## Contents

1  The Quantum Mechanics of Position  
   1.1  Probability and probability density: One particle in one dimension  
   1.2  Probability amplitude  
   1.3  How does wavefunction change with time?  
   1.4  Amplitude density: Two particles in one or three dimensions  
   1.5  Solving the Schrödinger time development equation for the infinite square well  
   1.6  What did we learn by solving the Schrödinger time development equation for the infinite square well?  
   1.7  Other potentials  
   1.8  Energy loss  
   1.9  Averages and operators  
   1.10 The classical limit of quantum mechanics  
   1.11 Position plus spin  

2  Identical Particles  
   2.1  Two or three identical particles  
   2.2  Symmetrization and antisymmetrization  
   2.3  Energy states for two identical, noninteracting particles  
   2.4  Spin plus space, two electrons  
   2.5  Spin plus space, three electrons, ground state
CONTENTS

3  Atoms 43

3.1 Central potentials in two dimensions 43
3.2 Central potentials in three dimensions 43
3.3 The hydrogen atom 45
3.4 The helium atom 51
3.5 The lithium atom 55
3.6 All other atoms 55

A  Catalog of Misconceptions 56
Chapter 1

The Quantum Mechanics of Position

We've been studying the quantum mechanics of a silver atom’s magnetic moment, and we’ve gotten a lot out of it: we uncovered the phenomena of quantization and interference and entanglement; we found how to use amplitude as a mathematical tool to predict probabilities; we learned about quantum mechanical states. As profitable as this has been, we knew from the start that eventually we would need to treat the quantum mechanics of position. Now is the time.

1.1 Probability and probability density:

One particle in one dimension

A single particle moves restricted to one dimension. You know the story of quantum mechanics: If we measure the position (say, by shining a lamp), we will find that it has a single position. But there is no way to predict that position beforehand: instead, quantum mechanics predicts probabilities.

But what exactly does this mean? There are an infinite number of points along any line. If there were a finite probability at each of these points, the total probability would be infinity. But the total probability must be one! To resolve this essential technical issue, we’ll look at a different case of probability along a line.

You are studying the behavior of ants in an ant farm. (Ant farm: two panes of glass close together, with sand and ants and ant food between the panes.) The ant farm is 100.0 cm long. You paint one ant red, and 9741 times you look at the ant farm and measure (to the nearest millimeter) the distance of the red ant from the left edge of the farm.

You are left with 9741 raw numbers, and a conundrum: how should you present these numbers to help draw conclusions about ant behavior?

The best way is to conceptually divide the ant farm into, say, five equal bins: locations from 0.0 cm to 20.0 cm, from 20.0 cm to 40.0 cm, from 40.0 cm to 60.0 cm, from 60.0 cm to 80.0 cm, and from 80.0 cm to
100.0 cm. Find the number of times the ant is in the first bin, and divide by the total number of observations (9741) to find the probability that the ant was in the first bin. Similarly for the other bins. You will produce a graph like this:

![Probability of finding ant in bin](image)

The five probabilities sum to 1, as they must.

Now you want more detail about the ant’s location. Instead of dividing the ant farm into 5 conceptual bins each of width 20 cm, divide it into 10 bins of width 10 cm. The probabilities now look like

![Probability of finding ant in bin](image)

Now there are ten probabilities, yet they still sum to 1, so the probabilities are smaller. (For example, the first graph shows a probability of 0.28 for the ant appearing between 0.0 cm and 20.0 cm. The second graph shows probability 0.18 for the bin between 0.0 cm and 10.0 cm and probability 0.10 for the bin between 10.0 cm and 20.0 cm. Sure enough 0.28 = 0.18 + 0.10.)
If you want still more detail, you can divide the ant farm into 50 bins, each of width 2 cm, as in

These fifty probabilities must still sum to 1, so they are smaller still.

You could continue this process, making the bins smaller and smaller, and every bin probability would approach zero. In symbols, if the probability for the bin surrounding point \( x_0 \) is called \( P_0 \), and the width of each bin is called \( \Delta x \), we have that

\[
\lim_{\Delta x \to 0} P_0 = 0
\]

for all points \( x_0 \). This is true but gives no information whatsoever about ant behavior!

To get information, focus not on the bin probability but on the so-called probability density, defined as

\[
\lim_{\Delta x \to 0} \frac{P_0}{\Delta x} \equiv \rho(x_0).
\]

In terms of probability density, we can say that “the probability of finding the ant between \( x_A \) and \( x_B \) is

\[
\int_{x_A}^{x_B} \rho(x) \, dx.
\]

And that “the probability of finding the ant in a small window of width \( w \) centered on \( x_0 \) is approximately \( \rho(x_0)w \), and this approximation grows better and better as the window grows narrower and narrower.”

The fact that the bin probabilities, summed over all bins, is unity, or in symbols

\[
\sum_i P_i = 1,
\]

becomes, in the limit \( \Delta x \to 0 \),

\[
\sum_i P_i \approx \sum_i \rho(x_i) \Delta x \to \int_{0.0 \text{ cm}}^{100.0 \text{ cm}} \rho(x) \, dx = 1.
\]

This property of probability densities is called “normalization".
1.2 Probability amplitude

The probability considerations for one ant in one dimension is directly analogous to the probability considerations for one quantal particle in one dimension. You know the drill of quantum mechanics: The bin probability \( P_0 \) will be related to some sort of bin amplitude \( \psi_0 \) through \( P_0 = |\psi_0|^2 \). How does bin amplitude behave as \( \Delta x \to 0 \)? Because

\[
\frac{|\psi_0|^2}{\Delta x} \to \rho(x_0), \quad \text{we will have} \quad \frac{\psi_0}{\sqrt{\Delta x}} \to \psi(x_0),
\]

where, for any point \( x_0 \), the probability density \( \rho(x_0) = |\psi(x_0)|^2 \).

What would be a good name for this function \( \psi(x) \)? I like the name “amplitude density”. It’s not really a density: a density would have dimensions \( 1/\text{length} \), whereas \( \psi(x) \) has dimensions \( 1/\sqrt{\text{length}} \). But it’s closer to a density than it is to anything else. Unfortunately, someone else (namely Schrödinger) got to name it before I came up with this sensible name, and that name has taken hold. It’s called “wavefunction”.

The normalization condition for wavefunction is

\[
\int_{-\infty}^{+\infty} |\psi(x)|^2 \, dx = 1.
\]

You should check for yourself that this equation is dimensionally consistent.

1.3 How does wavefunction change with time?

I’m going to throw down three equations. First, the classical formula for energy,

\[
E = \frac{p^2}{2m} + V, \tag{1.3}
\]

where \( V \) is the potential energy. Second, the de Broglie relations for energy and momentum

\[
E = \hbar \omega \quad \text{and} \quad p = \hbar k. \tag{1.4}
\]

Third, the particular wavefunction

\[
\psi(x, t) = Ae^{i(kx - \omega t)}. \tag{1.5}
\]

Plugging equations (1.4) mindlessly in equation (1.3) we obtain

\[
\hbar \omega = \frac{\hbar^2 k^2}{2m} + V. \tag{1.6}
\]

\(^1\)Erwin Schrödinger (1887–1961) had interests in physics, biology, philosophy, and Eastern religion. Born in Vienna, he held physics faculty positions in Germany, Poland, and Switzerland. In 1926 he discovered the time-development equation that now bears his name. This led, in 1927, to a prestigious appointment in Berlin. In 1933, disgusted with the Nazi regime, he left Berlin for Oxford, England. He held several positions in various cities before ending up in Dublin. There, in 1944, he wrote a book titled What is Life? which is widely credited for stimulating interest in what had been a backwater of science: biochemistry.
and multiplying both sides by $\psi(x, t)$ gives

$$\hbar \omega \psi(x, t) = \frac{\hbar^2 k^2}{2m} \psi(x, t) + V \psi(x, t). \quad (1.7)$$

However, for the particular wavefunction (1.5) we directly see

$$\omega \psi(x, t) = 1 - \frac{\partial \psi}{i \partial t}; \quad k \psi(x, t) = 1 \frac{\partial \psi}{i \partial x}; \quad k^2 \psi(x, t) = 1 \frac{-\partial^2 \psi}{\partial x^2} = -\frac{\partial^2 \psi}{\partial x^2}. \quad (1.8)$$

Plugging these into equation (1.7) gives

$$\hbar \left( 1 - \frac{\partial \psi}{i \partial t} \right) = \frac{\hbar^2}{2m} \left( -\frac{\partial^2 \psi}{\partial x^2} \right) + V \psi(x, t), \quad (1.9)$$

which rearranges to

$$\frac{\partial \psi(x, t)}{\partial t} = -\frac{i}{\hbar} \left[ \frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V(x) \right] \psi(x, t). \quad (1.10)$$

I concede from the very start that this is a stupid argument, and that if any of you had proposed it to me I would have gone ballistic. First, equation (1.3) is a classical fact plopped mindlessly into a quantal argument. Second, the de Broglie relations (1.4) are themselves suspect: de Broglie never said what it was that was waving, never said whether $E$ ought to mean kinetic energy or total energy. Third, there are many possible wavefunctions other than equation (1.5). The unjustified change of potential energy value $V$ in equation (1.9) to potential energy function $V(x)$ in equation (1.10) merely adds insult to injury. About the only thing I can say on this equation’s behalf is that it’s dimensionally consistent.

Oh, and one more thing. The equation is correct. Despite its dubious provenance, experimental tests have demonstrated to everyone’s satisfaction that wavefunction really does evolve in time this way. (I should qualify: wavefunction evolves this way in many situations: non-relativistic, no magnetic field, no friction or any other non-conservative force, and where the particle’s magnetic moment is unimportant.)

This equation for time development in quantal systems plays the same central role in quantum mechanics that $F = ma$ does in classical mechanics. And just as $F = ma$ cannot be proven, only motivated and then tested experimentally, so this time-development result cannot be proven. The motivation is lame, but the experimental tests are impressive and cannot be ignored.

This time-development equation has a name: it is “the Schrödinger equation”.

CHAPTER 1. THE QUANTUM MECHANICS OF POSITION
1.4 Amplitude density: Two particles in one or three dimensions

We will soon work on solving the Schrödinger equation for one particle moving in one dimension, but first we will consider how to describe two particles moving in one dimension.

Two particles, say an electron and a neutron, move in one dimension. As before, we start with a grid in one-dimensional space:

\[ x \]

\[ i \]

\[ j \]

\[ \Delta x \]

We ask for the probability that the electron will be found in bin \( i \) and the neutron will be found in bin \( j \), and call the result \( P_{i,j} \). Although this is a one-dimensional situation, this question generates a two-dimensional array of probabilities.

To produce a probability density, we must divide by \( (\Delta x)^2 \), and then take the limit as \( \Delta x \to 0 \), resulting in

\[ P_{i,j} \to \rho(x_e, x_n). \]

So the probability of finding an electron within a narrow window of width \( w \) centered on \( x_e = 5 \) and finding the neutron within a narrow window of width \( u \) centered on \( x_n = 9 \) is approximately \( \rho(5, 9)wu \).

The bin amplitude is \( \psi_{i,j} \) with \( P_{i,j} = |\psi_{i,j}|^2 \). To turn a bin amplitude into a wavefunction, divide by \( \Delta x = \sqrt{(\Delta x)^2} \) and take the limit

\[ \lim_{\Delta x \to 0} \frac{\psi_{i,j}}{\Delta x} = \psi(x_e, x_n). \]  (1.11)
This wavefunction has dimensions $1/[\text{length}]$.

The generalization to more particles and higher dimensionality is straightforward. If a single electron moves in three-dimensional space, the wavefunction $\psi(\vec{x})$ has dimensions $1/[\text{length}]^{3/2}$. If an electron and a neutron move in three-dimensional space, the wavefunction $\psi(\vec{x}_e, \vec{x}_n)$ has dimensions $1/[\text{length}]^3$. Note carefully: For a two-particle system, the state is specified by one function $\psi(\vec{x}_e, \vec{x}_n)$ of six variables. It is not specified by two functions of three variables, $\psi_e(\vec{x})$ giving the state of the electron, and $\psi_n(\vec{x})$ giving the state of the neutron. There are four consequences of this simple yet profound observation.

First, wavefunction is associated with a system, not with a particle. If you’re interested in a single electron and you say “the wavefunction of the electron”, then you’re technically incorrect — you should say “the wavefunction of the system consisting of a single electron” — but no one will go ballistic and say that you are in thrall of a deep misconception. However if you’re interested in a pair of particles (an electron and a neutron, for instance) and you say “the wavefunction of the electron”, then someone (namely me) will go ballistic and point out that you are in thrall of a deep misconception.

Second, the wavefunction (and amplitude in general) is a mathematical tool for calculating the results of experiments, it is not physically “real”. We have emphasized this before, but it particularly stands out here. Even for a system as simple as two particles, the wavefunction does not exist in ordinary three-dimensional space, but in the six-dimensional “configuration space”, as it is called. I don’t care how clever or talented an experimentalist you are: you cannot insert an instrument into six-dimensional space in order to measure wavefunction.

Third, it might happen that the wavefunction factorizes:

$$\psi(\vec{x}_e, \vec{x}_n) = \psi_e(\vec{x}_e)\psi_n(\vec{x}_n).$$

In this case the electron has state $\psi_e(\vec{x}_e)$ and the neutron has state $\psi_n(\vec{x}_n)$. Such a peculiar case is called “non-entangled”. But in all other cases the state is called “entangled” and the individual particles making up the system do not have states. The system has a state, namely $\psi(\vec{x}_e, \vec{x}_n)$, but there is no state for the electron and no state for the neutron, in exactly the same sense that there is no position for a silver atom ambivating through an interferometer. Leonard Susskind puts it this way: If entangled states existed in auto mechanics as well as quantum mechanics, then an auto mechanic might say to you “I know everything about your car but . . . I can’t tell you anything about any of its parts.” (Leonard Susskind and Art Friedman, Quantum Mechanics: The Theoretical Minimum, page xii.)

Fourth, quantum mechanics is complex/hard/rich. (See The Physics of Quantum Mechanics, section 11.4 “The phenomena of quantum mechanics” and section 14.1, “Many-particle systems in quantum mechanics”.)
1.5 Solving the Schrödinger time development equation for the infinite square well

Setup. A single particle moves restricted to one dimension. In classical mechanics, the state of the particle is given through position and velocity: That is, we want to know the functions of time

\[ x(t); \quad v(t). \]

These stem from the solution to the ordinary differential equation \( \sum \vec{F} = m \vec{a} \), or, in this case,

\[ \frac{d^2x(t)}{dt^2} = \frac{1}{m} F(x) \]

subject to the given initial conditions

\[ x(0) = x_0; \quad v(0) = v_0. \]

In quantum mechanics, the state of the particle is given through the wavefunction: That is, we want to know the function

\[ \psi(x, t). \]

This is the solution of the Schrödinger partial differential equation (PDE)

\[ \frac{\partial \psi(x, t)}{\partial t} = -\frac{i}{\hbar} \left[ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x, t), \]

subject to the given initial condition

\[ \psi(x, 0) = \psi_0(x). \]

Infinite Square Well. Since this is our first solution of a quantal time development problem, we’ll start out cautiously. The potential we’ll choose is the easiest one: the so-called infinite square well or “particle in a box”:

\[ V(x) = \begin{cases} \infty & \text{for } x \leq 0 \\ 0 & \text{for } 0 < x < L \\ \infty & \text{for } L \leq x \end{cases} \]

This is an approximate potential for an electron added to a hydrocarbon chain molecule, or for an atom trapped in a capped nanotube.

It is reasonable (although not rigorously proven) that for the infinite square well

\[ \psi(x, t) = \begin{cases} 0 & \text{for } x \leq 0 \\ \text{something} & \text{for } 0 < x < L \\ 0 & \text{for } L \leq x \end{cases} \]

We will adopt these conditions.
Strategy. The PDE is linear, so if we find some special solutions \( f_1(x,t), f_2(x,t), f_3(x,t), \ldots \), then we can generate many more solutions through

\[
\sum_n D_n f_n(x,t).
\]

Because the \( D_n \) are adjustable, this is a big set of solutions, indeed it might be a big enough set to be the most general solution. Once we have the most general solution, we will need to find the values of \( D_n \) that correspond to the particular initial condition \( \psi(x,0) \).

Casting about for special solutions: separation of variables. So, how do we find even one solution of the PDE? Let’s try a solution \( f(x,t) \) that is a product of a function \( X(x) \) of position alone and \( T(t) \) of time alone, that is, try a solution of the form

\[
f(x,t) = X(x)T(t).
\]

Plugging this guess into the PDE, we find

\[
X(x) \frac{dT(t)}{dt} = -\frac{i}{\hbar} \left[ -\frac{\hbar^2}{2m} \frac{d^2X(x)}{dx^2} T(t) + V(x)X(x)T(t) \right],
\]

where the partial derivatives have become ordinary derivatives because they now act upon functions of a single variable. Divide both sides by \( X(x)T(t) \) to find

\[
\frac{1}{T(t)} \frac{dT(t)}{dt} = -\frac{i}{\hbar} \left[ -\frac{\hbar^2}{2m} \frac{d^2X(x)}{dx^2} \frac{1}{X(x)} + V(x) \right].
\]

In this equation, there is a function of time alone on the left-hand side, and a function of position alone on the right-hand side. But time and position are independent variables. It seems like the left side will vary with time, even while the position is held constant so the right side stays constant! Similarly the other way around. There is only one way a function of \( t \) alone can always be equal to a function of \( x \) alone, and that’s if both sides are equal to the same constant.

We don’t yet know what that constant is, or how many such constants there might be. To allow for the possibility that there might be many such constants, we call the constant value of the quantity in square brackets by the name \( E_n \). (The name suggests, correctly, that this constant must have the dimensions of energy.) We conclude that

\[
\frac{1}{T(t)} \frac{dT(t)}{dt} = -\frac{i}{\hbar} E_n,
\]

\[
-\frac{\hbar^2}{2m} \frac{d^2X(x)}{dx^2} \frac{1}{X(x)} + V(x) = E_n.
\]

We started with one partial differential equation in two variables but (for solutions of the form \( f(x,t) = X(x)T(t) \)) ended with two ordinary differential equations. And we know a lot about how to solve ordinary differential equations! This technique for finding special solutions is called “separation of variables".
1.5. SOLVING THE SCHRÖDINGER TIME DEVELOPMENT EQUATION FOR THE INFINITE SQUARE WELL

Solving the first ODE. When faced with solving two equations, I always solve the easy one first. That way, if the result is zero, I won’t have to bother solving the second equation.

So first we try to solve

\[
\frac{1}{T(t)} \frac{dT(t)}{dt} = -\frac{i}{\hbar} E_n
\]

\[
\frac{dT(t)}{T(t)} = -\frac{i}{\hbar} E_n dt
\]

\[
\int \frac{dT(t)}{T(t)} = -\frac{i}{\hbar} E_n \int dt
\]

\[
\ln T = -\frac{i}{\hbar} E_n (t + \text{constant})
\]

\[
T_n(t) = T_0 e^{-i(E_n/\hbar)t}
\]

Well, that went well. I don’t know about you, but it was easier than I expected. In the last step, I changed the name of \(T(t)\) to \(T_n(t)\) to reflect the fact that we’ll get different solutions for different values of \(E_n\).

Solving the second ODE. We move on to

\[
-\frac{\hbar^2}{2m} \frac{d^2X(x)}{dx^2} \frac{1}{X(x)} + V(x) = E_n.
\]

Remembering the form of the infinite square well potential, and the boundary conditions \(\psi(0,t) = 0\) plus \(\psi(L,t) = 0\), the problem to solve is

\[
-\frac{\hbar^2}{2m} \frac{d^2X(x)}{dx^2} = E_n X(x) \quad \text{with} \quad X(0) = 0; X(L) = 0.
\]

Perhaps you regard this sort of ordinary differential equation as unfair. After all, I don’t yet know the permissible values of \(E_n\). I’m not just asking you to solve an ODE with given coefficients, I’m asking you find find out what the coefficients are! Fair or not, we plunge ahead.

We are used to solving differential equations of this form. If I wrote

\[
m \frac{d^2f(t)}{dt^2} = -kf(t),
\]

you’d respond: “Of course, this is the ODE for a mass on a spring! The solution is

\[
f(t) = C \cos(\omega t) + D \sin(\omega t) \quad \text{where} \quad \omega = \sqrt{k/m}.
\]

Well, then, the solution for \(X(x)\) has to be

\[
X_n(x) = C_n \cos(\omega x) + D_n \sin(\omega x) \quad \text{where} \quad \omega = \sqrt{2mE_n/\hbar^2},
\]

where again I have taken to calling \(X(x)\) by the name \(X_n(x)\) to reflect the fact there there are different solutions for different values of \(E_n\). Writing it out neatly,

\[
X_n(x) = C_n \cos((\sqrt{2mE_n/\hbar})x) + D_n \sin((\sqrt{2mE_n/\hbar})x).
\]
When you solved the problem of a mass on a spring in your classical mechanics course, you had to supplement the ODE solution with the initial values \( f(0) = x_0, \quad f'(0) = v_0 \), to find the constants \( C \) and \( D \). This is called an “initial value problem”. For the problem of a particle in a box, we don’t have an initial value problem, instead we are given \( X_n(0) = 0 \) and \( X_n(L) = 0 \) which is called a “boundary value problem”.

Plugging in \( x = 0 \) gives

\[
X_n(0) = C_n \cos(0) + D_n \sin(0) = C_n,
\]

so the boundary value \( X_n(0) = 0 \) means that \( C_n = 0 \) — for all values of \( n \)!

Thus

\[
X_n(x) = D_n \sin((\sqrt{2mE_n/\hbar})x).
\]

Plugging in \( x = L \) gives

\[
X_n(L) = D_n \sin((\sqrt{2mE_n/\hbar})L),
\]

so the boundary value \( X_n(L) = 0 \) means that

\[
\frac{\sqrt{2mE_n}}{\hbar} L = n\pi \quad \text{where} \quad n = 0, \pm 1, \pm 2, \pm 3, \ldots
\]

so

\[
X_n(x) = D_n \sin((n\pi/L)x).
\]

If you think about it for a minute, you’ll realize that \( n = 0 \) gives rise to \( X_0(x) = 0 \). True, this is a solution to the differential equation, but it’s not an interesting one. Similarly, the solution for \( n = -3 \) is just the negative of the solution for \( n = +3 \), so we get the same effect by changing the sign of \( D_3 \). We don’t have to worry about negative or zero values for \( n \).

In short, the solutions for the boundary value problem are

\[
X_n(x) = D_n \sin(n\pi x/L) \quad \text{where} \quad n = 1, 2, 3, \ldots
\]

and with

\[
E_n = n^2 \frac{\pi^2 \hbar^2}{2mL^2}.
\]

We have accomplished the “unfair”: we have not only solved the differential equation, we have also determined the permissible values of \( E_n \).

**Pulling things together.** We now know that a solution to the Schrödinger time development equation for the infinite square well of width \( L \) is

\[
\psi(x, t) = \sum_{n=1}^{\infty} D_n e^{-(i/\hbar)E_nt} \sin(n\pi x/L),
\]

where

\[
E_n = n^2 \frac{\pi^2 \hbar^2}{2mL^2}.
\]

This is a lot of solutions — there are an infinite number of adjustable parameters, after all! — but the question is whether it is the most general solution. In fact it is, although that’s not obvious. The branch of
mathematics devoted to such questions, called Sturm-Liouville theory,\(^2\) is both powerful and beautiful, but this is not the place to explore it.

**Fitting the initial conditions.** Remember that our problem is not to simply solve a PDE, it is to find how a given initial wavefunction

\[ \psi(x, 0) = \psi_0(x) \]

changes with time. To do this, we fit our solution to the given initial conditions.

To carry out this fitting, we must find \(D_n\) such that

\[ \psi(x, 0) = \sum_{n=1}^{\infty} D_n \sin(n\pi x / L) = \psi_0(x). \]

The problem seems hopeless at first glance, but there’s a valuable trick, worth remembering, that renders it straightforward.

The trick relies on the fact that, for \(n, m\) integers,

\[ \int_0^L \sin(n\pi x / L) \sin(m\pi x / L) \, dx = \begin{cases} L/2 & \text{for } n = m \\ 0 & \text{for } n \neq m \end{cases} \]

You can work this integral out for yourself, using either

\[ \sin A \sin B = \frac{1}{2} [\cos(A - B) - \cos(A + B)] \]

or

\[ \sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}, \]

whichever you like better.

To employ this fact, start with

\[ \sum_{n=1}^{\infty} D_n \sin(n\pi x / L) = \psi_0(x), \]

multiply both sides by \(\sin(m\pi x / L)\), and integrate from 0 to \(L\):

\[ \sum_{n=1}^{\infty} D_n \int_0^L \sin(n\pi x / L) \sin(m\pi x / L) \, dx = \int_0^L \psi_0(x) \sin(m\pi x / L) \, dx. \]

This looks even worse, until you realize that all but one of the terms on the left vanish! Once you do make that realization,

\[ D_m (L/2) = \int_0^L \psi_0(x) \sin(m\pi x / L) \, dx \]

and we have a formula for \(D_m\).

---

\(^2\)Charles-François Sturm (1803–1855) was a French mathematician who also helped make the first experimental determination of the speed of sound in water. Joseph Liouville (1809–1882), another French mathematician, made contributions in complex analysis, number theory, differential geometry, and classical mechanics. He was elected to the French Constituent Assembly of 1848 which established the Second Republic.
CHAPTER 1. THE QUANTUM MECHANICS OF POSITION

Pulling all things together. The time development of a particle of mass $m$ in an infinite square well of width $L$, with initial wavefunction $\psi_0(x)$, is

$$\psi(x,t) = \sum_{n=1}^{\infty} D_n e^{-iE_nt} \sin(n\pi x/L),$$

(1.12)

where

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2}$$

(1.13)

and

$$D_n = \frac{2}{L} \int_0^L \psi_0(x) \sin(n\pi x/L) \, dx.$$

(1.14)

1.6 What did we learn by solving the Schrödinger time development equation for the infinite square well?

The solution. In one sense, we learned that the time development of a particle of mass $m$ in an infinite square well of width $L$, with initial wavefunction $\psi_0(x)$, is given by the three equations above. But we should delve more deeply than simply saying “There’s the answer, let’s go to sleep now.”

Quantal revivals. Does this jumble of symbols tell us anything about nature? Does it have any peculiar properties? Here’s one. Suppose there were a time $T_{rev}$ such that

$$e^{-iE_nT_{rev}} = 1 \quad \text{for } n = 1, 2, 3, \ldots$$

(1.15)

What would the wavefunction $\psi(x,t)$ look like at time $t = T_{rev}$? It would be exactly equal to the initial wavefunction $\psi_0(x)$! If there is such a time, it’s called the “revival time”.

But it’s not clear that such a revival time exists. After all, equation (1.15) lists an infinite number of conditions to be satisfied for revival to occur. Let’s investigate. The revival conditions (1.15) are equivalent to

$$-(1/\hbar)E_nT_{rev} = 2\pi \text{(an integer)} \quad \text{for } n = 1, 2, 3, \ldots$$

Combined with the energy eigenvalues (1.13), these conditions are

$$-n^2 \frac{\pi \hbar}{4mL^2} T_{rev} = \text{(an integer)} \quad \text{for } n = 1, 2, 3, \ldots$$

And, looked at this way, it’s clear that yes, there is a time $T_{rev}$ that satisfies this infinite number of conditions. The smallest such time is

$$T_{rev} = \frac{4mL^2}{\pi \hbar}.$$  

(1.16)

Cute and unexpected! This behavior was packed into equations (1.12) and (1.13), but no one would have known this from a glance.
1.6. WHAT DID WE LEARN BY SOLVING THE SCHRÖDINGER TIME DEVELOPMENT EQUATION FOR THE INFINITE SQUARE WELL?


*Investigating the solution technique.* But I want to do more than investigate the properties of the solution, I want to investigate the characteristics of the solution technique. In his book Mathematics in Action, O. Graham Sutton writes that “A technique succeeds in mathematical physics, not by a clever trick, or a happy accident, but because it expresses some aspect of a physical truth.” What aspect of physical truth is exposed through the techniques we developed to solve this time development problem?

First, let’s review the problem we solved, then the techniques we used. The problem was solving the partial differential equation

\[
\frac{\partial \psi(x,t)}{\partial t} = -\frac{i}{\hbar} \left[ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x,t),
\]

subject to the initial condition

\[
\psi(x,0) = \psi_0(x).
\]

The three techniques used were:

1. Finding many particular solutions of the PDE that happen to factorize: \( f(x,t) = X(x)T(t) \) (“separation of variables”).
2. Summing all of these particular solutions to find a more general (and as it turns out, the most general) PDE solution: \( \sum D_n X_n(x) T_n(t) \).
3. Finding the coefficients \( D_n \) that match up to initial value \( \psi_0(x) \) (“Fourier\(^3\) sine series”).

*Fourier sine series.* Let’s look at the last step first. The technique of Fourier sine series is generally powerful. Any function \( f(x) \) with \( f(0) = 0 \) and \( f(L) = 0 \) can be expressed as

\[
f(x) = \sum_{n=1}^{\infty} f_n \sin(n\pi x/L) \quad \text{where} \quad f_n = \frac{2}{L} \int_0^L f(x) \sin(n\pi x/L) \, dx.
\]

This seems paradoxical: Complete information about the function is obtained through knowing \( f(x) \) at every real number \( 0 \leq x \leq L \). Alternatively, complete information about the function is obtained through knowing the coefficients \( f_n \) for every positive integer \( n \). But there are more real numbers between 0 and \( L \) than there are positive integers! I have no resolution for this paradox — I’ll just remark that in knowing the function through its Fourier coefficients \( f_n \), it seems that we’re getting something for nothing.

Well, there are lots of times when we want to get something for nothing! Fourier sine series are useful in data compression. For example, suppose we are interested in recording a sound that starts with silence at time 0, proceeds through several notes, then ends with silence at time \( L \). We could do this by keeping track

---

\(^3\)Joseph Fourier (1768–1830) was French, so his name is pronounced “Four - e - a” with a silent “r”. He arrived at the series which bears his name through studies of heat flow. He was the first to propose the phenomenon that we today call “the greenhouse effect.”
of the air pressure \( f(t) \) at every instant from 0 to \( L \), or we could do it by keeping track of the corresponding Fourier coefficients \( f_n \). In either case there are an infinite amount of data to keep track of, and some will have to be thrown out to enable it to fit within a finite computer. You might expect that it would be more efficient to store this information in the form of \( f_n \) than in the form of \( f(t) \), and your expectation is correct: for a given amount of storage space, the \( f_n \) give a more accurate reproduction of the sound than the \( f(t) \). There are many schemes for the details of exactly when the Fourier series should be truncated: one such scheme is called “MP3”.

Or, for pictures rather than sounds: A black-and-white photograph is a two-dimensional intensity function \( f(x, y) \). You could store the image on a computer by breaking space \((x, y)\) into a grid (“pixels”) and storing a value for the intensity at each grid point (the so-called TIFF format) or you could store the information through Fourier coefficients \( f_{n,m} \) (the so-called JPEG format). For a given level of image quality, the JPEG file is considerably smaller than the TIFF file.

Stationary states. Okay, this is fun and profitable, but it tells us about how clever humans are; it doesn’t tell us anything about nature. I’m going to probe in this direction: We see that, as far as time development is concerned, functions like \( \sin(3\pi x/L) \) play a special role. What if the initial wavefunction \( \psi_0(x) \) happens to have this form?

So the initial wavefunction is

\[
\psi_0(x) = A \sin(3\pi x/L).
\]

We need the constant \( A \) so that the initial wavefunction will (1) have dimensions and (2) be normalized. For all wavefunctions, the probability of being somewhere is 1, that is

\[
\int_{-\infty}^{+\infty} |\psi(x)|^2 \, dx = 1.
\]

This requirement is called “normalization”. Applying the general normalization requirement to this initial wavefunction for our particle in a box results in

\[
\int_0^L A^2 \sin^2(3\pi x/L) \, dx = 1,
\]

whence

\[
A^2(L/2) = 1 \quad \text{so} \quad A = \sqrt{2/L}.
\]

Notice that then

\[
\psi_0(x) = \sqrt{\frac{2}{L}} \sin(3\pi x/L)
\]

has the proper dimensions.

Well, for this initial wavefunction, what are the values of

\[
D_n = \frac{2}{L} \int_0^L \psi_0(x) \sin(n\pi x/L) \, dx?
\]
They are

\[ D_n = \frac{2}{L} \sqrt{\frac{2}{L}} \int_0^L \sin(3\pi x/L) \sin(n\pi x/L) \, dx \]
\[ = \frac{2}{L} \sqrt{\frac{2}{L}} \times \begin{cases} 
L/2 & \text{for } n = 3 \\
0 & \text{for } n \neq 3 
\end{cases} \]
\[ = \sqrt{\frac{2}{L}} \delta_{n,3}, \]

so

\[ \psi(x, t) = \sqrt{\frac{2}{L}} e^{-\left(i/\hbar\right)E_3 t} \sin(3\pi x/L) \]

(1.17)

That’s it! For this particular initial wavefunction, the system remains always in that same wavefunction, except multiplied by an overall phase factor of \( e^{-\left(i/\hbar\right)E_3 t} \). This overall phase factor has no effect whatsoever on the probability density! Such states are called “stationary states”.

**Generic states.** Contrast the time development of stationary states with the time development of generic states. For example, suppose the initial wavefunction were

\[ \psi_0(x) = \frac{4}{5} \sqrt{\frac{2}{L}} \sin(3\pi x/L) + \frac{3}{5} \sqrt{\frac{2}{L}} \sin(7\pi x/L). \]

How does this state change with time? You should check two things: First, the wavefunction \( \psi_0(x) \) given here is normalized. Second, it evolves in time to

\[ \psi(x, t) = \frac{4}{5} \sqrt{\frac{2}{L}} e^{-\left(i/\hbar\right)E_3 t} \sin(3\pi x/L) + \frac{3}{5} \sqrt{\frac{2}{L}} e^{-\left(i/\hbar\right)E_7 t} \sin(7\pi x/L). \]

(1.18)

It takes a little effort to see exactly how the probability density changes with time, but it’s clear at a glance that the probability density does change with time. This is not a stationary state.

Let’s go back to the Einstein-de Broglie relation

\[ E = \hbar \omega. \]

de Broglie was never clear about what was “oscillating” with frequency \( \omega \). But now we have a better idea. In stationary state (1.17), the amplitude at every point oscillates with frequency \( E_3/\hbar \). Using the Einstein-de Broglie relation, we say this state has energy \( E_3 \).

In contrast, the amplitude in generic state (1.18) has no single oscillation: there’s a combination of frequency \( E_3/\hbar \) and frequency \( E_7/\hbar \). This state doesn’t have an energy, in the same way that a silver atom with \( \mu_x = +\mu_B \) doesn’t have a value of \( \mu_z \), in the same way that an atom in state \( |z+\rangle \) passing through an \( x \) analyzer loop doesn’t have a position, in the same way that love doesn’t have a color. Instead, we say that this state has amplitude \( \frac{4}{5} \) to have energy \( E_3 \) and amplitude \( \frac{3}{5} \) to have energy \( E_7 \).

We have uncovered the physical significance of the “separation constant” \( E_n \).
**Energy eigenstates.** How did the remarkable stationary states come about? Remember how they arose mathematically: We looked for solutions to

$$\left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] X_n(x) = E_n X_n(x),$$

and the solutions we found (for the infinite square well) were those functions

$$X_n(x) = \sin(\frac{n\pi x}{L})$$

that we later used as building blocks to build up *any* wavefunction. These now seem important enough that they warrant their own name. Because each is associated with a particularly energy $E_n$ we call them “energy states”. Because wavefunctions are usually denoted by Greek letters we give them the name $\eta_n(x)$ where the Greek letter $\eta$, pronounced “eta”, is supposed to suggest “energy” through alliteration. We write

$$\left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \eta_n(x) = E_n \eta_n(x),$$

and recognize this as one of those “unfair” problems where you must find not only the ODE solution $\eta_n(x)$, but you must also find the value of $E_n$. Such problems are given a half-German, half-English name: eigenproblems. We say that $\eta_n(x)$ represents a “stationary state”, or an “energy state”, or an “energy eigenstate”. (The German word *eigen* means “characteristic of” or “peculiar to” or “belonging to”. The eigenstate $\eta_3(x)$ is the state “belonging to” energy $E_3$.) [The piece in square brackets is called the “energy operator” or (for historical reasons) the “Hamiltonian operator”.

In contrast, the state

$$\psi(x,t) = \frac{4}{5} e^{-i\frac{E_3}{\hbar}t} \eta_3(x) + \frac{3}{5} e^{-i\frac{E_7}{\hbar}t} \eta_7(x)$$

does not “belong to” a particular energy, because it involves both $E_3$ and $E_7$. Instead, as already mentioned, this state has amplitude $\frac{4}{5}$ to have energy $E_3$ and amplitude $\frac{3}{5}$ to have energy $E_7$.

A particle trapped in a one-dimensional infinite square well problem cannot have any old energy: the only energies possible are $E_1$, $E_2$, $E_3$, … Some people gain from this the misimpression that the only states possible are the energy eigenstates $\eta_1(x)$, $\eta_2(x)$, $\eta_3(x)$, …, and that’s false. We’ve just dealt with a state that’s a superposition of two energy states with different energies.

---

4William Rowan Hamilton (1805–1865) made important contributions to mathematics, optics, classical mechanics, and astronomy. At the age of 22 years, while still an undergraduate, he was appointed professor of astronomy at his university and the Royal Astronomer of Ireland. He was the first person to point out the central role that energy can play in time development in the formal theory of classical mechanics. He was not related to the American founding father Alexander Hamilton.
Our journey into quantum mechanics started with the experimental fact of quantized energies of black-body radiation or of an atom. This inspired a search for quantized values of $\mu_z$, which in turn prompted discovery of the new phenomena of interference and entanglement. Interference experiments suggested the mathematical tool of amplitude, and generalizing amplitude from magnetic moment to position prompted the mathematical tool of wavefunction. We asked the obvious question of how wavefunction changed with time, and answering that question brought us back to energy quantization with deeper insight. As T.S. Eliot wrote

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.

1.1 Problem: Fourier sine series
Suppose the initial wavefunction is a pure real tent:

$$\psi_0(x) = \begin{cases} 0 & x < 0 \\ ax & 0 < x < L/2 \\ a(L - x) & L/2 < x < L \\ 0 & L < x \end{cases}$$

a. Sketch this initial wavefunction.

b. Show that, to insure normalization, you must use

$$a = \frac{2}{L} \sqrt{\frac{3}{L}}.$$

c. Verify that $\psi_0(x)$ has the proper dimensions.

d. The Fourier expansion coefficients are

$$D_n = \frac{2}{L} \int_0^L \psi_0(x) \sin(n\pi x/L) \, dx.$$

Before rushing in to evaluate this integral, pause to think! Show that, for this particular $\psi_0(x)$, when $n$ is even $D_n = 0$, while when $n$ is odd

$$D_n = \frac{4}{L} \int_0^{L/2} \psi_0(x) \sin(n\pi x/L) \, dx.$$

e. Now go ahead and evaluate $D_n$ for $n$ odd. (I used the substitution $\theta = n\pi x/L$, but there are other ways of doing it.)

f. Write out the Fourier sine series representation for $\psi_0(x)$. Check that it has the proper dimensions, that it satisfies $\psi_0(0) = \psi_0(L) = 0$, and that it satisfies $\psi_0(L/2) = aL/2$. For the last check, use the result (see, for example, H.B. Dwight, *Tables of Integrals and Other Mathematical Data*, fourth edition, Macmillan, 1961, equation 48.12)

$$1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots = \frac{\pi^2}{8}.$$
1.7 Other potentials

Is quantization peculiar to the infinite square well? No. We don’t have the mathematical tools to prove it, but in fact the infinite square well is entirely generic.

For any potential, there will be energy eigenvalues $E_n$ and energy eigenstates $\eta_n(x)$ that satisfy

$$\left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \eta_n(x) = E_n \eta_n(x),$$

and with the property that

$$\int_{-\infty}^{+\infty} \eta_m^*(x) \eta_n(x) \, dx = \delta_{m,n}. \quad (1.20)$$

Any wavefunction can be expressed as a “superposition” of energy states:

$$\psi_0(x) = \sum_{n=1}^{\infty} c_n \eta_n(x). \quad (1.21)$$

We say that the energy states “span the set of wavefunctions” or that they constitute a “basis”, but don’t let these fancy names fool you: they just mean that starting from the energy states, you can use linear combinations to build up any wavefunction. The basis states in quantum mechanics play the same role as building blocks in a child’s construction. Just as a child can build castles, roadways, or trees — anything she wants — out of building blocks, so we can build any wavefunction we want out of energy states.

Superposition is the mathematical reflection of the physical phenomenon interference. An example of quantal interference is “the atom passing through an interferometer doesn’t take either path; instead it has amplitude $c_a$ to take path $a$ and amplitude $c_b$ to take path $b$”. An example of superposition is “the particle with wavefunction $\psi_0(x)$ given above doesn’t have an energy; instead it has amplitude $c_n$ to have energy $E_n$, with $n = 1, 2, 3, \ldots$.” It requires no great sophistication to see that these are parallel statements.

Furthermore, this wavefunction evolves in time to

$$\psi(x, t) = \sum_{n=1}^{\infty} c_n e^{-iE_n t} \eta_n(x). \quad (1.22)$$

It should be clear by now that it’s important to solve eigenproblems: they are needed to find the “allowed” energies, and it’s also the easiest way to solve the time development problem. There are oodles of analytic and numerical techniques for solving them. Most of these are effective but merely technical: they find the answer, but don’t give any insight into the character of the resulting energy eigenstates. In contrast, the rough-and-ready technique of Tom Moore for sketching energy eigenstates doesn’t give rigorous solutions, but does provide a lot of insight. Read his book, *Six Ideas That Shaped Physics, Unit Q: Particles Behave Like Waves* (second edition, 2003) chapter Q11, “Energy Eigenfunctions”.
1.7. OTHER POTENTIALS

1.2 Problem: Error

1.3 Problem: Scaling
We’ve seen that a normalizable solution of the energy eigenequation for a square well relies on a tradeoff between the well width and the candidate energy \( E \). It makes sense that a suitable change in width can be offset by a suitable change in \( E \). This problem explores that tradeoff.

Suppose \( \eta(x) \) is a solution to
\[
-\frac{\hbar^2}{2m} \frac{d^2 \eta}{dx^2} + V(x) \eta(x) = E \eta(x)
\]
with a particular energy \( E \). Now consider a different potential energy function \( U(x) = s^2 V(sx) \). For example, if \( s = 2 \) then \( U(x) \) has the same shape as \( V(x) \), but has half the width and four times the height. Call \( \sigma(x) = \eta(sx) \). Show that \( \sigma(x) \) solves the energy eigenproblem for potential energy \( U(x) \), but with a different energy. Find that energy.
1.8 Energy loss

I said earlier that the energy eigenstate $\eta_7(x)$ was a “stationary state”: that if you started off in state $\eta_7(x)$ you would remain in that state forever. This seems to contradict the experimental fact that most of the atoms we find lying about are in their ground states. Why don’t they just stay in state $\eta_7(x)$ for ever and ever?

The answer is that if the Hamiltonian were exactly given by equation (1.19) then the atom would stay in that stationary state forever. But real atoms are subject to collisions and radiation which means that the Hamiltonian (1.19) is not exactly correct. The effect of these extra terms is to cause the atom to fall into the ground state of Hamiltonian (1.19).

1.9 Averages and operators

*Recall the interference experiment.* A single atom ambivates through the two branches of an interferometer.

Suppose the interferometer has width $w$, so that branch $a$ has position $x_a = +w/2$ while branch $b$ has position $x_b = -w/2$.

You know the drill: The atom has amplitude $c_a$ of taking branch $a$, amplitude $c_b$ of taking branch $b$. If a lamp is turned on while an interference experiment is proceeding, the probability of the atom appearing in branch $a$ is $|c_a|^2$, the probability of the atom appearing in branch $b$ is $|c_b|^2$. In other words, if the atom’s position is measured while the interference experiment is proceeding, the result would be $+w/2$ with probability $|c_a|^2$, $-w/2$ with probability $|c_b|^2$. Hence the average position is

$$+(w/2)|c_a|^2 - (w/2)|c_b|^2.$$  \hfill (1.23)

It seems weird to say “The atom doesn’t have a position but its average position is given by equation (1.23).” Indeed, it would be more accurate to say “The atom doesn’t have a position but if a light is turned on — or if position can be determined in some other way — then it does have a position, and the average position found will be given by equation (1.23).” This more accurate sentence is such a mouthful that it’s rarely said: People say the first, inaccurate sentence as shorthand for the second, correct sentence. (Sometimes people call this, not an “average”, but an “expected value” to emphasize this distinction.)
A particle moves in a potential. A particle is in state
\[ \psi(x,t) = \sum_{n=1}^{\infty} c_n e^{-(i/\hbar)E_n t} \eta_n(x). \] (1.24)

Here \( c_n \) is the amplitude that the particle has energy \( E_n \), so \( |c_n|^2 \) is the probability that, if the energy were measured, the result \( E_n \) would be found. The average energy is thus clearly
\[ \langle E \rangle = \sum_{n=1}^{\infty} |c_n|^2 E_n. \] (1.25)

Once again, it seems weird to have a formula for the average energy of a particle that doesn’t have an energy... sort of like finding the population density of unicorns. The meaning is that if the energy were measured, \( \langle E \rangle \) is the average of the energy that would be found.

**New expression for average energy.** The above expression for average energy is correct, but it’s difficult to use. Suppose a particle with wavefunction \( \psi(x) \) is subject to a potential energy function \( V(x) \). To find the average energy you must: (1) Solve the energy eigenproblem to find the eigenenergies \( E_n \) and the energy eigenstates \( \eta_n(x) \). (2) Write the wavefunction \( \psi(x) \) in the form \( \psi(x) = \sum_n c_n \eta_n(x) \). (3) Now that you know the energies \( E_n \) and the amplitudes (expansion coefficients) \( c_n \), execute the sum \( \sum_n |c_n|^2 E_n \). Whew! Isn’t there an easier way?

There is. The wavefunction (1.24) has time derivative
\[ \frac{\partial \psi(x,t)}{\partial t} = \sum_{n=1}^{\infty} c_n \left( -\frac{i}{\hbar} E_n \right) e^{-(i/\hbar)E_n t} \eta_n(x) \] (1.26)

and complex conjugate
\[ \psi^*(x,t) = \sum_{m=1}^{\infty} c_m^* e^{+(i/\hbar)E_m t} \eta_m^*(x). \] (1.27)

Thus
\[ \psi^*(x,t) \frac{\partial \psi(x,t)}{\partial t} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m^* c_n \left( -\frac{i}{\hbar} E_n \right) e^{-(i/\hbar)(E_n-E_m) t} \eta_m^*(x) \eta_n(x) \] (1.28)

and
\[ \int_{-\infty}^{+\infty} \psi^*(x,t) \frac{\partial \psi(x,t)}{\partial t} \ dx = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m^* c_n \left( -\frac{i}{\hbar} E_n \right) e^{-(i/\hbar)(E_n-E_m) t} \int_{-\infty}^{+\infty} \eta_m^*(x) \eta_n(x) \ dx. \] (1.29)

But the integral on the right is zero unless \( m = n \), in which case it is 1. Thus
\[ \int_{-\infty}^{+\infty} \psi^*(x,t) \frac{\partial \psi(x,t)}{\partial t} \ dx = \sum_{n=1}^{\infty} c_n^* c_n \left( -\frac{i}{\hbar} E_n \right) = -\frac{i}{\hbar} \langle E \rangle \] (1.30)

whence
\[ \langle E \rangle = \int_{-\infty}^{+\infty} \psi^*(x,t) \left( i\hbar \frac{\partial \psi(x,t)}{\partial t} \right) \ dx. \] (1.31)
Now, according to Schrödinger’s equation

\[
\hat{H} \frac{\partial \psi(x, t)}{\partial t} = \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(x, t),
\]

so

\[
\langle E \rangle = \int_{-\infty}^{+\infty} \psi^*(x, t) \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(x, t) \, dx.
\]

This is the expression for average energy that does not require solving the energy eigenproblem or expanding \( \psi(x) \) in energy eigenstates.

This expression generalizes immediately to other characteristics: For example the average position is

\[
\langle x \rangle = \int_{-\infty}^{+\infty} \psi^*(x, t) [x] \psi(x, t) \, dx,
\]

the average of position squared is

\[
\langle x^2 \rangle = \int_{-\infty}^{+\infty} \psi^*(x, t) [x^2] \psi(x, t) \, dx.
\]

1.4 Problem: Wavefunction vs. probability density

The wavefunction \( \psi(x) \) is not directly measurable, but can be inferred (up to an overall phase) through a number of position and interference experiments. The probability density \( |\psi(x)|^2 \) is measurable through a number of position experiments alone. These facts lead some to the misconception that the probability density tells the “whole story” of a quantal state. This problem demonstrates the falsehood of that misconception by presenting a series of wavefunctions, all with the same probability density, but each with a different expected energy. The wavefunctions are

\[
\psi(x) = Ae^{-x^2/2\sigma^2}e^{ikx},
\]

where \( A \) is a normalization constant.

a. (Mathematical preliminary.) Use integration by parts to show that

\[
\int_{-\infty}^{+\infty} e^{-x^2} \, dx = 2 \int_{-\infty}^{+\infty} x^2 e^{-x^2} \, dx.
\]

b. When the particle is free, \( V(x) = 0 \), find the expected energy. If you use the above result, you will not need to evaluate any integral nor find the normalization constant.

1.10 The classical limit of quantum mechanics

If the theory of quantum mechanics is correct, then when quantum mechanics is applied to big things, it must give the results of classical mechanics.
1.10. **THE CLASSICAL LIMIT OF QUANTUM MECHANICS**

Because quantum mechanics emphasizes potential energy $V(x)$, and classical mechanics emphasizes force $F$, let’s remind ourselves how they’re related. The definition of potential energy (in one dimension) is

$$V(x) - V(x_0) = - \int_{x_0}^{x} F(x') \, dx',$$

(1.36)

where $F(x)$ is the force function — i.e., the force that would be exerted on the particle if the particle were at position $x$. (I’m not saying that the particle *is* at $x$, I’m saying that’s what the force *would be* if the particle *were* at $x$.) Taking the derivative of both sides with respect to $x$ (and using the fundamental theorem of calculus on the right)

$$\frac{dV(x)}{dx} = -F(x).$$

(1.37)

The classical time development equation is

$$F(x) = m \frac{d^2 x(t)}{dt^2}$$

which, of course, contains no references to $\hbar$.

In the classical limit, the quantal spread in $x$ will be small. We ask how the expected position (the “center of the wavepacket”) moves:

$$\langle x \rangle = \int_{-\infty}^{+\infty} x \psi^*(x,t) \psi(x,t) \, dx,$$

(1.38)

so

$$\frac{d\langle x \rangle}{dt} = \int_{-\infty}^{+\infty} x \frac{\partial \psi^*(x,t)}{\partial t} \psi(x,t) \, dx + \int_{-\infty}^{+\infty} x \psi^*(x,t) \frac{\partial \psi(x,t)}{\partial t} \, dx.$$

(1.39)

But the Schrödinger equation tells us how wavefunction $\psi(x,t)$ changes with time:

$$\frac{\partial \psi(x,t)}{\partial t} = -\frac{i}{\hbar} \left[ -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x,t)}{\partial x^2} + V(x) \psi(x,t) \right]$$

(1.40)

and

$$\frac{\partial \psi^*(x,t)}{\partial t} = +\frac{i}{\hbar} \left[ -\frac{\hbar^2}{2m} \frac{\partial^2 \psi^*(x,t)}{\partial x^2} + V(x) \psi^*(x,t) \right].$$

(1.41)

(From here on I’m going to write $\psi(x,t)$ as $\psi$ and $V(x)$ as $V$.) Thus

$$\frac{d\langle x \rangle}{dt} = -\frac{i}{\hbar} \left\{ \int_{-\infty}^{+\infty} x \left[ -\frac{\hbar^2}{2m} \frac{\partial^2 \psi^*}{\partial x^2} + V \psi^* \right] \psi \, dx - \int_{-\infty}^{+\infty} x \psi^* \left[ -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V \psi \right] \, dx \right\}$$

(1.42)

$$= -\frac{i}{\hbar} \left\{ -\frac{\hbar^2}{2m} \int_{-\infty}^{+\infty} x \frac{\partial^2 \psi^*}{\partial x^2} \psi \, dx - \int_{-\infty}^{+\infty} x \psi^* \frac{\partial^2 \psi}{\partial x^2} \, dx \right\} + \int_{-\infty}^{+\infty} xV \psi^* \psi \, dx - \int_{-\infty}^{+\infty} x \psi^* V \psi \, dx.$$

(1.43)

$$= -\frac{i}{2m} \left\{ \int_{-\infty}^{+\infty} xV \psi^* \psi \, dx - \int_{-\infty}^{+\infty} x \psi^* V \psi \, dx \right\}$$

(1.44)
It seems a bit odd that the two terms involving potential energy cancel, so no explicit dependence on $V(x)$ appears in this result, but we’ll just push on.

Can we say anything about integrals such as the second integral in square brackets above? Surprisingly, the answer is yes. If we define

$$ f(x) = x\psi^* \quad \text{and} \quad g(x) = \frac{\partial \psi}{\partial x} $$

then

$$ \int_{-\infty}^{+\infty} x\psi^* \frac{\partial^2 \psi}{\partial x^2} \, dx = \int_{-\infty}^{+\infty} f(x)g'(x) \, dx $$

which suggests integration by parts:

$$ \int_{-\infty}^{+\infty} f(x)g'(x) \, dx = [f(x)g(x)]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} f'(x)g(x) \, dx. \quad (1.47) $$

Now remember what the wavefunction is normalized, so it has to fall to zero at both infinity and negative infinity. Typically the slope $\partial \psi/\partial x$ also falls to zero at both infinity and negative infinity, and does so very rapidly — much more rapidly than linearly. (There are exceptions to these typical behaviors, such as scattering wave functions, and in these atypical cases this argument has to be rethought.) The upshot is that in typical situations

$$ [f(x)g(x)]_{-\infty}^{+\infty} = 0 \quad (1.48) $$

so

$$ \int_{-\infty}^{+\infty} x\psi^* \frac{\partial^2 \psi}{\partial x^2} \, dx = -\int_{-\infty}^{+\infty} \frac{\partial (x\psi^*)}{\partial x} \frac{\partial \psi}{\partial x} \, dx. \quad (1.49) $$

We’ll use this trick several times... I’ll just call it the “integration-by-parts trick”.

Applying this trick to both integrals of equation (1.44) gives

$$ \frac{d\langle x \rangle}{dt} = -\frac{i}{2m} \left[ \int_{-\infty}^{+\infty} \frac{\partial (x\psi)}{\partial x} \frac{\partial \psi^*}{\partial x} \, dx - \int_{-\infty}^{+\infty} \frac{\partial (x\psi^*)}{\partial x} \frac{\partial \psi}{\partial x} \, dx \right] $$

$$ = -\frac{i}{2m} \left[ \int_{-\infty}^{+\infty} x \frac{\partial \psi}{\partial x} \frac{\partial \psi^*}{\partial x} \, dx + \int_{-\infty}^{+\infty} \psi \frac{\partial \psi^*}{\partial x} \frac{\partial \psi}{\partial x} \, dx - \int_{-\infty}^{+\infty} x \frac{\partial \psi^*}{\partial x} \frac{\partial \psi}{\partial x} \, dx - \int_{-\infty}^{+\infty} \psi^* \frac{\partial \psi}{\partial x} \, dx \right] $$

$$ = -\frac{i}{2m} \left[ +\int_{-\infty}^{+\infty} \psi \frac{\partial \psi^*}{\partial x} \, dx - \int_{-\infty}^{+\infty} \psi^* \frac{\partial \psi}{\partial x} \, dx \right] $$

$$ = -\frac{\hbar}{m} \Im \{ \int_{-\infty}^{+\infty} \psi \frac{\partial \psi^*}{\partial x} \, dx \}. \quad (1.50) $$

Notice that $d\langle x \rangle/dt$ is real, as it must be. And notice that the dimensions are the same on both sides. (This isn’t proof that we’ve made no algebra errors, but if our expression for $d\langle x \rangle/dt$ had been complex, or had been dimensionally incorrect, then that would have been proof that we had made algebra errors.)

All this is fine and good, but it takes us only part way to our goal. This is clearly not a classical equation... it contains $\hbar$ right there! Since the classical $F = ma$ involves the second derivative of position with respect to time, we take one more derivative of $\langle x \rangle$, finding

$$ \frac{d^2\langle x \rangle}{dt^2} = -\frac{\hbar}{m} \Im \left\{ \int_{-\infty}^{+\infty} \frac{\partial \psi^*}{\partial t} \frac{\partial \psi}{\partial x} \, dx + \int_{-\infty}^{+\infty} \psi^* \frac{\partial \psi}{\partial t} \frac{\partial \psi}{\partial x} \, dx \right\}. \quad (1.51) $$
The second-order derivative on the right looks particularly grotesque, so we use the integration-by-parts trick to get rid of it:

\[
\frac{d^2}{dt^2} \langle x \rangle = \frac{\hbar}{m} \Im \left\{ \int_{-\infty}^{+\infty} \frac{\partial^2 \psi^*}{\partial x^2} \frac{\partial \psi}{\partial x} dx \right\} - \frac{2\hbar}{m} \Im \left\{ \int_{-\infty}^{+\infty} \frac{\partial \psi^*}{\partial x} \frac{\partial \psi}{\partial x} dx \right\}.
\] (1.52)

Now use Schrödinger’s equation:

\[
\frac{d^2}{dt^2} \langle x \rangle = -2\hbar m \Re \left\{ \int_{-\infty}^{+\infty} \frac{\partial \psi^*}{\partial x} \left\{ -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V \psi \right\} dx \right\},
\] (1.53)

but let’s apply the integration-by-parts trick to the first integral:

\[
\int_{-\infty}^{+\infty} \frac{\partial \psi^*}{\partial x} \frac{\partial^2 \psi}{\partial x^2} dx = -\int_{-\infty}^{+\infty} \frac{\partial^2 \psi^*}{\partial x^2} \frac{\partial \psi}{\partial x} dx.
\] (1.54)

Look at that... two of the \(\hbar\)s have canceled out! We’re not home yet because there’s still an \(\hbar\) within the square brackets, but we’re certainly making progress. We have that

\[
\frac{d^2}{dt^2} \langle x \rangle = -\frac{\hbar}{m} \Re \left\{ \int_{-\infty}^{+\infty} \frac{\partial \psi^*}{\partial x} V \psi dx \right\},
\] (1.55)

but let’s apply the integration-by-parts trick to this integral:

\[
\int_{-\infty}^{+\infty} \frac{\partial \psi^*}{\partial x} V \psi dx = -\frac{\hbar}{m} \int_{-\infty}^{+\infty} \frac{\partial \psi^*}{\partial x} V \psi dx + \int_{-\infty}^{+\infty} \frac{\partial V}{\partial x} \psi dx.
\] (1.56)

Think about this for a minute: if the integral on the left is \(z\), this equation says that \(z = -z^*\), whence \(z\) is pure imaginary or \(\Re\{z\} = 0\). Thus

\[
\frac{d^2}{dt^2} \langle x \rangle = -\frac{1}{m} \int_{-\infty}^{+\infty} \frac{\partial \psi^*}{\partial x} V \psi dx,
\] (1.57)

an expression devoid of \(\hbar\)s! Apply the integration-by-parts trick to this integral:

\[
\int_{-\infty}^{+\infty} \frac{\partial \psi^*}{\partial x} V \psi dx = -\int_{-\infty}^{+\infty} \psi^* \frac{\partial V}{\partial x} \psi dx - \int_{-\infty}^{+\infty} \psi \frac{\partial \psi}{\partial x} \frac{\partial V}{\partial x} \psi dx,
\]

Plugging this result back into equation (1.57) gives

\[
\frac{d^2}{dt^2} \langle x \rangle = -\frac{1}{m} \int_{-\infty}^{+\infty} \psi \frac{\partial V}{\partial x} \psi dx.
\] (1.59)

But remember that the force function is \(F(x) = -\partial V/\partial x\), so

\[
\frac{d^2}{dt^2} \langle x \rangle = \frac{1}{m} \int_{-\infty}^{+\infty} F(x) \psi^* (x, t) \psi (x, t) dx = \frac{1}{m} (F(x)).
\] (1.60)
CHAPTER 1. THE QUANTUM MECHANICS OF POSITION

There it is —

\[ \langle F(x) \rangle = m \frac{d^2(x)}{dt^2}. \]  

(1.61)

This result is called *Ehrenfest’s theorem.*

Note that the result is not

\[ F(\langle x \rangle) = m \frac{d^2(\langle x \rangle)}{dt^2}. \]  

(1.62)

1.11 Position plus spin

We investigated particles with magnetic moment, like the silver atom, doing our best to treat the quantum mechanics of spin and ignore the quantum mechanics of position. And then we investigated the position of particles, ignoring the quantum mechanics of magnetic moment. (For historical reasons, the magnetic moment is said to reflect the intrinsic “spin” of the particle.) It’s time to weld the two pieces together.

This is achieved in a straightforward way. Think back to our discussion of bin amplitudes for a single particle in one dimension. We asked the question “What is the amplitude for the particle to be found in bin \( i \)” But if the particle has a two-state spin, like a silver atom, we have to ask the question “What is the amplitude for the particle to be found in bin \( i \) with spin up?” Or “What is the amplitude for the particle to be found in bin \( i \) with spin down?” (Alternatively you could ask for \( x \)-spin positive or \( x \)-spin negative, or \( 17^\circ \)-spin positive or negative, or for the projection on any other axis.) We’ll call the answer to both questions \( \psi_j(m) \), where \( m = \pm 1 \). Do the standard thing: divide by the square root of bin size and take the limit as bin size shrinks to zero. This quantity becomes an amplitude density (wavefunction)

\[ \psi(x, m) \quad \text{where } -\infty < x < +\infty \text{ and } m = \pm 1. \]

Sometimes people write this as two separate wavefunctions:

\[ \psi_+(x) = \psi(x, +1) \quad \text{and} \quad \psi_-(x) = \psi(x, -1). \]

And sometimes they write it as a “spatial part” times a “spin part”:

\[ \psi_+(x) = \phi(x)\chi_+ \quad \text{and} \quad \psi_-(x) = \xi(x)\chi_. \]

But don’t let the notation fool you: all of these expressions represent the same thing.

It might happen that the two spatial parts are equal, \( \phi(x) = \xi(x) \), in which case we can say

\[ \psi(x, m) = \phi(x)\chi(m) \quad \text{where } \chi(+1) = \chi_+ \text{ and } \chi(-1) = \chi_. \]

But, this is not always the case.

---

5Paul Ehrenfest (1880–1933) contributed to relativity theory, quantum mechanics, and statistical mechanics, often by pointing out concrete difficulties in these theories. As a result, several telling arguments have names like “Ehrenfest’s paradox” or “Ehrenfet’s urn” or “the Ehrenfest dog-flea model”.


If the atom were nitrogen, with four possible spin projections, then you’d have to ask “What is the amplitude for the nitrogen atom to be found in bin \( i \) with vertical spin projection +2?” or +1 or −1 or −2. Or you might ask “What is the amplitude for the nitrogen atom to be found in bin \( i \) with a projection on the 67° axis of −1?” After asking these questions and taking the appropriate limits, the relevant wavefunction will be

\[ \psi(x, m) \quad \text{where} \ -\infty < x < +\infty \ \text{and} \ m = -3, -1, +1, +3. \]

If the atom moves in three dimensions, the wavefunction will take the form

\[ \psi(x, y, z, m) \equiv \psi(x), \]

where the single symbol \( x \) stands for the four arguments \( x, y, z, m \).

**Epilogue**

The quantum mechanics of position is very strange, yes.
And it’s very difficult, yes.
But it’s also very wonderful.
Chapter 2

Identical Particles

2.1 Two or three identical particles

Please review section 1.4, “Amplitude density: Two particles in one or three dimensions”. In that section we talked about two particles, say an electron and a neutron. We set up a grid, discussed bin amplitudes $\psi_{i,j}$, and talked about the limit as the width of each bin shrank to zero.

There is a parallel argument for two identical particles, with one twist. Here is the situation when one particle is found in bin 5, the other in bin 8:

No difference, of course . . . that’s the meaning of “identical”. And of course this holds not only for bins 5 and 8, but for any pair of bins $i$ and $j$, even if $i = j$. (The two particles might not interact, in which case it would be perfectly possible for them to occupy the same bin at the same time.)

In terms of a given grid of bin amplitudes, $\psi_{i,j}$, realize that $\psi_{j,i}$ represents the same state as $\psi_{i,j}$. That is $\psi_{j,i} = s\psi_{i,j}$, where the number $s$ is a possibly complex number with square modulus unity. (The name $s$ comes from “swap”. We’ve swapped the subscripts.) The quantity $s$ is a number . . . not a function of $i$ or $j$. 

32
Thus, for example, $\psi_{8,5} = s\psi_{5,8}$, $\psi_{7,9} = s\psi_{9,7}$, $\psi_{5,8} = s\psi_{8,5}$. ... Wait. What was that last one? Let’s put together the first and last examples:

$$\psi_{8,5} = s\psi_{5,8} = s(s\psi_{8,5}) = s^2\psi_{8,5}.$$ 

Clearly, $s^2 = 1$, so $s$ can’t be any old complex number with square modulus unity. It can be only $s = +1$ or $s = -1$.

Execute the now-familiar program of turning bin amplitudes into amplitude density, that is wavefunction, to find that

$$\psi(x_A, x_B) = +\psi(x_B, x_A) \quad \text{or} \quad \psi(x_A, x_B) = -\psi(x_B, x_A).$$

The first kind of wavefunction is called “symmetric under swapping”, the second is called “antisymmetric under swapping”. This requirement for symmetry or antisymmetry is called the Pauli principle.

It might distress you to see variables like $x_A$: Doesn’t $x_A$ mean the position of particle “A” while $x_B$ means the position of particle “B”? So doesn’t this terminology label the particles as “A” and “B”, and doesn’t that violate our initial requirement that the particles be identical? The answer is that this terminology does not label one particle “A” and the other particle “B”. Instead, it labels one point “A” and the other point “B”. For that reason I like to call these wavefunctions “(anti)symmetric under swap of coordinates”. But you’ll hear people using terms like “(anti)symmetric under particle swapping” or “under particle interchange” or “under particle exchange”.

What if the particles move in three-dimensional space, and what if they have spin? In that case, you must swap all the coordinates: That is

$$\psi(x_A, x_B) = +\psi(x_B, x_A) \quad \text{or} \quad \psi(x_A, x_B) = -\psi(x_B, x_A).$$

From our argument so far, two identical electrons might be in a symmetric or an antisymmetric state. Similarly for two identical neutrons, two identical silver atoms, etc. But it’s an empirical fact that the swap symmetry depends only on the kind of particle involved: Two electrons are always antisymmetric under swapping. Two $^4$He atoms are always symmetric. Particles that are always symmetric under swapping are called bosons; those that are always antisymmetric under swapping are called fermions.

---

1. Wolfgang Pauli (1900–1958), a Vienna-born Swiss physicist, was one of the founders of quantum mechanics. In 1924 he proposed the “exclusion principle”, ancestor of today’s symmetry/antisymmetry requirement; in 1926 he produced the first solution for the energy eigenproblem for atomic hydrogen; in 1930 he proposed the existence of the neutrino, a prediction confirmed experimentally in 1956; in 1934 he and “Viki” Weisskopf discovered how to make sense of relativistic quantum mechanics by realizing that the solutions to relativistic equations were not wavefunctions but instead creation and annihilation operators (check this last claim). He originated the insult, applied to ideas that cannot be tested, they they are “not even wrong”.

2. Satyendra Bose (1894–1974) of India made contributions in fields ranging from chemistry to school administration, but his signal contribution was elucidating the statistics of photons. Remarkably, he made this discovery in 1922, three years before Schrödinger developed the concept of wavefunction.

3. Enrico Fermi (1901–1954) of Italy excelled in both experimental and theoretical physics. He directed the building of the first nuclear reactor and produced the first theory of the weak nuclear interaction. The Fermi surface in the physics of metals was named in his honor. He elucidated the statistics of what are now called fermions in 1926. He produced so many thoughtful
What about three particles? The wavefunction of three bosons must be completely symmetric, that is, symmetric under swaps of any coordinate pair:

\[ +\psi(x_A, x_B, x_C) \]
\[ = +\psi(x_A, x_C, x_B) \]
\[ = +\psi(x_C, x_A, x_B) \]
\[ = +\psi(x_C, x_B, x_A) \]
\[ = +\psi(x_B, x_C, x_A) \]
\[ = +\psi(x_B, x_A, x_C). \]

(These 6 = 3! permutations are listed in the sequence called “plain changes” or “the Johnson-Trotter sequence”. This sequence has the benefit that each permutation differs from its predecessor by a single swap of adjacent letters.\(^4\)) Whereas the wavefunction of three fermions must be completely antisymmetric, that is, antisymmetric under swaps of any coordinate pair:

\[ +\psi(x_A, x_B, x_C) \]
\[ = -\psi(x_A, x_C, x_B) \]
\[ = +\psi(x_C, x_A, x_B) \]
\[ = -\psi(x_C, x_B, x_A) \]
\[ = +\psi(x_B, x_C, x_A) \]
\[ = -\psi(x_B, x_A, x_C). \]

Consequences of the Pauli principle

Does the requirement of symmetry or antisymmetry under coordinate swapping have any consequences? Here’s an immediate one for fermions: Take both \( x_A = X \) and \( x_B = X \). Now when these coordinates are swapped, you get back to where you started:

\[ \psi(X, X) = -\psi(X, X) \quad \text{so} \quad \psi(X, X) = 0. \]

That is, the probability of two identical fermions having the same coordinates is zero.

This illustrates the more general rule that:

Symmetric implies “huddled together”, antisymmetric implies “spread apart”.

---

Many more illustrations of this general rule are given through my computer program Ident.

In our everyday experience, when two particles tend to huddle together, it’s because they’re attracted to each other through a force. When two particles tend to spread apart, it’s because they’re repelled from each other through a force. This quantal case is vastly different. The cause of the huddling and spreading is not a force — it occurs for identical particles whether they interact or not. The cause is instead the symmetry/antisymmetry requirement: not a force like a hammer blow, but a piece of mathematics!

In our everyday experience, when two people tend to huddle together or to spread apart, it’s for emotional reasons. Therefore it’s difficult to come up with terms for the behavior of identical particles that don’t suggest either forces or emotions ascribed to particles: attract, repel; congregate, avoid; flock, scatter. “Huddled together” and “spread apart” are the best words I’ve been able to devise.

### 2.2 Symmetrization and antisymmetrization

Given the importance of wavefunctions symmetric or antisymmetric under coordinate swapping, it makes sense to investigate the mathematics of such “permutation symmetry”.

Suppose I start with a garden-variety function \( f(x_A, x_B, x_C) \). Can I use that function as a “seed” to build a completely symmetric or antisymmetric function? I can. The function

\[
+f(x_A, x_B, x_C) \\
+f(x_A, x_C, x_B) \\
+f(x_C, x_A, x_B) \\
+f(x_C, x_B, x_A) \\
+f(x_B, x_C, x_A) \\
+f(x_B, x_A, x_C)
\]

is completely symmetric while the function

\[
+f(x_A, x_B, x_C) \\
-f(x_A, x_C, x_B) \\
+f(x_C, x_A, x_B) \\
-f(x_C, x_B, x_A) \\
+f(x_B, x_C, x_A) \\
-f(x_B, x_A, x_C)
\]

is completely antisymmetric.

This trick is often used when the seed function is a product,

\[
f(x_A, x_B, x_C) = f_1(x_A)f_2(x_B)f_3(x_C),
\]
in which case you may think of the symmetrization/antisymmetrization process as being the sum over all permutations of the coordinates $x_A$, $x_B$, and $x_C$, as above, or as the sum over all permutations of the functions $f_1(x)$, $f_2(x)$, and $f_3(x)$: The function

$$
+f_1(x_A)f_2(x_B)f_3(x_C)
$$

is completely symmetric while the function

$$
+f_1(x_A)f_2(x_B)f_3(x_C)
- f_1(x_A)f_3(x_B)f_2(x_C)
+f_3(x_A)f_1(x_B)f_2(x_C)
- f_3(x_A)f_2(x_B)f_1(x_C)
+f_2(x_A)f_3(x_B)f_1(x_C)
- f_2(x_A)f_1(x_B)f_3(x_C)
$$

is completely antisymmetric. Some people write this last expression as the determinant of a matrix

$$
\begin{vmatrix}
  f_1(x_A) & f_2(x_A) & f_3(x_A) \\
  f_1(x_B) & f_2(x_B) & f_3(x_B) \\
  f_1(x_C) & f_2(x_C) & f_3(x_C)
\end{vmatrix},
$$

and call it the “Slater determinant”. I personally think this terminology confuses the issue (the expression works only if the seed function is a product of one-variable functions, it suppresses the delightful and useful “plain changes” sequence of permutations, plus I never liked determinants to begin with), but it’s widely used.

### 2.1 Problem: Antisymmetrizing the symmetric

a. There is one function that is both completely symmetric and completely antisymmetric. What is it?

b. Suppose the seed function is symmetric under swaps of the first two coordinates

$$
f(x_A, x_B, x_C) = f(x_B, x_A, x_C)
$$

and the antisymmetrization process (2.4) is executed. What is the result?

---

5 John C. Slater (1900–1976), American theoretical physicist who made major contributions to our understanding of atoms, molecules, and solids. Also important as a teacher, textbook author, and administrator.
c. Repeat part (b) for a seed function symmetric under swaps of the last two coordinates.

d. Repeat part (b) for a seed function symmetric under swaps of the first and third coordinates.

e. Suppose the seed function is a product as in equation (2.5), and two of the functions happen to be equal. What is the result of the antisymmetrization process?

2.3  \textbf{Energy states for two identical, noninteracting particles}

A particle moves subject to some potential energy. There are $M$ energy eigenstates (where usually $M = \infty$)

\[ \eta_1(x), \eta_2(x), \eta_3(x), \ldots, \eta_M(x). \]

Now two non-identical particles move subject to the same potential energy. They have the same mass, and do not interact with each other. You can see what the energy eigenstates are: State $\eta_3(x_A)\eta_8(x_B)$ has energy $E_3 + E_8$. There’s necessarily a degeneracy, because the different state $\eta_8(x_A)\eta_3(x_B)$ has the same energy. The energy eigenbasis has $M^2$ elements. Any state can be represented as a combination of these elements. The states are normalized. I could go on, but the picture is clear: The fact that there are two rather than one particles is unimportant: this energy eigenbasis has all the properties you expect of an energy eigenbasis.

This basis consists of unentangled states, but of course that’s just a coincidence. You could replace the two basis states

\[ \eta_3(x_A)\eta_8(x_B) \quad \text{and} \quad \eta_8(x_A)\eta_3(x_B) \]

with, for example,

\[ \frac{4}{5}\eta_3(x_A)\eta_8(x_B) + \frac{3}{5}\eta_8(x_A)\eta_3(x_B) \quad \text{and} \quad -\frac{3}{5}\eta_3(x_A)\eta_8(x_B) + \frac{4}{5}\eta_8(x_A)\eta_3(x_B). \]

Or, for that matter, with the two states

\[ \cos(\theta)\eta_3(x_A)\eta_8(x_B) + \sin(\theta)\eta_8(x_A)\eta_3(x_B) \quad \text{and} \quad -\sin(\theta)\eta_3(x_A)\eta_8(x_B) + \cos(\theta)\eta_8(x_A)\eta_3(x_B) \]

for any value of $\theta$.

Now two identical, non-interacting particles move subject to the same potential energy. In this case we don’t want a basis from which we can build any wavefunction: We want a basis from which we can build any symmetric wavefunction, or a basis from which we can build any antisymmetric wavefunction.

The basis for antisymmetric wavefunctions has elements like

\[ \frac{1}{\sqrt{2}} [-\eta_3(x_A)\eta_8(x_B) + \eta_8(x_A)\eta_3(x_B)]. \]

The basis for symmetric wavefunctions has elements like

\[ \frac{1}{\sqrt{2}} [+\eta_3(x_A)\eta_8(x_B) + \eta_8(x_A)\eta_3(x_B)]. \]
plus elements like 

\[ \eta_8(x_A) \eta_8(x_B). \]

You should convince yourself that there are \( \frac{1}{2} M(M - 1) \) elements in the antisymmetric basis and \( \frac{1}{2} M(M + 1) \) elements in the symmetric basis. Notice that entangled states are a strict necessity in this basis.

The basis for antisymmetric wavefunctions united with the basis for symmetric wavefunctions produces a basis for any wavefunction. This is a peculiarity of the two-particle case, and reflects the fact that any two-variable function is the sum of a completely symmetric function and a completely antisymmetric function. It is not true that any three-variable function is the sum of a completely symmetric and a completely antisymmetric function. For three noninteracting particles, the general basis has \( M^3 \) elements, the basis for antisymmetric wavefunctions has \( \frac{1}{6} M(M - 1)(M - 2) \) elements, and the basis for symmetric wavefunctions has \( \frac{1}{6} M(M + 1)(M + 2) \) elements. If you enjoy mathematical puzzles, you will enjoy proving these statements for yourself. But we won’t need them for this course.

### 2.4 Spin plus space, two electrons

Two electrons move subject to the same potential. The Hamiltonian has no spin-dependent terms. Pretend the two electrons don’t interact. (Perhaps a better name for this section would be “Spin plus space, two noninteracting spin-\(\frac{1}{2}\) fermions” but yikes, how long do you want this section’s title to be? Should I add “non-relativistic”?)

The spatial energy eigenstates for one electron are \( \eta_n(\vec{x}) \) for \( n = 1, 2, \ldots M/2 \). Thus the full energy eigenstates for one electron are the \( M \) states \( \eta_n(\vec{x})\chi_+ \) and \( \eta_n(\vec{x})\chi_- \). Now the question: What are the energy eigenstates for the two non-interacting electrons?

Well, what two-particle states can we build from the one-particle states with \( n = 1 \) and \( n = 3 \)? These correspond to four one-particle states:

\[
\begin{align*}
&\eta_1(\vec{x})\chi_+, \quad (2.9) \\
&\eta_1(\vec{x})\chi_-, \quad (2.10) \\
&\eta_3(\vec{x})\chi_+, \quad (2.11) \\
&\eta_3(\vec{x})\chi_. \quad (2.12)
\end{align*}
\]

The antisymmetric combination of (2.9) with itself vanishes. The antisymmetric combination of (2.9) with (2.10) is a combination of \( n = 1 \) with \( n = 1 \), not of \( n = 1 \) with \( n = 3 \). The antisymmetric combination of (2.9) with (2.11) is

\[
\eta_1(\vec{x}_A)\chi_+(A)\eta_3(\vec{x}_B)\chi_+(B) - \eta_3(\vec{x}_A)\chi_+(A)\eta_1(\vec{x}_B)\chi_+(B).
\] (2.13)

The antisymmetric combination of (2.9) with (2.12) is

\[
\eta_1(\vec{x}_A)\chi_+(A)\eta_3(\vec{x}_B)\chi_-(B) - \eta_3(\vec{x}_A)\chi_-(A)\eta_1(\vec{x}_B)\chi_+(B).
\] (2.14)
The antisymmetric combination of (2.10) with (2.11) is
\[ \eta_1(\vec{x}_A)\chi_-(A)\eta_3(\vec{x}_B)\chi_+(B) - \eta_3(\vec{x}_A)\chi_+(A)\eta_1(\vec{x}_B)\chi_-(B). \] (2.15)

The antisymmetric combination of (2.10) with (2.12) is
\[ \eta_1(\vec{x}_A)\chi_-(A)\eta_3(\vec{x}_B)\chi_-(B) - \eta_3(\vec{x}_A)\chi_-(A)\eta_1(\vec{x}_B)\chi_+(B). \] (2.16)

Finally, the antisymmetric combination of (2.11) with (2.12) is a combination of \( n = 3 \) with \( n = 3 \), not of \( n = 1 \) with \( n = 3 \).

State (2.13) factorizes into a convenient space times spin form:
\[ \eta_1(\vec{x}_A)\chi_-(A)\eta_3(\vec{x}_B)\chi_-(B) = \left[ \eta_1(\vec{x}_A)\eta_3(\vec{x}_B) - \eta_3(\vec{x}_A)\eta_1(\vec{x}_B) \right] \chi_+(A)\chi_+(B). \] (2.17)

Before I proceed any further I confess that I’m sick and tired of writing all these \( \eta \)s and \( \chi \)s and As and Bs that convey no information. I always write the \( \eta \) in front of the \( \chi \). I always write the As in front of the Bs. You’ll never confuse an \( \eta \) with a \( \chi \), because the \( \eta \)s are labeled 1, 3 while the \( \chi \)s are labeled +, −. I introduce a notation that takes all this for granted, so neither you nor I have to write the same thing out over and over again. In this notation, the above equation reads
\[ |1+, 3+\rangle - |3+, 1+\rangle = \left[ |1, 3\rangle - |3, 1\rangle \right] |+, +\rangle. \] (2.18)

In this new notation the states (2.13) through (2.16) read
\[ \begin{align*}
|1+, 3+\rangle - |3+, 1+\rangle &= |1, 3\rangle - |3, 1\rangle |+, +\rangle, \\
|1+, 3-\rangle - |3-, 1+\rangle &= |1-, 3+\rangle - |3+, 1-\rangle, \\
|1-, 3+\rangle - |3+, 1-\rangle &= |1, 3\rangle - |3, 1\rangle |-, -\rangle.
\end{align*} \] (2.19-2.22)

Well, this is cute. Two of the four states have this convenient space times spin form... and furthermore these two have the same spatial wavefunction! Two other states, however, don’t have this convenient form.

One thing to do about this is nothing. There’s no requirement that states have a space times spin form. But in this two-electron case there’s a slick trick that enables us to put the states into that form.

Notice that all four states (2.19) through (2.22) have the same energy, namely \( E_1 + E_3 \). So I can make linear combinations of the states for form other equally good states. Can I make a combination of states (2.20) and (2.21) that does factorize into space times spin? Nothing ventured, nothing gained. Let’s try it.

\[ \alpha |1+, 3-\rangle - |3-, 1+\rangle + \beta |1-, 3+\rangle - |3+, 1-\rangle \]

\[ = |1, 3\rangle \left[ \alpha |+, -\rangle + \beta |-, +\rangle \right] - |3, 1\rangle \left[ \alpha |-, +\rangle + \beta |+, -\rangle \right] \]
This will factorize only if the left term in square brackets is proportional to the right term in square brackets:

\[ \alpha|+,-\rangle + \beta|-,+,\rangle = c \left( \beta|+,\rangle + \alpha|-,\rangle \right), \]

that is only if

\[ \alpha = c\beta \quad \text{and} \quad \beta = c\alpha. \]

Combining these two equations results in \( c = \pm 1 \). If \( c = +1 \) then the combination results in the state

\[ |1,3⟩ - |3,1⟩ \alpha \left[ |+,\rangle + |-,\rangle \right] \]

whereas when \( c = -1 \) the result is

\[ |1,3⟩ + |3,1⟩ \alpha \left[ |+,\rangle - |-,\rangle \right]. \]

Putting all this together and, for the sake of good form, insuring normalized states, we find that that two-electron energy states can be written as

\[ \frac{1}{\sqrt{2}} \left( |1,3⟩ - |3,1⟩ \right) \alpha \left[ |+,\rangle + |-,\rangle \right] \]

The first three of these states have spatial wavefunctions that are antisymmetric under coordinate swaps and spin wavefunctions that are symmetric under coordinate swaps — these are called “ortho states”. The last one has a symmetric spatial wavefunction and an antisymmetric spin wavefunction — these are called “para states”. Our discussion on page 34 demonstrates that in ortho states, the two electrons tend to spread apart; in para states, they tend to huddle together.

2.2 Problem: Combining a spatial one-particle state with itself

What two-particle states can we build from the one-particle spatial state with \( n = 3 \)? How many of the resulting states are ortho, how many para?

2.3 Problem: Ortho and para accounting

Show that in our case with \( M/2 \) spatial energy states, the two-electron energy basis has \( \frac{1}{2}M(M-1) \) elements, of which \( \frac{3}{2}(M/2)[(M/2) - 1] \) are ortho (antisymmetric in space and symmetric in spin) and \( \frac{1}{2}(M/2)[(M/2) + 1] \) are para (symmetric in space and antisymmetric in spin).
2.5 Spin plus space, three electrons, ground state

Three electrons are in the situation described above (spin-independent Hamiltonian, non-interacting electrons). The full listing of states has been done, but it’s an accounting nightmare, so I ask a simpler question: What is the ground state?

Call the one-particle spatial energy eigenfunctions $\eta_1(\vec{x}), \eta_2(\vec{x}), \eta_3(\vec{x}), \ldots$. The ground state will be the antisymmetrized combination of

$$\eta_1(\vec{x}_A)\chi_+(A)\eta_1(\vec{x}_B)\chi_-(B)\eta_2(\vec{x}_C)\chi_+(C)$$

or the antisymmetrized combination of

$$\eta_1(\vec{x}_A)\chi_+(A)\eta_1(\vec{x}_B)\chi_-(B)\eta_2(\vec{x}_C)\chi_-(C).$$

These two states are degenerate: both have energy $2E_1 + E_2$.

Write out the first one in detail. It is

$$\frac{1}{\sqrt{6}}[ +\eta_1(\vec{x}_A)\chi_+(A)\eta_1(\vec{x}_B)\chi_-(B)\eta_2(\vec{x}_C)\chi_+(C)$$

$$-\eta_1(\vec{x}_A)\chi_+(A)\eta_2(\vec{x}_B)\chi_+(B)\eta_1(\vec{x}_C)\chi_-(C)$$

$$+\eta_2(\vec{x}_A)\chi_+(A)\eta_1(\vec{x}_B)\chi_+(B)\eta_1(\vec{x}_C)\chi_+(C)$$

$$-\eta_2(\vec{x}_A)\chi_+(A)\eta_1(\vec{x}_B)\chi_-(B)\eta_1(\vec{x}_C)\chi_+(C)$$

$$+\eta_1(\vec{x}_A)\chi_-(A)\eta_2(\vec{x}_B)\chi_+(B)\eta_1(\vec{x}_C)\chi_+(C)$$

$$-\eta_1(\vec{x}_A)\chi_-(A)\eta_1(\vec{x}_B)\chi_+(B)\eta_2(\vec{x}_C)\chi_+(C) ].$$ 

This morass is another good case for the abbreviated notation of page 39. I’m not concerned with normalization for the moment, so I’ll write this first state as

$$+|1+, 1-, 2+\rangle$$

$$-|1+, 2+, 1-\rangle$$

$$+|2+, 1+, 1-\rangle$$

$$-|2+, 1-, 1+\rangle$$

$$+|1-, 2+, 1+\rangle$$

$$-|1-, 1+, 2+\rangle$$

and the second one as

$$+|1+, 1-, 2-\rangle$$

$$-|1+, 2-, 1-\rangle$$

$$+|2-, 1+, 1-\rangle$$

$$-|2-, 1-, 1+\rangle$$

$$+|1-, 2-, 1+\rangle$$

$$-|1-, 1+, 2-\rangle.$$
It’s clear that both of these states are antisymmetric, but that neither of them factorize into a neat “space part times spin part”. If, following our approach used with two electrons, you attempt to find a linear combination of these two that does so factorize, you will fail. The ground state wavefunction cannot be made to factor into a space part times a spin part.

2.4 Problem: A doomed attempt

Show that a linear combination of the two states has the form

\[
\begin{align*}
+|1, 1, 2\rangle & = \alpha |+, -, +\rangle + \beta |+, -, -\rangle \\
-|1, 2, 1\rangle & = \alpha |+, +, -\rangle + \beta |+, +, -\rangle \\
+|2, 1, 1\rangle & = \alpha |+, +, -\rangle + \beta |-, +, -\rangle \\
-|2, 1, 1\rangle & = \alpha |+, -, +\rangle + \beta |-, -, +\rangle \\
+|1, 2, 1\rangle & = \alpha |-, +, +\rangle + \beta |-, -, +\rangle \\
-|1, 1, 2\rangle & = \alpha |-, +, +\rangle + \beta |-, -, -\rangle,
\end{align*}
\] (2.30)

and that this form can never be factorized into a space part times a spin part.
Chapter 3

Atoms

All this is fine and good — lovely, in fact. But we have to apply it to experimentally accessible systems, and while things like carbon nanotubes exist, the most readily accessible systems are atoms.

3.1 Central potentials in two dimensions

See *The Physics of Quantum Mechanics*, section 13.1 “Energy eigenproblem in two dimensions”.

3.2 Central potentials in three dimensions

The method used for central potentials in two dimensions works in three dimensions as well. The details are (as expected) messier, and I’m not going to present them. Instead, I’ll just quote the result:

To solve the three-dimensional energy eigenproblem for a spherically-symmetric potential energy $V(r)$, namely

$$\frac{-\hbar^2}{2M} \nabla^2 \eta(\vec{r}) + V(r) \eta(\vec{r}) = E \eta(\vec{r}),$$

(3.1)

first solve the one-variable radial energy eigenproblem

$$-\frac{\hbar^2}{2M} \frac{d^2 u(r)}{dr^2} + \left[ V(r) + \frac{\hbar^2 \ell(\ell + 1)}{2M} \right] u(r) = Eu(r)$$

(3.2)

for $\ell = 0, 1, 2, \ldots$. For a given $\ell$, call the resulting energy eigenfunctions and eigenvalues $u_{n,\ell}(r)$ and $E_{n,\ell}$ for $n = 1, 2, 3, \ldots$. Then the three-dimensional solutions are

$$\eta_{n,\ell,m}(r, \theta, \phi) = \frac{u_{n,\ell}(r)}{r} Y_{\ell}^m(\theta, \phi) \quad \text{with energy } E_{n,\ell},$$

(3.3)
where the “spherical harmonics” $Y^m_\ell(\theta, \phi)$ are particular special functions of the angular variables that you could look up if you needed to. Here $m$ takes on the $2\ell + 1$ values $-\ell, -\ell + 1, \ldots, 0, \ldots, \ell - 1, \ell$. Notice that the $2\ell + 1$ different solutions for a given $n$ and $\ell$, but with different $m$, are degenerate.

As in the two dimensional case, the energy eigenvalues come about from solving the one-variable energy eigenproblem with effective potential

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2 \ell(\ell + 1)}{2M} \frac{1}{r^2}.$$  \hspace{1cm} (3.4)

Now, it’s clear just from looking at it that at any point $r$, $V_{\text{eff}}(r)$ increases with increasing $\ell$. It’s reasonable then that the energy eigenvalues will also increase with increasing $\ell$: that the fifth eigenvalue, for example, will always satisfy $E_{5,0} < E_{5,1} < E_{5,2}$ and so forth. This guess is in fact correct, and it can be proven mathematically, but it’s so reasonable that I won’t interrupt this chapter to prove it.

Finally, there’s a strange terminology that you need to know. You’d think that the states with $\ell = 0$ would be called “$\ell = 0$ states”, but in fact they’re called “s states”. You’d think that the states with $\ell = 1$ would be called “$\ell = 1$ states”, but in fact they’re called “p states”. States with $\ell = 2$ are called “d states” and states with $\ell = 3$ are called “f states”. (I am told that these names come from a now-obsolete system for categorizing atomic spectral lines as “sharp”, “principal”, “diffuse”, and “fundamental”. States with $\ell \geq 4$ are not frequently encountered, but they are called g, h, i, k, l, m, … states. For some reason j is omitted. “Sober physicists don’t find giraffes hiding in kitchens.”)
In summary, the energy eigenvalues for some generic potential will look sort of like this:

<table>
<thead>
<tr>
<th>$l = 0$ (s)</th>
<th>$l = 1$ (p)</th>
<th>$l = 2$ (d)</th>
<th>$l = 3$ (f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = 0$</td>
<td>$m = -1, 0, +1$</td>
<td>$m = -2 \ldots +2$</td>
<td>$m = -3 \ldots +3$</td>
</tr>
<tr>
<td>degen = 1</td>
<td>degen = 3</td>
<td>degen = 5</td>
<td>degen = 7</td>
</tr>
</tbody>
</table>

This graph shows only the four lowest energy eigenvalues for each value of $\ell$. A single horizontal line in the “$\ell = 0$ (s)” column represents a single state, whereas a single horizontal line in the “$\ell = 2$ (d)” column represents five states, each with the same energy (“degenerate states”).

### 3.3 The hydrogen atom

An electron and a proton interact through the classical electrostatic potential energy function

$$V(r) = -\frac{1}{4\pi\epsilon_0} \frac{e^2}{r},$$

so you might think that the exact energy eigenproblem for the hydrogen atom is

$$-\frac{\hbar^2}{2M} \nabla^2 \eta(\vec{r}) - \frac{1}{4\pi\epsilon_0} \frac{e^2}{r} \eta(\vec{r}) = E \eta(\vec{r}).$$
That’s not exactly correct. This Hamiltonian treats the proton as stationary while the electron does all
the moving: in fact, although the proton is almost 2000 times more massive than the electron, it’s not
infinitely massive and it does do some moving. This Hamiltonian assumes the proton is a point particle:
in fact, although the nucleus is small compared to an atom, it does have some size. This Hamiltonian is
non-relativistic and, in particular, it treats the electromagnetic field as purely classical: both false. This
Hamiltonian ignores the electron’s spin. All of these are good approximations, but this is a model for a
hydrogen atom, not the exact thing.

But let’s go with what we have, rather than holding out for an exact solution of an exact Hamiltonian
that will never come. What happens if we solve the three-dimensional central potential problem with the
model potential (3.5)? We don’t yet have the mathematical tools to actually perform this solution, but we
are in a position to appreciate the character of the solution.

First of all, because the Columbic potential is a particular kind of central potential, it will have all the
properties listed in the last section for three-dimensional central potentials: Each solution will be character-
ized by an \( \ell \) and an \( m \), where \( \ell = 0, 1, 2, \ldots \) and where \( m = -\ell, \ldots, +\ell \). The eigenenergy will be independent
of \( m \), resulting in degeneracy. And the eigenenergies will increase with increasing \( \ell \).
Here is what the eigenenergies for the Coulomb potential turn out to be:

<table>
<thead>
<tr>
<th>$l$ = 0 (s)</th>
<th>$l$ = 1 (p)</th>
<th>$l$ = 2 (d)</th>
<th>$l$ = 3 (f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$ = 0</td>
<td>$m$ = -1, 0, +1</td>
<td>$m$ = -2 ... +2</td>
<td>$m$ = -3 ... +3</td>
</tr>
<tr>
<td>degen = 1</td>
<td>degen = 3</td>
<td>degen = 5</td>
<td>degen = 7</td>
</tr>
</tbody>
</table>

What a surprise! The eigenenergies for $l = 1$ are pushed up so much that they exactly line up with all but the lowest eigenenergies for $l = 0$. The eigenenergies for $l = 2$ are pushed up so much that they exactly line up with all but the lowest eigenenergies for $l = 1$. And so forth.

Normally eigenstates are labeled by $n = 1, 2, 3, \ldots$. But this surprising line up of energies encourages a different notation for the Coulomb problem. For “$l = 0$ (s)” the eigenstates are labeled as usual by $n = 1, 2, 3, 4, \ldots$. But for “$l = 1$ (p)” the eigenstates are labeled by $n = 2, 3, 4, \ldots$. For “$l = 2$ (d)” they are labeled by $n = 3, 4, \ldots$. And so forth. With this labeling scheme the eigenenergies turn out to be given by

$$E_n = -\frac{M}{2\hbar^2} \left( \frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{1}{n^2}. \quad (3.7)$$

The surprising line up is called “accidental degeneracy”.

### 3.1 Problem: Counting hydrogen states

Show that the degeneracy for states characterized by $n$ is $n^2$. 
3.2 Problem: Dimensional analysis for energy eigenvalues

The Hamiltonian (3.6) contains only two parameters:

\[ \frac{\hbar^2}{M} \quad \text{and} \quad \frac{e^2}{4\pi\epsilon_0}. \]

Use dimensional analysis to show that these two parameters can come together to form an energy in only one way.

The coefficient in equation (3.7) is called the “Rydberg\(^1\) energy”, and this equation is usually written

\[ E_n = -\frac{\text{Ry}}{n^2}, \quad \text{where} \quad \text{Ry} = 13.6 \text{ eV}. \]

(I recommend that you memorize this energy 13.6 eV, which sets the scale for typical energies in atomic physics. Much to my embarrassment, I forgot it during my graduate qualifying oral exam.)

It’s a triumph to know the eigenenergies, but we should know also something about the energy eigenstates, which are labeled \( \eta_{n,\ell,m}(\vec{r}) \). Terminology: a state with \( n = 3, \ell = 2, \) and any value of \( m \) is called a “3d state”.

\(^1\)Johannes Rydberg (1854–1919), Swedish spectroscopist, discovered a closely-related formula empirically in 1888.
3.3. THE HYDROGEN ATOM

\[ V_{\text{eff}}(r) = -\frac{1}{r} + \left( l(l+1)/r^2 \right) \]

\[ V_{\text{eff}}(r) \]

\[ r \]

\[ + l(l+1)/r^2 \]
This graph suggests that for a given value of $n$, the states with larger $\ell$ will have larger expectation values for $r$, the distance from the proton to the electron.

3.3 **Problem**: Which states are distant, which are close?

Argue also, on the basis of the graph, that for a given value of $\ell$, states with larger $n$ will have larger expectation values for $r$. 
Transitions

If the Hamiltonian of equation (3.6) were exactly correct, then a hydrogen atom starting in energy state \( \eta_{3,2,-1}(\vec{r}) \) would remain in that state forever. Furthermore, a hydrogen atom starting in a linear combination with probability 0.8 of being in energy state \( \eta_{3,2,-1}(\vec{r}) \) and probability 0.2 of being in energy state \( \eta_{2,1,0}(\vec{r}) \) would maintain those probabilities forever.

But the Hamiltonian of equation (3.6) is not exactly correct. It ignores collisions, it ignores external electromagnetic field (e.g., incident light), and it ignores coupling to the electromagnetic field (e.g., radiated light). These effects mean that the state \( \eta_{3,2,-1}(\vec{r}) \) is a stationary state of the model Hamiltonian, but it is not a stationary state of the exact Hamiltonian. In other words, these effects result in transitions between stationary states of the model Hamiltonian.

To understand these transitions you need to understand the transition-causing effects, and at this point in your education you’re not ready to do that. But I’ll tell you one thing right now: A transition involving a single photon (either absorbing or emitting a single photon) will result in a transition with \( \Delta \ell = \pm 1 \). So, for example, a hydrogen atom in a 2p state (that is, one with \( n = 2, \ell = 1 \), and any legal value of \( m \)) could transition to the 1s ground state by emitting a single photon. A hydrogen atom in a 2s state (that is, one with \( n = 2, \ell = 0 \), and \( m = 0 \)) cannot transition to the ground state by emitting a single photon. (It could do so by emitting two photons, or through a collision.) An atom in the 1s ground state, exposed to a light source with photons of energy \( \frac{3}{4}\text{Ry} \), can be excited to 2p state by absorbing a single photon, but it cannot be excited to the 2s state by absorbing a single photon.

I regard this fact as a picky detail appropriate for an advanced course, but the people who write the graduate record exam in physics seem to think it’s important so you should probably remember it.

3.4 The helium atom

Here’s the energy eigenproblem for the helium atom, at the same level of approximation as the eigenproblem (3.6) for the hydrogen problem:

\[
\begin{align*}
-\frac{\hbar^2}{2M} \nabla_A^2 \eta(\vec{r}_A, \vec{r}_B) & - \frac{\hbar^2}{2M} \nabla_B^2 \eta(\vec{r}_A, \vec{r}_B) \\
& - \frac{1}{4\pi \epsilon_0} \frac{2e^2}{r_A} \eta(\vec{r}_A, \vec{r}_B) - \frac{1}{4\pi \epsilon_0} \frac{2e^2}{r_B} \eta(\vec{r}_A, \vec{r}_B) + \frac{1}{4\pi \epsilon_0} \frac{e^2}{|\vec{r}_A - \vec{r}_B|} \eta(\vec{r}_A, \vec{r}_B)
\end{align*}
\]

\( = E \eta(\vec{r}_A, \vec{r}_B). \)  (3.8)

We have no chance of solving this two-electron problem exactly. Our approach instead is to solve the one-electron problem for a different potential, and then use those one-electron solutions as building blocks for the two-electron problem through the antisymmetrization process. The strategy may seem crude, but in practice it can produce highly accurate results.
Instead of focusing on two electrons, interacting with the nucleus and with each other, focus on one electron interacting with the nucleus and with the average of the other electron. I don’t yet know exactly how the “other” electron is averaged, but I assume it spread out in a spherically symmetric fashion.

Finding the potential. Remember the shell theorem for spherically symmetric charge distributions from electricity and magnetism: When the electron under focus is close to the nucleus, it feels only the electric field due to the nucleus, so the potential energy is

\[
V(r) \approx -\frac{1}{4\pi\epsilon_0} \frac{2e^2}{r}.
\]

(3.9)

Whereas when the electron under focus is far from the nucleus, it feels the electric field due to the nucleus, plus the electric field due to the cloud collapsed into the nucleus, so the potential energy is

\[
V(r) \approx -\frac{1}{4\pi\epsilon_0} \frac{e^2}{r}.
\]

(3.10)

This phenomenon is called “shielding”. The shielded potential interpolates between the known limits of small \(r\) and large \(r\) behavior. The exact character of that interpolation is unclear: if you were doing high accuracy calculations, you would need to find it. But we will not be doing high accuracy calculations, and it will be enough for us just to know the two limits.

\[2\text{Using a technique called the Hartree-Fock approximation.}\]
3.4. THE HELIUM ATOM

This approach is called a “mean field approximation” and is said to “ignore correlations” between the two electrons. This approach is not exact and cannot be made exact, but it enables progress to be made.

Finding the one-electron eigenvalues. What will the one-electron energy eigenvalues be for a shielded potential energy? If the potential energy were

\[ V(r) = -\frac{1}{4\pi\epsilon_0} \frac{2e^2}{r}, \]  

then the energy eigenvalues would be

\[ E_n = -4 \frac{\text{Ry}}{n^2}. \]  

(3.11)

(3.12)

If the potential energy were

\[ V(r) = -\frac{1}{4\pi\epsilon_0} \frac{e^2}{r}, \]  

then the energy eigenvalues would be

\[ E_n = -\frac{\text{Ry}}{n^2}. \]  

(3.13)

(3.14)

But in fact, the potential energy interpolates between these two forms, so the energy eigenvalues interpolate between these two possibilities. Let’s examine this interpolation, first for the s states:

\[ n = 1 \quad \text{shielded} \quad Z = 1 \]

\[ n = 2 \]

\[ n = 3 \]

\[ n = 4 \]

energy eigenvalue
To the left, under the heading “Z = 2”, are the s state eigenvalues (3.12), which are four times deeper than those under the heading “Z = 1”, the s state eigenvalues (3.14). I’ve drawn a vertical line midway between them and dashed lines connecting the two sets of eigenvalues. If the eigenvalues for the shielded potential were exactly halfway between the eigenvalues for the two limits, then they would fall where the dashed lines cross the vertical line. But they don’t fall exactly there. For states that are mostly near the nucleus, the energies are closer to (3.12). For states that are mostly far from the nucleus, the energies are closer to (3.14).

This process can be repeated for p states, d states, and f states, showing that a shielded potential energy function will give rise to a set of energy eigenvalues like this:

\[
\begin{array}{cccc}
\text{l = 0 (s)} & \text{l = 1 (p)} & \text{l = 2 (d)} & \text{l = 3 (f)} \\
\text{m = 0} & \text{m = -1, 0, +1} & \text{m = -2 ... +2} & \text{m = -3 ... +3} \\
\text{degen = 1} & \text{degen = 3} & \text{degen = 5} & \text{degen = 7}
\end{array}
\]

Building two-electron states from one-electron states. Now that we have one-electron eigenstates, we can combine them through the antisymmetrization process to produce two-electron eigenstates. This process was described in section 2.4.

Misconception. Perhaps you learned in high-school chemistry that the ground state of the helium atom has ‘one electron in the ‘1s, spin up’ state, and one electron in the ‘1s, spin down’ state”. That’s just plain
wrong — if were correct, then you’d be able to distinguish between the two electrons (the one with spin up, the one with spin down) and then the two electrons wouldn’t be identical. What’s correct instead is that the two electrons are in one state, namely the entangled state

\[ \eta_{100}(\vec{x}_A)\eta_{100}(\vec{x}_B) \frac{1}{\sqrt{2}} [\chi_{+}(A)\chi_{-}(B) - \chi_{-}(A)\chi_{+}(B)]. \]

3.5 The lithium atom

The ground state of hydrogen is two-fold degenerate. The ground state of helium is non-degenerate. The ground state of lithium is two-fold degenerate.

The ground states of hydrogen and helium can (within the shielded potential approximation) be written as a spatial part times a spin part. The ground state of lithium cannot be so written.

3.6 All other atoms

Atoms larger than lithium are a real headache. Specialists have done them in exquisite detail. But for this course we’re not going to try to find the energy spectrum, we’re not going to try to find the ground state degeneracy, we’re not even going to try to write down a ground state. Instead, we’re only going to list the one-electron states that are thrown together through the antisymmetrization machinery to make the many-electron ground state.

Try this for carbon \((Z = 6)\). In the figure on page 53, pertaining to helium, the left-hand energies are four times deeper than the right-hand energies. If I were to draw a parallel figure for carbon, the left-hand energies would be 36 times deeper than the right-hand energies!

Build up the periodic table using shielded potential energies like the figure on page 54.
Appendix A

Catalog of Misconceptions

Effective teaching does not simply teach students what is correct — it also insures that students do not believe what is incorrect. There are a large number of prevalent misconceptions concerning quantum mechanics. The catalog below presents some misconceptions mentioned in this book, together with the page number where that misconception is pointed out and corrected.

- all states are energy states, 20
- identical particles attract/repel through a force, 35
- identical particles in ground state of helium, 55
- identical particles, label particles vs. coordinates, 33
- identical particles, three particles, 38
- probability density is all that matters, 26
- state of a two-particle system, 10
- transition to ground state, 24, 51
- wavefunction associated not with system but with particle, 10
- wavefunction exists in position space, 10
- wavefunction is dimensionless, 7
- wavefunction must factorize into space $\times$ spin, 42
Index

amplitude, 7
amplitude density, 7, 30
antisymmetric under swapping, 33
basis, 22
Bose, Satyendra, 33
boson, 33
boundary value problem, 14
classical limit of quantum mechanics, 26–30
configuration space, 10
data compression, 17
degeneracy, 37, 44
    accidental, 47
derivation
    sloppy, 8
dimensional analysis, 7, 8, 10, 12, 18, 28, 48
Ehrenfest’s theorem, 30
Ehrenfest, Paul, 30
eigenproblem, 13, 20
Eliot, T.S., 21
energy eigenproblem, 13, 20
energy loss, 24
entanglement, 10, 37, 38, 55
exact vs. model, 45
exchange, 33
exclusion principle, 33
Fermi, Enrico, 33
fermion, 33
Fourier sine series, 15, 17–18
Fourier, Joseph, 17
GRE, 51
Hamilton, Alexander, 20
Hamilton, William Rowan, 20
Hamiltonian, 20
Hartree-Fock approximation, 52
huddle together, 34, 40
identical particles, 32–42
ignore correlations, 53
infinite square well, 11
initial value problem, 14
interchange, 33
interference, 22
Johnson-Trotter sequence of permutations, 34
language
    average vs. expected value, 24
    for huddle/spread, 35
    how precise is necessary?, 38
    wavefunction, 7
Liouville, Joseph, 15
mean field approximation, 53
misconceptions, catalog of, 56
model vs. exact, 45
normalization, 6, 7, 18
not even wrong, 33
ortho, 40
para, 40
particle in a box, 11
Pauli principle, 33
Pauli, Wolfgang, 33
permutation symmetry, 35
permutations, 34
photon, 51
plain changes, 36
plain changes sequence of permutations, 34
potential energy functions
  arbitrary, 22
  infinite square well, 11
probability density, 6, 9
quantal revivals, 16
revival time, 16
Rydberg energy, 48
Rydberg, Johannes, 48
Schrödinger equation, 8
Schrödinger, Erwin, 7
separation of variables, 12
shielding, 52
shorthand, dangerous: say it but don’t think it, 24
Slater determinant, 36
Slater, John, 36
sloppy derivation, 8
span, 22
spherical harmonics, 44
spin, 30
spread apart, 34, 40
stationary state, 19
Sturm, Charles-François, 15
Sturm-Liouville theory, 15
superposition, 22
Susskind, Leonard, 10
Sutton, O. Graham, 17
swap, 33
symmetric under swapping, 33
symmetrization and antisymmetrization, 35–37
symmetry
  permutation, 35
symmetry/antisymmetry requirement, 33
terminology
  s p d f, 44, 48
three identical particles, 34, 38
Trotter-Johnson sequence of permutations, 34
two-particle systems, 9–10, 32–40
wavefunction, 7, 9, 30
wonder, 31, 43