & V_{10} \\ V_{01} & \epsilon_0 \end{bmatrix}. \quad (12.149)$$

Of course, the true energy levels are not ϵ_1 nor ϵ_2 , and the eigenstates are mixtures of $|0\rangle$ and $|i\rangle$. Altering the position of the “special” level affects the energies of the nearby “ordinary” levels.

When the spectrum of ordinary levels forms a continuum, the sum is promoted to an integral, and it is convenient to break up the denominator into a principle value part and a delta function

$$\frac{1}{E - \epsilon} \rightarrow \mathcal{P} \frac{1}{E - \epsilon} + i\pi\delta(E - \epsilon). \quad (12.150)$$

The eigenvalue equation has real and imaginary parts, with

$$\text{Re}(E) = \epsilon_0 + \mathcal{P} \int d\epsilon \frac{|V_{0\epsilon}|^2}{(E - \epsilon)} \quad (12.151)$$

and

$$\text{Im}(E) = \int d\epsilon |V_{0\epsilon}|^2 \pi \delta(E - \epsilon) \quad (12.152)$$

The self-consistent solution of the real part of the equation would give the energy levels. The imaginary part of the expression gives the decay width of the state.

Chapter 13

Scattering

Scattering via perturbation theory

Scattering experiments are among the most important tools of the physicist. A beam of particles bombards a target. Particles emerge from the collision, moving at some angle with respect to the beam, and are seen by detectors. The rate of scattering, as parameterized by cross sections, angular distributions, energy loss, and so on, give information about the nature of the interactions and about the internal structure of the target.

Indeed, in my own field of elementary particle physics, almost all the experimental information we have comes from scattering experiments. Even in those cases where we observe the decay of excited states, the states whose spectra we study are produced in collisions of other particles.

There are two ways to study scattering. The first is in the context of perturbation theory and leads to the “Born approximation” and its higher order generalizations. This is an appropriate approach if the forces causing the scattering are relatively weak and the wavelength of the particle is relatively small. The second methodology works directly with the time independent Schrödinger equation $(H - E)\psi = 0$, with $E > 0$ and boundary conditions such that the solution corresponds to an incoming plus an outgoing wave. These two pictures are complementary.

Kinematic preliminaries

Suppose we have an object whose area A we wish to determine. A way to measure this area would be to use the object as a target and to bombard it with particles. One could measure the area as the counts per second or particle hits per second, divided by the beam current I (particles in the beam per unit area per second). To generalize this idea, think of the scattering of the beam from a portion of the target a distance s from its center, which scatters the beam in a direction labeled by a solid angle Ω where it is seen by a counter of angular size $d\Omega$. The counting rate $dN(\theta)$ is

$$dN(\theta) = I(sds)d\phi = Id\sigma \quad (13.1)$$

where $d\sigma$ is a differential area element, so

$$\frac{dN(\theta)}{I} = d\sigma. \quad (13.2)$$

The counting rate may depend on θ so we write the scattering rate as

$$\frac{dN(\theta, \phi)}{d\Omega} = I \frac{d\sigma}{d\Omega} \quad (13.3)$$

where $d\sigma/d\Omega$ is called the “differential cross section.” We also define the “total cross section” to be the counts per unit time in a detector with 4π solid angle resolution, divided by the beam current

$$\sigma = \frac{1}{I} \int d\Omega \frac{dN}{d\Omega} \equiv \int \frac{d\sigma}{d\Omega} d\Omega \quad (13.4)$$

The total cross section σ may not have any physical meaning. For example, we will see that in the case of Coulomb scattering, the differential cross section behaves as

$$\frac{d\sigma}{d\Omega} \simeq \frac{1}{\sin^4 \theta/2} \quad (13.5)$$

as θ vanishes. In this case the total cross section is divergent.

Let us specialize to two-body scattering (two particles, beam and target, going in and coming out). We already know that we can write the coordinates in terms of relative and center of mass variables, and that in the center of mass frame the two bodies can be treated as a single body of reduced mass μ . Scattering is easiest to visualize in the center of mass frame. Both before and after the collision, the particles' momenta are equal and oppositely directed. The particles emerge back to back at some angle θ_c with respect to the original direction. We will develop scattering theory in the center of mass frame and defer the complicated (though trivial) problem of converting results from this frame to the laboratory frame to any good book about classical mechanics.

Scattering experiments are usually performed in the laboratory in either of two situations. The first is the frame with one particle (call its mass M) at rest and the other (of mass m) moving with velocity v_0 . For historical reasons this is called the “laboratory frame.” This label gives short shrift to modern particle physics experiments, which use colliding beams. Here the laboratory and center of mass frames coincide. In the center of mass frame the two particles have momenta \vec{p}_1 and $\vec{p}_2 = -\vec{p}_1$, and this frame moves at a velocity \bar{v} with respect to the lab frame. Then the components of the two momenta along \bar{v} are $p_1 = m(v_0 - \bar{v})$ and $p_2 = -M\bar{v}$. Setting $p_1 + p_2 = 0$ yields the velocity of the center of mass frame as $\bar{v} = m/(M + m)v_0$. Then $p_2 = -Mm\bar{v}/(M + m) = -\mu v_0$ where $\mu = mM/(M + m)$ is the reduced mass. The other particle's momentum is $p_1 = mv_0(1 - m/(M + m)) = \mu v_0$. In the center of mass, the kinetic energy of the two particles is $p_2^2/(2M) + p_1^2/(2m) = (1/2)\mu v_0^2$. In the lab frame, the kinetic energy is $(1/2)mv_0^2 = (1/2)\mu v_0^2 + (1/2)(M + m)\bar{v}^2$, where the

second term is the kinetic energy due to the motion of the frame. There are of course similar formulas for relativistic kinematics. (In that case, it is much more convenient to use kinematic variables with simple Lorentz transformation properties.) Notice that in the limit that M/m becomes very large, $\mu = m$ and $\bar{v} = 0$.

The Born approximation

Let us consider the scattering potential as a perturbation. The two particles are assumed to be noninteracting at large separation and approach each other as plane waves. They interact weakly, scatter through some angle θ and emerge as free particles or plane waves. The transition probability per unit time is given by the Golden Rule.

In the center of mass frame we treat the two scatterers as a single particle of reduced mass μ . Its initial state has momentum \vec{p}_i and its final state has momentum \vec{p}_f . In elastic scattering the initial and final energies are equal, so $p_i^2/(2\mu) = p_f^2/(2\mu)$. The Golden Rule tells us that the transition probability per unit time Γ due to a perturbing potential V is

$$d\Gamma = \frac{2\pi}{\hbar} |\langle \vec{p}_f | V | \vec{p}_i \rangle|^2 \delta(E_f - E_i) d\rho(E_f) \quad (13.6)$$

where the phase space factor is

$$d\rho(E_f) = V \frac{d^3 p_f}{(2\pi\hbar)^3} \quad (13.7)$$

for free particles in a box of volume V . This is identical to the phase space factor we found for photons, Eq. 12.55, because the derivation just involved counting plane wave states in a box. What is different is the energy - momentum dispersion relation, $p_f^2/(2\mu) = E_f$ for non-relativistic motion as opposed to $E_f = cp_f$ for the massless photon.

The zeroth order initial and final states are noninteracting free particle states

$$\psi_i = \frac{1}{\sqrt{V}} \exp(i\vec{p}_i \cdot \vec{r}/\hbar) \quad (13.8)$$

and

$$\psi_f = \frac{1}{\sqrt{V}} \exp(i\vec{p}_f \cdot \vec{r}/\hbar). \quad (13.9)$$

Note that we adopt a “box normalization” for the states,

$$\int_V d^3x |\psi|^2 = 1 \quad (13.10)$$

to adapt to the same conventions that are used in the phase space definition.

We must now convert $d\Gamma$ into a differential cross section. The counting rate (particles per unit time) dN is just $d\Gamma$. The differential cross section is dN divided by the particle current I , which is just the density of particles in the beam (one particle per volume V , in our convention for the initial state wave function) times the particle velocity $|\vec{p}_i|/\mu$. Thus

$$d\sigma = \frac{2\pi}{\hbar} \frac{\mu V}{|p_i|} |\langle \vec{p}_f | V | \vec{p}_i \rangle|^2 \delta(E_f - E_i) V \frac{d^3 p_f}{(2\pi\hbar)^3}. \quad (13.11)$$

Note that the volume factors cancel between the phase space factor, the current, and the normalization of the initial and final states.

It is expedient to rewrite the phase space factor to make the integral over the energy-conserving delta function trivial. Observing that

$$d^3 p_f \delta(E_f - E_i) = p_f^2 dp_f d\Omega \delta(E_f - E_i) = p_f (p_f dp_f) d\Omega \delta(E_f - E_i), \quad (13.12)$$

we use $E_f = p_f^2/(2\mu)$ to make the replacement $dE_f = p_f dp_f/\mu$. Then the phase space factor becomes $\mu p_f dE_f d\Omega \delta(E_f - E_i)$. Integrating the energy over the delta function gives the differential cross section as

$$\frac{d\sigma}{d\Omega} = \frac{2\pi}{\hbar} \frac{\mu V^2}{|p_i|} |\langle \vec{p}_f | V | \vec{p}_i \rangle|^2 \frac{\mu |p_f|}{(2\pi\hbar)^3}. \quad (13.13)$$

For elastic scattering $|p_i| = |p_f|$ and the momentum factors cancel, leaving

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \left(\frac{\mu V}{2\pi\hbar^2}\right)^2 |\langle \vec{p}_f | V | \vec{p}_i \rangle|^2 \\ &\equiv |f(\theta)|^2. \end{aligned} \quad (13.14)$$

We have introduced the “scattering amplitude,” $f(\theta)$. It is conventionally defined with a minus sign,

$$f(\theta) = -\frac{\mu V}{2\pi\hbar^2} \langle \vec{p}_f | V | \vec{p}_i \rangle. \quad (13.15)$$

In Born approximation

$$f(\theta) = -\frac{\mu}{2\pi\hbar^2} \int d^3 r e^{-i\vec{p}_f \cdot \vec{r}/\hbar} V(\vec{r}) e^{i\vec{p}_i \cdot \vec{r}/\hbar}. \quad (13.16)$$

or

$$f(\theta) = -\frac{\mu}{2\pi\hbar^2} \int V(r) d^3 r e^{i\vec{q} \cdot \vec{r}/\hbar}. \quad (13.17)$$

where $\vec{q} = \vec{p}_f - \vec{p}_i$ is the momentum transfer. This is a very important formula, as much because of its simplicity as for anything else. It says that the scattering amplitude is proportional to the three-dimensional Fourier transform of the scattering potential.

There is a similar formula in classical electrodynamics: the amplitude of an electromagnetic wave scattered from a small fluctuation in the dielectric constant is given by the Fourier transform of the shape of fluctuation into wave number space. The quantum mechanical formula is actually simpler, because there are no polarization factors to track.

In elastic scattering, the magnitudes of the initial and final momenta are equal, so the momentum transfer \vec{q} has a magnitude $q = 2|p| \sin \theta/2$, where θ is the scattering angle.

Now let us evaluate $f(\theta)$ in the useful special case of a central potential, $V(\vec{r}) = V(r)$. Evaluate the integral in a coordinate system where \vec{q} defines the \hat{z} axis. Then $\vec{q} \cdot \vec{r} = qr \cos \theta$ and an easy passage of steps yields a one-dimensional integral,

$$f(\theta) = -\frac{2\mu}{q\hbar} \int_0^\infty r dr V(r) \sin \frac{qr}{\hbar}. \quad (13.18)$$

As a specific example, consider the so-called Yukawa potential

$$V(r) = \frac{A}{r} e^{-mr} \quad (13.19)$$

where m is a parameter characterizing the inverse range of the potential. The choice of m is motivated by Yukawa's proposal for the origin of the nuclear force in the exchange of mesons of mass m_π : $m = m_\pi c^2 / (\hbar c)$. In condensed matter physics and for light scattering, this potential is called the "Orenstein-Zernicke correlation function," and $1/m$ is called the "correlation length." In the limit of vanishing m and for $A = Z_1 Z_2 e^2$, this potential describes Coulomb scattering, or the famous Rutherford experiment.

A few lines of algebra yield the differential cross section

$$\frac{d\sigma}{d\Omega} = \frac{4\mu^2 A^2}{(q^2 + m^2 \hbar^2)^2}. \quad (13.20)$$

Writing $q^2 = 4p^2 \sin^2 \theta/2$, we expose the angular dependence of the differential cross section,

$$\frac{d\sigma}{d\Omega} = \frac{4\mu^2 A^2}{(4p^2 \sin^2 \theta/2 + m^2 \hbar^2)^2}. \quad (13.21)$$

This formula has several interesting limits. If the energy is very low, $p/m \ll 1$, the differential cross section is isotropic: $d\sigma/d\cos\theta$ is a constant. A variation on this limit is scattering from a delta function potential. To convert the Yukawa potential to delta function form, we chose A so that $\int d^3r \delta^3(r) = 1$ in the large m limit. We can choose A so that V has the correct limit (as m goes to infinity): Set

$$1 = A \int d^3r \frac{e^{-mr}}{r} = 4\pi A \int_0^\infty r dr e^{-mr} = \frac{4\pi A}{m^2} \quad (13.22)$$

so that

$$\lim_{m \rightarrow \infty} V_0 \frac{m^2 e^{-mr}}{4\pi r} \rightarrow V_0 \delta(\vec{r}). \quad (13.23)$$

The differential cross section becomes

$$\frac{d\sigma}{d\Omega} = \frac{\mu^2 V_0^2}{4\pi^2 \hbar^4} \quad (13.24)$$

The range of the delta function's potential is so small that the scattering wave cannot resolve structure in it and scatters "minimally."

The second useful limit is that of Coulomb scattering, where $A = Z_1 Z_2 e^2$ and $m = 0$:

$$\frac{d\sigma}{d\Omega} = \frac{\mu^2 Z_1^2 Z_2^2 e^4}{4p^4 \sin^4 \theta/2} = \frac{\mu^2 Z_1^2 Z_2^2 e^4}{q^4} \quad (13.25)$$

This is the famous Rutherford result. The cross section diverges as θ vanishes. This is because the Coulomb interaction is long range – no matter how far away the two particles are, they still interact electrically.

Before continuing with examples, let us summarize a few general results about scattering in Born approximation:

- $d\sigma/d\Omega$ and $f(\theta)$ are functions of q alone. That is, $f(\theta)$ depends on $p^2/(2\mu)$ and θ only through the combination $q = 2p(1 - \cos \theta)$.
- $f(\theta)$ is always real
- $d\sigma/d\Omega$ is independent of the sign of V .
- $f(\theta)$ becomes small for large q . This is due to the $\exp(-\vec{q} \cdot \vec{r}/\hbar)$ term in the integrand.
- For small p , so that q is small, and the potential is short range, the scattering amplitude is isotropic.

Of these results, only the last one is generally true beyond Born approximation.

Because it is perturbative, the Born approximation is generally valid if the scattering amplitude is small. Physically, we want the true wave function in the scattering region not to be too different from a plane wave. Since $f(\theta)$ has units of area, "small" must mean with respect to some other quantity with dimensions of length. That could be the inverse wave number of the beam. We could imagine our small ratio to be

$$k^2 |f(\theta)| \ll 1, \quad (13.26)$$

for example.

An alert reader might wonder: We have considered scattering on a bare point charge, and re-derived Rutherford's result. However, Rutherford actually did his scattering on atoms, which are electrically neutral. How does that change the answer? And why do we read about "big scattering at large angle?" We have found $d\sigma/d\Omega \simeq 1/q^4$, which certainly becomes small at large angle.

The solution to this question introduces some interesting physics. Use the fact that $\nabla^2 \exp(i\vec{q} \cdot \vec{r}/\hbar) = -q^2/\hbar^2 \exp(i\vec{q} \cdot \vec{r}/\hbar)$ to write the scattering amplitude, Eq. 13.17, as

$$f(\theta) = \frac{\mu}{2\pi\hbar^2 q^2} \int V(r) d^3r \nabla^2 e^{i\vec{q} \cdot \vec{r}/\hbar}. \quad (13.27)$$

Then integrate twice by parts to pull the differential operator onto the potential term:

$$f(\theta) = \frac{\mu}{2\pi\hbar^2 q^2} \int d^3r e^{i\vec{q} \cdot \vec{r}/\hbar} \nabla^2 V(r). \quad (13.28)$$

Now assume that the interaction is Coulombic. The beam particle has charge Ze , and the potential arises from the charge distribution of the target, $\rho(r)$. In that case Poisson's equation is

$$\nabla^2 V(r) = Ze \times (-4\pi\rho(r)) \quad (13.29)$$

and

$$f(\theta) = -\frac{2Ze\mu}{q^2} \int d^3r e^{i\vec{q} \cdot \vec{r}/\hbar} \rho(r). \quad (13.30)$$

The object in the integral is the Fourier transform of the charge density, and is called the "form factor" $F(q)$:

$$f(\theta) = -2Ze\mu \frac{F(q)}{q^2}. \quad (13.31)$$

For Coulomb scattering on a point charge, $\rho(r) = Ze\delta^3(r)$ so $F(q) = Ze$ and $d\sigma/d\Omega = 4Z^2 Z'^2 e^4 \mu^2 / q^4$ as before.

Now we can understand "big scattering at large angles." Rutherford had in mind a model for the atom with had some smooth, extended charge density $\rho(r)$ extending over the entire size of the atom – about an Angstrom in range. The form factor for such a charge distribution will fall off at large q – for $qa_0/\hbar \gg 1$ it would be tiny. The form factor of the delta function is a constant, so at any sufficiently large angle (or q) it will give a scattering amplitude which is greater than that from any smooth charge distribution.

As an example, imagine scattering on a hydrogen atom, when the electron is in its 1S state. Then the total charge density, proton and electron, is

$$\rho(r) = e[\delta^3(r) - |\psi_{1S}(r)|^2] \quad (13.32)$$

where $|\psi_{1S}|^2 = \exp(-2r/a_0)/(\pi a_0^3)$. The form factor is $F(q) = e[1 - F_e(q)]$, and a short passage of steps gives

$$F_e(q) = \frac{1}{(1 + \frac{q^2 a_0^2}{4\hbar^2})^2}. \quad (13.33)$$

As expected, $F_e(q=0) = 1$. In this limit the form factor is just the total charge (in units of the electronic charge). Thus, the differential cross section for charged particle scattering from a hydrogen atom is

$$\frac{d\sigma}{d\Omega} = \frac{4Z^2 e^4 \mu^2}{q^4} \left(1 - \frac{1}{(1 + \frac{q^2 a_0^2}{4\hbar^2})^2}\right)^2. \quad (13.34)$$

Observe that as $qa_0/\hbar \rightarrow \infty$ (high energy or large angle), the electron contribution is negligible compared to the proton. The Rutherford result is recovered. In the other limit, $qa_0/\hbar \rightarrow 0$, the cross section goes to a constant:

$$\frac{d\sigma}{d\Omega} = \frac{4Z^2 e^4 \mu^2}{q^4} \times 4 \left(\frac{q^2 a_0^2}{4\hbar^2}\right)^2 = \frac{Z^2 \alpha^4 (\mu c^2)^2}{(\hbar c)^2} a_0^4. \quad (13.35)$$

The atom looks neutral at long distance—there is no $1/q^4$ peaking.

Inelastic electron-atom scattering

It is easy to generalize the case of elastic Coulomb scattering to the situation where the scattering excites the target. The process of interest is electron + ground state atom in state $\psi_0 \rightarrow$ electron + excited state atom in state ψ_n . The process is called “inelastic” because the energy of the final state electron is less than the energy of the initial electron. Neglecting symmetrization, the initial state of beam electron plus target electrons is

$$|i\rangle = \frac{1}{\sqrt{V}} e^{ip_i x} \psi_0(x_1, \dots, x_N) \quad (13.36)$$

while the final state is

$$|f\rangle = \frac{1}{\sqrt{V}} e^{ip_f x} \psi_n(x_1, \dots, x_N) \quad (13.37)$$

The un-subscripted x labels the bombarding electron's coordinate while the subscripted ones are for the bound electrons. In principle, one should worry about statistics for the identical electrons. However, if the bombarding energy is large, the exchange terms will be small and we can treat the bombarding electron as distinguishable from the target ones.

As before, the potential is that between the atom and the beam, a sum of Coulomb attraction and repulsions

$$V(x) = -\frac{Ze^2}{r} + \sum_{i=1}^N \frac{e^2}{|\vec{x} - \vec{x}_i|}. \quad (13.38)$$

The differential cross section is

$$\frac{d\sigma(0 \rightarrow n)}{d\Omega} = \frac{m}{p_i} \frac{2\pi}{\hbar} |\langle f|V|i \rangle|^2 \frac{p_f m}{(2\pi\hbar)^3} \quad (13.39)$$

where the first term is the beam flux, the last term is the phase space factor, and the matrix element is

$$\langle f|V|i \rangle = \int d^3x e^{iqx/\hbar} \langle n| \left(-\frac{Ze^2}{r} + \sum_{i=1}^N \frac{e^2}{|\vec{x} - \vec{x}_i|} \right) |i \rangle. \quad (13.40)$$

Specifically,

$$\begin{aligned} \langle n| \left(-\frac{Ze^2}{r} + \sum_{i=1}^N \frac{e^2}{|\vec{x} - \vec{x}_i|} \right) |i \rangle &= \int \prod_{i=1}^N d^3x_i \\ &\psi_n(x_1, \dots, x_N)^* \left(-\frac{Ze^2}{r} + \sum_{i=1}^N \frac{e^2}{|\vec{x} - \vec{x}_i|} \right) \psi_0(x_1, \dots, x_N). \end{aligned} \quad (13.41)$$

In the first term of this expression, r is just the separation between the bombarding electron and the nucleus. It does not involve the bound electrons. Thus, the initial and final atomic electron states must coincide if the matrix element is not to vanish.

$$\langle n| -\frac{Ze^2}{r} |0 \rangle = -\frac{Ze^2}{r} \langle n|0 \rangle = -\frac{Ze^2}{r} \delta_{n,0}. \quad (13.42)$$

The second term can be turned into another kind of form factor. Write

$$\int d^3x e^{iqx/\hbar} \langle n| \sum_{i=1}^N \frac{1}{|\vec{x} - \vec{x}_i|} |0 \rangle = \langle n| \sum_i \int \frac{e^{iqx/\hbar}}{|x - x_i|} d^3x |0 \rangle = \langle n| \sum_i \int \frac{e^{iq(x+x_i)/\hbar}}{|x|} |0 \rangle \quad (13.43)$$

(shifting the origin), and the integral becomes

$$\langle n| \sum_i e^{iqx_i/\hbar} |0 \rangle \int d^3x \frac{e^{iqx/\hbar}}{|x|} = \frac{4\pi\hbar^2}{q^2} \langle n| \sum_i e^{iqx/\hbar} |0 \rangle. \quad (13.44)$$

We define a transition form factor

$$ZF_n(q) = \langle n | \sum_i e^{iqx_i/\hbar} | 0 \rangle. \quad (13.45)$$

(Compare this formula to that for the usual (one particle) form factor, $\langle 0 | \sum_i e^{iqx_i/\hbar} | 0 \rangle$.)

Thus

$$\langle f | V | 0 \rangle = \frac{4\pi^2 \hbar^2 Z e^2}{q^2} (\delta_{n0} - F_n(q)) \quad (13.46)$$

and the differential cross section is

$$\frac{d\sigma(0 \rightarrow n)}{d\Omega} = \frac{p_f}{p_i} \left| \frac{4Z^2 e^4 \mu^2 \hbar^4}{q^4} |\delta_{n0} - F_n(q)|^2 \right|. \quad (13.47)$$

In an inelastic collision, the δ_{n0} term gives zero and the overall cross section is rescaled by the ratio of phase space factor to beam current, $p_f \neq p_i$. In an elastic collision, $p_f/p_i = 1$ and the result recovers the ordinary (elastic) form factor.

Photoelectric effect

As a final example of “Born physics” we consider the process (photon + atom \rightarrow electron + ion), the photoelectric effect. The physical process we have in mind is to compute the cross section for absorption of X-rays in an atom. For simplicity, we assume that the electron begins in the 1S state of an atom and is knocked out, the final state is a plane wave with momentum $\vec{P} = \hbar\vec{K}$. The energy of the ejected electron is E_f . The Golden Rule gives the differential transition probability

$$d\Gamma = \frac{2\pi}{\hbar} |\langle \hbar\vec{K} | \frac{e}{mc} \vec{A} \cdot \vec{p} | 1S \rangle|^2 \rho(E_f) \delta(E_f - E_{1S} - E_\gamma) \quad (13.48)$$

As before, the phase space factor is

$$\rho(E_f) \delta(E_f - E_{1S} - E_\gamma) = m |\hbar K| V \frac{d\Omega}{(2\pi\hbar)^3} \quad (13.49)$$

Here $m = m_e = \mu$ the reduced mass. We convert $d\Gamma$ into a differential cross section by dividing by the beam intensity, (1 photon/volume V) times c . Thus

$$\frac{d\sigma}{d\Omega} = \frac{2\pi}{\hbar c} V^2 |\hbar K| \frac{\mu}{(2\pi\hbar)^3} |\langle f | V | i \rangle|^2. \quad (13.50)$$

We compute the matrix element with the following ingredients: The outgoing electron’s wave function is

$$\psi_f = \frac{1}{\sqrt{V}} e^{iK \cdot r}, \quad (13.51)$$

the vector potential is

$$\vec{A} = \left(\frac{2\pi\hbar c^2}{\omega V} \right)^{1/2} e^{ik \cdot r} \vec{\epsilon}, \quad (13.52)$$

(absorbing the frequency dependence into the delta-function: note: we can't use the dipole approximation!) and the initial electronic wave function is taken to be hydrogenic,

$$\psi_i = \left(\frac{\lambda^3}{\pi} \right)^{1/2} e^{-\lambda r} \quad (13.53)$$

where $\lambda = Z/a_0$. Then the matrix element is

$$\langle f|V|i\rangle = \frac{e}{mc} \left(\frac{2\pi\hbar c^2 \lambda^3}{\omega V} \right)^{1/2} \frac{1}{V} \int d^3r e^{-iK \cdot r} \vec{p} \cdot \vec{\epsilon} e^{ik \cdot r} e^{-\lambda r} \quad (13.54)$$

The operator \vec{p} is $(\hbar/i)\vec{\nabla}$. We can push \vec{p} to the left to get

$$\langle f|V|i\rangle = \frac{e}{mc} \left(\frac{2\pi\lambda^3}{\omega} \right)^{1/2} (\hbar\vec{K} \cdot \vec{\epsilon}) \int d^3r e^{iq \cdot r} e^{-\lambda r} \quad (13.55)$$

where $\vec{q} = \vec{k} - \vec{K}$ is the wave number (scaled momentum) of the recoiling atom. We visualize the scattering as follows: the electron and atom emerge back to back, with \vec{K} at an azimuthal angle of ϕ to the polarization direction. \vec{K} has a polar angle of θ with respect to the beam.

We did the integral when we evaluated the form factor of Hydrogen. It is just

$$\frac{8\pi\lambda}{(q^2 + \lambda^2)^2}. \quad (13.56)$$

Putting all the pieces together, we have

$$\frac{d\sigma}{d\Omega} = \frac{2\pi}{\hbar c} \frac{\hbar K m}{(2\pi\hbar)^3} \frac{e^2}{m^2} \frac{2\lambda^3 \hbar^3}{ck} (K \sin \theta \cos \phi)^2 \frac{64\pi^2 \lambda^2}{(q^2 + \lambda^2)^4} \quad (13.57)$$

$$= 32 \frac{e^2}{\hbar c} \frac{\sin^2 \theta \cos^2 \phi}{k^2} \frac{\hbar k}{mc} \left(\frac{\lambda^2}{q^2 + \lambda^2} \right)^4 \left(\frac{K}{\lambda} \right)^3 \quad (13.58)$$

The answer is not terribly illuminating. Let's make some approximations.

First, call the atom's ionization potential I . Energy conservation says $\hbar ck = I + \hbar^2 K^2/(2m)$. I is equal to $\hbar^2 \lambda^2/(2m)$. Go to high energies, so that $\hbar^2 K^2/(2m)$ and $\hbar ck$ are much greater than I . But don't go too high in energy, or you will have to deal with special relativity: keep $mc^2 > \hbar ck$. Energy conservation neglecting I tells us that

$$\hbar ck = \hbar^2 K^2/(2m), \quad (13.59)$$

so $K^2 = (2mc^2)/(\hbar ck)k^2$. Putting all the inequalities together,

$$2mc^2 \gg \hbar ck \simeq \hbar^2 K^2/(2m) \gg I \quad (13.60)$$

or $K^2 \gg k^2$ and $K^2 \gg \lambda^2$. Then

$$q^2 = (\vec{K} - \vec{k})^2 \simeq K^2 - 2Kk \cos \theta = K^2(1 - 2k/K \cos \theta) = 2mck/\hbar(1 - \hbar K/(mc) \cos \theta). \quad (13.61)$$

Now $\hbar K/(mc) = pc/(mc^2) = v/c \equiv \beta$ for the electron, so finally $q^2 = 2mck/\hbar(1 - \beta \cos \theta)$. $q^2 \gg \lambda^2$ so $q^2 + \lambda^2$ can be replaced by q^2 . The last alteration we make is to write

$$\left(\frac{K}{\lambda}\right)^3 = \left(\frac{K^2}{\lambda^2}\right)^{3/2} = \left(\frac{\hbar^2 K^2/(2m)}{\hbar^2 \lambda^2/(2m)}\right)^{3/2} = \left(\frac{\hbar ck}{I}\right)^{3/2} \quad (13.62)$$

and we have re-introduced the ionization potential. We can rewrite

$$\frac{\lambda^2}{q^2 + \lambda^2} \simeq \frac{\lambda^2}{q^2} = \frac{2mI}{2mck\hbar(1 - \beta \cos \theta)} = \frac{I}{\hbar ck(1 - \beta \cos \theta)}. \quad (13.63)$$

After all these substitutions, the differential cross section is

$$\frac{d\sigma}{d\Omega} = \frac{32\alpha \sin^2 \theta \cos^2 \phi}{(1 - \beta \cos \theta)^4} \left(\frac{I}{\hbar ck}\right)^4 \left(\frac{\hbar ck}{I}\right)^{7/2} \frac{\hbar ck}{mc^2} \frac{1}{k^2}. \quad (13.64)$$

The last two terms are

$$\frac{\hbar ck}{mc^2} \frac{1}{k^2} = \frac{\hbar ck}{mc^2} \frac{(\hbar c)^2}{(\hbar ck)^2} = \frac{(\hbar c)^2}{mc^2} \frac{I}{\hbar ck} \frac{1}{I}, \quad (13.65)$$

so

$$\frac{d\sigma}{d\Omega} = 32\alpha \left(\frac{\hbar ck}{I}\right)^{7/2} \frac{(\hbar c)^2}{Imc^2} \frac{\sin^2 \theta \cos^2 \phi}{(1 - \beta \cos \theta)^4}. \quad (13.66)$$

That was a lot of work, but we have separated the angular and energy dependences in a useful way. Now we can understand a number of features of the photoelectric effect: First, the angular factors show that the ejected electron is thrown dominantly forward, from the $(1 - \beta \cos \theta)^4$ in the denominator. Near threshold, the antenna pattern for the electron is a dipole ($\sin^2 \theta$) but as the energy rises, the pattern is folded up. The $\sin^2 \theta$ in the numerator is from the $\vec{\epsilon} \cdot \vec{K}$ factor in the interaction: the electron's momentum follows the direction of polarization. This pattern is superficially quite similar to that for bremsstrahlung or other radiation process, but note, this is the pattern for an ejected electron, not a photon.

Notice that the cross section falls steeply with $E_\gamma = \hbar ck$: $\sigma \simeq (I/E_\gamma)^{-7/2}$. The cross section switches off at $E_\gamma = I$ due to energy conservation (the ejected electron must have

positive energy). In a real atom, there are of course many energy levels (s, p, d) with $I_s > I_p > I_d \dots$, and so the total cross section is a set of overlapping shoulders, one for each threshold. It is often customary to plot the cross section as a function of wavelength; then the cross section rises with wavelength and there are a series of “notches” as the energy falls below the threshold to ionize a particular atomic level. The notches serve, of course, as markers for the energy levels of electrons in the atom.

At very high energy, $E_\gamma > 2mc^2$, the photoelectric cross section falls to zero, but a new process takes over and dominates: the photon can convert to an electron-positron pair in the presence of the electric field of the nucleus. This is called the “Bethe-Heitler” process after the two physicists who first calculated its rate.

Photo-disassociation of the deuteron

A problem related to the photoelectric effect is a nice venue to discuss phase space in more generality. Let us examine the photo-disassociation of the deuteron. In this problem a photon with momentum $\hbar\vec{K}$ and a deuteron of momentum \vec{P}_D collide to produce an outgoing proton and neutron, with momenta \vec{p}_p and \vec{p}_n , respectively. The interaction Hamiltonian is (as usual)

$$H = \frac{1}{2m_p}(\vec{p}_p - \frac{e}{c}\vec{A})^2 + \frac{p_n^2}{2m_n} + V(\vec{r}_p - \vec{r}_n), \quad (13.67)$$

with the inter-nucleon potential described by $V(r)$. We expand the Hamiltonian and work to first order in e to write

$$H_0 = \frac{p_p^2}{2m_p} + \frac{p_n^2}{2m_n} + V(r); \quad (13.68)$$

the interaction term is the usual electromagnetic expression

$$H_I = -\frac{e}{m_p c} \vec{p}_p \cdot \vec{A}. \quad (13.69)$$

We define center of mass and relative coordinates through $M\vec{R} = m_p\vec{r}_p + m_n\vec{r}_n$, with $M = m_p + m_n$, and $\vec{r} = \vec{r}_p - \vec{r}_n$. The reduced mass is $\mu = m_p m_n / M$. Then

$$H_0 = \frac{P_R^2}{2M} + \frac{p_r^2}{2\mu} + V(r). \quad (13.70)$$

The form of $V(r)$ is un-needed, other than to specify the wave function of the proton and neutron in the deuteron $\psi(r)$ in the initial state

$$|i\rangle = \frac{e^{i\vec{P}_D \cdot \vec{R}}}{\sqrt{V}} \psi(r). \quad (13.71)$$

The final state is assumed to be a product of plane waves for the proton and neutron,

$$|f\rangle = \frac{1}{\sqrt{V}} \exp(i\vec{p}_p \cdot \vec{r}_p / \hbar) \frac{1}{\sqrt{V}} \exp(i\vec{p}_n \cdot \vec{r}_n / \hbar). \quad (13.72)$$

Let us assemble our ingredients for the Golden Rule. The photon, deuteron, proton and neutron energies are $E_\gamma = \hbar ck$, $E_D = m_D c^2 + E_{kin}(D)$, $E_p = m_p c^2 + E_{kin}(p)$, and $E_n = m_n c^2 + E_{kin}(n)$. Writing $m_d c^2 = m_p c^2 + m_n c^2 - B$, where B is the binding energy of the deuteron, energy conservation becomes $E_{kin}(p) + E_{kin}(n) = E_{in} = \hbar ck + E_{kin}(D) - B$. Physically, the energy of the photon plus the deuteron's kinetic energy equals the sum of the two outgoing projectile's energies plus the energy to disassociate the deuteron.

The transition amplitude becomes

$$M = \langle f | H_I | i \rangle = \int d^3 R d^3 r \frac{e^{-i\vec{p}_p \cdot \vec{r}_p / \hbar} e^{-i\vec{p}_n \cdot \vec{r}_n / \hbar}}{V^{3/2}} \frac{e}{m_p c} (\vec{p}_p \cdot \vec{A}) e^{i\vec{P}_D \cdot \vec{R} / \hbar} \psi(r), \quad (13.73)$$

where

$$\vec{A} = \left(\frac{2\pi\hbar c^2}{\omega_k V} \right)^{1/2} \epsilon e^{i\vec{k} \cdot \vec{r}_p} \quad (13.74)$$

is the vector potential at the location of the proton. Again letting the momentum operator $\vec{p}_p = \hbar/i\vec{\nabla}_p$ act to the left on the outgoing proton wave function, the transition amplitude becomes

$$M = \frac{1}{V^{3/2}} \frac{e}{m_p c} \left(\frac{2\pi\hbar c^2}{\omega_k V} \right)^{1/2} \vec{p}_p \cdot \vec{\epsilon} \int d^3 R d^3 r e^{-i\vec{p}_p \cdot \vec{r}_p / \hbar} e^{-i\vec{p}_n \cdot \vec{r}_n / \hbar} e^{i\vec{P}_D \cdot \vec{R} / \hbar} \psi(r). \quad (13.75)$$

We now replace r_n and r_p by the relative and center of mass coordinates, $r_p = R + (m_n/M)r$, $r_n = R - (m_p/M)r$. The exponential factor in Eq. 13.75 becomes

$$i\vec{R} \cdot (\vec{P}_D + \hbar\vec{k} - \vec{p}_p - \vec{p}_n) + i\vec{r} \cdot \left(\frac{m_p}{M} \vec{p}_n - (\vec{p}_p - \hbar\vec{k}) \frac{m_n}{M} \right). \quad (13.76)$$

Notice that M is a product of an integral over R which is independent $\psi(r)$ times an integral over r which contains all the physics. Call that term M_r :

$$M_r = \frac{1}{V^{3/2}} \frac{e}{m_p c} \left(\frac{2\pi\hbar c^2}{\omega_k V} \right)^{1/2} \vec{p}_p \cdot \vec{\epsilon} \int d^3 r \exp(-i(-\frac{m_p}{M} p_n + \frac{m_p}{M} (\vec{p}_p + \hbar\vec{k})) \cdot \vec{r}) \psi(r). \quad (13.77)$$

Let's break away from the calculation for a moment and focus on the R part of M . We can integrate over R , working in a box of volume $V = L^3$. In one dimension it would be an integral of the form

$$\int_0^L dR e^{iQR/\hbar} = \frac{e^{iQL/\hbar} - 1}{iQ/\hbar}. \quad (13.78)$$

The Golden Rule needs the square of M . To deal with Eq. 13.78 we first square the result of the integral and then take the $L \rightarrow \infty$ limit. Just as the time integral in the square of the Golden Rule amplitude

$$\int_0^T dt e^{i\omega T} = \frac{e^{i\omega T} - 1}{i\omega} \quad (13.79)$$

squares to give an energy-conserving delta function,

$$\lim_{T \rightarrow \infty} \left| \frac{e^{i\omega T} - 1}{i\omega} \right|^2 \rightarrow 2\pi T \delta(\omega), \quad (13.80)$$

so the spatial integral squares to give a momentum conserving delta function,

$$\lim_{L \rightarrow \infty} \left| \frac{e^{iQL/\hbar} - 1}{iQ/\hbar} \right|^2 \rightarrow 2\pi L \delta(Q/\hbar) = 2\pi \hbar L \delta(Q). \quad (13.81)$$

Thus the squared center of mass integral contributes a factor

$$L^3 (2\pi \hbar)^3 \delta^3(\vec{P}_D + \hbar \vec{k} - \vec{p}_p - \vec{p}_n) \quad (13.82)$$

to the Golden Rule expression:

$$d\Gamma = \frac{2\pi}{\hbar} |M_r|^2 (2\pi \hbar)^3 V \delta^3(\vec{P}_D + \hbar \vec{k} - \vec{p}_p - \vec{p}_n) \delta(E_d + \hbar ck - B - E_p - E_n) \rho(E). \quad (13.83)$$

The phase space factor must count all the states for both outgoing particles, so it is

$$\rho(E) = V \frac{d^3 p_p}{(2\pi \hbar)^3} V \frac{d^3 p_n}{(2\pi \hbar)^3}. \quad (13.84)$$

Had we had N particles in the final state, $\rho(E)$ would have been a product of N terms, one for each particle. One often sees this written in the literature without explanation:

$$d\Gamma(i \rightarrow j_1, j_2 \dots j_n) = |M|^2 \frac{2\pi}{\hbar} \delta\left(\sum_i E_i - \sum_j E_j\right) (2\pi \hbar)^3 \delta^3\left(\sum_i \vec{p}_i - \sum_j \vec{p}_j\right) \prod_j \frac{V d^3 p_j}{(2\pi \hbar)^3}. \quad (13.85)$$

Eq. 13.83 clearly indicates that both momentum and energy are conserved in the process, a very sensible result.

Including the flux factor for the beam, $J = v_{rel} \times (\text{one photon in volume } V) = (c/V)$, gives the differential cross section

$$d\sigma = \frac{V^2}{c} \frac{2\pi}{\hbar} |M_r|^2 (2\pi \hbar)^3 \delta^3(\vec{P}_D + \hbar \vec{k} - \vec{p}_p - \vec{p}_n) \delta(E_d + \hbar ck - B - E_p - E_n) \frac{d^3 p_p}{(2\pi \hbar)^3} \frac{d^3 p_n}{(2\pi \hbar)^3}. \quad (13.86)$$

All the volume factors neatly cancel, as they should. To groom this expression, we can just integrate over one of the outgoing particles' momentum. This will impose overall momentum conservation on the process and give a familiar-looking result,

$$d\sigma = \frac{2\pi}{\hbar c} |\tilde{M}|^2 \frac{d^3 p_p}{(2\pi \hbar)^3} \delta(E_d + \hbar ck - B - E_p - E_n) \quad (13.87)$$

where

$$\tilde{M} = \frac{e}{m_p c} \left(\frac{2\pi \hbar c^2}{\omega_k} \right)^{1/2} \vec{p}_p \cdot \vec{\epsilon} \int d^3 r \exp(-iq \cdot \vec{r}) \psi(r) \quad (13.88)$$

and

$$\hbar \vec{q} = -\frac{m_p}{M} \vec{p}_n + \frac{m_p}{M} (\vec{p}_p - \hbar \vec{k}) \quad (13.89)$$

Notice that in Eq. 13.86 we do not have to specify any particular reference frame. The center of mass frame is a convenient one, however. In it, $\vec{P}_D + \hbar\vec{K} = 0$, $\vec{P}_p = -\vec{p}_n = \vec{p} = \hbar\vec{K}$, and $E_p + E_n = p^2/(2\mu)$. Performing the phase space integrals leaves

$$\frac{d\sigma}{d\Omega} = \frac{2\pi}{\hbar c} \frac{p_p \mu}{(2\pi\hbar)^3} |\tilde{M}|^2. \quad (13.90)$$

If we write $\vec{p}_p = \hbar\vec{K}$, $\vec{p}_n = -\hbar\vec{K}$, we find

$$\vec{q} = \vec{K} \frac{m_p + m_n}{M} - \vec{k} \frac{m_N}{M} = \vec{K} - \vec{k} \frac{m_N}{M}. \quad (13.91)$$

This is essentially the same calculation as for the atomic photo-effect. Psychologically, it is easier to deal with the atomic case, because the kinematics in that case “naturally” seems to be one-body (photon in, electron out), although of course both situations have two particles in the final state. In the case of the deuteron, $m_n/M = 1/2$ and $\vec{q} = \vec{K} - \vec{k}/2$ while in the atomic case $m_n/M \rightarrow 1$ and $\vec{q} = \vec{K} - \vec{k}$. In the deuteron, the reduced mass is $\mu = m_p/2$; in the atom, $\mu = m_e$ the electron mass. If we had more than two bodies in the final state, we would have had a $3N$ dimensional phase space integral over all the particle’s momenta, and four delta functions as constraints. The $3N - 4$ dimensional integral would have bounds induced by the requirement that all the particles’ energies must be positive. It’s a bit of a mess, but completely comprehensible (at least in principle).

We have not completed the calculation, of course. What should we take for the wave function of the deuteron, $\psi(r)$? The deuteron is barely bound, $B \sim 2.2$ MeV. The range of the nuclear potential is a few times 10^{-13} cm – the size of the proton or neutron. Because the deuteron is so weakly bound, its wave function must have an enormous tail in the classically forbidden region. A good model for the deuteron is to put the particles in an attractive square well potential and adjust the parameters in the potential so that the binding energy is nearly at zero. Then the wave function is nearly all outside the well, and

$$\psi(r) = \sqrt{\frac{\kappa}{2\pi}} \frac{e^{-\kappa r}}{r} \quad (13.92)$$

Of course, κ is $\sqrt{2\mu B/\hbar^2}$. With this choice for $\psi(r)$ the integral is pretty much identical to the one for the photoelectric effect.

Scattering theory from the Schrödinger equation

The Lippmann-Schwinger equation

A second approach to scattering theory treats it as a special case of the time-independent Schrödinger equation

$$\left(\hat{H}_0 + V\right) |\psi\rangle = E |\psi\rangle. \quad (13.93)$$

The zeroth-order Hamiltonian is $\hat{H}_0 = \frac{p^2}{2m}$. Its free space solution (an eigenstate of \hat{H}_0) is

$$\langle x|\phi\rangle = \frac{e^{i\vec{p}\cdot\vec{x}/\hbar}}{(2\pi\hbar)^{3/2}} \quad (13.94)$$

If one were to reduce the strength of the potential to zero, one would expect that $|\psi\rangle$ would reduce to $|\phi\rangle$. To capture that behavior, rewrite the Schrödinger equation as an integral equation,

$$|\psi\rangle = \frac{1}{E - \hat{H}_0} V |\psi\rangle + |\phi\rangle. \quad (13.95)$$

As it stands, the equation is not well defined because H_0 has a continuous spectrum; one must avoid the pole in the first term. This is done by shifting the singularity slightly off the real axis, modifying Eq. 13.95 to

$$|\psi^\pm\rangle = \frac{1}{E - H_0 \pm i\epsilon} V |\psi^\pm\rangle + |\phi\rangle. \quad (13.96)$$

This is called the Lippmann-Schwinger equation. To understand what we have done, evaluate the Lippmann-Schwinger equation in coordinate space. It is

$$\langle x|\psi^\pm\rangle = \int d^3x' \langle x|\frac{1}{E - H_0 \pm i\epsilon}|x'\rangle \langle x'|V|\psi^\pm\rangle + \langle x|\phi\rangle \quad (13.97)$$

where

$$G_\pm(x, x') \equiv \langle x|\frac{1}{E - H_0 \pm i\epsilon}|x'\rangle \quad (13.98)$$

is a free particle Green's function. One may compute it by passing to momentum space,

$$\begin{aligned} G_\pm(x, x') &= \int d^3p \int d^3p' \langle x|p\rangle \delta^3(p - p') \frac{1}{E - \frac{p^2}{2m} \pm i\epsilon} \langle p'|x'\rangle \\ &= \int \frac{d^3p}{(2\pi\hbar)^3} \frac{e^{i\vec{p}\cdot(\vec{x}-\vec{x}')/\hbar}}{E - \frac{p^2}{2m} \pm i\epsilon} \end{aligned} \quad (13.99)$$

$$\begin{aligned}
&= \frac{2\pi}{(2\pi\hbar)^3} \int_0^\infty p^2 dp \int_{-1}^1 d(\cos\theta) \frac{e^{ip|\vec{x}-\vec{x}'|\cos\theta/\hbar}}{E - \frac{p^2}{2m} \pm i\epsilon} \\
&= \frac{1}{i(2\pi\hbar)^2 |\vec{x} - \vec{x}'|} \int_0^\infty p dp \frac{\left[e^{ip|\vec{x}-\vec{x}'|/\hbar} - e^{-ip|\vec{x}-\vec{x}'|/\hbar} \right]}{E - \frac{p^2}{2m} \pm i\epsilon}.
\end{aligned}$$

To simplify notation, let $p = \hbar q$ and $2mE = \hbar^2 k^2$. Making these substitutions, one finds poles at $q \simeq k \pm i\epsilon$,

$$G_{\pm}(x, x') = \frac{im}{4\pi^2 |\vec{x} - \vec{x}'| \hbar^2} \int_{-\infty}^{\infty} q dq \frac{\left[e^{iq|\vec{x}-\vec{x}'|} - e^{-iq|\vec{x}-\vec{x}'|} \right]}{q^2 - k^2 \pm i\epsilon} \quad (13.100)$$

which can be evaluated by closing the contours separately for each exponential, and then taking the $\epsilon \rightarrow 0$ limit. This gives

$$G^{\pm}(x, x') = -\frac{1}{4\pi} \frac{2m}{\hbar^2} \frac{e^{\pm ik|\vec{x}-\vec{x}'|}}{|\vec{x} - \vec{x}'|}. \quad (13.101)$$

This is the Green's function for the Helmholtz equation,

$$(\nabla^2 + k^2)G^{\pm}(x, x') = \frac{2m}{\hbar^2} \delta^3(\vec{x} - \vec{x}'). \quad (13.102)$$

Now one can understand the $\pm\epsilon$: the $+i\epsilon$ solution represents an outgoing spherical wave, and the $-i\epsilon$ solution is an incoming spherical wave.

To check this result, one can return to the Lippmann-Schwinger equation. Inserting the Green's function gives

$$\langle x|\psi^{\pm}\rangle = \langle x|\phi\rangle - \frac{m}{2\pi\hbar^2} \int d^3x' \frac{e^{\pm ik|\vec{x}-\vec{x}'|}}{|\vec{x} - \vec{x}'|} \langle x'|V|\psi^{\pm}\rangle. \quad (13.103)$$

For usual coordinate-space potentials, $\langle x'|V|x''\rangle = V(x')\delta^3(x' - x'')$, so

$$\langle x|\psi^{\pm}\rangle = \langle x|\phi\rangle - \frac{m}{2\pi\hbar^2} \int d^3x' \frac{e^{\pm ik|\vec{x}-\vec{x}'|}}{|\vec{x} - \vec{x}'|} V(x') \langle x'|\psi^{\pm}\rangle. \quad (13.104)$$

We can confirm the meaning of the $\pm i\epsilon$ solutions by considering a short range scattering potential and computing the wave function far away from it. In that limit, $|\vec{x} - \vec{x}'| \simeq r - \hat{r} \cdot \vec{x}'$. Inserting the homogeneous (plane wave) solution to the Schrödinger equation for $\langle x|\phi\rangle$, the Lippmann-Schwinger equation becomes

$$\begin{aligned}
\langle x|\psi^{\pm}\rangle &\longrightarrow \frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi\hbar)^{3/2}} - \frac{m}{2\pi\hbar^2} \frac{e^{\pm ikr}}{r} \int d^3x' V(x') e^{\mp i\vec{k}'\cdot\vec{x}'} \langle x'|\psi^{\pm}\rangle \\
&\longrightarrow \frac{e^{i\vec{k}\cdot\vec{x}}}{(2\pi\hbar)^{3/2}} + \frac{e^{\pm ikr}}{r} f_{\pm}(k', k)
\end{aligned} \quad (13.105)$$

where $\vec{k}' \equiv k\hat{r}$. This confirms the statement that the $(+i\epsilon)$ convention corresponds to an outgoing wave and the $-i\epsilon$ one to an incoming wave. The physical situation of scattering uses the outgoing wave solution.

Explicitly, the scattering amplitude can be given entirely in terms of a matrix element involving $|\psi_+\rangle$,

$$\begin{aligned} f(\vec{k}, \vec{k}') &= \frac{1}{4\pi} \frac{2m}{\hbar^2} (2\pi)^3 \int \frac{d^3x'}{(2\pi)^{3/2}} e^{-ik'x} V(x') \langle x' | \psi_+ \rangle \\ &= -\frac{(2\pi)^3}{4\pi} \frac{2m}{\hbar^2} \langle k' | V | \psi_+ \rangle. \end{aligned} \quad (13.106)$$

Let us assume that the potential is weak and solve the Lippmann-Schwinger equation by iteration:

$$\begin{aligned} \langle x | \psi_+ \rangle &= \langle x | \phi \rangle + \int d^3x' \langle x | \frac{1}{E - H_0 + i\epsilon} V | x' \rangle \langle x' | \psi_+ \rangle \\ &= \langle x | \phi \rangle + \int d^3x' \langle x | \frac{1}{E - H_0 + i\epsilon} V | x' \rangle \langle x' | \phi \rangle + \dots \end{aligned} \quad (13.107)$$

Keeping the first term in this series corresponds to the first order Born approximation. To go further, it is useful to define an operator $T|\phi\rangle = V|\psi_+\rangle$. The operator T is labeled as the T-matrix, in analogy with the the T-matrix used in time-dependent calculations. With this definition,

$$V|\psi_+\rangle = V|\phi\rangle + V \frac{1}{E - H_0 + i\epsilon} V|\psi_+\rangle \quad (13.108)$$

and so

$$T|\phi\rangle = V|\phi\rangle + V \frac{1}{E - H_0 + i\epsilon} T|\phi\rangle \quad (13.109)$$

or

$$T = V + V \frac{1}{E - H_0 + i\epsilon} T. \quad (13.110)$$

The formula for the scattering amplitude

$$f(\vec{k}, \vec{k}') = -\frac{1}{4\pi} \frac{2m}{\hbar^2} (2\pi)^3 \langle k' | V | \psi_+ \rangle \quad (13.111)$$

is equivalent to

$$f(\vec{k}, \vec{k}') = -\frac{1}{4\pi} \frac{2m}{\hbar^2} (2\pi)^3 \langle k' | T | k \rangle. \quad (13.112)$$

These results suggest that one can approach scattering problems by solving the operator equation for T directly. A simple approach is just to iterate:

$$T = V + V \frac{1}{E - H_0 + i\epsilon} V + V \frac{1}{E - H_0 + i\epsilon} V \frac{1}{E - H_0 + i\epsilon} V + \dots \quad (13.113)$$

As an example, the second order contribution to the scattering amplitude is

$$\begin{aligned}
 f^{(2)}(\vec{k}, \vec{k}') &= -\frac{1}{4\pi} \frac{2m}{\hbar^2} (2\pi)^3 \langle k' | V \frac{1}{E - H_0 + i\epsilon} V | k \rangle \\
 &= -\frac{m(2\pi)^3}{2\pi\hbar^2} \int d^3x \int d^3x' \langle k' | x \rangle \langle x | V \frac{1}{E - H_0 + i\epsilon} V | x' \rangle \langle x' | k \rangle \\
 &= -\frac{m(2\pi)^3}{2\pi\hbar^2} \int d^3x \int d^3x' \langle k' | x \rangle V(x) \langle x | \frac{1}{E - H_0 + i\epsilon} | x' \rangle V(x') \langle x' | k \rangle \\
 &= -\frac{m(2\pi)^3}{2\pi\hbar^2} \int d^3x \int d^3x' \langle k' | x \rangle V(x) G(x, x') V(x') \langle x' | k \rangle.
 \end{aligned} \tag{13.114}$$

Here $G(x, x')$ is the free particle propagator. The physics of Eq. 13.113 is simple, and is illustrated by the second order formula we have just written: the particle comes into the region of the potential and interacts with it at location x' . It then propagates as a free particle (through the Green's function) to location x and interacts again. This evolution would continue in higher order. We must sum over all the locations x' and x . It is quite reminiscent of the interpretation of the evolution operator in time dependent perturbation theory, except that here we have no explicit time dependence.

The optical theorem

The optical theorem relates the imaginary part of the forward scattering amplitude to the total elastic cross section,

$$\frac{k\sigma_{tot}}{4\pi} = \text{Im}f(\theta = 0), \tag{13.115}$$

where $f(\vec{k}, \vec{k}) = f(\theta = 0)$ and

$$\sigma_{tot} = \int \frac{d\sigma}{d\Omega} d\Omega. \tag{13.116}$$

To verify this result, begin with

$$f(\theta = 0) = -\frac{1}{4\pi} \frac{2m}{\hbar^2} (2\pi)^3 \langle k | T | k \rangle = -(2\pi)^2 \frac{m}{\hbar^2} \langle k | T | k \rangle. \tag{13.117}$$

Replacing $|k\rangle$ by $|\psi^+\rangle$ using the Lippmann - Schwinger equation, the imaginary part of this expression can be written as

$$\begin{aligned}
 \text{Im} \langle k | T | k \rangle &= \text{Im} \langle \psi^+ | V^\dagger | k \rangle \\
 &= \text{Im} \left[\langle \psi_+ | V^\dagger | \psi^+ \rangle - \langle \psi_+ | V^\dagger \frac{1}{E - H_0 + i\epsilon} V | \psi_+ \rangle \right].
 \end{aligned} \tag{13.118}$$

Since this expression is to be evaluated under an integral, the principal part expansion allows us to write

$$\frac{1}{E - H_0 + i\epsilon} = \mathcal{P} \frac{1}{E - H_0} + i\pi\delta(E - H_0). \quad (13.119)$$

If V is Hermitian, the first term in Eq. 13.119 and the \mathcal{P} term are real. Only the delta-function piece contributes to the imaginary part, so

$$\text{Im} \langle k|T|k \rangle = -\pi \langle \psi_+ | V^\dagger \delta(E - H_0) V | \psi_+ \rangle. \quad (13.120)$$

Furthermore, the definition for the T-matrix, $T|\phi\rangle = V|\psi_+\rangle$, tells us that this expression can be written as

$$\text{Im} \langle k|T|k \rangle = -\pi \langle k|T\delta(E - H_0)T|k \rangle \quad (13.121)$$

$$\begin{aligned} &= -\pi \int d^3k' \langle k|T|k' \rangle \langle k'|T|k \rangle \delta\left(E - \frac{\hbar^2 k'^2}{2m}\right) \\ &= -\pi \int d\Omega k^2 dk' |\langle k|T|k' \rangle|^2 \frac{m}{\hbar^2 k'} \delta(k - k'), \end{aligned}$$

$$(13.122)$$

inserting a complete set of $|k\rangle$ states and picking up the Jacobian from the change of variables in the delta function. (In the second line, $H_0|k'\rangle = \hbar^2 k'^2/(2m)$ and of course $E = \hbar^2 k^2/(2m)$.) Evaluating the integral and inserting the necessary numerical factors to replace $|\langle k|T|k'\rangle|^2$ by $|f(k, k')|^2$, we find

$$\begin{aligned} \text{Im} f(\theta = 0) &= \left[-\frac{(2\pi)^2 m}{\hbar}\right] \left[-\pi \left(\frac{mk}{\hbar^2}\right)\right] \int d\Omega |\langle k'|T|k \rangle|^2 \\ &= \frac{k}{4\pi} \int d\Omega \left|\frac{(2\pi)^2 m}{\hbar} \langle k'|T|k \rangle\right|^2 \\ &= \frac{k}{4\pi} \int d\Omega |f(k', k)|^2 \\ &= \frac{k}{4\pi} \int d\Omega \frac{d\sigma}{d\Omega} \\ &= \frac{k}{4\pi} \sigma_{tot}. \end{aligned}$$

$$(13.123)$$

This is the desired result. Note that it is an exact relation, independent of any properties of the scattering interaction.

Several remarks are in order.

First, the optical theorem is really a consequence of unitarity. Being very general, quantum mechanical evolution can be thought of as the action of an operator S (this is basically the evolution operator U , but the literature of scattering uses the letter S for what is called the “scattering matrix” or “S-matrix”) – $|\psi_{out}\rangle = S|\psi_{in}\rangle$. S is unitary, so $S^\dagger S = 1$. We are usually not interested in the case where $|\psi_{out}\rangle$ is identical to $|\psi_{in}\rangle$ so we write $S = 1 + iT$; the “1” leaves the state unchanged and the T changes it. Then $S^\dagger S = 1 = 1 + i(T - T^\dagger) + |T|^2$. The optical theorem is basically the cancellation of the last two terms of this expression.

Sometimes having an exact relation can be used even when the physics at hand is imperfectly known. We will use the optical theorem when we construct the general form of the scattering amplitude as an expansion in Legendre polynomials or “partial waves.”

The optical theorem is only exact if the scattering amplitude is computed exactly. The relation will not be satisfied if $f(k, k')$ is computed approximately. In our simple example, suppose $S = \exp(i\delta) = 1 + i\delta - 1/2\delta^2 + \dots$. The constraint $S^\dagger S = 1$ is only true when all the terms in the series are kept. A place where this is encountered is the (first order) Born approximation, where $f(k, k')$ is real. Higher order terms in the expansion will (gradually) unitarize the scattering amplitude.

Sometimes people use the optical theorem as a check on their approximate calculation: how far can I trust what I did? Sometimes it is a guide to model building. The author is old enough to remember the use of this formula, with an approximation of the scattering amplitude which was believed to be very accurate at small angles, to predict total cross sections – today’s physics example would be the total cross section for proton - proton collisions at very high energies, such as at the Large Hadron Collider.

There is an optical theorem in electrodynamics, where energy conservation replaces unitary evolution. It is a bit more annoying to derive than what we did because the objects in the calculation (the \vec{E} and \vec{B} fields) carry vector indices. Our derivation implicitly assumed that we were dealing with spinless particles – $\psi(x)$ is simply a complex field. There is an extensive literature of scattering with spin, or scattering where the initial and final states are very different, but perhaps you should go to the specialized literature to read about that.

And finally, unitarity imposes some pretty remarkable constraints on the scattering amplitude, which we will encounter below.

Spherical wave states

Suppose that the scattering potential is rotationally symmetric. In this case it is very convenient to work with states which are simultaneous eigenstates of the free Hamiltonian H_0 , the total squared angular momentum, \hat{L}^2 , and the projection of the angular momentum along the z -axis, \hat{L}_z . These states are called spherical wave states; we will denote them as $|E, l, m\rangle$. They must satisfy the orthogonality relation

$$\langle E', l', m' | E, l, m \rangle = \delta_{ll'} \delta_{mm'} \delta(E - E'). \quad (13.124)$$

We expect that the projection of these states into a momentum or wave number basis will be proportional to the spherical harmonics,

$$\langle \vec{k} | E, l, m \rangle = g_{lE}(k) Y_l^m(\hat{k}). \quad (13.125)$$

If the particle is assumed to be traveling along the z -axis (i.e. the wave number vector points along the z -axis) there should be no z -projected angular momentum associated with the state. This result follows from

$$\begin{aligned} \hat{L}_z &= (xp_y - yp_x) |k\hat{z}\rangle \\ &= (xp_y - yp_x) |k_x = 0, k_y = 0, k_z = k\rangle = 0. \end{aligned} \quad (13.126)$$

We can expand $|k\hat{z}\rangle$ in the spherical wave state basis by using completeness,

$$|k\hat{z}\rangle = \sum_{l'} \int dE' |E', l', m' = 0\rangle \langle E', l', m' = 0 | k\hat{z}\rangle. \quad (13.127)$$

An arbitrary momentum state $|\vec{k}\rangle$ can be constructed by performing a rotation from the momentum state along the z -axis, which can be characterized by Euler angles

$$|\vec{k}\rangle = \mathcal{D}(\alpha = \phi, \beta = \theta, \gamma = 0) |k\hat{z}\rangle. \quad (13.128)$$

With this relation, a general spherical wave state in the momentum basis is

$$\langle E, l, m | \vec{k} \rangle = \sum_{l'} \int dE' \langle E, l, m | \mathcal{D}(\phi, \theta, 0) | E', l', m' = 0 \rangle \langle E', l', m' = 0 | k\hat{z} \rangle. \quad (13.129)$$

The rotation matrix is an independent entity from the spherical wave states, so

$$\langle E, l, m | \vec{k} \rangle = \sum_{l'} \int dE' \delta(E - E') \delta_{ll'} \mathcal{D}'_{m0} \langle E', l', m' = 0 | k\hat{z} \rangle \quad (13.130)$$

or

$$\langle E, l, m | \vec{k} \rangle = \mathcal{D}_{m0}^l(\phi, \theta, 0) \langle E, l, 0 | k \hat{z} \rangle. \quad (13.131)$$

Recalling that that \mathcal{D}_{m0}^l is proportional to Y_l^{-m} , Eq. 13.131 becomes

$$\langle \vec{k} | E, l, m \rangle = \sqrt{\frac{4\pi}{2l+1}} Y_l^m(\hat{k}) g_{lE}(k). \quad (13.132)$$

Our goal is now to find the functional form of the proportionality factor $g_{lE}(k)$. Begin with

$$\langle \vec{k} | \hat{H}_0 - E | E, l, m \rangle = \left(\frac{\hbar^2 k^2}{2m} - E \right) \langle \vec{k} | E, l, m \rangle = 0, \quad (13.133)$$

which tells us that $g_{lE}(k) = N \delta \left(\frac{\hbar^2 k^2}{2m} - E \right)$. Now perform the expansion

$$\begin{aligned} \langle E', l', m' | E, l, m \rangle &= \int d^3 k'' \langle E', l', m' | \vec{k}'' \rangle \langle \vec{k}'' | E, l, m \rangle \\ &= \int k^2 dk d\Omega N^2 \delta \left(\frac{\hbar^2 k''^2}{2m} - E \right) \delta \left(\frac{\hbar^2 k''^2}{2m} - E' \right) Y_l^{m'*}(\Omega) Y_l^m(\Omega) \\ &= |N|^2 \frac{mk'}{\hbar^2} \delta(E - E') \delta_{l'l'} \delta_{mm'}. \end{aligned} \quad (13.134)$$

The normalization constraint is $N = \frac{\hbar}{\sqrt{mk}}$ and so the momentum spherical wave state is

$$\langle \vec{k} | E, l, m \rangle = \frac{\hbar}{\sqrt{mk}} \delta \left(\frac{\hbar^2 k^2}{2m} - E \right) Y_l^m(\hat{k}). \quad (13.135)$$

The coordinate space spherical wave state is computed straightforwardly. We expect that $\langle \vec{x} | E, l, m \rangle \propto j_l(kr) Y_l^m(\hat{r})$. To find the constant of proportionality we use

$$\begin{aligned} \langle \vec{x} | \vec{k} \rangle &= \frac{e^{i\vec{k} \cdot \vec{x}}}{(2\pi)^{3/2}} = \sum_{lm} \int dE \langle \vec{x} | E, l, m \rangle \langle E, l, m | \vec{k} \rangle \\ &= \sum_{lm} \int dE c_l j_l(kr) Y_l^m(\hat{r}) \frac{\hbar}{\sqrt{mk}} \delta \left(\frac{\hbar^2 k^2}{2m} - E \right) Y_l^{m*}(\hat{k}). \end{aligned} \quad (13.136)$$

Exploiting the identity between the spherical harmonics and Legendre polynomials

$$\sum_m Y_l^m(\hat{r}) Y_l^{m*}(\hat{k}) = P_l(\hat{k} \cdot \hat{r}) \frac{2l+1}{4\pi}, \quad (13.137)$$

we find

$$\langle \vec{x} | \vec{k} \rangle = \sum_l \frac{2l+1}{4\pi} P_l(\hat{k} \cdot \hat{r}) j_l(kr) c_l \frac{\hbar}{\sqrt{mk}}. \quad (13.138)$$

Furthermore, the expansion

$$\frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^{3/2}} = \sum_l \frac{2l+1}{4\pi} P_l(\hat{k}\cdot\hat{r}) j_l(kr) \frac{1}{(2\pi)^{3/2}} i^l, \quad (13.139)$$

yields the relation $c_l = \frac{i^l}{\hbar} \sqrt{\frac{2mk}{\pi}}$. The final results are

$$\langle \vec{k} | E, l, m \rangle = \frac{\hbar}{\sqrt{mk}} \delta \left(E - \frac{\hbar^2 k^2}{2m} \right) Y_l^m(\hat{k}) \quad (13.140)$$

$$\langle \vec{x} | E, l, m \rangle = \frac{i^l}{\hbar} \sqrt{\frac{2mk}{\pi}} j_l(kr) Y_l^m(\hat{r}). \quad (13.141)$$

We will use all this machinery in the next section to derive some extremely useful results.

Partial wave analysis of scattering

If the potential is spherically symmetric then \hat{T} is a rotational scalar. Consequently,

$$[\hat{T}, \hat{L}^2] = 0 \quad [\hat{T}, \hat{L}_z] = 0. \quad (13.142)$$

The Wigner-Eckhart theorem then tells us that

$$\langle E, l, m | \hat{T} | E', l', m' \rangle = \delta_{ll'} \delta_{mm'} T_l(E). \quad (13.143)$$

Expanding the scattering amplitude in the spherical wave state basis gives

$$\begin{aligned} f(\vec{k}, \vec{k}') &= -\frac{1}{4\pi} \frac{2m}{\hbar^2} (2\pi)^3 \frac{\hbar^2}{mk} \sum_{lm} \sum_{l'm'} \int dE \int dE' \langle \vec{k}' | E', l', m' \rangle T_l(E) \delta_{ll'} \delta_{mm'} \langle E, l, m | \vec{k} \rangle \\ &= -\frac{4\pi^2}{k} \sum_{lm} T_l(E) Y_l^{m*}(\hat{k}') Y_l^m(\hat{k}). \end{aligned} \quad (13.144)$$

We now define the “partial wave amplitude,”

$$f_l(k) = \frac{\pi T_l(E)}{k}. \quad (13.145)$$

Taking \vec{k} to lie along the z-axis,

$$Y_l^m(\hat{z}) = \delta_{m0} \sqrt{\frac{2l+1}{4\pi}} \quad (13.146)$$

and in general

$$Y_l^0(\hat{k}) = \sqrt{\frac{2l+1}{4\pi}} P_l(\cos\theta). \quad (13.147)$$

The scattering amplitude becomes

$$f(\vec{k}, \vec{k}') = \sum_{l=0}^{\infty} (2l+1) f_l(k) P_l(\cos \theta). \quad (13.148)$$

At large distances from the scattering source, the coordinate space wave function was found to be

$$\langle x | \psi_+ \rangle = \frac{1}{(2\pi)^{3/2}} \left(e^{ikz} + \frac{e^{ikr}}{r} f(\theta) \right). \quad (13.149)$$

Our goal is to combine these expressions into a compact formula for the outgoing wave. To begin, we use

$$e^{ikz} = \sum_l (2l+1) P_l(\cos \theta) (e^{i\pi/2})^l j_l(kr). \quad (13.150)$$

In the limit $kr \gg 1$, the Bessel functions take their asymptotic forms

$$j_l(kr) \rightarrow \frac{\sin(kr - \frac{l\pi}{2})}{kr} = \frac{e^{i(kr - l\pi/2)} - e^{-i(kr - l\pi/2)}}{2ikr}, \quad (13.151)$$

and so the wave function is

$$\begin{aligned} \langle x | \psi_+ \rangle &= \frac{1}{(2\pi)^{3/2}} \left(\sum_l (2l+1) P_l(\cos \theta) \left(\frac{e^{ike} - e^{-i(kr - l\pi)}}{2ikr} \right) + \sum_l (2l+1) P_l(\cos \theta) f_l(k) \frac{e^{ikr}}{r} \right) \\ &= \frac{1}{(2\pi)^{3/2}} \sum_l (2l+1) \frac{P_l(\cos \theta)}{2ik} \left((1 + 2ik f_l(k)) \frac{e^{ikr}}{r} - \frac{e^{i(kr - l\pi)}}{r} \right). \end{aligned} \quad (13.152)$$

This is not a terribly illuminating expression, but it conceals a very useful nugget of information – a completely general parameterization of the partial wave amplitude f_l . To extract the nugget, notice that the first term in this expression is an outgoing spherical wave and the second piece is an incoming spherical wave. Now recall that probability must be conserved in a scattering process. Probability conservation implies the existence of a conserved current, or that $\vec{\nabla} \cdot \vec{J} = \frac{\partial \rho}{\partial t}$. If the process is time-independent, $\vec{\nabla} \cdot \vec{J} = 0$, and

$$\int \vec{J} \cdot d\vec{A} = 0 \quad (13.153)$$

where the integration is implicitly taken over a spherical surface at infinity. For this expression to hold, the squared amplitudes of the incoming and outgoing spherical waves in Eq. 13.152 must be equal, forcing

$$1 = |1 + 2ik f_l(k)|^2, \quad (13.154)$$

which tells us that $1 + 2ikf_l(k)$ must be a pure phase,

$$1 + 2ikf_l(k) = e^{2i\delta_l(k)}. \quad (13.155)$$

The quantity $\delta_l(K)$ is called the “phase shift for the l th partial wave.” (The 2 in the exponent is a convention.) Notice that there are three equivalent expressions for the scattering amplitude,

$$f_l(k) = \frac{e^{2i\delta_l} - 1}{2ik} = \frac{e^{i\delta_l} \sin \delta_l}{k} = \frac{1}{k \cot \delta_l - ik}. \quad (13.156)$$

You will see all of these expressions in the literature.

Written in terms of phase shifts, scattering amplitude becomes

$$f(\theta) = \sum_{l=0}^{\infty} (2l+1) \left(\frac{e^{2i\delta_l} - 1}{2ik} \right) P_l(\cos \theta). \quad (13.157)$$

This formula is very often used to analyze the results of scattering experiments. It is completely general; its derivation only depended on unitarity and rotational invariance. Note that it says that only one energy dependent parameter (δ_l) per partial wave is needed to parameterize the scattering amplitude.

We write the differential cross section as

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 = \frac{1}{k^2} \sum_{l,l'} (2l+1)(2l'+1) e^{i(\delta_l - \delta_{l'})} \sin \delta_l \sin \delta_{l'} P_l(\cos \theta) P_{l'}(\cos \theta). \quad (13.158)$$

The total cross section is simply the differential cross section integrated over solid angle. Using

$$\int d\Omega P_l(\cos \theta) P_{l'}(\cos \theta) = \frac{4\pi}{2l+1} \delta_{ll'}, \quad (13.159)$$

it is

$$\sigma = \frac{4\pi}{k^2} \sum_l (2l+1) \sin^2 \delta_l. \quad (13.160)$$

Let us check the optical theorem:

$$\begin{aligned} f(\theta = 0) &= \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) e^{i\delta_l} \sin \delta_l P_l(1) \\ &= \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) \sin \delta_l (\cos \delta_l + i \sin \delta_l) \\ \text{Im}(f(\theta = 0)) &= \frac{1}{k} \sum_{l=0}^{\infty} (2l+1) \sin^2 \delta_l = \frac{1}{4\pi} \sigma \end{aligned} \quad (13.161)$$

The derivation of these results assumed potential scattering and, therefore, that the scattering process was unitary. A variation on this language can be used to parameterized a particular non-unitary variation of this picture. We imagine that a scattering process can either involve an emerging final state particle, and further that this particle has the same energy as the incident one, or that the incident particle is completely absorbed by the target and nothing comes out. We refer to the two processes as “elastic scattering” or “absorption.” This is clearly a phenomenological description, but the general case of scattering where many different final states might emerge can be cast into a generalization of this language. This is done as follows:

Recall that the flux into the origin per partial wave is

$$\left| \frac{1}{(2\pi)^{3/2}} \frac{1}{2ik} \frac{e^{-ikr}}{r} \right|^2 \quad (13.162)$$

and the outgoing flux per partial wave is

$$\left| \frac{1}{(2\pi)^{3/2}} \frac{1}{2ik} (1 + 2ikf_l(k)) \frac{e^{ikr}}{r} \right|^2. \quad (13.163)$$

We define an “inelasticity factor” η_l , with $0 \leq \eta_l \leq 1$, from the ratio of the outgoing and incoming fluxes,

$$\eta_l \equiv \frac{\text{flux out}}{\text{flux in}} \quad (13.164)$$

With it, we write

$$1 + 2ikf_l(k) = \eta_l e^{2i\delta_l} \quad (13.165)$$

so that the elastic scattering amplitude is

$$f_l(k) = \frac{\eta_l e^{2i\delta_l} - 1}{2ik}. \quad (13.166)$$

The total elastic cross section is

$$\sigma_{el} = \int d\Omega |f(\theta)|^2 = 4\pi \sum_l (2l+1) |f_l|^2 \quad (13.167)$$

where

$$|f_l|^2 = \frac{1 + \eta_l^2 - 2\eta_l \cos(2\delta_l)}{4k^2}. \quad (13.168)$$

The absorption or inelastic cross section is

$$\sigma_{inel} = \sigma_{abs} = \frac{\text{flux lost}}{\text{flux in}} = \frac{\pi}{k^2} \sum_l (2l+1) (1 - \eta_l^2). \quad (13.169)$$

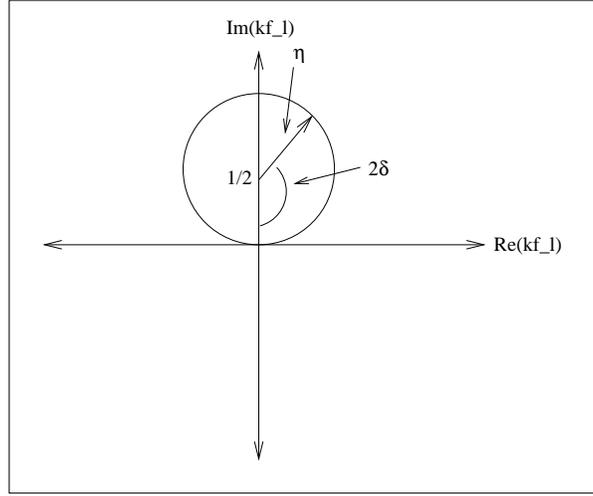


Figure 13.1: Variables in an Argand diagram.

The total cross section is the sum,

$$\begin{aligned}\sigma_{tot} &= \sigma_{el} + \sigma_{inel} = \frac{\pi}{k^2} \sum_l (2l+1) (1 + \eta_l^2 - 2\eta_l \cos(2\delta_l) + 1 - \eta_l^2) \\ &= \frac{\pi}{k^2} \sum_l (2l+1) 2(1 - \eta_l \cos(2\delta_l)).\end{aligned}\quad (13.170)$$

The formalism preserves the optical theorem:

$$\text{Im}(f(\theta=0)) = \sum_l (2l+1) \frac{1 - \eta_l \cos(2\delta_l)}{2k} = \frac{k}{4\pi} \sigma_{tot}.\quad (13.171)$$

The Argand diagram, Fig. 13.1, is a convenient way to present information about scattering amplitudes: it is a plot of

$$k f_l(k) = \frac{\eta_l e^{2i\delta_l} - 1}{2i}\quad (13.172)$$

in the complex plane: as we vary k , $k f_l(k)$ traces a circle of radius $\frac{1}{2}$ when $\eta_l = 1$. The circle touches the real axis of $k f_l$ when $\delta_l = 0$. We represent scattering data as point on or within the circle. A typical Argand diagram would look like Fig. 13.2.

How can we compute phase shifts from a given potential $V(r)$? If the potential is identically zero beyond some distance, $V(r) = 0$ for $r > R$, then for $r > R$, the wave function corresponds to that of a free outgoing wave. This solution can be constructed from linear

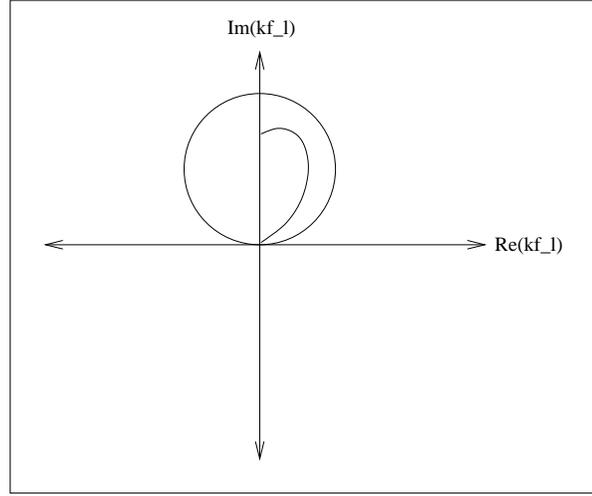


Figure 13.2: A typical Argand diagram is a path in the complex plane. The scattering amplitude $kf(k)$ follows a path in the complex plane as k is varied.

combinations of the spherical Bessel and Neumann functions, the Hankel functions.

$$h_l^{(1)} = j_l + in_l \rightarrow \frac{\exp(i(kr - l\pi/2))}{ikr} \quad (13.173)$$

$$h_l^{(2)} = j_l - in_l \rightarrow -\frac{\exp(-i(kr - l\pi/2))}{ikr}. \quad (13.174)$$

The arrow shows the asymptotic limiting form for large value of the argument. The first Hankel function thus represents an outgoing wave while the second Hankel function corresponds to an incoming wave. The general solution to Schrödinger's equation with $V(r) = 0$ will be the superposition

$$\langle x|\psi_+\rangle = \frac{1}{(2\pi)^{3/2}} \sum_l i^l (2l+1) A_l(r) P_l(\cos\theta) \quad (13.175)$$

where

$$A_l(r) = c_l^{(1)} h_l^{(1)}(kr) + c_l^{(2)} h_l^{(2)}(kr). \quad (13.176)$$

Its limiting form must match the one we had previously found, in Eq. 13.152, so that $c_l^{(1)} = \frac{1}{2}e^{2i\delta_l}$ and $c_l^{(2)} = \frac{1}{2}$. Thus, for $r \gg R$ we have

$$A_l(r) = e^{i\delta_l} (\cos\delta_l j_l(kr) - \sin\delta_l n_l(kr)). \quad (13.177)$$

Next, we patch together the solutions from $r < R$ to those for $R > r$ at the boundary of $r = R$. Matching functions and derivatives is most economically done by matching

logarithmic derivatives at the boundary. So define for the interior solution the quantity

$$\beta_l = \frac{r}{A_l} \frac{dA_l}{dr} \Big|_{r=R} = \frac{r}{\psi_{inside}} \frac{d\psi_{inside}}{dr} \Big|_{r=R}. \quad (13.178)$$

From the outside wave function we have

$$\beta_l = kR \frac{\cos \delta_l \frac{dj_l}{dr}(kR) - \sin \delta_l \frac{dn_l}{dr}(kR)}{\cos \delta_l j_l(kR) - \sin \delta_l n_l(kR)}. \quad (13.179)$$

Equating the expressions and grooming a bit gives

$$\tan \delta_l = \frac{j_l(kR)}{n_l(kR)} \left[\frac{kR \frac{j_l'(kR)}{j_l(kR)} - \beta_l}{kR \frac{n_l'(kR)}{n_l(kR)} - \beta_l} \right]. \quad (13.180)$$

(Recall, parenthetically, that the most convenient expression for the interior Schrödinger equation is the one-dimensional form

$$\frac{d^2 u_l(r)}{dr^2} - \left(k^2 - \frac{2mV}{\hbar^2} - \frac{l(l+1)}{\hbar^2} \right) u_l = E u_l \quad (13.181)$$

where $u_l(r) = rA_l(r)$.)

An example might be helpful: consider a hard sphere potential,

$$V(r) = \begin{cases} 0 & r > R \\ \infty & r < R \end{cases} \quad (13.182)$$

Obviously, the potential acts as a boundary condition: $A_l(R) = 0$, and so

$$\tan \delta_l = \frac{j_l(kR)}{n_l(kR)}. \quad (13.183)$$

Consider first s-wave scattering:

$$\tan \delta_0 = \frac{\frac{\sin(kR)}{kR}}{-\frac{\cos(kR)}{kR}} = -\tan(kR) \quad (13.184)$$

or $\delta_0 = -kR$. The wave function has just been shifted in space. At low energy (i.e. for long wavelength, $kR \ll 1$) the phase shifts for the higher partial waves are small,

$$\tan \delta_l = \frac{j_l(kR)}{n_l(kR)} \rightarrow \frac{-(kR)^{2l+1}}{(2l+1)!!(2l-1)!!} \quad (13.185)$$

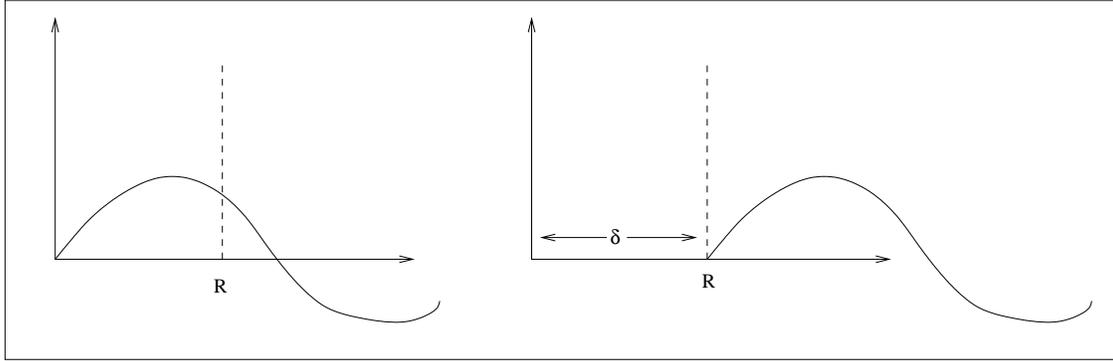


Figure 13.3: The hard sphere potential and the phase shift for s-waves.

and only s-wave scattering will be important. This is a general result, as can be seen from Eq. 13.180 – the expression in the square brackets is only order unity for small kR . The cross section can then be approximated as

$$\frac{d\sigma}{d\Omega} = \frac{\sin^2 \delta_0}{k^2} \approx R^2 \quad (13.186)$$

or

$$\sigma = 4\pi R^2 \quad (13.187)$$

This is four times the expected geometric cross section, πR^2 , but when the wavelength is small compared to the geometric size of a scatterer, we do not expect geometry to give a correct answer.

Very high energy scattering, with $kR \gg 1$, can be treated as scattering from a classical target. Recalling that

$$|\vec{L}| = |\vec{r} \times \vec{p}| = rp = r\hbar k, \quad (13.188)$$

we can define a maximum angular momentum $l_{max} = kR$. The cross section can then be written as

$$\sigma = \frac{4\pi}{k^2} \sum_{l=0}^{l=kR} (2l+1) \sin^2 \delta_l. \quad (13.189)$$

Using the asymptotic forms valid at large kR ,

$$j_l \rightarrow \frac{\sin(kR - l\pi/2)}{kR} \quad (13.190)$$

$$n_l \rightarrow \frac{\cos(kR - l\pi/2)}{kR} \quad (13.191)$$

the phase shifts approach the limiting value

$$\tan \delta_l = \tan \left(kR - \frac{l\pi}{2} \right) \quad (13.192)$$

or

$$\delta_l = kR - \frac{l\pi}{2}. \quad (13.193)$$

Inserting this into the formula for the cross section and, because many partial waves contribute, converting the sum to an integral, the cross section becomes

$$\begin{aligned} \sigma &= \frac{4\pi}{k^2} \sum_{l=0}^{l_{max}=kR} (2l+1) \sin^2 \left(kR - \frac{l\pi}{2} \right) \\ &\approx \frac{4\pi}{k^2} \int_0^{kR} (2l+1) \sin^2 \left(kR - \frac{l\pi}{2} \right) \\ &\approx \frac{4\pi}{k^2} (kR)^2 \frac{1}{2} = 2\pi R^2. \end{aligned} \quad (13.194)$$

The \sin^2 term has averaged to $\frac{1}{2}$ because it is fluctuating so rapidly.

We might have expected that the cross section would just have been the area of the target, πR^2 , so what about the extra factor of two? Write the scattering amplitude as the sum of two terms,

$$f(\theta) = \frac{1}{2ik} \sum_{l=0}^{kR} (2l+1) e^{2i\delta_l} P_l(\cos \theta) + \frac{i}{2k} \sum_{l=0}^{kR} (2l+1) P_l(\cos \theta) \equiv f_1 + f_s. \quad (13.195)$$

The cross section contribution from f_1 is

$$\sigma = \int d\Omega |f_s|^2 = \frac{\pi}{k^2} \sum_{l=0}^{kR} (2l+1) = \pi R^2. \quad (13.196)$$

The f_s amplitude makes no reference to the potential, is pure imaginary, and is large at $\theta = 0$, because that is where the Legendre polynomials are peaked. It is called the shadow amplitude,

$$f_s = \frac{i}{2k} \sum_{l=0}^{kR} (2l+1) P_l(\cos \theta). \quad (13.197)$$

For small angles and big l 's the Legendre polynomials can be approximated as the first Bessel function of $l\theta$ argument,

$$P_l(\cos \theta) \rightarrow J_0(l\theta). \quad (13.198)$$

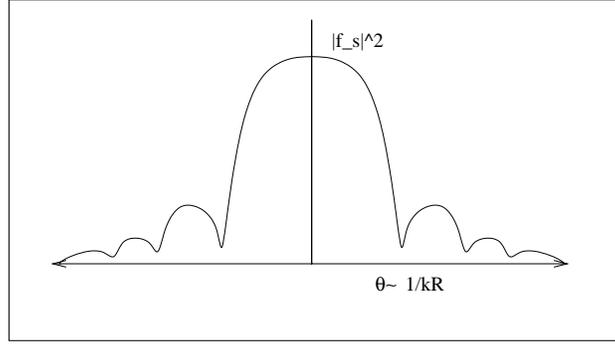


Figure 13.4: Schematic plot of the shadow scattering differential cross section.

Then the shadow amplitude is approximately

$$f_s = \frac{2i}{2k} \int_0^{kR} dl J_0(l\theta) l = \frac{iR J_1(kR\theta)}{\theta} \quad (13.199)$$

Graphically, the differential cross section $|f_s|^2$ is shown in Fig. 13.4. The corresponding cross section is

$$\sigma_{sh} = \int d\Omega |f_s|^2 = 2\pi \int \frac{R^2 J_1^2(kR\theta) d(\cos\theta)}{\theta} = 2\pi R^2 \int_0^\infty \frac{|J_1(\xi)|^2 d\xi}{\xi} = \pi R^2, \quad (13.200)$$

where we have set $\xi = kR\theta$ to do the integral. The interference term between f_1 and f_s averages to zero because $2\delta_{l+1} - 2\delta_l = \pi$ while f_s is always imaginary. So the result is $2\pi R^2$, half from the actual phase shift, and half from the shadow term.

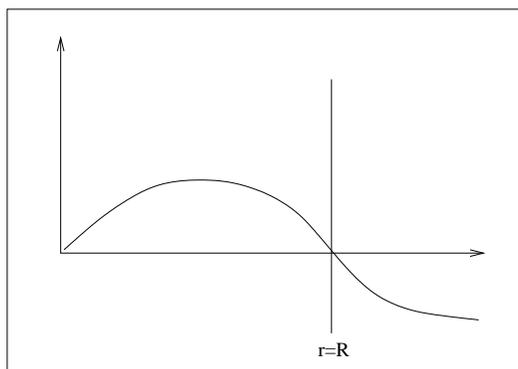
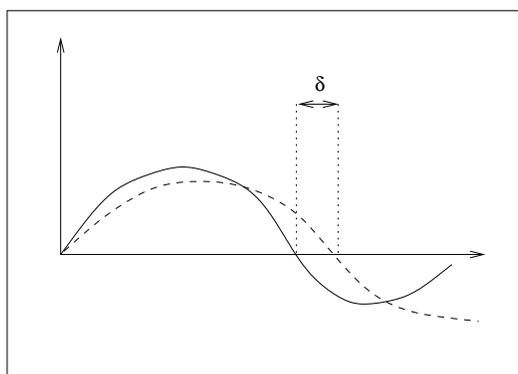
That they are equal is easy to see if we look immediately behind the scattering center, where nothing should be observed, and a shadow must be created. In fact, the shadow term can be thought of as giving the entire cross section, from the optical theorem. This occurs because $\text{Im } f_1$ averages to zero, while the partial waves of f_s add coherently,

$$\frac{4\pi}{k} \text{Im } f_s(0) = \frac{4\pi}{k} \text{Im} \left(\frac{i}{2k} \sum_{l=0}^{kR} (2l+1) \right) = 2\pi R^2. \quad (13.201)$$

Universal features of scattering at low energy

At very low energy, only s-wave scattering is important, since the phase shift for the l th partial wave goes like k^{2l+1} and the cross section scales as $\sigma_l \simeq k^{4l}$. Outside the region where the potential is nonzero, the wave function is

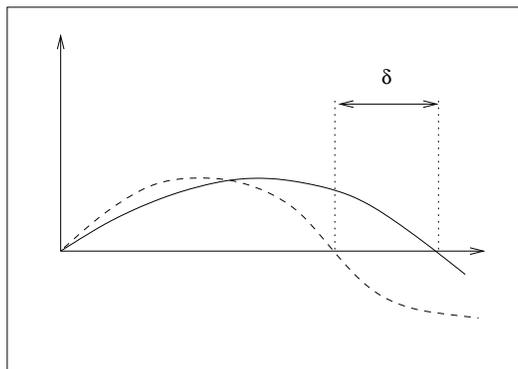
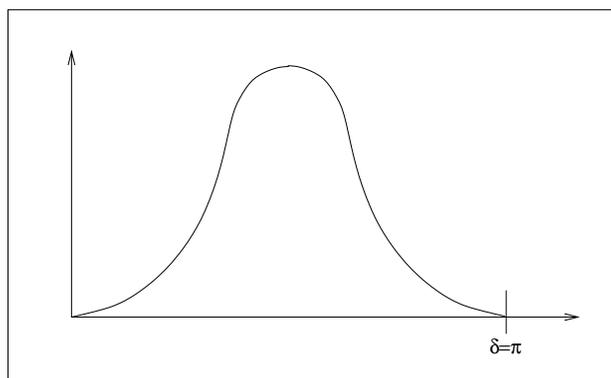
$$u(r) = \sin(kr + \delta). \quad (13.202)$$

Figure 13.5: Wave function when $V = 0$.Figure 13.6: Wave function when $V < 0$ and $\delta > 0$.

How is the sign of δ related to the sign of the potential? We have three cases, illustrated in Figs. 13.5-13.7:

- 1) $V = 0$: In this case the phase shift $\delta = 0$
- 2) $V < 0$: In this case the interior wave number, $k' = \sqrt{\frac{2m}{\hbar^2}(E - V)}$ is greater than the exterior wave number $k = \sqrt{\frac{2mE}{\hbar^2}}$. The location of the first node occurs closer to the origin than in the $V = 0$ case. Thus $\delta > 0$.
- 3) $V > 0$: In this case $k' = \sqrt{\frac{2m}{\hbar^2}(E - V)}$ is less than $k = \sqrt{\frac{2mE}{\hbar^2}}$. This implies that $\delta < 0$.

As seen in Fig. 13.6, we expect that δ grows larger as the potential becomes deeper and deeper. (Compare Fig. 13.8.) Eventually the point might reach $\delta = \pi$. At this point the

Figure 13.7: Wave function when $V > 0$ and $\delta < 0$.Figure 13.8: Variation of the phase shift as $V(r)$ becomes more and more negative.

cross section will fall to zero. This is quite odd – the cross section is zero even though the potential is very attractive. But it has been observed, in electron scattering from noble gas atoms, where it is called the Ramsauer-Townsend effect.

Let us continue to assume that k is very small. For $V = 0$ the Schrödinger equation

$$\frac{d^2u}{dr^2} + k^2u = 0 \quad (13.203)$$

becomes

$$\frac{d^2u}{dr^2} = 0 \quad (13.204)$$

or

$$u(r) = \text{const.} \times (r - a). \quad (13.205)$$

Another way to see this result is from the limiting form of the exact solution (recall,

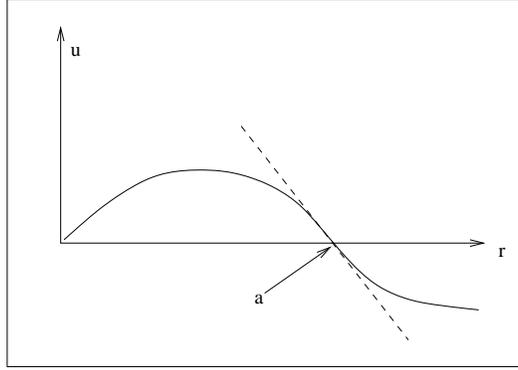


Figure 13.9: Linearizing the wave function about its zero.

$\delta_l \simeq k^{2l+1}$):

$$\lim_{k \rightarrow 0} \sin(kr + \delta_0) \approx k \left(r + \frac{\delta_0}{k} \right). \quad (13.206)$$

Comparing the two expressions,

$$-a = \frac{\delta_0}{k} \quad (13.207)$$

Recall that it is the logarithmic derivative of the wave function which enters into the expression for the phase shift, Eq. 13.180. With $u = \sin(kr + \delta)$, the logarithmic derivative at the matching point is

$$\frac{u'}{u} = k \cot(kr + \delta) \Big|_R = k \cot \left(k \left(r + \frac{\delta}{k} \right) \right) \Big|_R \approx \frac{1}{r-a} \Big|_R. \quad (13.208)$$

Furthermore, if the range of the potential is small then

$$\frac{1}{r-a} \approx -\frac{1}{a} \quad (13.209)$$

or

$$\lim_{k \rightarrow 0} k \cot(kr + \delta) = -\frac{1}{a}. \quad (13.210)$$

The scattering amplitude is

$$f = \frac{1}{k \cot \delta - ik} = -\frac{1}{1/a + ik}. \quad (13.211)$$

The cross section itself is

$$\sigma = 4\pi \lim_{k \rightarrow 0} \frac{1}{|k \cot \delta - ik|^2} = 4\pi a^2. \quad (13.212)$$

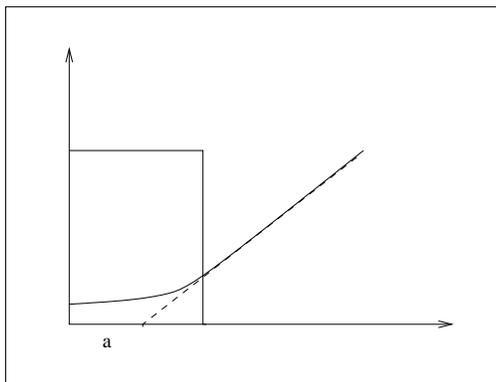


Figure 13.10: A repulsive potential generally implies a positive scattering length.

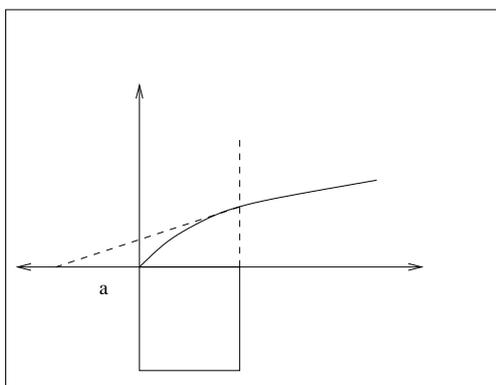


Figure 13.11: An attractive potential might imply a negative scattering length...

The quantity “ a ” is called the scattering length. Notice that it completely characterizes the cross section at low energy, regardless of the shape of the potential.

What can we say about the sign of the scattering length? If $V = 0$, then the wave function rises linearly from the origin and $a = 0$. Now suppose $V > 0$: A repulsive V suppresses the wave function under it. If we were extrapolate a linear wave function to the origin, we would find an $a > 0$. This is illustrated by Fig. 13.10. An attractive potential ($V < 0$) might give a positive a (Fig. 13.11) or a negative one (Fig. 13.12).

We can take the analysis a bit farther. We have assumed that k is small, but have not

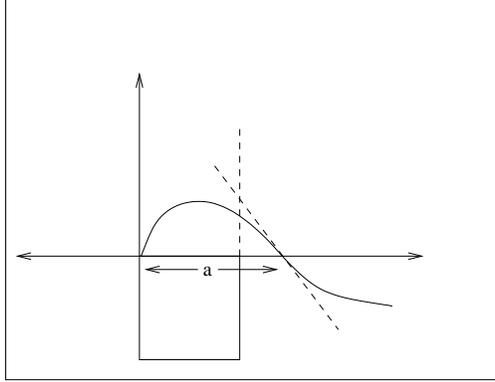


Figure 13.12: Or a positive scattering length.

found any dependence of σ on k . To remedy that, recall Eq. 13.180, specialized to s-waves:

$$\tan \delta_0 = \frac{j_0(kR)}{n_0(kR)} \left[\frac{kR \frac{j_0'}{j_0} - \beta}{kR \frac{n_0'}{n_0} - \beta} \right] \quad (13.213)$$

Substituting for the spherical Bessel functions gives

$$\tan \delta_0 = -\tan(kR) \left[\frac{-1 - \beta + kR \cot kR}{-1 - \beta - kR \tan kR} \right] = \frac{kR - (1 + \beta) \tan kR}{1 + \beta + kR \tan kR}. \quad (13.214)$$

For $kR \ll 1$ we can expand this expression and pick up the leading k dependence (note $\tan x = x(1 + x^2/3)$),

$$\tan \delta_0 = -(kR) \left[\frac{\beta + \frac{(kR)^2}{3}(1 + \beta)}{1 + \beta + (kR)^2} \right]. \quad (13.215)$$

This says that

$$\frac{1}{k} \tan \delta_0 = -R (C_1 + C_2 k^2 + \dots) \quad (13.216)$$

or, more conventionally,

$$k \cot \delta_0 \approx -\frac{1}{a} + \frac{1}{2} r_0 k^2. \quad (13.217)$$

Again, the scattering length is a . The quantity r_0 is a new parameter, called the effective range. The cross section is characterized by two parameters:

$$\sigma = \frac{4\pi a^2}{1 + \left(1 - \frac{r_0}{a}\right) k^2 a^2} + O(k^4). \quad (13.218)$$

The $O(k^4)$ represents both higher order terms in the Taylor expansion and the still - uncomputed p-wave contribution.

This is a remarkable result: regardless of the actual shape of $V(r)$, the scattering amplitude at low energy is completely characterized by two parameters. This represents a decoupling of short and long distance physics: the detailed behavior of the system at short distance ($\ll 1/k$) does not affect long distance ($\simeq 1/k$) probes. This is good news: to understand long distance physics, one does not have to understand short distances. For example, consider cold trapped atoms: how do they interact? Atoms are hopelessly complicated and it would be a very difficult problem to compute the scattering amplitude for two atoms at any moderate energy. However, to understand atomic interactions in a trap, one only needs two numbers, a and r_0 . If they cannot be computed from first principles, they can be determined from one or two experimental measurements. Then all other properties of the trapped atoms which depend on how they interact (for example, the shift in the critical temperature T_c for Bose-Einstein condensation, away from the ideal gas result) can be expressed in terms of a and r_0 , plus the mass of the atom.

There is corresponding bad news: one can't learn much about short distance physics with long distance experiments. An old illustration of this fact comes from the 1930's. After the neutron was discovered, a natural question to ask was, what is the neutron-proton potential $V(r)$? One could do scattering of protons on deuterium (at energies of a few MeV) and determine the associated $p - n$ cross sections. Then one would pick a parameterization for $V(r)$, calculate $d\sigma/d\Omega$, fit the data, and determine the parameters. This project failed, because all potentials which give the same a and r_0 produce identical scattering amplitudes, regardless of their detailed shape.

The modern generalization of this story is called "effective field theory." We describe low energy behavior of Nature using Lagrangians or Hamiltonians which consist of only as many operators as there are relevant parameters (here a and r_0), regardless of whether the field variables describe fundamental physics, or not. In the strong interactions, the relevant degrees of freedom at low energy are the pions, plus static protons and neutrons. They are not the fundamental constituents of strongly interacting matter, quarks and gluons. If desired, one could use the more fundamental theory to predict the parameter values of the low energy effective theory. However, all of low energy physics will be contained in them, and one would use the low energy theory to do calculations.

Physics at low energies depends only on the symmetries which are manifest at low energies, and the values of a few low energy parameters. In a BEC, the important degrees of freedom are the (local) wave function of the condensate, and the symmetries are invariances under global phase rotations (and local ones, if couplings to electromagnetism are included).

That the condensate has an atomic substructure is irrelevant to its low energy, long distance properties.

Let us illustrate how the scattering length and effective range appear in the “spherical square well,” $V(r) = -V_0$ for $r < R$, $V = 0$ otherwise. Recall the bound state solution

$$u(r) = \begin{cases} \sin(Gr) & r < R \\ e^{-\lambda r} & r > R \end{cases} \quad (13.219)$$

where $\frac{\hbar^2 \lambda^2}{2m} = |E_b|$ and $\frac{\hbar^2 G^2}{2m} = V_0 - |E_b|$. Recall also that $G^2 + \lambda^2 = \frac{2mV_0}{\hbar^2}$. Matching solutions and derivatives at the boundary (or matching logarithmic derivatives) gives us

$$GR \cot(GR) = -\lambda R \quad (13.220)$$

which can be solved to give the bound state energies. In the scattering situation, the energy of the incoming state is positive, and so the wave function is

$$u(r) = \begin{cases} \sin(Kr) & r < R \\ \sin(kr + \delta) & r > R \end{cases} \quad (13.221)$$

We again have the familiar relations $\frac{\hbar^2 K^2}{2m} = V_0 + E_{scat}$ and $\frac{\hbar^2 k^2}{2m} = E_{scat}$. Matching wave functions and derivatives at the boundary gives

$$KR \cot(KR) = kR \cot(kR + \delta), \quad (13.222)$$

which will determine the phase shift.

Now suppose that $E_{scat} \approx 0$. Additionally, suppose that the potential is such it supports one bound state, which is also close to zero: $|E_b| \approx 0$. This corresponds to $K \approx G$. The interior solution can be expanded in terms of bound state parameters as

$$K \cot(KR) = G \cot(GR) + (K - G) \frac{\partial}{\partial G} (G \cot(GR)) \quad (13.223)$$

$$\begin{aligned} &= -\lambda + (K - G) [\cot GR - GR(1 + \cot^2 GR)] \\ &= -\lambda + (K - G) \left(\frac{\lambda}{G} - GR \left(1 + \frac{\lambda^2}{G^2} \right) \right). \end{aligned} \quad (13.224)$$

If λ/G is tiny, this is

$$K \cot KR = -\lambda - (K - G)GR. \quad (13.225)$$

Furthermore, Taylor expanding $K - G$ in the assumed small $k^2 + \lambda^2$, we find

$$K - G = \left(\frac{2mV_0}{\hbar^2} - \lambda^2 + \lambda^2 + k^2 \right)^{\frac{1}{2}} - \left(\frac{2mV_0}{\hbar^2} - \lambda^2 \right)^{\frac{1}{2}} \quad (13.226)$$

$$\begin{aligned} &\approx \frac{1}{2} \frac{k^2 + \lambda^2}{\sqrt{\frac{2mV_0}{\hbar^2} - \lambda^2}} \\ &= \frac{1}{2} \frac{k^2 + \lambda^2}{G}. \end{aligned} \quad (13.227)$$

Then

$$K \cot(KR) = -\lambda - \frac{1}{2}(k^2 + \lambda^2)R. \quad (13.228)$$

For small kR compared to δ we can expand the expression

$$k \cot(kR + \delta) = k \cot \delta - k^2 R (1 + \cot^2 \delta). \quad (13.229)$$

Then we can combine the two equations to rebuild Eq. 13.222,

$$k \cot \delta - k^2 R (1 + \cot^2 \delta) = -\lambda - \frac{1}{2}(k^2 + \lambda^2)R. \quad (13.230)$$

Replace the $\cot^2 \delta$ by λ^2/k^2 , the leading behavior in the expression (we only need to work to order k^2) and we have

$$k \cot \delta - k^2 R \left(1 + \frac{\lambda^2}{k^2}\right) = -\lambda - \frac{1}{2}(k^2 + \lambda^2)R, \quad (13.231)$$

or

$$k \cot \delta = -\lambda + \frac{1}{2}(\lambda^2 + k^2)R. \quad (13.232)$$

Since $k \cot \delta = -1/a + \frac{1}{2}r_0 k^2$, we can read off the scattering length as

$$\frac{1}{a} = \lambda - \frac{1}{2}R\lambda^2 \quad (13.233)$$

and the effective range is just $r_0 = R$. The scattering length has little to do with the length scale in the problem (R) – the direct connection of the scattering length is to $1/\lambda$, the size of the wave function in the classically forbidden region.

We began this derivation by making the assumption that the bound state energy was close to zero. Notice that as the potential is tuned so that the bound state energy is driven to zero, the scattering length diverges as $1/\lambda$: the cross section can become enormous. There is a contemporary use of this relation for trapped cold atoms: their interactions can be tuned by tuning parameters in the potential, and can be made large by bringing the energy of a bound state close to zero energy. The words ‘‘Feshbach resonance’’ are used to describe this situation.

Analyticity properties of scattering amplitudes

Far from the origin, the general form of the (un-normalized and $l = 0$) wave functions are

$$\psi_{scat} = e^{2i\delta} \frac{e^{ikr}}{r} - \frac{e^{-ikr}}{r} \quad (13.234)$$

$$\equiv S(k) \frac{e^{ikr}}{r} - \frac{e^{-ikr}}{r} \quad (13.235)$$

for the scattering solution, and

$$\psi_{bound} = \frac{e^{-Kr}}{r} \quad (13.236)$$

for the bound state. For the bound state, only particular values of $K = K_n$ are allowed, while we have a continuum of possibilities for k for the scattering solution. Recall that the bound state wave function the scattering wave functions are solutions of the same Schrödinger equation (only the boundary conditions at infinity are different). They must be connected somehow. To make the connection, imagine evaluating the scattering solution at one of the bound state energies. It is

$$\psi(k \equiv iK_n) = S(iK) \frac{e^{-K_n r}}{r} - \frac{e^{K_n r}}{r} \quad (13.237)$$

It would reproduce the bound state solution if the incoming wave's amplitude were zero, that is, if

$$\frac{\psi_{incoming}}{\psi_{outgoing}} = \frac{1}{S(iK_n)} \rightarrow 0. \quad (13.238)$$

The connection between the bound state and the scattered wave function will only hold if this ratio goes to zero at the special values of $k = iK_n$. This will happen only if the scattering amplitude has poles at all these values,

$$S(k) \propto \prod_n \frac{1}{k - iK_n} \quad (13.239)$$

This is a very general and profound statement. It says that one can determine bound state properties from very careful scattering experiments. Recognizing that the connection we have discovered applies to all partial waves, we put an index on S_l , the scattering amplitude in the l th partial wave, We add to our list of properties of $S_l = e^{2i\delta_l}$:

- 1) S_l has poles at $k = iK$ for all bound states.

2) $S_l = e^{2i\delta_l}$ for all $k > 0$.

3) Because $\tan \delta_l$ or $\delta_l \sim k^{2l+1}$ at small k , $S = 1$ at $k = 0$

Let us check that these results can be obtained from the effective range formula:

$$k \cot \delta_l = ik \left[\frac{e^{i\delta} + e^{-i\delta}}{e^{i\delta} - e^{-i\delta}} \right] = ik \left[\frac{e^{2i\delta} + 1}{e^{2i\delta} - 1} \right] = ik \left[\frac{S + 1}{S - 1} \right] = -\frac{1}{a} + \frac{1}{2}r_0k^2 \quad (13.240)$$

If S has a pole at some k , then $\frac{S+1}{S-1} \rightarrow 1$ and

$$\frac{1}{a} = -ik + \frac{1}{2}r_0k^2 \quad (13.241)$$

At the bound state, we recall the definition of Eq. 13.219, that $k = i\lambda$, or

$$\frac{1}{a} = \lambda - \frac{1}{2}r_0\lambda^2 \quad (13.242)$$

This agrees with Eq. 13.233.

In the scattering length limit, the scattering amplitude is

$$f = -\frac{1}{1/a + ik} \quad (13.243)$$

Notice that we need $a > 0$ to have a true bound state, so that $ik = -1/a$ and e^{ikr} becomes $e^{-r/a}$. The pole at $ik = -1/a$ with $a < 0$ does not correspond to a bound state. People speak of a “virtual bound state,” but that is just a shorthand for having $a < 0$. These situations occur in low energy proton-neutron scattering. The total spin $S = 1$ channel has a bound state (the deuteron) which is at a tiny energy, 2.2 MeV compared to the mass of the neutron or proton; the singlet state has no bound state, but its scattering length is negative and about four times as large as the triplet’s.

Resonance and scattering

Recall that the total cross section can be written as

$$\sigma = \frac{4\pi}{k^2} \sum_{l=0}^{\infty} (2l+1) \sin^2 \delta_l. \quad (13.244)$$

There is a maximum value each partial wave can contribute to the cross section, which obviously occurs when $|\sin \delta_l| = 1$ or δ_l is an integer multiple of $\pi/2$. This is called the “unitarity bound.”

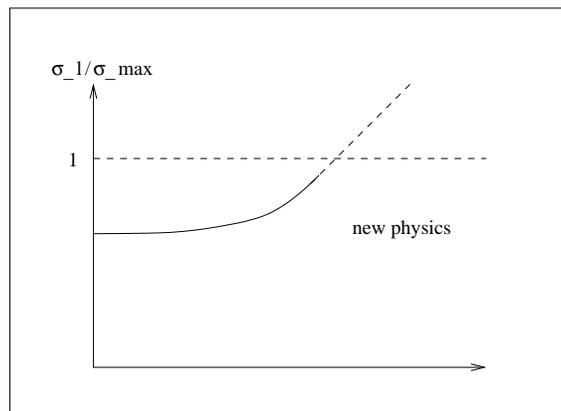


Figure 13.13: Projecting a scattering cross section above the energy where the unitarity bound is satisfied.

The unitarity bound is often invoked as a harbinger of new physics. Imagine that an experiment only measure a certain region of the energy spectrum, and in that region the cross section in some partial wave is observed to be rising. (Compare Fig. 13.13.). A linear projection the data might carry the cross section above the unitarity bound. Since this is not allowed, some new physics must appear, which will modify the scattering amplitude and preserve unitarity.

One kind of new physics is a resonance. To introduce the idea of a resonance in a perhaps too-unphysical way, suppose that the phase shift δ_l happens to increase with energy and crosses the value $\pi/2$ at an energy E_0 , as shown in Fig. 13.14. This behavior can be parameterized as

$$\tan \delta_l(E) \sim \frac{\frac{1}{2}\Gamma}{E_0 - E}. \quad (13.245)$$

Then the cross section can be re-expressed in terms of E_0 and Γ , as

$$\sigma_l = \frac{4\pi}{k^2}(2l+1) \frac{\tan^2 \delta_l}{1 + \tan^2 \delta_l} = \frac{\pi}{k^2}(2l+1) \frac{\Gamma^2}{(E_0 - E)^2 + \frac{\Gamma^2}{4}}. \quad (13.246)$$

The familiar Lorentzian line shape is characteristic of resonant behavior. (Often, the names “Breit-Wigner” are attached to this shape.) From the formula, we see the physics of resonance: the cross section becomes large because a state whose energy is close to the center of mass energy in the scattering experiment is easily excited.

Resonances play an important role in scattering experiments: they are a way to observe the spectrum of excitations in the target. It’s worth spending some time looking more closely

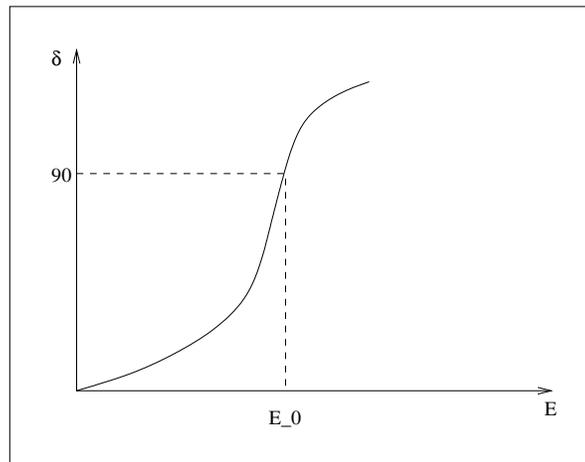


Figure 13.14: Resonant behavior in the scattering phase shift. The resonance is at energy E_0 .

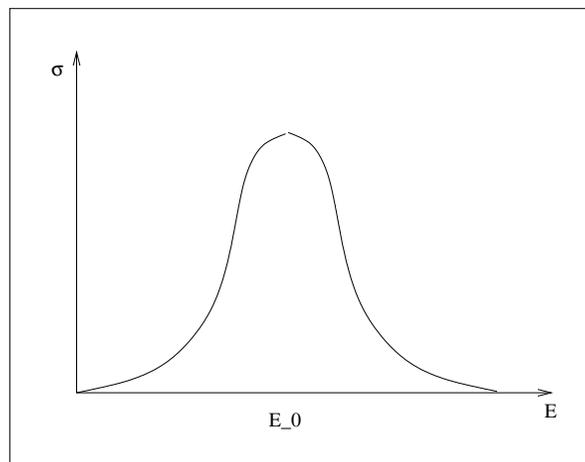


Figure 13.15: Energy dependence of a resonant cross section.

at what is going on. Let us continue the exposition by returning to a perturbative T-matrix perspective. The second order expression involves a sum over all intermediate states

$$\langle f|T|i\rangle = \sum_n \frac{\langle f|V|n\rangle \langle n|V|i\rangle}{E_i - E_n} \quad (13.247)$$

In this expression, E_i is the CM energy for the scattering process. If there is an intermediate state $|j\rangle$ with energy E_j , it will dominate the T-matrix when E_i is near E_j . Then the T-matrix can be approximated by a single term

$$\langle f|T|i\rangle = \frac{\langle f|V|j\rangle \langle j|V|i\rangle}{E_i - E_j}. \quad (13.248)$$

Now consider the first order transition probability per unit time for state $|j\rangle$ to decay into state $|f\rangle$ under the influence of the perturbation V :

$$\gamma_{jf} = \frac{2\pi}{\hbar} \rho(E_j) |\langle f|V|j\rangle|^2 \quad (13.249)$$

This has units of inverse time. The corresponding quantity with dimensions of energy is $\Gamma_{jf} = \hbar\gamma_{jf}$, and

$$\Gamma_{jf} = 2\pi\rho(E_j) |\langle f|V|j\rangle|^2. \quad (13.250)$$

The two body phase space factor is

$$\rho(E_j) = \frac{d^3p}{(2\pi\hbar)^3} \delta(E_j - E). \quad (13.251)$$

If the squared matrix element has no angular dependence then we can perform the angular integral and the phase space factor becomes

$$\rho(E_j) = \frac{4\pi\mu p}{(2\pi\hbar)^3} dE \delta(E_j - E) = \frac{\mu p}{2\pi^2\hbar^3}. \quad (13.252)$$

We can substitute the Γ 's for the matrix elements of V into the relevant places in Eq. 13.248, so

$$|\langle f|T|i\rangle|^2 = \frac{|\langle f|V|j\rangle|^2 |\langle j|V|i\rangle|^2}{(E_i - E_j)^2} = \frac{\Gamma_{jf}\Gamma_{ji}}{(E_i - E_j)^2 (2\pi\rho(E_j))^2}. \quad (13.253)$$

Now we return to the cross section for the reaction $i \rightarrow f$. It is proportional to the square of the T-matrix element, weighted by phase space factors. (Recall, everything is s-wave, still.)

$$\begin{aligned} \sigma &= \frac{2\pi}{\hbar} \rho(E) |\langle f|T|i\rangle|^2 \frac{1}{v_{rel}} \\ &= \frac{2\pi}{\hbar} \frac{\mu}{p} \rho(E) \frac{1}{(E - E_j)^2} \frac{\Gamma_{jf}\Gamma_{ji}}{(2\pi\rho(E_j))^2} \end{aligned} \quad (13.254)$$

When $E \sim E_j$, or when the c. m. energy is close to the resonance energy, the $\rho(E)$'s cancel and we have

$$\sigma = \frac{2\pi \mu}{\hbar} \frac{1}{p} \frac{1}{4\pi^2} \frac{1}{\rho(E)} \frac{\Gamma_{jf}\Gamma_{ji}}{(E - E_j)^2} \quad (13.255)$$

$$= \frac{\pi}{k^2} \frac{\Gamma_{jf}\Gamma_{ji}}{(E - E_j)^2} \quad (13.256)$$

This is resonant behavior. The answer is a bit too extreme, because the cross section diverges at $E = E_j$. This just is an unphysical consequence of using lowest order perturbation theory. We can fix this with the following simple argument: Suppose that the state $|j\rangle$ has a lifetime of $\frac{1}{\gamma}$. (We know it can decay; we have just computed its decay rate into the state $|f\rangle$ in Eq. 13.249.) Then

$$\langle j(t)|j(t)\rangle \sim e^{-\gamma t} \quad (13.257)$$

and so

$$\begin{aligned} |j(t)\rangle &= |j\rangle e^{-\gamma t/2} e^{iE_j t/\hbar} \\ &= |j\rangle e^{-i/\hbar(E_j - i\Gamma/2)t} \end{aligned} \quad (13.258)$$

We then recall how the second order T-matrix formula was derived. There was a factor of $\exp(-iE_j t/\hbar)$ which upon integration gave a factor of E_j in a denominator. We can just replace E_j by $E_j - \frac{i\Gamma}{2}$ in the denominator,

$$\frac{1}{E - E_j} \rightarrow \frac{1}{E - E_j + \frac{i\Gamma}{2}}, \quad (13.259)$$

and we arrive at our final form for the resonance cross section:

$$\sigma(i \rightarrow f) \rightarrow \frac{\pi}{k^2} \frac{\Gamma_{jf}\Gamma_{ji}}{(E - E_j)^2 + \frac{\Gamma^2}{4}}. \quad (13.260)$$

Although the derivation was done in the context of perturbation theory, it is true in general. The formula has many practical applications. The Γ 's have units of energy. Quantities such as Γ_{jf} are called the ‘‘partial decay width’’ for the resonance to decay into final state f . Γ with no subscript is the ‘‘total decay width’’ and of course

$$\Gamma = \sum_f \Gamma_{jf}. \quad (13.261)$$

The physical interpretation of a scattering process parameterized by Eq. 13.260 is that the initial state couples to the resonance with a strength proportional to Γ_{ji} and the final state

couples with a strength Γ_{jf} (or, equivalently, the intermediate state couples to the final state with the same factor). The range of energies over which the resonance contributes significantly to scattering, Γ , gives a measure of the lifetime of the state. Sometimes, your literature will convert Γ into an inverse lifetime; more often the quantity is quoted as an energy. The form of the expression “on resonance,” (at $E = E_j$) is also worth viewing:

$$\sigma = \frac{4\pi}{k^2} \frac{\Gamma_{ji}\Gamma_{jf}}{\Gamma^2} = \frac{4\pi}{k^2} \left(\frac{\Gamma_{ji}}{\Gamma} \right) \left(\frac{\Gamma_{jf}}{\Gamma} \right) \quad (13.262)$$

This suggests a useful experimental technique: while sitting on the resonance, one can measure “branching ratios” Γ_{jk}/Γ for the resonant state either from the relative rate of production of different final states at fixed initial state, or from the relative rates of production of the resonance with different initial states.

The whole derivation assumed s-wave scattering. If the resonance occurs in the l th partial wave, the expression just picks up an overall factor of $2l + 1$.

Many real-life experimental situations are more complicated. There can be multiple, overlapping resonances, in which case one must add the amplitudes for the different resonances. Often, scattering can proceed through the resonance and through non-resonant channels (“background”). In that case we parameterize the scattering amplitude as

$$kf_l(E) = \frac{-\frac{1}{2}\Gamma}{E - E_0 + \frac{i\Gamma}{2}} + R(E) \quad (13.263)$$

where $R(E)$ is the smooth background. The resonance will interfere with the background. The phase of the resonant scattering amplitude depends strongly on energy, while the background amplitude generally varies slowly. A typical interference one would see in a resonance would look like Fig. 13.16 or, cleaning the graph, Fig. 13.17. The background may be some known physics, and then the interference pattern will be a diagnostic for understanding the resonance. This was used at the time of the first experiments for Z^0 production in electron-positron annihilation. The Z^0 has a mass of about 91 GeV and its production $e^+e^- \rightarrow Z^0 \rightarrow \mu^+\mu^-$ interfered with the (known) QED process $e^+e^- \rightarrow \gamma \rightarrow \mu^+\mu^-$.

Unitarity arguments were quite important in the development of the Standard Model of the weak and electromagnetic interactions. Low energy cross sections for various weak processes were observed to rise as a positive power of the energy, with an extrapolated energy for the loss of unitarity estimated to be several hundred GeV. Something new had to occur. That new physics was, of course, the W and Z particles. The same arguments are made today, involving the scattering cross section for particular polarization states of W's, to infer

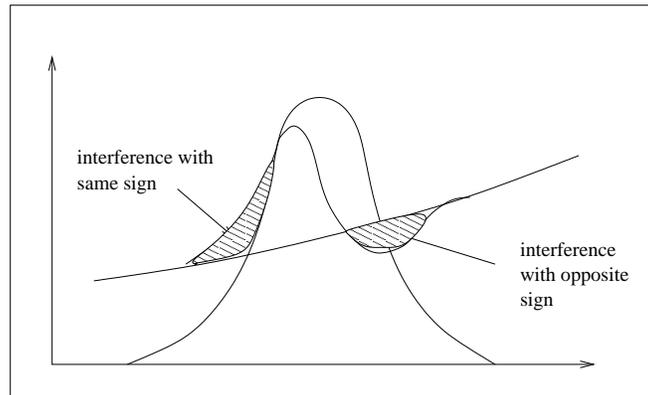


Figure 13.16: A resonance interfering with background, with regions of constructive and destructive interference labeled.

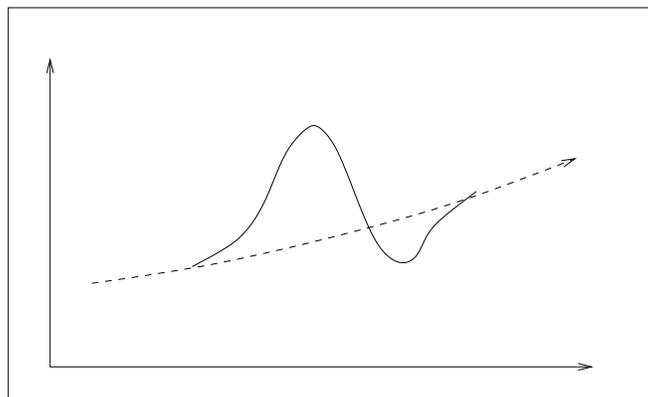


Figure 13.17: A resonance interfering with background.

the existence of some new physics beyond the Higgs boson within the energy reach of the Large Hadron Collider.

Scattering with identical particles

Consider a scattering process involving two identical spinless bosons. The wave function is symmetric under exchange of the particles' coordinates. In the absence of symmetry considerations, the wave function of the final state is a function of the relative spatial coordinate $\vec{x} = \vec{x}_1 - \vec{x}_2$. To generate a correct result, we must explicitly symmetrize the scattering amplitude,

$$\psi_s(\vec{x}) = e^{ikx} + e^{-ikx} + (f(\theta) + f(\pi - \theta)) \frac{e^{ikr}}{r}. \quad (13.264)$$

The differential cross section then becomes

$$\frac{d\sigma}{d\Omega} = |f(\theta) + f(\pi - \theta)|^2 \quad (13.265)$$

$$= |f(\theta)|^2 + |f(\pi - \theta)|^2 + 2\text{Re}[f^*(\theta)f(\pi - \theta)]. \quad (13.266)$$

Note that it is symmetric under the replacement $\theta \rightarrow \pi - \theta$.

Fermion wave functions are antisymmetric. Two spin- $\frac{1}{2}$ fermions can be in a relative spin state $S = 0$, state, in which the spin wave function is antisymmetric and the spatial wave function is symmetric, or in a $S = 1$ state with an antisymmetric space wave function. The scattering amplitudes in the two cases are then

$$S = 0 \rightarrow f(\theta) + f(\pi - \theta) \quad (13.267)$$

$$S = 1 \rightarrow f(\theta) - f(\pi - \theta) \quad (13.268)$$

where of course $f(\theta)$ is the unsymmetrized scattering amplitude.

Scattering experiments are often performed using an unpolarized beam and target. A mixture of the two scattering channels is observed. For spin- $\frac{1}{2}$ fermions the initial - spin - averaged and final - spin - summed differential cross section is

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{1}{4}|F_S|^2 + \frac{3}{4}|F_A|^2 \\ &= |f(\theta)|^2 + |f(\pi - \theta)|^2 - \text{Re}(f^*(\theta)f(\pi - \theta)). \end{aligned} \quad (13.269)$$

Symmetry has an immediate effect on the partial wave expansion

$$f(\theta) = \sum_l f_l P_l(\cos \theta). \quad (13.270)$$

Under parity $\hat{P}P_l(\cos\theta) \rightarrow (-1)^l P_l(\cos\theta)$. Therefore the space-symmetric scattering amplitude is a sum restricted to even partial waves,

$$F_S(\theta) = \sum_{l \text{ even}} f_l P_l(\cos\theta) \quad (13.271)$$

while the antisymmetric amplitude is restricted to the odd partial waves

$$F_A(\theta) = \sum_{l \text{ odd}} f_l P_l(\cos\theta). \quad (13.272)$$

As an interesting consequence, there is no s-wave scattering amplitude for a pair of fermions in a spin-triplet state, and no p-wave amplitude for the scattering of two identical spinless bosons.

Chapter 14

Classical waves – quantum mechanical particles

The quantum electromagnetic field

By now, you know very well that systems which have a classical description as particles have a quantum mechanical description which is wave-like: at the very least, their behavior is described by the Schrödinger equation, and we naturally expect to see wave phenomena such as interference in the normal state of events.

But what about systems whose classical limit is a system of waves, such as the electromagnetic field, vibrations of a crystalline solid, and the like? You colloquially know that these systems show particle-like behavior, in that their energies are quantized in multiples of a fundamental frequency, $E = n\hbar\omega_k$ with $n = 0, 1, 2, \dots$: For the electromagnetic field, this is a description in terms of photons. We have yet to explore what this particle-like behavior really means. It happens that it is ubiquitous: any classical system which is a set of coupled (or uncoupled) harmonic oscillators shows photon-like behavior. Let us describe this, now.

This will be a minimalist development of the quantum mechanical electromagnetic field. I will not worry about special relativity per se, or about setting up the formalism to deal with high orders of perturbation theory. The goal is to improve on the semi-classical description of radiation given in Chapter 12, to make it a fully quantum mechanical one.

We recall that the quantum mechanics of a single simple harmonic oscillator is described using raising and lowering operators

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} + \frac{i\hat{p}}{m\omega} \right) \quad \hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} - \frac{i\hat{p}}{m\omega} \right) \quad (14.1)$$

such that $[a, a^\dagger] = 1$ and

$$H = \frac{\hbar\omega}{2}(a^\dagger a + a a^\dagger). \quad (14.2)$$

Eigenstates of H are the number states $|n\rangle$ where

$$a^\dagger a |n\rangle = n |n\rangle; \quad a |n\rangle = \sqrt{n} |n-1\rangle; \quad a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle \quad (14.3)$$

and $n = 0, 1, 2, \dots$. Of course, $E_n = \hbar\omega(n + 1/2)$. Finally recall the Heisenberg equations of motion for the raising and lowering operators are equivalent to the classical ones for the complex amplitudes

$$i\hbar\dot{a} = [H, a] \rightarrow \dot{a} = -i\omega a \quad (14.4)$$

and similarly for a^\dagger .

So much for one oscillator. Suppose we have many uncoupled oscillators. The Hamiltonian is

$$H = \sum_i H_i = \sum_i \frac{p_i^2}{2m_i} + \frac{m_i \omega_i^2 x_i^2}{2}. \quad (14.5)$$

“Uncoupled” means that $[x_i, p_j] = 0$ if $i \neq j$. Clearly, we can solve this system by introducing separate creation and annihilation operators for each oscillator, just putting a subscript on the raising and lowering operators. The only nontrivial commutator will be $[a_i, a_j^\dagger] = \delta_{i,j}$. The Hamiltonian becomes

$$H = \sum_i \hbar \omega_i \left(a_i^\dagger a_i + \frac{1}{2} \right) \quad (14.6)$$

Our states are product states, with an integer for each index:

$$|\{n\}\rangle = |n_1, n_2, n_3 \dots\rangle \quad (14.7)$$

where

$$\begin{aligned} a_i |\{n\}\rangle &= \sqrt{n_i} |\dots, n_{i-1}, n_i - 1, n_{i+1} \dots\rangle \\ a_i^\dagger |\{n\}\rangle &= \sqrt{n_i + 1} |\dots, n_{i-1}, n_i + 1, n_{i+1} \dots\rangle \end{aligned} \quad (14.8)$$

The energy is a sum of terms, each of which has its own integer.

The problem of quantizing the electromagnetic field, or any linear wave system, can be solved if we can rewrite the classical system as a set of uncoupled harmonic oscillators. Before explicitly considering electrodynamics, let's begin by looking at a slightly simpler problem, a set of coupled oscillators. Suppose our classical Hamiltonian is a generalized quadratic function of the coordinates and momenta:

$$H = \sum_j \frac{p_j^2}{2} + \sum_{j,k} \frac{1}{2} x_j V_{jk} x_k \quad (14.9)$$

with $V_{jk} = V_{kj}$. (This is not the most general quadratic Hamiltonian we can consider, but it will suffice for our needs.) The equations of motion, which tie all the x 's and p 's together,

$$\begin{aligned} \dot{p}_j &= -V_{jk} x_k \\ \dot{x}_j &= p_j, \end{aligned} \quad (14.10)$$

can be uncoupled by defining linear combinations $X_j = R_{jl} x_l$ or $x_j = R_{jl}^{-1} X_l$. Then

$$\dot{p}_j = -V_{jk} R_{jl}^{-1} X_l \quad (14.11)$$

$$R_{nj}[p_n = R_{jl}^{-1}\dot{X}_l] \rightarrow \dot{X}_n = R_{nj}p_j \equiv P_n \quad (14.12)$$

so

$$\dot{P}_n = -R_{nj}V_{jk}R_{jl}^{-1}X_l \quad (14.13)$$

Now choose R so that $R_{nj}V_{jk}R_{jl}^{-1}$ is diagonal, equal to $-\omega_l^2\delta_{nl}$. The system decouples:

$$\begin{aligned} \dot{P}_j &= -\omega_j X_j \\ \dot{X}_j &= P_j \end{aligned} \quad (14.14)$$

and

$$H = \sum_j \left(\frac{P_j^2}{2} + \frac{\omega_j^2 X_j^2}{2} \right). \quad (14.15)$$

Our coupled system has been re-written as a set of uncoupled oscillators. The ω_n 's are the frequencies of the normal modes.

Now for the quantum mechanical case. We could begin with the original Hamiltonian Eq. 14.9 and replace the p_i 's and x_i 's by operators obeying the usual commutation relations. However, the Hamiltonian couples the coordinates, and hence the raising and lowering operators. It is better to jump immediately to the diagonal case, Eq. 14.15, and work with the P_n 's and X_n 's: define $A_n = X_n + iP_n$ and $A_n^\dagger = X_n - iP_n$. Operators associated with different normal modes commute among themselves. They obey the standard commutation relation of number operators, the states are given by Eq. 14.7 and the energies are integer-spaced multiples of $\hbar\omega_k$ for each normal mode k .

Finally, note that if we have a classical system with a set of dynamical variables v_k , and if we can write the classical equations of motion for a system as

$$\dot{v}_k = -i\omega v_k \quad \dot{v}_k^* = i\omega v_k^* \quad (14.16)$$

then it is a collection of uncoupled oscillators – and so is the quantum system. The classical v_k 's refer to the normal modes; their quantum mechanical operators act on states whose frequencies are the eigenfrequencies of the classical system. The R_{kl} matrices give us the translation dictionary between the normal mode states and the local degrees of freedom.

So, are the classical Maxwell equations oscillator-like? Recall that in Coulomb gauge ($\vec{\nabla} \cdot \vec{A} = 0$ and $\phi = 0$) the vector potential is a solution of the wave equation

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} = 0. \quad (14.17)$$

In a box of volume V , with appropriate boundary conditions, the wave numbers of A are quantized and we can write the vector potential as a superposition of eigenmodes

$$\vec{A}(\vec{x}, t) = \sum_k \sum_{\sigma=1,2} \left(\frac{2\pi\hbar c}{\omega V} \right)^{1/2} \vec{e} \left(a_{k\sigma}(t) e^{i(\vec{k}\cdot\vec{x})} + a_{k\sigma}^*(t) e^{-i(\vec{k}\cdot\vec{x})} \right) \quad (14.18)$$

The σ index labels the polarization, and if $\vec{\nabla} \cdot \vec{A} = 0$, then $\vec{e}_{k\sigma} \cdot \vec{k} = 0$. There are of course two such polarizations, which can be chosen rather arbitrarily. The factors $a_{k\sigma}(t)$ are the Fourier coefficients or amplitudes of the wave number \vec{k} part of the superposition. The pre-factor is our old one-photon normalization, chosen purely for convenience because the a 's could take on any classical value.

Now, do the classical Fourier coefficients have Eq. 14.16 as their equation of motion? If they do, we are done: the quantum electromagnetic field is a direct product of oscillator states; that is, it is composed of photons.

This is straightforwardly done, modulo an annoying factor of two. If we keep both a and a^* in the equation for A , we over-count the number of k modes by a factor of two. We can take care of this over-counting if we clip out half the k modes, perhaps by restricting $k_z > 0$ in the sum.

Inserting Eq. 14.18 into the wave equations gives us

$$\frac{d^2 a_{k\sigma}}{dt^2} + \omega_k^2 a_{k\sigma} = 0 \quad (14.19)$$

where $\omega_k = ck$. The solution is, of course,

$$a_{k\sigma}(t) = a_{k\sigma}^{(1)}(0) e^{-i\omega_k t} + a_{k\sigma}^{(2)}(0) e^{i\omega_k t}. \quad (14.20)$$

The annoying factor of two now returns, usefully. We can remove the restriction $k_z > 0$ if we define $a_{k\sigma} = a_{k\sigma}^{(1)}(0)$ if $k_z > 0$ and $a_{k\sigma}^{(2)}(0)$ if $k_z < 0$. This amounts to just keeping the solution

$$a_{k\sigma}(t) = a_{k\sigma}^{(0)} e^{-i\omega_k t} \quad (14.21)$$

or

$$\frac{da_{k\sigma}}{dt} = -i\omega_k a_{k\sigma}. \quad (14.22)$$

This gives us a set of classical variables with the desired first-order equation of motion. The energy in the field is also that of a set of oscillators:

$$U = \frac{1}{8\pi} \int (E^2 + B^2) = \sum_{k\sigma} \frac{\hbar\omega_k}{2} (a_{k\sigma}(t)^* a_{k\sigma}(t) + a_{k\sigma}(t) a_{k\sigma}^*(t)) \quad (14.23)$$

And there we are: the classical mechanical electromagnetic field is a collection of oscillators, one for each value of (k, σ) . We replace the Fourier modes by a set of number operators. The Hamiltonian becomes

$$H = \sum_{k\sigma} \frac{\hbar\omega_k}{2} (a_{k\sigma} a_{k\sigma}^\dagger + a_{k\sigma}^\dagger a_{k\sigma}). \quad (14.24)$$

The quantum vector potential is an operator (as are all dynamical variables). Looking at it again, we see that it is a superposition of raising operators and lowering operators:

$$\vec{A}(\vec{x}, t) = \sum_k \sum_{\sigma=1,2} \left(\frac{2\pi\hbar c}{\omega V} \right)^{1/2} \vec{e} \left(a_{k\sigma} e^{i\vec{k}\cdot\vec{x}} + a_{k\sigma}^\dagger e^{-i\vec{k}\cdot\vec{x}} \right). \quad (14.25)$$

It acts on our photon-counting (number basis) Hilbert space to increase or decrease $n_{k\sigma}$ – to increase or decrease the number of photons in a mode. We can make a suggestive change in nomenclature: $a_{k\sigma}$ is conventionally called an “annihilation operator” since it decreases the number of photons in a state; $a_{k\sigma}^\dagger$ is called a “creation operator.”

The Hamiltonian of the electromagnetic field,

$$H_{rad} = \sum_{k\sigma} \hbar\omega \left[a_{k\sigma}^\dagger a_{k\sigma} + \frac{1}{2} \right] \quad (14.26)$$

has an annoying feature – the second term. The energy of an oscillator in its ground state is $\hbar\omega/2$. Since there are an infinite number of modes of the electromagnetic field, this “zero point energy” is infinite. Fortunately, we do not have to think about it. We only care about energy differences in calculations, and so we could only imagine seeing the zero point energy in very unusual circumstances (quantum gravity, anyone?) or under changes in the boundary conditions, the Casimir effect. So we’ll ignore it henceforth and just write

$$H_{rad} = \sum_{k\sigma} \hbar\omega a_{k\sigma}^\dagger a_{k\sigma}. \quad (14.27)$$

Let’s use this formalism to reconsider emission and absorption of light by an atom, finally treating the electromagnetic field quantum mechanically. As in Chapter 12, we write

$$H = H_0 + H_{int} \quad (14.28)$$

where $H_0 = H_{atom} + H_{rad}$ is the zeroth order Hamiltonian, whose eigenstates we use in our perturbative expansion. The (first order) interaction term is

$$H_{int} = -e \frac{\vec{p} \cdot \vec{A}}{mc}. \quad (14.29)$$

For the case of emission, our initial and final states are

$$\begin{aligned} |i\rangle &= |a\rangle_{atom} |\dots, n_{k\sigma}, \dots\rangle_{rad} \\ |f\rangle &= |b\rangle_{atom} |\dots, n_{k\sigma} + 1, \dots\rangle_{rad} \end{aligned} \quad (14.30)$$

and we need to write the interaction term explicitly, in order to evaluate it,

$$H_{int} = -\frac{e}{mc} \sum_{k\sigma} \left(\frac{2\pi\hbar c}{\omega V} \right)^{1/2} \vec{p} \cdot \vec{\epsilon} [a_{k\sigma} e^{ikx} + a_{k\sigma}^\dagger e^{-ikx}]. \quad (14.31)$$

Its matrix element is

$$\langle f | H_{int} | i \rangle = -\frac{e}{mc} \sum_{k\sigma} \left(\frac{2\pi\hbar c}{\omega V} \right)^{1/2} \langle b | \vec{p} \cdot \vec{\epsilon} e^{-ikx} | a \rangle \langle n_{k\sigma} + 1 | a_{k\sigma}^\dagger | n_{k\sigma} \rangle. \quad (14.32)$$

Note that our choice of $|f\rangle$ and $|i\rangle$ means that $\langle f | a_{k\sigma} | i \rangle = 0$. Only the creation term contributes. The matrix element becomes

$$\langle f | H_{int} | i \rangle = -\frac{e}{mc} \sum_{k\sigma} \left(\frac{2\pi\hbar c}{\omega V} \right)^{1/2} \sqrt{n_{k\sigma} + 1} \langle b | \vec{p} \cdot \vec{\epsilon} e^{-ikx} | a \rangle. \quad (14.33)$$

This is formally a time-independent perturbation. Therefore, the Golden Rule's delta function constrains the energy difference between the initial and final states, $\delta(E_b + \hbar\omega_k - E_a)$. Here the $\hbar\omega_k$ comes from the energy of the state of the electromagnetic field. It compensates mathematically for the $e^{\pm i\omega t}$ of the semi-classical calculation. The Golden Rule, including one-body phase space for the states of the electromagnetic field, gives the differential decay rate

$$d\Gamma = \frac{2\pi}{\hbar} \left(\frac{2\pi\hbar c}{\omega V} \right) (n_{k\sigma} + 1) |\langle b | \vec{p} \cdot \vec{\epsilon} e^{-ikx} | a \rangle|^2 \delta(E_b + \hbar\omega_k - E_a) \times V \frac{d^3 p_\gamma}{(2\pi\hbar)^3}. \quad (14.34)$$

Note that apart from the $(n_{k\sigma} + 1)$, this is exactly the same result we had in semi-classical approximation. (This means that all of our discussion about selection rules and intensities goes through unchanged.) But now our calculation is completely quantum mechanically consistent. The operator \vec{A} has taken us from an $n_{k\sigma}$ photon state to an $n_{k\sigma} + 1$ photon state.

For absorption, the final state is

$$|f\rangle = |b\rangle_{atom} |\dots, n_{k\sigma} - 1, \dots\rangle_{rad} \quad (14.35)$$

and the Golden Rule gives us

$$d\Gamma = \frac{2\pi}{\hbar} \left(\frac{2\pi\hbar c}{\omega V} \right) n_{k\sigma} |\langle b | \vec{p} \cdot \vec{\epsilon} e^{-ikx} | a \rangle|^2 \delta(E_c - \hbar\omega_k - E_a). \quad (14.36)$$

The n or $n + 1$ factors are sensible. Recall that if we had had an external beam of intensity $I(\omega)\Delta\omega$ (energy per unit volume), then

$$|a_{ext}|^2 = \frac{2\pi c^2}{\omega^2} I(\omega)\Delta\omega. \quad (14.37)$$

If the beam is actually N photons of energy $\hbar\omega$, we'd expect

$$N \left(\frac{2\pi\hbar c}{\omega V} \right) = \frac{2\pi c^2}{\omega^2} I(\omega)\Delta\omega \quad (14.38)$$

or

$$N\hbar\omega = VI(\omega)\Delta\omega. \quad (14.39)$$

This seems consistent.

Note that the emission rate is proportional to $n_{k\sigma} + 1$. If there were no photons in the initial state, this would be unity. But this case is what we meant by “spontaneous emission” – the atom just decays. The $n_{k\sigma}$ part of the expression corresponds to what we meant by induced or stimulated emission. Just as in the case of absorption, it is proportional to the intensity of radiation already present in the to-be-occupied state. There is of course a more holistic way to regard our result: if $|i\rangle$ has $n_{k\sigma}$ photons already present, the decay rate for the atom is enhanced by a factor $n_{k\sigma} + 1$. The photonic state becomes even more populated. Hmm...light amplification by stimulated emission of radiation? Perhaps there are practical applications. Note that the effect has a rather trivial origin, from the algebra of the raising operator, $\langle n + 1 | a^\dagger | n \rangle = \sqrt{n + 1}$. The same thing will happen for all oscillator-analog systems.

By the way, do you recall the Einstein relations? We had N_a atoms in state a , N_b in state b , in a black body cavity,

$$\frac{N_a}{N_b} = e^{-\frac{E_b - E_a}{kT}} \quad (14.40)$$

In equilibrium, and with $E_b < E_a$)

$$\begin{aligned} \frac{dN_b}{dt} &= -N_b(\text{absorption}(b \rightarrow a)) + N_a(\text{emission}(a \rightarrow b)) \\ \frac{dN_b}{dt} &= -N_a(\text{emission}(a \rightarrow b)) + N_b(\text{absorption}(b \rightarrow a)) \end{aligned} \quad (14.41)$$

In equilibrium $dN_i/dt = 0$ for either i , meaning that $N_b/N_a = (\text{emission rate})/(\text{absorption rate}) = (n_{k\sigma} + 1)/n_{k\sigma}$. Thus

$$\frac{n_{k\sigma}}{n_{k\sigma} + 1} = e^{-\frac{E_b - E_a}{kT}} = e^{-\frac{\hbar\omega}{kT}} \quad (14.42)$$

or

$$n_{k\sigma} = \frac{1}{e^{\frac{\hbar\omega}{kT}} - 1} \quad (14.43)$$

which is the Planck formula, mode by mode. Phase space just counts the number of available modes per wave number interval.

Particles to waves to particles again

Our discussion about the free electromagnetic field gave us a language to talk about the creation and annihilation of photons. It also gave us a Hamiltonian, Eq. 14.27, which was additive in the photon number. Finally, it gave us a vector potential operator, Eq. 14.25. Notice its structure: each operator ($a_{k\sigma}$, for example) is multiplied by a spatially dependent function ($a_k(x) = \left(\frac{2\pi\hbar c}{\omega V}\right)^{1/2} \vec{\epsilon} e^{i(\vec{k}\cdot\vec{x})}$) which is a solution to the classical wave equation (or “field equation”) $(\nabla^2 + k^2)a_k(x) = 0$. Here ∇^2 is just a differential operator, with no connection to momentum or any quantum operator.

Let’s think about some analog situations in Nature. In particular, consider a collection of non-interacting particles. If they have a set of energy levels ϵ_n , then the energy of a many-particle system is

$$E = \sum_{j=1}^n \epsilon_j n_j \quad (14.44)$$

where n_j is the number of particles in energy state j . Note the resemblance to the photon formula.

Next, think more closely about conservation laws. Generally, quantities like charge are conserved, but the number of particles might not be conserved. For example, in a gas of electrons and positrons, the number of electrons or positrons is not conserved because they can annihilate. Only the difference in their number, the net charge, is conserved. Also, it is often useful to think about particle creation and annihilation in a more extended context: imagine the situation when an electron in an atom leaves the 1S state to go to the 2P state: $\Delta n_{1S} = -1$, $\Delta n_{2P} = +1$. The total particle number is conserved, the number of particles in any given state is not.

Finally, why are all electrons alike? We could have asked before, why are all photons alike? A crazy but consistent answer: they are all states created by $A(x, t)$, and there is only one operator $A(x, t)$.

The combination of all these ideas, that particles are field quanta which can be created and annihilated, is called Quantum Field Theory. It is a more general formulation of quantum dynamics – the quantum mechanics of a one or a few particles in a potential contains a subset of quantum field theoretic ideas – and all of (quantum) Nature is described at the most fundamental level we know about by one quantum field theory or another. In addition, many “effective field theories” are used to describe particular natural systems.

Let’s formulate quantum field theory for Schrödinger particles. The idea is to imagine thinking of the Schrödinger equation as a differential equation (∇^2 is just a differential operator, nothing more)

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(x)\psi. \quad (14.45)$$

Let $\psi_n(x)$ solve

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(x)\right] \psi_n(x) = E_n \psi_n(x) \quad (14.46)$$

and write the general solution to Eq. 14.45 as

$$\psi(x, t) = \sum_n b_n(t) \psi_n(x) \quad (14.47)$$

where

$$i\hbar \frac{\partial b_n}{\partial t} = E_n b_n \quad (14.48)$$

or

$$\frac{\partial b_n}{\partial t} = -i\omega_n b_n. \quad (14.49)$$

This is just a “classical” differential equation, but we appear to have two thirds (the wave equation, a time dependent b) of the photon game ingredients in hand. If we could write

$$H = \sum_n \epsilon_n b_n^* b_n \quad (14.50)$$

we could re-interpret the expansion coefficient as an operator in an abstract Hilbert space. So we need to find an H which can be used to give Eq. 14.49 as an equation of motion. A natural guess for H is

$$H = \int d^3x \psi^*(x, t) \left[-\frac{\hbar^2}{2m} \nabla^2 + V(x)\right] \psi(x, t) \quad (14.51)$$

which reduces to

$$H = \sum_n E_n b_n^* b_n, \quad (14.52)$$

just from orthogonality. Now boldly interpret b_n as an operator with commutation relations

$$[b_n, b_m^\dagger] = \delta_{nm} \quad [b_n, b_m] = 0 \quad [b_n^\dagger, b_m^\dagger] = 0 \quad (14.53)$$

(so that different levels n don't interact). The Heisenberg equation of motion is

$$i\hbar \frac{\partial b_n}{\partial t} = [b_n, H] = \sum_j [b_n, b_j^\dagger b_j] E_j = E_n b_n. \quad (14.54)$$

We have a theory of “field quanta of the Schrödinger equation.” Our Hilbert space is a number - operator - diagonal space

$$|\{n\}\rangle = |n_1, n_2, n_3, \dots\rangle \quad (14.55)$$

where the subscript labels levels of the Schrödinger differential operator. The quantity

$$\psi(x, t) = \sum_n \psi_n(x) b_n(t) \quad (14.56)$$

is an operator, a sum of a product of a c-number function of the spatial coordinates times an (annihilation) operator acting on the space of states. In contrast to photons, for Schrödinger particles, different operators create or annihilate particles:

$$\psi^\dagger(x, t) = \sum_n \psi_n^*(x) b_n^\dagger(t) \quad (14.57)$$

Note that the b 's act on states like harmonic oscillator operators: $b_n^\dagger b_n |n\rangle = n |n\rangle$. The quantity n is an integer $n = 0, 1, 2, \dots$. Also $b_n^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle$, $b_n |n\rangle = \sqrt{n} |n-1\rangle$. This follows immediately from the commutation relations. As we can have any number of particles in the state, we have just invented a quantum field theory for bosons.

What about fermions? We want to keep

$$H = \sum_n E_n b_n^\dagger b_n, \quad (14.58)$$

but we want to restrict n to be only zero or unity. It turns out that we can do that – and pick up the minus signs needed in many - fermion wave functions – if we replace the commutation

relation for the b 's by an anti-commutation relation

$$\begin{aligned}
 [b_n, b_m^\dagger]_+ &= b_n b_m^\dagger + b_m^\dagger b_n = \delta_{nm} \\
 [b_n, b_m]_+ &= b_n b_m + b_m b_n = 0 \\
 [b_n^\dagger, b_m^\dagger]_+ &= b_n^\dagger b_m^\dagger + b_m^\dagger b_n^\dagger = 0
 \end{aligned}
 \tag{14.59}$$

Then

$$\begin{aligned}
 i\hbar \frac{\partial b_n}{\partial t} &= \sum_m E_m [b_n, b_m^\dagger b_m] \\
 &= \sum_m E_m (b_n b_m^\dagger b_m - b_m^\dagger b_m b_n) \\
 &= \sum_m E_m ((\delta_{nm} - b_m^\dagger b_n) b_m - b_m^\dagger b_m b_n) \\
 &= \sum_m E_m (\delta_{nm} b_m + b_m^\dagger b_m b_n - b_m^\dagger b_m b_n) \\
 &= E_n b_n.
 \end{aligned}
 \tag{14.60}$$

The desired equation of motion is preserved.

Now we want eigenstates of $b_n^\dagger b_n$. Note that $(b_n^\dagger b_n)(b_n^\dagger b_n) = b_n^\dagger (1 - b_n^\dagger b_n) b_n = b_n^\dagger b_n - b_n^\dagger b_n^\dagger b_n b_n$. The second term is zero applied to any state because $b_n b_n = -b_n b_n$. So if $|\lambda\rangle$ is an eigenstate of $b_n^\dagger b_n$,

$$\begin{aligned}
 b_n^\dagger b_n |\lambda\rangle &= \lambda |\lambda\rangle \\
 (b_n^\dagger b_n)(b_n^\dagger b_n) |\lambda\rangle &= \lambda^2 |\lambda\rangle = \lambda |\lambda\rangle
 \end{aligned}
 \tag{14.61}$$

or $\lambda^2 = \lambda$, meaning that $\lambda = 0, 1$. Only zero or one particle can occupy a state.

Finally, we need matrix elements of b_n and b_n^\dagger . We have

$$\begin{aligned}
 b_n^\dagger b_n |N_n\rangle &= N_n |N_n\rangle; \quad N_n = 0, 1 \\
 (b_n^\dagger b_n) b_n^\dagger |N_n\rangle &= b_n^\dagger (1 - b_n^\dagger b_n) |N_n\rangle = (1 - N_n) |N_n\rangle
 \end{aligned}
 \tag{14.62}$$

so $b_n^\dagger |N_n\rangle = C_n |1 - N_n\rangle$. To find C_n , square this:

$$|C_n|^2 = \langle N_n | b_n b_n^\dagger | N_n \rangle = \langle N_n | (1 - b_n^\dagger b_n) | N_n \rangle = 1 - N_n
 \tag{14.63}$$

Thus $C_n = \theta_n \sqrt{1 - N_n}$ where θ_n is a phase factor. Similarly,

$$b_n |N_n\rangle = \theta_n \sqrt{N_n} |1 - N_n\rangle \quad (14.64)$$

For most applications, the value of the phase factor is irrelevant.

Although the formalism looks different, everything is the same as in ordinary quantum mechanics. Let's see how that works out. Note that

$$\psi(x, t) = \sum_n b_n(t) \psi_n(x) \quad (14.65)$$

is again an operator which annihilates particles in states n . Let's look at the (anti)commutation relations for ψ and ψ^\dagger :

$$[\psi(x, t)\psi^\dagger(x', t)]_\pm = \sum_n \sum_m \psi_n(x)\psi_m(x')[b_n, b_m^\dagger]_\pm = \delta^3(x - x') \quad (14.66)$$

from a combination of the commutation relations for the b 's and completeness for the ψ 's. Similarly, $[\psi(x, t)\psi(x', t)]_\pm = [\psi^\dagger(x, t)\psi^\dagger(x', t)]_\pm = 0$. You can show that when the Hamiltonian is

$$H = \int d^3x \psi^\dagger(x, t) \left[-\frac{\hbar^2}{2m} \nabla^2 + V \right] \psi(x, t), \quad (14.67)$$

the Heisenberg operator equation

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = [\psi(x, t), H] \quad (14.68)$$

gives the time-dependent Schrödinger equation for ψ . The number density operator is $n(x, t) = \psi^\dagger(x, t)\psi(x, t)$ and the total particle number is

$$N = \int d^3x n(x, t). \quad (14.69)$$

We can build states beginning with the "vacuum state" $|0\rangle$, which has no particles in it. It obeys the relation $\psi(x, t)|0\rangle = 0$. Following this line, $\psi^\dagger(x, t)|0\rangle$ should be a state with one particle at x . Is it? Look at the application of the number density operator

$$\begin{aligned} n(x', t)\psi^\dagger(x, t)|0\rangle &= \psi^\dagger(x', t)\psi(x', t)\psi^\dagger(x, t)|0\rangle \\ &= \psi^\dagger(x', t)(\delta^3(x' - x) \mp \psi^\dagger(x, t)\psi(x', t))|0\rangle \\ &= \delta^3(x' - x)\psi^\dagger(x', t)|0\rangle. \end{aligned} \quad (14.70)$$

The state is an eigenfunction of the number density operator with an eigenvalue which is a delta function. This is the mathematical equivalent of the statement that there is a particle located at x . Integrating, we find that the state is an eigenstate of the number operator,

$$N\psi^\dagger(x, t)|0\rangle = \int d^3x' n(x', t)\psi^\dagger(x, t)|0\rangle = \psi^\dagger(x, t)|0\rangle, \quad (14.71)$$

so the state does have one particle in it. Similarly

$$\psi^\dagger(x_1, t)\psi^\dagger(x_2, t)|0\rangle \quad (14.72)$$

is a state with two particles in it, one at x_1 , the other at x_2 .

Electrodynamics (yet) again

We return one last time to quantum electrodynamics, this time with number operators for particles in addition to those for the electromagnetic field. The interaction Hamiltonian is

$$H_I = \int d^3x \psi^\dagger(x, t) \left[-\frac{e\hbar}{imc} \vec{A} \cdot \vec{\nabla} \right] \psi(x, t) \quad (14.73)$$

and the zeroth-order Hamiltonian is a sum of a particle term

$$H_p = \int d^3x \psi^\dagger(x, t) \left[-\frac{\hbar^2}{2m} \nabla^2 + V \right] \psi(x, t) = \sum_n E_n b_n^\dagger b_n \quad (14.74)$$

and a radiation term

$$H_{rad} = \sum_{k\sigma} \hbar\omega_k a_{k\sigma}^\dagger a_{k\sigma}. \quad (14.75)$$

To actually perform calculations, we have to insert the explicit formulas for the field operators into H_I , arriving at

$$H_I = \sum_{nn'} \sum_{k\sigma} [b_n^\dagger b_{n'} a_{k\sigma} M(\vec{k}, \sigma, n, n') + b_n^\dagger b_n' a_{k\sigma}^\dagger M(-\vec{k}, \sigma, n, n')] \quad (14.76)$$

where

$$M(\vec{k}, \sigma, n, n') = -\frac{e\hbar}{imc} \sum_{k\sigma} \left(\frac{2\pi\hbar c}{\omega V} \right)^{1/2} \int d^3x \psi_n^*(x) [e^{ikx} \vec{\epsilon}_{k\sigma} \cdot \vec{\nabla}] \psi_{n'}(x) \quad (14.77)$$

is the complex amplitude we have seen so many times before. However, look more closely: Our initial and final states are

$$\begin{aligned} |i\rangle &= |\dots N_{n'}, \dots\rangle_{particle} |\dots n_{k'\sigma'} \dots\rangle_{rad} \\ |f\rangle &= |\dots N_n, \dots\rangle_{particle} |\dots n_{k\sigma} \dots\rangle_{rad} \end{aligned} \quad (14.78)$$

and H_I induces transitions among these states. Its first term destroys a photon of momentum k , destroys the particle in state n' , and creates a particle in state n . The second term again destroys n' and creates a particle in n along with a photon in k . The first term obviously represents absorption of the photon while the second one is for emission. Let's examine the matrix element for absorption:

$$\langle f|H_I|i\rangle = M(\vec{k}, \sigma, n, n')(\delta_{n_{k\sigma}, n_{k'\sigma'}-1} \sqrt{n_{k'\sigma'}}) \sqrt{N'_n} \sqrt{N_n \mp 1}, \quad (14.79)$$

again the same answer we've gotten many times before plus something new: extra Bose or Fermi counting factors associated with previous occupancy in our initial and final states. In particular, if our final state is already occupied, and we are a fermion, we can't go there. This is new physics. Apart from this, all the formalism of quantum field theory is overkill for this problem. But, in dealing with many particle states, and in cases where the analogs of the M 's can be pre-computed, it is a very convenient language. We need only think about the many-body part of the physics.

Light scattering

Most students who take a graduate quantum mechanics class also take a graduate electricity and magnetism class and learn a little bit about the scattering of light on matter. The story is entirely classical and starts with electric dipole scattering from a single scattering source. The story can proceed to considering many scatterers, or extended scatterers, simply summing or integrating over the locations of the individual scatterers. This allows one to understand a wide variety of phenomena: the blue sky, the red sunset, Bragg peaks in X-ray crystallography, critical opalescence, and much more. If you can find a copy, the old article "How light interacts with matter," by V. F. Weisskopf (Scientific American **219** 3,60, September 1968) is an evocative description of this physics.

The classical description of microscopic dynamics models the atom as an electron on a spring of natural frequency ω_0 with a damping force $F = -\Gamma d\vec{x}/dt$ in an external electric

field $\vec{E} \exp(-i\omega t)$:

$$m \frac{d^2 \vec{x}}{dt^2} + \Gamma \frac{d\vec{x}}{dt} + m\omega_0^2 \vec{x} = e\vec{\epsilon}_0 E_0 \exp(-i\omega t). \quad (14.80)$$

For a steady state solution $\vec{x}(t) = \vec{\epsilon}_0 x \exp(-i\omega t)$, so there is an induced time dependent dipole moment, which serves as an antenna. After a short passage of algebra we discover that the differential cross section for scattering from an initial polarization $\vec{\epsilon}_0$ into an outgoing polarization direction $\vec{\epsilon}$ is

$$\frac{d\sigma}{d\Omega} = \left(\frac{e^2}{mc^2}\right)^2 |\vec{\epsilon} \cdot \vec{\epsilon}_0|^2 \frac{\omega^4}{|\omega_0^2 - \omega^2 + i\frac{\omega\Gamma}{m}|^2}. \quad (14.81)$$

The prefactor

$$r_0 = \frac{e^2}{mc^2} \quad (14.82)$$

is the classical electron radius, about 10^{-13} cm, and it sets the scale for the cross section, far tinier than the physical size of the atom. Notice how Rayleigh scattering ($\sigma \propto \omega^4$) is recovered for $\omega \ll \omega_0$. At high frequency, the cross section becomes independent of frequency. This is called Thomson scattering. (It is the same formula as for scattering off a free electron.)

Of course, this is all classical, and it is a natural question to ask: how much of this classical result remains, with a fully quantum mechanical description? The answer: quite a lot! Let's do the calculation. (Note: this is a very old subject. Apparently Heisenberg and Kramers guessed the answer before the development of quantum mechanics. The description that follows is based on the one in Dirac's book.)

Light scattering is a second order process: the initial photon is annihilated, the final photon is created. We have to take the interaction Hamiltonian

$$H_I = -\frac{1}{m} \left(\frac{e}{c} \vec{A} \cdot \vec{p}\right) + \frac{e^2}{2mc^2} \vec{A} \cdot \vec{A} \quad (14.83)$$

and work to second order in the first term and first order in the second term.

Label the incoming photon's momentum as \vec{q}_1 and the outgoing photon's momentum as \vec{q}_2 (with energies $cq_1 = \hbar\omega_1$ and $cq_2 = \hbar\omega_2$); give the incident particle a momentum \vec{p}_1 and energy E_1 ; the final state particle has momentum \vec{p}_2 and energy E_2 .

Focus on the photon operators to start the calculation. All the terms in the T matrix operator will have to be sandwiched between initial and final states

$$|i_{photon}\rangle = a_{q_1\sigma_1}^\dagger |0\rangle$$

$$|f_{\text{photon}}\rangle = a_{q_2\sigma_2}^\dagger |0\rangle \quad (14.84)$$

and so the A^2 photon operator piece of T will be (temporarily suppressing lots of irrelevant factors)

$$\sum_{k_b, \sigma_b} \sum_{k_a, \sigma_a} \langle 0 | a_{q_2\sigma_2} (a_{k_a\sigma_a} + a_{k_a\sigma_a}^\dagger) (a_{k_b\sigma_b} + a_{k_b\sigma_b}^\dagger) a_{q_1\sigma_1} | 0 \rangle. \quad (14.85)$$

For a nonzero matrix element, there are two possibilities: $k_a\sigma_a = k_1\sigma_1$ and $k_b\sigma_b = k_2\sigma_2$, or $k_a\sigma_a = k_2\sigma_2$ and $k_b\sigma_b = k_1\sigma_1$. The A^2 term gives

$$\langle f | H_I | i \rangle = 2 \left[\frac{e^2}{2mc^2} \right] \left[\frac{2\pi\hbar c^2}{V\sqrt{\omega_1\omega_2}} \right] \delta_{if} \int \frac{d^3x}{V} \exp(i(p_1 - p_2)x/\hbar) \exp(i(q_1 - q_2)x/\hbar) \quad (14.86)$$

I have written the electron's (or atom's) wave function as a product of a relative piece (with state labeled as $|i\rangle$ or $|f\rangle$) and plane waves for the center of mass motion. These factors get stripped off (or integrated over) to give the phase space integral for the cross section. The derivation is identical to ones we have done before (for the photoelectric effect, for example). The result is that the Golden Rule formula for the differential cross section is

$$\begin{aligned} d\sigma &= \frac{1}{v_{\text{rel}}} \frac{d^3q_2}{(2\pi\hbar)^3} |T|^2 \left[\frac{d^3p_2}{(2\pi\hbar)^3} (2\pi\hbar)^3 \delta^3(\vec{q}_1 + \vec{p}_1 - \vec{q}_2 - \vec{p}_2) \right] \\ &\quad \times \frac{2\pi}{\hbar} \delta(cq_1 + E_1 - cq_2 - E_2). \end{aligned} \quad (14.87)$$

With $v_{\text{rel}} = c$ and $d^3q_2 = d\Omega q_2^2 dq_2$ a few integrals over delta functions give us

$$\frac{d\sigma}{d\Omega} = \frac{|T|^2}{(2\pi\hbar c^2)^2} \omega_2^2. \quad (14.88)$$

and it's obviously useful to define

$$T = \frac{2\pi\hbar c^2}{\sqrt{\omega_1\omega_2}} \bar{T} \quad (14.89)$$

for an even cleaner expression,

$$\frac{d\sigma}{d\Omega} = |\bar{T}|^2 \frac{\omega_2}{\omega_1}. \quad (14.90)$$

The constants multiplying \bar{T} are obviously the normalization factors for the two photon fields, which cancel against terms in the phase space formula.

Keeping only the A^2 term, and assuming that we simply have free electrons for our initial and final matter states, gives the Thomson cross section

$$\frac{d\sigma}{d\Omega} = \left(\frac{e^2}{mc^2} \right)^2 |\vec{\epsilon}_1 \cdot \vec{\epsilon}_2|^2. \quad (14.91)$$

Let's write the full expression for \bar{T} and make some comments before evaluating it. We have to work to second order in the $\vec{A} \cdot \vec{p}$ term. There are two possible time orderings: first, the initial photon could be absorbed before the final photon was emitted. The intermediate state consists only of matter states $|n\rangle$. The energy denominator is $\Delta E = \hbar\omega_1 + E_i - E_n = \hbar\omega_1 - \hbar\omega_{ni}$. But there is a second possibility: the atom could emit the final photon before absorbing the initial one. The intermediate state consists of the matter state $|n\rangle$ plus the two photons. The energy denominator is $\Delta E = \hbar\omega_1 + E_i - (E_n + \hbar\omega_1 + \hbar\omega_2) = -\hbar\omega_2 - \hbar\omega_{ni}$. We have

$$\bar{T} = \frac{e^2}{mc^2} [\vec{\epsilon}_2 \cdot \vec{\epsilon}_1 \delta_{fi} + \frac{1}{m\hbar} \sum_n (\frac{\langle f|\vec{p} \cdot \vec{\epsilon}_2|n\rangle \langle n|\vec{p} \cdot \vec{\epsilon}_1|i\rangle}{\omega_1 - \omega_{ni}} - \frac{\langle f|\vec{p} \cdot \vec{\epsilon}_1|n\rangle \langle n|\vec{p} \cdot \vec{\epsilon}_2|i\rangle}{\omega_2 + \omega_{ni}})]. \quad (14.92)$$

There is a lot of physics here! Notice that while the A^2 term does not change the state of the atom, the other two terms can. The final state $|f\rangle$ does not have to equal the state $|i\rangle$. Of course, energy is conserved overall. The physical process is called "fluorescence" (if $\omega_2 < \omega_1$) or "Raman scattering" with "Stokes" or "anti-Stokes" transitions, when the matter state gains or loses energy.

Back to Thompson scattering for a second. The two new terms do contribute to the differential cross section, but they are momentum dependent and for an initially static electron and an incident photon energy much less than the electron mass, they are sub-dominant. Only in that case is $\omega_2 = \omega_1$: the correct kinematics produces the familiar Compton energy or wavelength shift (or to be precise, the non relativistic form of that formula, since we are treating the electron nonrelativistically). The fully relativistic differential cross section is called the "Klein-Nishina" formula. If you take a quantum field theory class, you'll probably have to evaluate it at some point.

Back to atomic scattering. Eq. 14.92 is a bit awkward to use, but we can fix that with the usual substitution $\langle f|\vec{p}|i\rangle = im\omega_{fi} \langle f|\vec{x}|i\rangle$ where $\hbar\omega_{fi} = E_f - E_i$. But now, can we see Rayleigh scattering ($\sigma \propto \omega^4$) in some variation of Eq. 14.92? Amazingly, yes. Let's pull the polarization factors out and write the dot products as sums on repeated indices. Then

$$\bar{T} = \frac{e^2}{mc^2} \epsilon_2^l \epsilon_1^m [\delta_{lm} \delta_{fi} + \frac{1}{m\hbar} \sum_n \omega_{fn} \omega_{ni} (\frac{\langle f|x_l|n\rangle \langle n|x_m|i\rangle}{\omega_1 - \omega_{ni}} - \frac{\langle f|x_m|n\rangle \langle n|x_l|i\rangle}{\omega_2 + \omega_{ni}})]. \quad (14.93)$$

We can make the first term look like the other two by using $x_l p_m - p_l x_m = i\hbar \delta_{lm}$, so that

$$\delta_{lm} \delta_{fi} = \frac{1}{i\hbar} \sum_n (\langle f|x_l|n\rangle \langle n|p_m|i\rangle - \langle f|p_l|n\rangle \langle n|x_m|i\rangle)$$

$$= \frac{m}{\hbar} \sum_n (\langle f|x_l|n\rangle \langle n|x_m|i\rangle \omega_{ni} - \langle f|x_m|n\rangle \langle n|x_l|i\rangle \omega_{fn}) \quad (14.94)$$

This allows us to write

$$\bar{T} = \frac{e^2}{mc^2} \epsilon_2^l \epsilon_1^m \frac{m}{\hbar} \sum_n (\langle f|x_l|n\rangle \langle n|x_m|i\rangle (\omega_{ni} - \frac{\omega_{ni}\omega_{fn}}{\omega_1 - \omega_{ni}}) - \langle f|x_m|n\rangle \langle n|x_l|i\rangle (\omega_{fn} - \frac{\omega_{ni}\omega_{fn}}{\omega_2 + \omega_{ni}})). \quad (14.95)$$

Dirac adds in $0 = \omega_2 \sum_n (\langle f|x_l|n\rangle \langle n|x_m|i\rangle - \langle f|x_m|n\rangle \langle n|x_l|i\rangle)$ to give

$$\begin{aligned} \bar{T} &= \frac{e^2}{mc^2} \epsilon_2^l \epsilon_1^m \frac{m}{\hbar} \sum_n (\langle f|x_l|n\rangle \langle n|x_m|i\rangle (\omega_{ni} - \frac{\omega_{ni}\omega_{fn}}{\omega_1 - \omega_{ni}} + \omega_2) \\ &\quad - \langle f|x_m|n\rangle \langle n|x_l|i\rangle (\omega_{fn} - \frac{\omega_{ni}\omega_{fn}}{\omega_2 + \omega_{ni}} + \omega_2)). \end{aligned} \quad (14.96)$$

The energy conservation relation $\omega_1 - \omega_{ni} - \omega_{fn} = \omega_2$ collapses both terms in the sum to

$$\bar{T} = \frac{e^2}{mc^2} \epsilon_2^l \epsilon_1^m \frac{m}{\hbar} \omega_1 \omega_2 \sum_n (\frac{\langle f|x_l|n\rangle \langle n|x_m|i\rangle}{\omega_1 - \omega_{ni}} - \frac{\langle f|x_m|n\rangle \langle n|x_l|i\rangle}{\omega_2 + \omega_{ni}}). \quad (14.97)$$

so that

$$\frac{d\sigma}{d\Omega} = (\frac{e^2}{mc^2})^2 \omega_1 \omega_2^3 |\frac{m}{\hbar} \epsilon_2^l \epsilon_1^m \sum_n (\frac{\langle f|x_l|n\rangle \langle n|x_m|i\rangle}{\omega_1 - \omega_{ni}} - \frac{\langle f|x_m|n\rangle \langle n|x_l|i\rangle}{\omega_2 + \omega_{ni}})|^2. \quad (14.98)$$

This exposes the $\sigma \sim \omega^4$ Rayleigh form for elastic scattering, when $\omega_1 = \omega_2$. When all the ω_{ni} terms are much greater than ω_1 , the terms with the energy denominators are suppressed compared to the contact term and the Thomson formula re-emerges. On resonance, where a denominator vanishes, the formula breaks down. We know what to do in that case: somehow introduce the decay width into the denominator, as we did when studying resonant scattering.

The free electron gas at zero temperature

Let's use the language of quantum field theory to calculate some of the properties of a gas of non-interacting electrons at zero temperature. The ground state wave function has one filled state for every value of momentum up to the Fermi momentum p_F , and then all states are empty. This makes the momentum space properties of the Fermi gas pretty simple. However, the coordinate space properties are nontrivial, and those are our goal.

We should spend a few moments trying to make all the factors correct. The ground state wave function $|\Phi\rangle$ must be such that

$$\langle\Phi|a_{ps}^\dagger a_{ps}|\Phi\rangle = 1 \quad |p| < p_F \quad (14.99)$$

and zero otherwise, where

$$[a_{ps}, a_{p's'}^\dagger]_+ = \delta^3(p - p')\delta_{ss'}. \quad (14.100)$$

We will have a lot of momentum integrals to do, and it is convenient to simplify notation. Put the system in a box of volume V . Then the momentum spectrum is discrete and in fact there are V distinct p 's in all. Replacing the Dirac delta function by a Kronecker delta,

$$[a_{ps}, a_{p's'}^\dagger]_+ = \delta(p, p')\delta_{ss'}, \quad (14.101)$$

a properly normalized field variable is

$$\psi(r) = \frac{1}{\sqrt{V}} \sum_p e^{ipr} a_p. \quad (14.102)$$

To check (suppressing the spin label):

$$\begin{aligned} [\psi(r), \psi^\dagger(r')]_+ &= \frac{1}{V} \sum_{pp'} e^{ipr} e^{ip'r'} [a_p, a_{p'}]_+ \\ &= \frac{1}{V} \sum_p e^{ip(r-r')} \\ &= \frac{V}{V} \delta_{rr'} \end{aligned} \quad (14.103)$$

The number density operator is $n(x) = \psi^\dagger(x)\psi(x)$ and its expectation value is

$$\begin{aligned} \langle\Phi|\psi^\dagger(x)\psi(x)|\phi\rangle &= \frac{1}{V} \sum_{pp'} \langle\Phi|a_p^\dagger a_p|\Phi\rangle \\ &= \frac{1}{V} \sum_p n_p \\ &= \frac{N}{V}. \end{aligned} \quad (14.104)$$

The gas has uniform density – no surprise. (Incidentally, to make contact with the usual statistical mechanics story,

$$\sum_p = V \int \frac{d^3p}{(2\pi)^3} \quad (14.105)$$

per spin state.) The Fermi momentum is defined via the number density,

$$\frac{N}{V} = \sum_s \int_0^{p_F} \frac{4\pi p^2 dp}{(2\pi)^3} = \frac{p_F^3}{3\pi^2} \quad (14.106)$$

With our confidence high we move to more interesting observables. Consider the Green's function, a generalization of Eq. 14.104,

$$G_s(x - x') = \langle \Phi | \psi_s^\dagger(x) \psi_s(x') | \Phi \rangle \quad (14.107)$$

It is the amplitude to remove a particle at location x' and put it back in at location x . From the calculation of the density, it is

$$\begin{aligned} \langle \Phi | \psi_s^\dagger(x) \psi_s(x) | \Phi \rangle &= \frac{1}{V} \sum_{ps} n_{ps} p e^{ip(r-r')} \\ &= \frac{N \sum_p n_{ps} e^{ip(r-r')}}{\sum_p n_{ps}} \end{aligned} \quad (14.108)$$

The last step is done to pull in an overall factor of the density, $n = N/V$. Our technical problem is the integral

$$I = \int_0^{p_F} d^3 p e^{i\vec{p} \cdot (\vec{r} - \vec{r}')} \quad (14.109)$$

We do this in spherical coordinates, picking the \hat{z} axis along the $\vec{x} - \vec{x}'$ direction and calling $|\vec{x} - \vec{x}'| = R$. This gives

$$\begin{aligned} I &= 2\pi \int_0^{p_F} p^2 dp d\cos\theta e^{ipR\cos\theta} \\ &= 2\pi \int_0^{p_F} p^2 dp \frac{\sin pR}{pR} \\ &= 2\pi \int_0^{p_F} p^2 dp j_0(pR) \end{aligned} \quad (14.110)$$

where $j_0(x)$ is the spherical Bessel function. Dropping this result into Eq. 14.108 and grooming it a bit gives

$$G(r - r') = \frac{3}{2} n \frac{j_1(p_F R)}{p_F R} \quad (14.111)$$

and the limiting value of this expression at $R = 0$ is just $n/2$, as we expect. (Half the particles have one of the spins.)

A more interesting question: what is the probability to find a particle at location x' , given that there is one at location x ? One way to answer this question is to remove a particle of spin s at x , leaving behind $N - 1$ particles, and then measure the density of particles of spin s' in the new state, $|\Phi'(r, s)\rangle$:

$$\begin{aligned} \langle \Phi'(r, s) | \psi_{s'}^\dagger(r') \psi_{s'}(r') \Phi'(r, s) \rangle &= \langle \Phi | \psi_s^\dagger(r) \psi_s^\dagger(r') \psi_{s'}(r') \psi_s(r) \Phi \rangle \\ &\equiv \left(\frac{n}{2}\right)^2 g_{ss'}(r - r') \\ &= \sum_{pp'qq'} e^{-i(p-p')r} e^{-i(q-q')r'} \langle \Phi | a_{ps}^\dagger a_{qs'}^\dagger a_{q's'} a_{p's} | \Phi \rangle \end{aligned} \quad (14.112)$$

There are two cases. First, suppose $s \neq s'$. Then it must be that $p = p'$ and $q = q'$ to get a non-vanishing contraction of the a 's and a^\dagger 's. Then

$$\begin{aligned} \left(\frac{n}{2}\right)^2 g_{ss'}(r - r') &= \sum_{pq} \langle \Phi | n_{ps} n_{qs'} | \Phi \rangle \\ &= n_s n_{s'}. \end{aligned} \quad (14.113)$$

This means $g_{ss'}(r - r') = 1$. The different spins do not know about each other.

If $s = s'$ then either $p = p'$, $q = q'$ or $p = q'$, $q = p'$. The first term will contain a factor

$$\langle \Phi | a_{ps}^\dagger a_{qs}^\dagger a_{qs} a_{ps} | \Phi \rangle = \langle \Phi | n_{ps} n_{qs} | \Phi \rangle \quad (14.114)$$

after anti-commuting the operators through. The second term has a factor

$$\langle \Phi | a_{ps}^\dagger a_{qs}^\dagger a_{ps} a_{qs} | \Phi \rangle = - \langle \Phi | n_{ps} n_{qs} | \Phi \rangle \quad (14.115)$$

Putting the pieces together,

$$\begin{aligned} C(r, r') &= \frac{1}{V^2} \sum_{pq} n_{ps} n_{qs} (1 - e^{i(p-q)(r-r')}) \\ &= n^2 - \left| \frac{1}{V} \sum_p n_{ps} e^{i(r-r')} \right|^2 \\ &= n^2 - n^2 g(r - r')^2 \end{aligned} \quad (14.116)$$

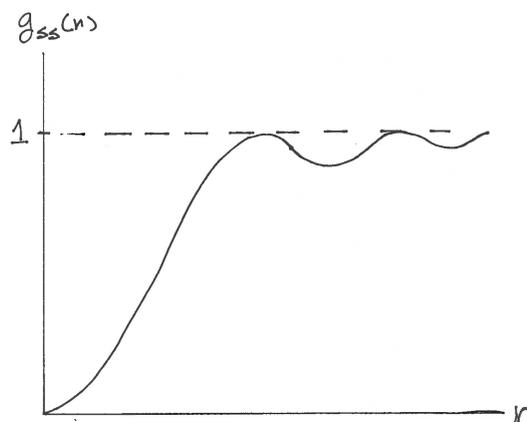


Figure 14.1: pair correlation function for the non-interacting Fermi gas.

recalling the expression for the Green's function. We can call $C(r, r') = n^2 \mathcal{G}(r - r')$ where $\mathcal{G}(r)$ parametrizes the interesting physics. From the explicit functional form of $g(r)$,

$$\mathcal{G}(r) = 1 - \left(\frac{3}{p_F r} j_1(p_F r) \right)^2. \quad (14.117)$$

Note that $\mathcal{G}(r)$ vanishes at $r = 0$. This is a consequence of Fermi statistics: if there was a particle already at a location r , there will not be a second one there, too.

Particles that interact among themselves

So far we have discussed “free particles” – particles in a potential $V(x)$, but otherwise not interacting with each other. How can we introduce two-body interactions? This is easy –

add to our one-body Hamiltonian

$$H_1 = \int d^3x \psi^\dagger(x, t) \left[-\frac{\hbar^2}{2m} \nabla^2 + V(x) \right] \psi(x, t) \quad (14.118)$$

a two-body term

$$H_2 = \int d^3x d^3x' \psi^\dagger(x, t) \psi^\dagger(x', t) v(x, x') \psi(x, t) \psi(x', t). \quad (14.119)$$

Note that we have chosen the same ordering of operators as we used for the pair correlation function. Here $v(x)$ is the potential between the particles. Now suppose that $v(x, x')$ is, in some sense, small. If it were zero, it would be natural to expand in plane wave states,

$$\psi(x, t) = \frac{1}{\sqrt{V}} \sum_k b_k(t) e^{ikx}. \quad (14.120)$$

If $v(x, x')$ is actually a function of the relative separation of the particles, $v(x - x')$, then we can write the Hamiltonian in a plane wave basis as

$$\begin{aligned} H &= \sum_{k_1 k_2} b_{k_1}^\dagger b_{k_2} \int \frac{d^3x}{V} e^{-ik_1 x} e^{-ik_2 x} \frac{\hbar^2 k^2}{2m} \\ &+ \sum_{k_1} \sum_{k_2} \sum_{k_3} \sum_{k_4} b_{k_1}^\dagger b_{k_2}^\dagger b_{k_3} b_{k_4} \int \frac{d^3x_1}{V} \frac{d^3x_2}{V} e^{i(k_3 - k_1)x_1} e^{i(k_4 - k_2)x_2} v(x_1 - x_2). \end{aligned} \quad (14.121)$$

The first term is

$$\sum_k \frac{\hbar^2 k^2}{2m} b_{k_1}^\dagger b_{k_1} \quad (14.122)$$

(from the delta function in the integral) and the second term can also be condensed: define $r = x' - x$, change variables from x, x' to x, r , and use

$$\int \frac{d^3x_1}{V} \frac{d^3x_2}{V} e^{i(k_3 - k_1)r} v(r) e^{i(k_4 - k_2 + k_3 - k_1)x} = \delta^3(k_4 - k_2 + k_3 - k_1) v(q) \quad (14.123)$$

where $\vec{q} = \vec{k}_3 - \vec{k}_1$ and

$$v(q) = \int \frac{d^3x}{V} e^{iqr} v(r). \quad (14.124)$$

Then the two-body Hamiltonian is

$$H_2 = \sum_{k_1} \sum_{k_2} \sum_q v(q) b_{k_1+q}^\dagger b_{k_2-q}^\dagger b_{k_2} b_{k_1}, \quad (14.125)$$

that is, the interaction scatters particles of wave number k_1 and k_2 into wave number $k_1 + q$ and $k_2 - q$ (Or, it annihilates initial state particles and creates final state ones.) We can compute the lowest order T-matrix by specifying initial and final states,

$$\begin{aligned} |i\rangle &= b_{p_1}^\dagger b_{p_2}^\dagger |0\rangle \\ |f\rangle &= b_{p_3}^\dagger b_{p_4}^\dagger |0\rangle \end{aligned} \tag{14.126}$$

and then the lowest order T-matrix element is

$$\langle f | H_2 | i \rangle = v(q) \delta^3(p_3 + p_4 - p_1 - p_2) \tag{14.127}$$

No surprise again, but we can do so much more.

For an off-beat example of the use of this formalism, consider a weakly interacting collection of bosons near absolute zero.

Excitation in a Bose-Einstein condensate

Recall your statistical mechanics for bosons: the number of particles in a state \vec{k} is given by

$$N(k) = \frac{1}{C e^{E(k)/T} - 1} \tag{14.128}$$

where T is the temperature (in energy units, $k = 1$) and C is related to the fugacity or chemical potential. In a normal system, C and the particle number are related:

$$\frac{N}{V} = \int \frac{d^3k}{(2\pi)^3} \frac{1}{C e^{E(k)/T} - 1}. \tag{14.129}$$

However, as T falls, there is no C which can solve the equation. You have to split off the zero energy state,

$$N = \frac{1}{C - 1} + V \int \frac{d^3k}{(2\pi)^3} \frac{1}{C e^{E(k)/T} - 1} \tag{14.130}$$

and it contributes a finite amount to the right hand side. The mathematics is telling us that the ($\vec{k} = 0$) ground state is macroscopically occupied – there is a condensate. This is in contrast to the microscopic occupation of each remaining phase space differential volume element d^3k .

Let's think about the condensed system in the language of second quantization. The ground state has order N particles in it,

$$b_0^\dagger b_0 |\psi\rangle = N_0 |\psi\rangle. \quad (14.131)$$

In addition, there is some occupation at $k \neq 0$

$$b_k^\dagger b_k |\psi\rangle = N(k) \quad (14.132)$$

given by Eq. 14.128 above.

To make life interesting, let's assume that we do not have an ideal Bose gas, but suppose that there is some interaction among the bosons. We can describe it by some potential $v(r)$. In fact, let's assume that $v(r) = v\delta^3(r)$ so that $v(q) = v$ is a constant. (This is old-fashioned language. Looking back at Eq. 14.127, we see that the T-matrix is a constant, and recalling the chapter about scattering, we are replacing the T-matrix by its scattering length approximation.)

Now for approximations. The ground state operators do not commute, $[b_0, b_0^\dagger] = 1$, but $b_0^\dagger b_0 |\psi\rangle = N_0 |\psi\rangle$ where $N_0 \gg 1$. In this sense, b_0 and b_0^\dagger "almost" commute. Let's treat them as classical objects, whose size is about \sqrt{N} . Then, in H_2 , there is a natural hierarchy of terms:

1. $b_0^\dagger b_0^\dagger b_0 b_0$, the scattering of condensate particles, has a size roughly $vb_0^4 \sim vN^2$
2. Terms like $b_k^\dagger b_0^\dagger b_0 b_0$ vanish – they do not conserve momentum.
3. Order N terms: $b_k^\dagger b_{-k}^\dagger b_0 b_0 + b_0^\dagger b_0^\dagger b_k b_{-k}$ for two particles either leaving or entering the condensate, and $2^2 b_k^\dagger b_k b_0^\dagger b_0$ which controls the scattering of a normal particle off the condensate
4. Ordinary particle scattering is an order (1) effect.

The Hamiltonian of item (3) is

$$H_I^{(3)} = b_0^2 \sum_{k \neq 0} (b_k^\dagger b_{-k}^\dagger + b_k b_{-k} + 4b_k^\dagger b_k). \quad (14.133)$$

We must be slightly careful with item (1): the total number of particles is

$$N = b_0^2 + \sum_{k \neq 0} b_k^\dagger b_k \quad (14.134)$$

and so

$$b_0^4 = N^2 - 2N \sum_{k \neq 0} b_k^\dagger b_k. \quad (14.135)$$

Then, our approximate Hamiltonian, which considers scattering into or out of the condensate, is

$$H = \sum_k \frac{\hbar^2 k^2}{2m} b_k^\dagger b_k + N^2 v + Nv \sum_{k \neq 0} (b_k^\dagger b_{-k}^\dagger + b_k b_{-k} + (4-2)b_k^\dagger b_k). \quad (14.136)$$

Now notice something important: this H is quadratic in the b 's. We can make a change of variables, to a new basis, a_k^\dagger and a_k , and write

$$H = N^2 v + \sum_k \epsilon(k) a_k^\dagger a_k; \quad (14.137)$$

that is, the system is described by a set of excitations of “quasi-particles” annihilated by a_k . These are the normal modes of the system. The transformation will be

$$a_k = \frac{b_k + L_k b_{-k}^\dagger}{\sqrt{1 - L_k^2}}; \quad a_k^\dagger = \frac{b_k^\dagger + L_k b_{-k}}{\sqrt{1 - L_k^2}} \quad (14.138)$$

or

$$b_k = \frac{a_k - L_k a_{-k}^\dagger}{\sqrt{1 - L_k^2}}; \quad b_k^\dagger = \frac{a_k^\dagger - L_k a_{-k}}{\sqrt{1 - L_k^2}}. \quad (14.139)$$

Of course, we have to find L_k and we do this by making the substitutions into H and choosing it to cancel the unwanted terms. Notice that

$$[a_k, a_{k'}^\dagger] = \frac{[b_k, b_{k'}^\dagger] + L_k^2 [b_{-k}^\dagger, b_{-k'}]}{1 - L_k^2} = \delta_{kk'} \quad (14.140)$$

so the correct creation-annihilation operator algebra is maintained. Physically, a_k annihilates an original boson carrying momentum k (for a momentum change $= -k$) or creates a boson with momentum $-k$, which is also a momentum change $-k$. A wee bit of algebra yields

$$L_k = \frac{\epsilon_k + 2Nv - E(k)}{2Nv} \quad (14.141)$$

where $\epsilon_k = \hbar^2 k^2 / (2m)$ and

$$E(k)^2 = \left(\frac{\hbar^2 k^2}{2m} \right)^2 + 2Nv \frac{k^2}{m} \quad (14.142)$$

$E(k)$ is the energy of an excitation with momentum k . As the equation and Fig. 14.2 show, the high momentum quasi-particles are just the original particles. However, the low-lying, long wavelength spectrum has a linear dispersion relation, $E(k) = C_s \hbar k$ where $C_s =$

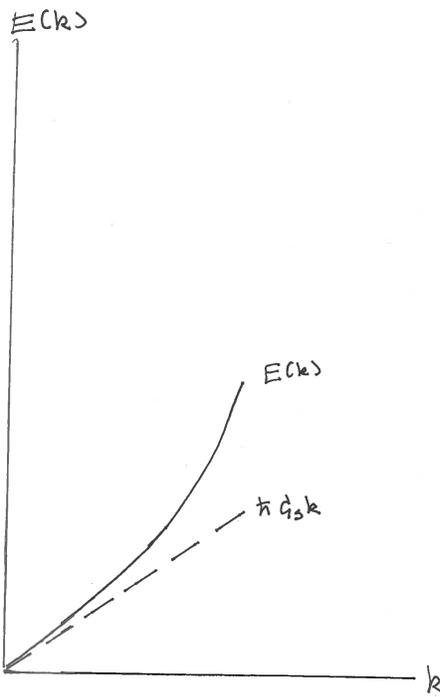


Figure 14.2: Quasi-particle (phonon) dispersion relation for the weakly-interacting condensed Bose gas.

$\sqrt{2Nv/m}$. These are sound waves – they are called “phonons,” in complete analogy with the quantized vibrational modes of a crystal lattice. The non-ideal weakly-interacting condensed Bose gas is best thought of as a gas of non-interacting phonons.

Finally, suppose we have an impurity atom (of mass M) moving through the condensate. The only way it can lose energy is to create an excitation in the condensate. Suppose it has initial momentum $\hbar\vec{q}$, and also suppose it creates an excitation of momentum $\hbar\vec{k}$ at an angle θ away from its direction of motion. Conservation of energy says

$$\begin{aligned}\frac{\hbar^2 q^2}{2M} &= \frac{\hbar^2}{2M} |\vec{q} - \vec{k}|^2 + E(k) \\ &= \frac{\hbar^2}{2M} [q^2 + k^2 - 2kq \cos \theta] + E(k)\end{aligned}\tag{14.143}$$

or

$$qk \cos \theta = \frac{k^2}{2} + \frac{ME(k)}{\hbar^2}.\tag{14.144}$$

Calling the impurity’s velocity $v = \hbar q/M$, this is

$$\cos \theta = \frac{k}{q} + \frac{ME(k)}{\hbar^2 k} = \frac{\hbar k}{2Mv} + \frac{E/\hbar k}{v}.\tag{14.145}$$

For phonons, $E(k)/\hbar k > C_s$, so we need $v > C_s$ for $\cos \theta < 1$, allowing a phonon to be emitted. This means that if the impurity atom is moving too slowly, it cannot emit a phonon, and it cannot lose energy. Now look at the process from the point of view of an observer riding along with the impurity. The condensate streams past without friction (if the velocity of the condensate is low enough). This is superfluidity – we have just discovered that the condensate is a superfluid!

Chapter 15

Atoms and molecules

Atoms

An exact treatment of the Z -body Schrödinger equation for electrons bound to a nucleus is very hard, and, in the days before supercomputers, it was impossible. Not wishing to do these calculations, we will have to make approximations from the very beginning. Let us begin our study of atoms keeping only the Coulomb interaction for all the electrons. Our many-electron Hamiltonian is

$$\hat{H} = \sum_{i=1}^Z \left(\frac{\hat{p}_i^2}{2m} - \frac{Ze^2}{r_i} \right) + \sum_{j=1}^Z \sum_{i<j} \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \quad (15.1)$$

At zeroth order, we choose to neglect the inter-electronic repulsion, and keep only the nuclear attraction. We might be able to justify this for heavier atoms, where $Z \gg 1$. It is certainly a simple approximation: our many-electron wave function becomes a product of hydrogenic wave functions. The additional imposition of the exclusion principle allows us to build atoms. So we begin, with Table 15.1.

Table 15.1: Filling the periodic table.

Atom	Highest e^-
H	$1s$
He	$1s^2$
Li	$1s^2 2s$
Be	$1s^2 2s^2$
B	$1s^2 2s^2 2p$
...	

But wait! In hydrogen, the $2s$ and $2p$ energies are degenerate at the order we are working. Why, in the periodic table, are the s levels filled first?

Clearly, we need a better approximation. This is provided by the “central field approximation:” each electron feels the Coulomb repulsion of all the others, but angle-averaged:

$$V(r_k) = -\frac{Ze^2}{r_k} + e\Phi_{CFk}(r_k) \quad (15.2)$$

where

$$\rho_{CFk}(r) = e \int d\Omega \sum_{i \neq k} |u_i(r)|^2, \quad (15.3)$$

$$\nabla^2 \Phi_{CFk}(r) = -4\pi \rho_{CFk}(r) \quad (15.4)$$

or

$$\frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial \Phi_{CFk}(r)}{\partial r} = -4\pi \rho_{CFk}(r). \quad (15.5)$$

How can this be done? Why might it be a good approximation?

To carry out this calculation, assume that the many-electron wave function is separable,

$$\psi(r_1, r_2, \dots, r_Z) = u_1(r_1)u_2(r_2) \dots u_Z(r_Z), \quad (15.6)$$

and pick some convenient functional form for all the u 's. From them, compute $\rho_{CFk}(r)$ and $\Phi_{CFk}(r)$. Then solve

$$\left[-\frac{\hbar^2}{2m} \nabla_k^2 - \frac{Ze^2}{r_k} + e\Phi_{CFk}(r_k) \right] u_k(r_k) = \epsilon_k u_k(r_k). \quad (15.7)$$

This is the Schrödinger equation for a particle in a radial potential, so the solution is

$$u_k(\vec{r}) = R(r)Y_l^m(\theta, \phi). \quad (15.8)$$

This reduces the problem to the solution of the radial part of the wave equation. Repeat this procedure for each k . Take these u_k 's, compute ρ_{CFk} and Φ_{CFk} , and repeat the procedure until it converges.

What we have described is also known as the “self consistent field” or “Hartree” approximation. The solution will show the following generic features: The innermost electrons will see the nuclear potential, shifted by a constant by the outer electrons. This can be understood without calculation, from Gauss' law: $\vec{E} = 0$ inside a spherical charge distribution, so the potential is a constant. The outermost electrons see the nucleus, screened by the other electrons, $V(r) = -(Z - 1)e^2/r$. Both of these are known to be good approximations.

Why does this procedure work? To give a partial answer, the spherical symmetry of the charge distribution is easy to understand, if all the levels of an orbital are filled. This follows from the identity

$$\sum_m |Y_l^m(\theta, \phi)|^2 = \text{constant}. \quad (15.9)$$

But there is a more complete answer: this is actually a variational calculation. Let us repeat the derivation with that idea in mind. We begin, again, with the ansatz Eq. 15.6. We will not antisymmetrize it; we simply force our allowed states to be consistent with the

exclusion principle. We use this wave function in a variational calculation of $H\psi = E\psi$ with the Hamiltonian of Eq. 15.1. We find

$$\begin{aligned}
 \langle H \rangle &= \int \dots \int d^3r_1 d^3r_2 \dots d^3r_Z \psi^* H \psi \\
 &= \sum_k \int u_k^*(r_k) \left[-\frac{\hbar^2}{2m} \nabla_k^2 - \frac{Ze^2}{r_k} \right] u_k(r_k) \\
 &\quad + \sum_k \sum_{j, \text{pairs}} \int \int u_k^*(r_k) \left[u_j^*(r_j) \frac{e^2}{r_{jk}} u_j(r_j) \right] u_k(r_k) \\
 &= \sum_k \int u_k^*(r_k) H_k u_k(r_k) d^3r_k
 \end{aligned} \tag{15.10}$$

where

$$H_k = -\frac{\hbar^2}{2m} \nabla_k^2 - \frac{Ze^2}{r_k} + \sum_{j \neq k} \int u_j^*(r_j) \frac{e^2}{r_{jk}} u_j(r_j) d^3r_j \tag{15.11}$$

Eq. 15.10 is the expectation value of H_k for a trial wave function u_k . From our variational discussion, the minimum energy is obtained when u_k is an eigenfunction of H_k ,

$$H_k u_k = \epsilon_k u_k \tag{15.12}$$

which is the eigenvalue equation for the Hartree wave function. The electronic energy of the atom is then

$$E = \sum_k \epsilon_k - \sum_{j > k} \int \int |u_j(r_j)|^2 |u_k(r_k)|^2 \frac{e^2}{r_{jk}} d^3r_j d^3r_k \tag{15.13}$$

The sum $\sum_k \epsilon_k$ counts the pairs' Coulomb repulsion twice; hence the correction term. Note that this does not give the central field approximation directly (it is still just an approximation). However, it arises naturally for all closed shells.

Now we can return to the periodic table. The $n = 1$ shell is as before. To handle $n = 2$, recall that the angular momentum barrier, $\hbar^2 l(l+1)/(2mr^2)$, pushes higher l states away from the nucleus. One an electron is at bigger r , other electrons can screen the nucleus. This makes it energetically favorable to fill lower l first. So we have the results in Table 15.2. Next we come to $n = 3$, where we have s , p and d levels. The energy difference between them is smaller than for $n = 2$. At some point a higher n 's s state has lower energy than a lower n 's d state. So we fill the $3s$ and $3p$ states, then pause to fill the $4s$, making K and Ca, before returning to fill the ten $3d$ states – Sc to Zn. The rare earths and the actinides take a break in filling d orbitals to fill f ones.

Table 15.2: Filling the periodic table.

Atom	Highest e^-
Li	$1s^2 2s$
Be	$1s^2 2s^2$
B	$1s^2 2s^2 2p$
\vdots	
Ne	$1s^2 2s^2 2p^6$

Atomic spectroscopy

The two largest terms left out of the central field approximation are the spin-orbital coupling

$$V_{SO} = \sum_k \left(\frac{1}{2mc^2} \frac{1}{r_k} \frac{\partial V}{\partial r_k} \right) \vec{L}_k \cdot \vec{S}_k \quad (15.14)$$

and the difference between the true inter-electron interaction (even “true” in the sense of the Hartree self consistent field calculation) and the central field approximate potential, which we will call V_c . In closed shells, $V_{SO} = 0$; the term in parentheses in Eq. 15.14 is the same for all, and terms with opposite m_l ’s and m_s ’s cancel among each other.

We can imagine two extreme cases. The first, $V_c > V_{SO}$, occurs for light atoms and gives rise to “L-S coupling” or “Russell-Saunders coupling.” Good quantum numbers are total \vec{J} (of course) and total \vec{S} . The goodness of \vec{S} arises from the same interplay of Coulomb interactions, Fermi statistics and spin that we saw for helium. Since $J = L + S$, the total L is also a good quantum number. Our states are then

$$|n's, J_{tot}, L_{tot}, S_{tot}\rangle. \quad (15.15)$$

We then treat V_{SO} perturbatively. We will omit a detailed calculation, and just remark that it is possible to parametrize the spin-orbital interaction as

$$V_{SO} = F(r) \vec{L} \cdot \vec{S} \quad (15.16)$$

where $F(r)$ is some (hard-to-evaluate) radial function, and \vec{L} and \vec{S} are the total orbital and spin angular momenta of the valence electrons. Then, squaring $J = L + S$,

$$\langle \vec{L} \cdot \vec{S} \rangle = \frac{\hbar^2}{2} (j(j+1) - l(l+1) - s(s+1)). \quad (15.17)$$

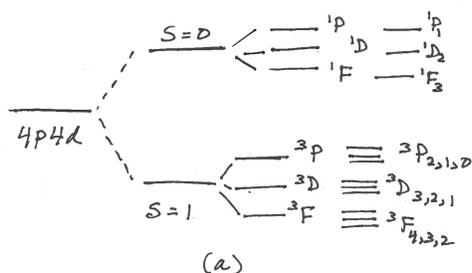


Figure 15.1: Pattern of splitting for a pd pair of electrons under Russell-Saunders coupling.

As an example, suppose we have two valence electrons; for the sake of argument take them as $4p4d$. As in the case of helium, the Coulomb interaction differs depending on whether the two electrons' spatial wave function is symmetric or antisymmetric, and hence on whether the total spin is 1 or 0. Usually, as we saw for helium, higher S lies lower. This is called "Hund's rule." The electrostatic energy depends on the particular values of n and l : in terms of the direct and exchange terms it will be $\Delta E = J_{nl} \pm K_{nl}$. For our pd pair, $L = 1 + 2 = 3, 2, 1$ for F, D, and P levels. The states of different L (for given S) will be non-degenerate. Usually, lower L lies lower. A classical argument for this result is that we want to minimize the Coulomb repulsion of the electrons, and if we put them all the same orbital, we can maintain maximum separation if they all rotate together. Fig. 15.1 shows the resulting splittings expected.

Finally, we have V_{SO} . We have states of given S and L , which combine in the usual way into all appropriate J 's, and then Eq. 15.17 gives us the final pattern of splittings. The

Lande interval rule

$$E(J + 1, L, S) - E(J, L, S) = C((J + 1)(J + 2) - J(J + 1)) = 2C(J + 1) \quad (15.18)$$

can be used to identify the J quantum number of a level.

What is the sign of the L-S coupling term? For less than a half-filled orbital, the situation is like hydrogen: higher J lies higher in energy. However, if the level is more than half filled, the situation reverses. As an example, consider four electrons in a p-orbital. Think of this as six electrons and two holes. The six electrons form a closed shell and are inert. Holes have positive charge, so their spin-orbit interaction flips sign compared to electrons'. (All other effects go as e^2 and are charge-symmetric.) So for more than half filled levels, higher J lies lower.

The author is a particle physicist and he is used to dealing with baryons, three-quark bound states. In principle, multi-electron states are simple: combine two, then add the third ... Angular momentum technology beyond what we have introduced is really necessary if you want to be professional about this.

And don't forget missing levels due to the Pauli principle. Two valence p electrons in different n states have the spectroscopy shown in Fig. 15.2(a). However, if they have the same n , their radial wave function is symmetric. Two electron states with even total L are symmetric ($|22\rangle = |11\rangle|11\rangle$, for example) while states of odd total L are antisymmetric. So some states (such as and $S=0$ P-wave) are missing. Compare Fig. 15.2(b).

This is all quite intricate, but it allows us to predict the electronic ground state configuration of atoms with one or two valence electrons. Examples are shown in Table 15.3.

Table 15.3: Ground state wave function for some two valence electron atoms.

Atom	electrons	quantum numbers	$^{2S+1}L_J$
B	$2p$	$S = 1/2, L = 1, J = 1/2$	$^2P_{1/2}$
C	$2p^2$	$S = 1, L = 1, J = 0$	3P_0
O	$2p^4$	$S = 1, L = 1, J = 2$	3P_2
F	$2p^5$	$S = 1/2, L = 1, J = 3/2$	$^2P_{3/2}$
Ne	$2p^6$	$S = 0, L = 0, J = 0$	1S_0

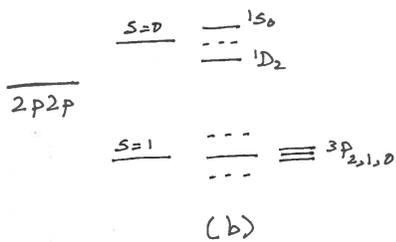
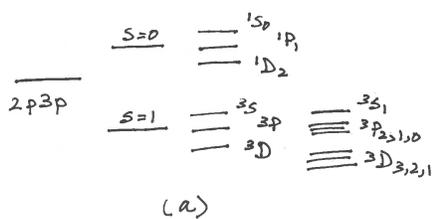


Figure 15.2: Pattern of splitting for a pp pair of electrons under Russell-Saunders coupling, in different radial states (a) or in the same radial state, with missing levels (b)

For heavy atoms, the spin-orbital term

$$\sum_k \zeta_k \vec{L}_k \cdot \vec{S}_k \quad (15.19)$$

is much greater than the residual electrostatic interaction. This means that each electron's j is a good quantum number and the system is said to be “j-j coupled.”

Because $J = L + S$ and because $S = 1/2$, $2\vec{L} \cdot \vec{S} = l$ for $j = l + 1/2$ and $-(l + 1)$ when $j = l - 1/2$. Two electrons are all I can do: we have four states whose energies are given by sums of the appropriate $\vec{L} \cdot \vec{S}$ values. We then add the two j 's to give the total j and, indirectly, the splitting due to V_c . An example of a pd state's splittings is shown in Fig. 15.3.

Molecules

Molecules are harder than atoms because, generally, we don't have a central potential, and rotational invariance is lost. This makes the solution of the Schrödinger equation much more difficult. To compensate, the ratio $m_e/M_{nucleus}$ is 10^{-4} or smaller. This means that in the molecule, the nucleus moves more slowly than the electron.

The electrons think that the nuclei are at rest, so they feel an almost static Coulomb potential from the nuclei. If the size of the molecule is roughly that of the Bohr radius, the electronic energy levels will be $E_{elec} \sim \hbar^2/(m_e a^2) \sim \text{eV's}$, like in atoms.

To the nuclei, the electrons move quickly. As the nuclei move, they distort the electronic cloud, which slightly changes the electronic energy. Thinking classically for a moment, the nucleons move in an elastic medium. There is a minimum energy configuration, where the energy I am describing is the sum of the Coulomb repulsion of the nuclei and the electronic energy $\epsilon(r)$, which depends on the coordinates of the nuclei.

The easiest way to describe the situation is to model it as a simple harmonic oscillator: $V(R) \sim M\omega_0^2(R - R_p)^2$ where M is a typical nuclear mass and R_p is the equilibrium separation. If the nuclei move a distance $\Delta R \sim a$, the potential V goes up by an amount $\hbar^2/(m_e a^2)$. Equating the two expressions, $M\omega_0^2 a^2 \sim \hbar^2/(m_e a^2)$ says $\omega_0^2 = \hbar^2/(a^4 m_e M)$ or

$$\hbar\omega_0 \sim \frac{\hbar^2}{m_e a^2} \sqrt{\frac{m_e}{M}} \quad (15.20)$$

The excitations are those of an oscillator, and are called “vibrational levels.” From this calculation, they have an energy of 1/10 to 1/100 eV.

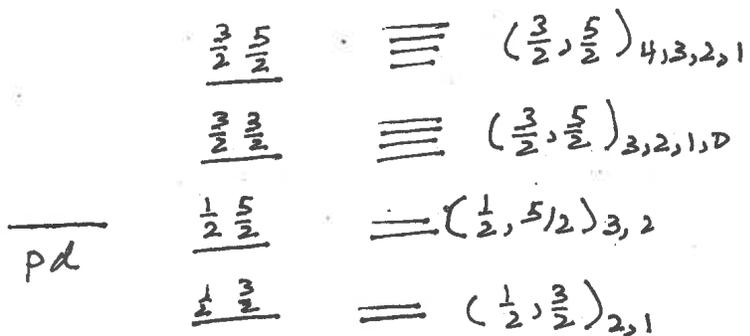


Figure 15.3: Schematic picture of the energy levels of a pd electron pair, jj coupled.

What is the nuclear velocity? We can get it from $p^2/(2m) \sim \hbar\omega/2$ or

$$v_{nuc}^2 \sim \frac{\hbar^2}{m_e^2 a^2} \left(\frac{m_e}{M}\right)^{3/2} \quad (15.21)$$

The factor $\hbar/(m_e a)$ is the electron velocity so this calculation shows that the nuclei move slowly. The nuclei also do not move over large distances. Again setting $\hbar\omega \sim M\omega^2\delta^2$, we discover $(\delta/a)^2 \sim \sqrt{m_e/M}$. This means that $\delta/a \sim 1/10$.

Finally, the whole molecule can rotate. Assume that the rotation is rigid. Then

$$E_{rot} \sim \hbar^2 \frac{l(l+1)}{2I} \quad (15.22)$$

where I is the moment of inertia. Because all of the mass of the molecule is in its nuclei, and because the molecule's size is a , $I \sim Ma^2$ and $E_{rot} \sim \hbar^2(m_e a^2)(m/M)$. This is about 10^{-4} eV; rotation transition lines sit in the far infrared. Notice the hierarchical ratio of energies,

$$E_{elec} : E_{vib} : E_{rot} \sim 1 : \sqrt{\frac{m_e}{M}} : \frac{m_e}{M} \quad (15.23)$$

The Born-Oppenheimer approximation encodes these simple ideas. The Hamiltonian is

$$H = \sum_{electrons} \frac{p_i^2}{2m} + \sum_{nuclei} \frac{P_\alpha^2}{2M_\alpha} + V_{ee} + V_{NN} + V_{eN} \quad (15.24)$$

where V_{ee} is the repulsive electron-electron Coulomb interaction, V_{NN} is the repulsive nuclear-nuclear interaction, and V_{eN} is the attractive electron-nuclear Coulomb interaction. To begin, treat the $\sum \frac{P_\alpha^2}{2M_\alpha}$ term as small. (We already saw, it is about $\sqrt{m/M}$ in size.) In fact, to really begin, drop it altogether, so the nuclear positions are frozen. Generically, call their positions R , and call the electronic coordinates r . The wave function $\phi(r, R)$ obeys the "electronic equation"

$$[T_e + V_{ee}(r) + V_{eN}(r, R)]\phi_n(r, R) = (\epsilon_n(R) - V_{NN}(R))\phi_n(r, R). \quad (15.25)$$

$V_{NN}(R)$ is just a constant, and $\epsilon_n(R)$ is the n th energy level of the electrons. The functions $\phi_n(r, R)$ form a complete set of states for any R . (Dropping the nuclear kinetic term, R is just a parameter.) So, use these states to expand the exact wave function,

$$\Psi(r, R) = \sum_m \phi_m(r, R)c_m = \sum_m \phi_m(r, R)\Phi_n(R). \quad (15.26)$$

The expansion coefficients c_n depend on R , so we have renamed them $\Phi_n(R)$. Now take the Schrödinger equation

$$(T_e + T_N + V_{ee} + V_{NN} + V_{eN})\Psi(r, R) = E\Psi(r, R) \quad (15.27)$$

and substitute,

$$\sum_m \Phi_m(R) ((T_e + V_{ee} + V_{NN} + V_{eN})\phi_m(r, R) + T_N\phi_m(r, R)\Phi_m(R) = E \sum_m \Phi_m(R)\phi_m(r, R) \quad (15.28)$$

or

$$\sum_m \Phi_m(R) (\epsilon_m(R)\phi_m(r, R) + T_N\phi_m(r, R)\Phi_m(R) = E \sum_m \Phi_m(R)\phi_m(r, R) \quad (15.29)$$

so

$$\sum_m ((\epsilon_m(R) + T_N)\Phi_m(R)\phi_m(r, R) = E \sum_m \Phi_m(R)\phi_m(r, R). \quad (15.30)$$

Note that T_N , which is proportional to $\partial^2/\partial R^2$, acts on both ϕ and Φ .

Multiply both sides by $\phi_n(r, R)^*$ and integrate over R :

$$(\epsilon_m(R) + T_N)\Phi_m(R) = E\Phi_m(R) - \sum_m A_{nm} \quad (15.31)$$

where

$$A_{nm} = \int dr \phi_n(r, R)^* \left[\sum_m \frac{1}{2M_\alpha} (\nabla_R^2 (\Phi_m(R)\phi_m(r, R)) - \phi_m(r)\nabla_R^2 \Phi_m(R)) \right]. \quad (15.32)$$

One can show that A_{nm} is $O(m/M)$ compared to the rest of the expression. Neglect it, therefore, leaving the “nuclear equation”

$$(\epsilon_m(R) + T_N)\Phi_m(R) = E\Phi_m(R). \quad (15.33)$$

The quantity $\epsilon_m(R)$ is the effective potential energy for the nuclei due to electrons in level m , plus the nuclear Coulomb interactions. For a bound state, it is a function with a minimum.

At this point, it's useful to specialize to diatomic molecules. Then there is one coordinate, the internuclear separation R . We have

$$T_N = -\frac{\hbar^2}{2M_N} \nabla_R^2 = -\frac{\hbar^2}{2M_N} \frac{\partial^2}{\partial R^2} + \frac{\hbar^2 L^2}{2M_R R^2} \quad (15.34)$$

where M_R is the reduced mass. The nuclear equation becomes

$$-\frac{\hbar^2}{2M_N} \frac{\partial^2}{\partial R^2} \Phi_m(R) + [\epsilon_m(R) + \frac{\hbar^2 K(K+1)}{2M_R R^2}]. \quad (15.35)$$

We have introduced the rotational quantum number K (an integer, 0, 1, 2, ...). The nuclei move in a potential which is a sum of the electronic term and a centrifugal barrier.

To first approximation, we simply expand the effective potential about its minimum:

$$V_m(R) = V_{n0} + \frac{1}{2}k_{n0}(R - R_0)^2 + \dots \quad (15.36)$$

The nuclear states are a set of vibrational levels, $E = (v + \frac{1}{2})\hbar\omega_{n0}$ where $\omega_{n0}^2 = k_{n0}/M_R$ and v is an integer. We can add in the centrifugal term and repeat:

$$V_m(R) = [V_{n0} + \frac{\hbar^2 K(K+1)}{2M_R R_K^2}] + B(R - R_K)^2 + \dots \quad (15.37)$$

where

$$B = (\frac{1}{2}k_{n0} + \frac{3\hbar^2 K(K+1)}{2M_R R_K^2} + \dots) \quad (15.38)$$

We have a “stretched oscillator,” with slightly different frequencies for the vibrational levels as K rises.

If we ignore the stretching, the wave function is a simple harmonic oscillator in r and a spherical harmonic in the angular variables

$$\psi(\vec{r}) = u_v(r - r_K)Y_K^m(\theta, \phi) \quad (15.39)$$

where $u_v(r - r_K) \sim \exp(-M_R\omega^2(r - r_K)^2)$. The energy is

$$E_{v,K} = \hbar\omega(v + \frac{1}{2}) + \frac{\hbar^2 K(K+1)}{2M_R R_K^2}, \quad (15.40)$$

a set of rotational bands beginning at a set of equally spaced energy intervals.

The resulting spectroscopy is simple: the interaction Hamiltonian is, in dipole approximation, $H = -e\vec{r} \cdot \vec{E}$ and from the Golden Rule, the transition probability per unit time is

$$d\Gamma \propto \frac{2\pi}{\hbar} |\langle f|\vec{r}|i\rangle|^2 \quad (15.41)$$

Because $\vec{r} = \hat{n}((r - r_K) + r_K)$ the matrix element

$$\int dr u_{v'}(r - r_K)^* [(r - r_K) + r_K] u_v(r - r_K) \quad (15.42)$$

connects states with $v' = v \pm 1$ (the $(r - r_K)$ piece) and states with $v' = v$ (the r_K piece). In addition we have the usual angular selection rules $\Delta K = \pm 1, \Delta m_K = \pm 1, 0$.

Thus there are two kinds of transitions involving nuclear motion. The first are pure rotational transitions, $\Delta v = 0$, $\Delta K = \pm 1$. The energy of the photon which is absorbed (for $K \rightarrow K + 1$) or emitted (for $K \rightarrow K - 1$)

$$E_\gamma = \frac{\hbar^2}{2M_R R_K^2} [K_f(K_f + 1) - K_i(K_i + 1)] \quad (15.43)$$

Since the difference $E(K + 1) - E(K) \propto K + 1$, this is a set of equally-spaced photon lines. As we remarked earlier, the typical energy for rotational transitions is in the far infrared. Vibrational-rotational spectra have selection rules $\Delta v = \pm 1$, $\Delta K = \pm 1$. An absorption spectrum $v \rightarrow v + 1$ can have two branches, $K \rightarrow K + 1$ and $K \rightarrow K - 1$. The two K differences are $(K + 1)(K + 2) - K(K + 1) = 2(K + 1)$ and $(K - 1)K - K(K + 1) = -2K$ (but $K \geq 1$). This is a “comb” of equally spaced lines with a gap at $\omega = \omega_0$, the natural vibrational frequency. Typically these lines sit in the near infrared. All of these features are shown in Fig. 15.4.

Finally, transitions can involve the electronic levels. The angular momentum of the electronic cloud enters the problem. Life becomes complicated.

We conclude this section with a brief description of the H_2^+ ion, the simplest molecule. The electronic Hamiltonian is

$$H + \frac{\hbar^2}{2m} \nabla^2 - \frac{e^2}{|\vec{r} - \vec{R}_A|} - \frac{e^2}{|\vec{r} - \vec{R}_B|} + \frac{e^2}{|\vec{R}_B - \vec{R}_A|} \quad (15.44)$$

where (see Fig. 15.5) the two protons are located at \vec{R}_B and \vec{R}_A . As there is no exact solution to this problem, we attempt a variational solution with a superposition of hydrogenic wave functions,

$$\psi(r) = \alpha\psi_A(r) + \beta\psi_B(r), \quad (15.45)$$

where

$$\psi_A(r) = \frac{1}{(\pi a_0^3)^{1/2}} \exp(-|\vec{r} - \vec{R}_A|/a_0) \quad (15.46)$$

is called a “molecular orbital state.” The system is reflection-symmetric about the midplane, so we choose $\alpha = \pm\beta$. Our parity eigenstates become

$$\psi_\pm(r) = C_\pm[\psi_A \pm \psi_B]. \quad (15.47)$$

An immediate complication is that ψ_A and ψ_B are not orthogonal. The normalization factor becomes

$$1 = C_\pm^2 \int d^3r |\psi_A \pm \psi_B|^2 = C_\pm^2 [1 + 1 + 2 \int d^3r \psi_A^* \psi_B]$$

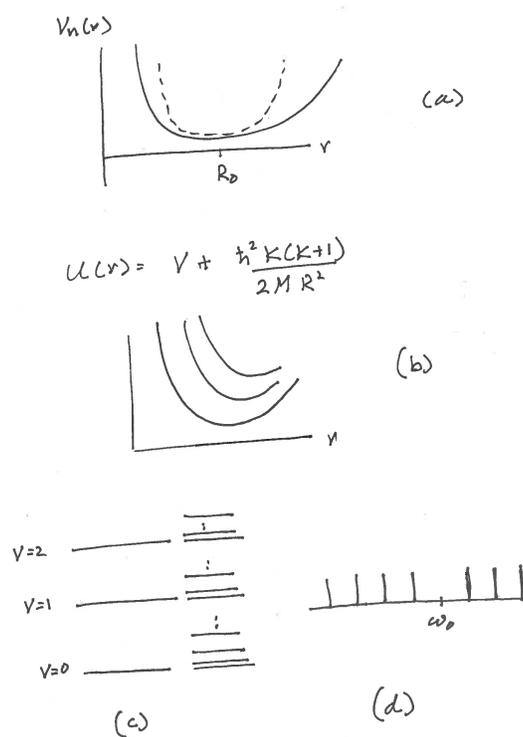


Figure 15.4: Figures for diatomic molecules: (a) Electronic energy $V_n(R)$ and its quadratic approximation (b) modification of the potential by the centrifugal barrier (c) Vibrational-rotational energy levels and (d) the vibrational-rotational “comb” with its missing line.

$$\equiv C_{\pm}^2(2 + 2S(R)). \quad (15.48)$$

$S(R)$ is called the overlap integral. It is roughly $S \sim (e^2/a_0) \exp(-R/a_0)$.

The energy expectation value is

$$\begin{aligned} \langle H_{\pm} \rangle = \epsilon(R) &= \frac{\langle A|H|A \rangle + \langle B|H|B \rangle \pm 2 \langle A|H|B \rangle}{2(1 \pm S)} \\ &= \frac{\langle A|H|A \rangle \pm \langle A|H|B \rangle}{1 \pm S} \end{aligned} \quad (15.49)$$

where $\langle A|H|A \rangle$ is a sum of three terms, ϵ_1 the ground state energy of hydrogen, e^2/R the nuclear repulsion, and

$$\epsilon_D = - \int d^3r |\psi_A(r)|^2 \frac{e^2}{|\vec{r} - \vec{R}_B|} \sim -\frac{e^2}{a_0} \exp(-2R/a_0) \quad (15.50)$$

the attraction to the other nucleus. A rough plot of $\epsilon(R)$ is shown in Fig. 15.5. The negative parity state has no minimum, and hence, no bound state. The positive parity state is bound. This makes sense: the odd parity state does not bind because it has a node at the midpoint. The electron in the even parity state feels an attraction from both nuclei. Further, if either Z is greater than unity, there is no binding: there is no He^{++} ion. The variational calculation gives a binding energy of $E = -1.76$ eV and a mean separation for the protons of $R = 1.3$ Å. The experimental values are -2.8 eV and 1.06 Å. But it is easy to point to a failure for our variational wave function: as R becomes small, ψ should look like the wave function for a He^+ atom, a hydrogenic wave function with $Z = 2$. Ours does not, so we know that we can do better. Chemistry is all about clever variational choices, so perhaps we should leave the discussion to the more specialized literature.

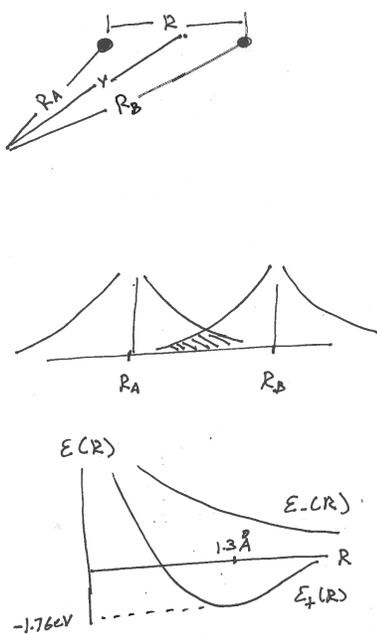


Figure 15.5: The H_2^+ molecule: coordinates, the overlap integral, schematic electronic energy vs R .