QUANTUM MECHANICS

Michael Fowler

University of Virginia

LibreTexts™

University of Virginia

Quantum Mechanics

Michael Fowler

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the hundreds of other texts available within this powerful platform, it is freely available for reading, printing and "consuming." Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects.

Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.

The LibreTexts mission is to unite students, faculty and scholars in a cooperative effort to develop an easy-to-use online platform for the construction, customization, and dissemination of OER content to reduce the burdens of unreasonable textbook costs to our students and society. The LibreTexts project is a multi-institutional collaborative venture to develop the next generation of open-access texts to improve postsecondary education at all levels of higher learning by developing an Open Access Resource environment. The project currently consists of 14 independently operating and interconnected libraries that are constantly being optimized by students, faculty, and outside experts to supplant conventional paper-based books. These free textbook alternatives are organized within a central environment that is both vertically (from advance to basic level) and horizontally (across different fields) integrated.

The LibreTexts libraries are Powered by NICE CXOne and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org. More information on our activities can be found via Facebook (https://facebook.com/Libretexts), Twitter (https://twitter.com/libretexts), or our blog (http://Blog.Libretexts.org).

This text was compiled on 04/15/2025

TABLE OF CONTENTS

Licensing

1: Reviewing Elementary Stuff

- 1.1: Breakdown of Classical Mechanics
- 1.2: Early Quantum Mechanics
- 1.3: Wave Equations, Wavepackets and Superposition
- 1.4: The Uncertainty Principle
- 1.5: Electron in a Box

2: Some Essential Math

- 2.1: Fourier Series and Integrals, the Dirac Function
- 2.2: Linear Algebra
- 2.3: Function Spaces
- 2.4: Complex Variable, Stationary Path Integrals

3: Mostly 1-D Quantum Mechanics

- 3.1: 1-D Schrödinger Equation - Example Systems
- 3.2: General Uncertainty Principal
- 3.3: Energy-Time Uncertainty Principle
- 3.4: The Simple Harmonic Oscillator
- 3.5: Propagators and Representations
- 3.6: Coherent States
- 3.7: Path Integrals
- 3.8: Path Integrals for the SHO
- 3.9: Appendix- Some Exponential Operator Algebra

4: Angular Momentum, Spin and the Hydrogen Atom

- 4.1: Angular Momentum Operator Algebra
- 4.2: Orbital Eigenfunctions- 2-D Case
- 4.3: Note on Curvilinear Coordinates
- 4.4: Orbital Eigenfunctions in 3-D
- 4.5: Spin
- 4.6: The Hydrogen Atom
- 4.7: Adding Angular Momenta
- 4.8: Tensor Operators

5: Interlude - The Nature of Electrons

- 5.1: Bosons and Fermions
- 5.2: Multielectron Atoms

6: Charged Particle in Magnetic Field

- 6.1: Charged Particle in a Magnetic Field

7: The Density Matrix

- 7.1: The Density Matrix

8: Approximate Methods

- 8.1: Variational Methods
- 8.2: The WKB Approximation
- 8.3: Note on the WKB Connection Formula

9: Perturbation Theory

- 9.1: Time-Independent Perturbation Theory
- 9.2: The Peierls Transition - an Unexpected Insulator
- 9.3: Van Der Waals Forces between Atoms
- 9.4: The Interaction Representation
- 9.5: Time-Dependent Perturbation Theory
- 9.6: The Photoelectric Effect in Hydrogen
- 9.7: Quantizing Radiation

10: Scattering Theory

- 10.1: Scattering Theory
- 10.2: More Scattering Theory - Partial Waves
- 10.3: Scattering Amplitudes, Bound States, Resonances
- 10.4: Identical Particles- Symmetry and Scattering

11: Other Material

- 11.1: Schrödinger Spreadsheets
- 11.2: Homeworks, exams, etc.

Index

Index

Glossary

Detailed Licensing

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing****.*

CHAPTER OVERVIEW

1: Reviewing Elementary Stuff

1.1: Breakdown of Classical Mechanics
1.2: Early Quantum Mechanics
1.3: Wave Equations, Wavepackets and Superposition
1.4: The Uncertainty Principle
1.5: Electron in a Box

Thumbnail: Propagation of de Broglie waves in 1d—real part of the complex amplitude is blue, imaginary part is green. The probability (shown as the color opacity) of finding the particle at a given point x is spread out like a waveform, there is no definite position of the particle. (Public Domain; Maschen)

1.1: Breakdown of Classical Mechanics

What was Wrong with Classical Mechanics?

Basically, classical statistical mechanics wasn't making sense...

Maxwell and Boltzmann evolved the equipartition theorem: a physical system can have many states (gas with particles having different velocities, or springs in different states of compression).

At nonzero temperature, energy will flow around in the system, it will constantly move from one state to another. So, what is the probability that at any instant it is in a particular state with energy E?

M&B proves it was proportional to $e^{-E/kT}$. This proportionality factor is also correct for any subsystem of the system: for example a single molecule.

Notice this means if a system is a set of oscillators, different masses on different strength springs, for example, then in thermal equilibrium *each oscillator has on average the same energy as all the others*. For three-dimensional oscillators in thermal equilibrium, the average energy of each oscillator is $3kT$, where k is Boltzmann's constant.

Black Body Radiation

Now put this together with Maxwell's discovery that light is an electromagnetic wave: inside a hot oven, Maxwell's equations can be solved yielding standing wave solutions, and the set of different wavelength allowed standing waves amount to an infinite series of oscillators, with no upper limit on the frequencies on going far into the ultraviolet. Therefore, from the classical equipartition theorem, an oven at thermal equilibrium at a definite temperature should contain an infinite amount of energy—of order kT in each of an infinite number of modes—and if you let radiation out through a tiny hole in the side, you should see radiation of all frequencies.

This is not, of course, what is observed: as an oven is warmed, it emits infrared, then red, then yellow light, etc. This means that the higher frequency oscillators (blue, etc.) are in fact *not* excited at low temperatures: equipartition is not true.

Planck showed that the experimentally observed intensity/frequency curve was exactly reproduced if it was assumed that the radiation was *quantized*: light of frequency f could only be emitted in quanta—now photons—having energy hf, h being Planck's constant. This was the beginning of quantum mechanics.

The Photoelectric Effect

Einstein showed the same quantization of electromagnetic radiation explained the photoelectric effect: a photon of energy hf knocks an electron out of a metal, it takes a certain work W to get it out, the rest of the photon energy goes to the kinetic energy of the electron, for the fastest electrons emitted (those that come right from the surface, so encountering no further resistance). Plotting the maximum electron kinetic energy as a function of incident light frequency confirms the hypothesis, giving the same value for h as that needed to explain radiation from an oven. (It had previously been assumed that more intense light would increase the kinetic energy—this turned out not to be the case.)

The Bohr Atom

Bohr put together this quantization of light energy with Rutherford's discovery that the atom had a nucleus, with electrons somehow orbiting around it: for the hydrogen atom, light emitted when the atom is thermally excited has a particular pattern, the observed emitted wavelengths are given by

$$\frac{1}{\lambda} = R_H \left(\frac{1}{4} - \frac{1}{n^2}\right) \tag{1.1.1}$$

with $n = 3, 4, 5\ldots$ R_H is now called the Rydberg constant.) Bohr realized these were photons having energy equal to the *energy difference between two allowed orbits* of the electron circling the nucleus (the proton), $E_n - E_m = hf$, leading to the conclusion that the allowed levels must be:

$$E_n = -\frac{hcR_H}{n^2} \tag{1.1.2}$$

How could the quantum hf restricting allowed radiation energies also restrict the allowed electron orbits? Bohr realized there must be a connection—because h has the dimensions of angular momentum! What if the electron were only allowed to be in circular orbits of angular momentum nKh, with n an integer? Bohr did the math for orbits under an inverse square law, and found that the observed spectra were in fact correctly accounted for by taking $K = 1/2\pi$.

But then he realized he did not even need the experimental results to find K: quantum mechanics *must* agree with classical mechanics in the regime where we know experimentally that classical mechanics (including Maxwell's equations) is correct, that is, for systems of macroscopic size. Consider a negative charge orbiting around a fixed positive charge at a radius of 10 cm., the charges being such that the speed is of order meters per second (we don't want relativistic effects making things more complicated). Then from classical E&M, the charge will radiate at the orbital frequency. Now imagine this is actually a hydrogen atom, in a perfect vacuum, in a high state of excitation. It must be radiating at this same frequency. But Bohr's theory can't just be right for small orbits, so the radiation must satisfy $E_n - E_m = hf$. The spacing between adjacent levels will vary slowly for these large orbits, so h times the orbital frequency must be the energy difference between adjacent levels. Now, that energy difference depends on the allowed angular momentum step between the adjacent levels: that is, on K. Reconciling these two expressions for the radiation frequency gives $K = 1/2\pi$.

This *classical limit argument, then, predicts the Rydberg constant* in terms of already known quantities:

$$R_H = \left(\frac{1}{4\pi\varepsilon_0}\right)^2 \cdot \frac{2\pi^2 me^4}{ch^3} \tag{1.1.3}$$

.

What's right about the Bohr atom?

1. It gives the Balmer series spectra.
2. The first orbit size is close to the observed size of the atom: and remember there are no adjustable parameters, the classical limit argument determines the spectra and the size.

What's *wrong* with the Bohr atom?

1. No explanation for why angular momentum should be quantized. (This was solved by de Broglie a little later.)
2. Why don't the circling electrons radiate, as predicted classically? Well, the fact that radiation is quantized means the classical picture of an accelerating charge smoothly emitting radiation cannot work if the energies involved are of order h times the frequencies involved.
3. The lowest state has nonzero angular momentum. This is a defect of the model, corrected in the truly quantum model (Schrödinger's equation).
4. In an inverse square field, orbits are in general elliptical.

This was at first a puzzle: why should there be only **circular orbits** allowed? In fact, the model does allow elliptical orbits, and they do not show up in the Balmer series because, as proved by Sommerfeld, if the allowed elliptical orbits have the same allowed angular momenta as Bohr's orbits, they have the same set of energies. This is a special property of the inverse square force. .

De Broglie Waves

The first explanation of why only certain angular momenta are allowed for the circling electron was given by de Broglie: just as photons act like particles (definite energy and momentum), but undoubtedly are wave like, being light, so particles like electrons perhaps have wave like properties. For photons, the relationship between wavelength and momentum is $p = h/\lambda$. Assuming this is also true of electrons, and that the allowed circular orbits are standing waves, Bohr's angular momentum quantization follows.

Schrödinger's Wave Equation

De Broglie's idea was clearly on the right track—but waves in space are three-dimensional, thinking of the circular orbit as a string under tension cannot be right, even if the answer is.

Photon waves (electromagnetic waves) obey the equation

$$\nabla^2 \vec{E} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = 0 \tag{1.1.4}$$

A solution of definite momentum is the plane wave

$$\left(\frac{\partial^2}{\partial x^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\vec{E}_0 e^{i(kx-\omega t)}=\left(k^2-\frac{\omega^2}{c^2}\right)\vec{E}_0 e^{i(kx-\omega t)}=0 \tag{1.1.5}$$

Notice that the last equality is essentially just $\omega=ck$, where for a plane wave solution the energy and momentum of the photon are translated into differential operators with respect to time and space respectively, to give a differential equation for the wave.

Schrödinger's wave equation is equivalently taking the (nonrelativistic) energy-momentum relation $E=p^2/2m$ and using the same recipe to translate it into a differential equation:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t}=-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} \tag{1.1.6}$$

Making the natural extension to three dimensions, and assuming we can add a potential term in the most naïve way possible, that is, going from $E=p^2/2m$ to $E=p^2/2m+V(x,y,z)$, we get

$$i\hbar\frac{\partial\psi(x,y,z,t)}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\psi(x,y,z,t)+V(x,y,z)\psi(x,y,z,t) \tag{1.1.7}$$

This is the equation Schrödinger wrote down and solved, the solutions gave the same set of energies as the Bohr model, but now the ground state had zero angular momentum, and many of the details of the solutions were borne out by experiment, as we shall discuss further later.

A Conserved Current

Schrödinger also showed that a conserved current could be defined in terms of the wave function ψ:

$$\frac{\partial\rho}{\partial t}+div\vec{j}=0 \tag{1.1.8}$$

where

- $\rho=\psi^*\psi=|\psi|^2$ and
- $\vec{j}=\frac{\hbar}{2mi}(\psi^*\vec{\nabla}\psi-\psi\vec{\nabla}\psi^*)$.

Schrödinger's interpretation of his equation was that the electron was simply a wave, not a particle, and this was the wave intensity. But thinking of electromagnetic waves in this way gave no clue to the quantum photon behavior—this could not be the whole story.

Interpreting the Wave Function

The correct interpretation of the wave function (due to Born) follows from analogy to the electromagnetic case. Let's review that briefly. The basic example is the two-slit diffraction pattern, as built up by sending through one photon at a time, to a bank of photon detectors. The pattern gradually emerges: solve the wave equation, then the predicted local energy density (proportional to $|E(x,y,z,t)|^2dxdydz$) gives the probability of one photon going through the system landing at that spot.

Born suggested that similarly $|\psi|^2$ at any point was proportional to the probability of detecting the electron at that point. This has turned out to be correct.

Localizing the Electron

Despite its wavelike properties, we know that an electron can behave like a particle: specifically, it can move as a fairly localized entity from one place to another. What's the wave representation of that? It's called a wave packet: a localized wave excitation. To see how this can come about, first remember that the Schrödinger equation is a linear equation, the sum of any two or more solutions is itself a solution. If we add together two plane waves close in wavelength, we get beats, which can be regarded as a string of wave packets. To get a single wave packet, we must add together a continuous range of wavelengths.

The standard example is the Gaussian wave packet, $\psi(x,t=0)=Ae^{ik_0x}e^{-x^2/2\Delta^2}$ where $p_0=\hbar k_0$

Using the standard result

$$\int_{-\infty}^{+\infty}e^{-ax^2}\,dx=\sqrt{\frac{\pi}{a}} \tag{1.1.9}$$

we find $|A|^2=(\pi\Delta^2)^{-1/2}$ so

$$\psi(x,t=0)=\frac{1}{(\pi\Delta^2)^{1/4}}e^{ik_0x}e^{-x^2/2\Delta^2}. \tag{1.1.10}$$

But how do we construct this particular wavepacket by superposing plane waves? That is to say, we need a representation of the form:

$$\psi(x)=\int_{-\infty}^{+\infty}\frac{dk}{2\pi}e^{ikx}\phi(k) \tag{1.1.11}$$

The function $\phi(k)$ represents the weighting of plane waves in the neighborhood of wavenumber k. This is a particular example of a *Fourier transform*—we will be discussing the general case in detail a little later in the course. Note that if $\phi(k)$ is a bounded function, any particular k value gives a vanishingly small contribution, the plane-wave contribution to $\psi(x)$ from a range dk is $\phi(k)dk/2\pi$. In fact, $\phi(k)$ is given in terms of $\psi(x)$ by

$$\phi(k)=\int_{-\infty}^{+\infty}dxe^{-ikx}\psi(x). \tag{1.1.12}$$

It is perhaps worth mentioning at this point that this can be understood *qualitatively* by observing that the plane wave prefactor e^{-ikx} will interfere destructively with all plane wave components of $\psi(x)$ except that of wavenumber k, where it may at first appear that the contribution is infinite, but recall that as stated above, any particular k component has a vanishingly small weight—and, in fact, this is the right answer, as we shall show in more convincing fashion later.

In the present case, the above handwaving argument is unnecessary, because both the integrals can be carried out exactly, using the standard result:

$$\int_{-\infty}^{\infty}e^{-ax^2+bx}\,dx=e^{b^2/4a}\sqrt{\frac{\pi}{a}} \tag{1.1.13}$$

giving

$$\phi(k)=(4\pi\Delta^2)^{\frac{1}{4}}e^{-\Delta^2(k-k_0)^2/2}. \tag{1.1.14}$$

.

The Uncertainty Principle

Note that the spreads in *x*-space and *p*-space are inversely related: Δx is of order Δ, $\Delta p=\hbar\Delta k\sim\hbar/\Delta$. This is of course the Uncertainty Principle, localization in *x*-space requires a large spread in contributing momentum states.

It's worth reviewing the undergraduate exercises on applications of the uncertainty principle. The help sharpen one's appreciation of the wave/particle nature of quantum objects.

There's a limit to how well the position of an electron can be determined: it is detected by bouncing a photon off of it, and the photon wavelength sets the limit on Δx. But if the photon has enough energy to create an electron-positron pair out of the vacuum, you can't be sure which electron you're seeing. This limits $\Delta x\sim\hbar/mc$ at best. (This is called the Compton wavelength, written λ_c – it appears in Compton scattering.) How much smaller that a hydrogen atom ground state wave function is this? $\lambda_c/a_0=e^2/\hbar c(CGS)=e^2/4\pi\varepsilon_0\hbar c(SI)=1/137$, known as the fine structure constant. This is also the ratio of the electron speed in the first Bohr orbit to the speed of light, and so is an indication of the importance of relativistic corrections to energies of electron states; these differences in electron orbit energies for circular and elliptical states having the same energy when calculated nonrelativistically lead to fine structure in the atomic spectra.

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

1.2: Early Quantum Mechanics

This is just a quick review of the experimental basis for quantum mechanics, and some of the early formulations. You don't need to know the historical facts, of course, but some of the physics arguments are worth recalling—for example, Bohr's derivation of the Rydberg constant from his model atom.

Why Do We Need Quantum Mechanics?

Just over 100 years ago, in the 1890's, physics looked in pretty good shape. The beautiful mathematical development of Newton's mechanics, coupled with increasingly sophisticated technology, predicted the movements of the solar system to incredible accuracy, apart from a tiny discrepancy in the orbit of Mercury. It had been less than a hundred years since it was realized that an electric current could exert a force on a magnet, but that discovery had led to power stations, electric trains and a network of telegraph wires across land and under the oceans. It had also been only a hundred years since it had been established that light was a wave, and only forty years since Maxwell's realization that the waves in a light signal were electric and magnetic fields, satisfying a wave equation he was able to derive purely by considering electric and magnetic field phenomena. In particular, he was able to predict the speed of light by measuring the electrostatic attractive forces between charges and the magnetic forces between currents.

At about the same time, in the 1860's, Maxwell and Boltzmann gave a brilliant account of the properties of gases by assuming that they were made up of weakly interacting molecules flying about in a container, bouncing off the sides, with a statistical distribution of energies so that the probability of a molecule having energy E was proportional to $e^{-E/kT}$, k being a universal constant known as Boltzmann's constant. Boltzmann generalized this result from a box of gas to any system. For example, a solid can be envisioned classically as a lattice of balls (the atoms) connected by springs, which can sustain oscillations in many different ways, each such mode can be thought of as a simple harmonic oscillator, with reasonable approximations concerning the properties of the springs, etc. Boltzmann's work leads to the conclusion that each such mode of oscillation, or degree of freedom, would at temperature T have average energy kT, made up of $\frac{1}{2}kT$ potential energy, $\frac{1}{2}kT$ kinetic energy. Notice that this average energy is independent of the strength of the springs, or the masses! *All* modes of vibration, which will vibrate at very different rates, contain the same energy at the same temperature. This equal sharing is called *the Equipartition of Energy*. It is not difficult to check this for a one-dimensional classical harmonic oscillator, averaging the energy by integrating over all displacements and momenta (independently) with the weighting factor $e^{-E/kT}$ (which of course needs to be normalized). The result doesn't depend on the spring constant or the mass. Boltzmann's result gave an excellent account of the specific heats of a wide range of materials over a wide temperature range, but there *were* some exceptions, for example hydrogen gas at low temperatures, and even solids at low enough temperatures. Still, it was generally felt these problems could be handled within the existing framework, just as the slightly odd behavior of Mercury was likely caused by a small planet, named Vulcan, closer to the sun, and so very hard to observe.

Blackbody Radiation

But there was one problem that was hard to get a grip on, an apparently blatant violation of the equipartition of energy. Consider an oven with a small hole in the door, through which the radiation inside is observed. This oven can be heated until it's white hot. The radiation inside is infrared at low temperatures, becoming visible light as the temperature increases. So, the oven's full of electromagnetic waves, satisfying Maxwell's wave equation, with boundary conditions at the walls of the oven, the electric field has to be essentially zero there, because the walls conduct currents. Of course, the radiation originates in oscillating charges in the walls, using the same analysis of Maxwell's equations that gives the radiation form an antenna. Anyway, there is a set of standing wave modes of electromagnetic vibrations inside the oven, just a three-dimensional version of the series of allowed standing wave modes of vibration of a string fixed at both ends. So, we should be able to find the energy density of these waves using the same ideas that worked pretty well for the specific heats of solids and gases, that is to say, assume there's kT of energy in each mode of vibration. (This is $\frac{1}{2}kT$ of kinetic energy, $\frac{1}{2}kT$ of potential energy for each independent direction of vibration.)

But—this leads to disaster. The problem is that there are infinitely many modes of vibration of the electromagnetic field in an oven. There is no upper limit to the number of wiggles the wave can have between the walls. So, if we take kT in each mode, we deduce that the oven contains an infinite amount of energy, and radiates an infinite amount through our small hole. Furthermore, this analysis gives no clue as to why the color we see changes with temperature. Evidently, equipartition of energy isn't working in this case. There's only a finite amount of energy in the oven—and at low temperatures there's no energy at all in the modes corresponding to visible light, although that changes as things get hotter.

In the 1890's, German experimentalists measured the energy density as a function of wavelength to great precision, it's called the blackbody radiation spectrum. A theorist, Planck, found a mathematical formula that fitted this curve exactly,

$$R_T(\nu)d\nu = \frac{8\pi hV\nu^3 df/c^3}{e^{h\nu/kT}-1} \tag{1.2.1}$$

He did not at first have any theoretical justification for this formula, but it was a very accurate fit to some very precise experiments for a suitable value of the constant h, which we discuss in a moment.

Factoring out the number of modes of oscillation in the frequency range $d\nu$, Planck's formula gives the average energy per mode to be

$$\frac{h\nu}{e^{h\nu/kT}-1} \tag{1.2.2}$$

For low frequencies, $h\nu << kT$, this correctly gives kT per mode.

But, for higher frequencies it's clear that the oscillators are not getting their "fair share" kT of energy. Somehow, the oscillating charges in the walls are not radiating so much energy at the high frequencies. The only way Planck could derive the formula theoretically was by making a weird assumption: he assumed that the oscillating charges in the walls could *not* just radiate energy continuously, as Maxwell's equations would predict (and as was known to be true for ordinary antennas) but were only allowed to radiate energy in chunks he called *quanta*. Furthermore, the amount of energy in one quantum depended on the frequency of the oscillation, in fact linearly: for frequency f, the quantum has energy hf, where h is the constant introduced into the formula above, now known as Planck's constant. It follows that the oscillators themselves could only be oscillating with energies that form a ladder with steps hf apart, above some lowest energy which would be their energy at absolute zero temperature.

The formula follows if we assume the oscillating field component in the oven having frequency f can only have a whole number of quanta of energy, that is to say, its energy must be one of: $0, hf, 2hf, 3hf, \ldots$If we further assume that the relative probability of it having energy E is $e^{-E/kT}$, then its relative probabilities of having energy $0, hf, 2hf, \ldots$are in the ratio 1: $e^{-hf/kT})$: $e^{-2hf/kT}$, etc.

The actual probabilities are given by dividing these relative probabilities by the sum of all of them. They clearly are the terms of a geometric series, so their sum is just $1/(1-e^{-hf/kT})$. So, to find the average energy in the oscillator, we take the possible energies $0, hf, 2hf, 3hf, \ldots$and weight each of them with their probability of occurring, that is, we must find

$$0\cdot 1 + hf\cdot e^{-hf/kT} + 2hf\cdot e^{-2hf/kT} + \ldots, \tag{1.2.3}$$

and divide the sum by $1/(1-e^{-hf/kT})$.

So, Planck's quantum assumption explains the observed blackbody radiation curve. It also gives a qualitative explanation of the change in color of the radiated light as the temperature is increased. The oscillators in the walls derive their energy from the heat vibrations of neighboring molecules: typically, such a vibration has energy of order kT, with probabilities of more energy going down as $e^{-E/kT}$. This means that if the potentially radiating oscillator can only absorb energies in quanta hf, if kT << hf, it will be very unlikely to absorb any energy, and therefore very unlikely to radiate. In the three-dimensional oven, the number of standing wave oscillations in a small frequency range Δf increases with f as f^2 so we find that the maximum radiation intensity occurs at a frequency f such that hf is of order kT. Therefore, as the temperature increases, the frequency at which the most intense radiation occurs increases, and hence the color moves from red to blue.

The Photoelectric Effect

If light shines on certain metals, electrons are emitted. This is the photoelectric effect. If the metal is in air, the electrons bounce off air molecules and are almost certainly rapidly reabsorbed, but if the metal surface is in a vacuum, the electrons can fly away, and in a vacuum tube they can be collected by another piece of metal, and light can cause a current to flow, the origin of the photoelectric cell.

In 1902, Lenard studied how the energy of the emitted photoelectrons varied with the intensity of the light. He used a carbon arc light, and could increase the intensity a thousand-fold. The ejected electrons hit another metal plate, the collector, which was connected to the cathode by a wire with a sensitive ammeter, to measure the current produced by the illumination. To measure the energy of the ejected electrons, Lenard charged the collector plate negatively, to repel the electrons coming towards it. Thus, only electrons ejected with enough kinetic energy to get up this potential hill would contribute to the current. Lenard discovered that there was a well-defined minimum voltage that stopped any electrons getting through, we'll call it V_{stop}. To his surprise, he found that V_{stop} did not depend at all on the intensity of the light! Doubling the light intensity doubled the number of electrons emitted, but did not affect the energies of the emitted electrons. He also discovered, by using light of different colors, that the maximum electron energy did increase as the frequency of the incident light increased.

Einstein Suggests an Explanation

In 1905 Einstein gave a very simple interpretation of Lenard's results. He just assumed that the incoming radiation should be thought of as quanta of frequency hf, with f the frequency. In photoemission, one such quantum is absorbed by one electron. If the electron is some distance into the material of the cathode, some energy will be lost as it moves towards the surface. There will always be some electrostatic cost as the electron leaves the surface, this is usually called the work function, W. The most energetic electrons emitted will be those very close to the surface, and they will leave the cathode with kinetic energy

$$E = hf - W \tag{1.2.4}$$

On cranking up the negative voltage on the collector plate until the current just stops, that is, to V_{stop}, the highest kinetic energy electrons must have had energy eV_{stop} on leaving the cathode. Thus,

$$eV_{stop} = hf - W \tag{1.2.5}$$

Thus Einstein's theory makes a very definite quantitative prediction: if the frequency of the incident light is varied, and V_{stop} plotted as a function of frequency, the slope of the line should be h/e. It is also clear that there is a minimum light frequency for a given metal, that for which the quantum of energy is equal to the work function. Light below that frequency, no matter how bright, will not cause photoemission.

Millikan's Attempts to Disprove Einstein's Theory

If we accept Einstein's theory, then, this is a *completely different* way to measure Planck's constant. The American experimental physicist Robert Millikan, who did not accept Einstein's theory, which he saw as an attack on the wave theory of light, worked for ten years, until 1916, on the photoelectric effect, to disprove Einstein's theory. He even devised techniques for scraping clean the metal surfaces inside the vacuum tube. For all his efforts he found disappointing results (for him!): he *confirmed* Einstein's theory, measuring Planck's constant to within 0.5% by this method. One consolation was that he did get a Nobel prize for this series of experiments.

The point to be emphasized is that the same value for Planck's constant, 6.6×10^{-34} Joule.sec, emerges from two completely different experiments: the measurement of blackbody radiation, and measuring energies of emitted electrons in the photoelectric effect. This is clearly a general property of electromagnetic radiation, and is confirmed by many later experiments, for example Compton scattering, in which light scatters off electrons. By measuring the energy change and momentum change of the electron, it is found that a single quantum of light was scattered. (At very high energies, more particles may be generated.)

The Nature of Light

It is firmly established experimentally that the propagation of light is well described by a wave equation, which in fact is not difficult to derive from Maxwell's equations:

$$\nabla^2 \vec{E} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = 0 \tag{1.2.6}$$

For a plane wave moving in the *x*-direction this reduces to

$$\frac{\partial^2 \vec{E}}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = 0 \tag{1.2.7}$$

The monochromatic solution to this wave equation has the form

$$\vec{E}(x,t) = \vec{E}_0 e^{i(kx-\omega t)} \tag{1.2.8}$$

(Another possible solution is proportional to $cos(kx-\omega t)$. We shall find that the exponential form, although a complex number, proves more convenient. The physical electric field can be taken to be the real part of the exponential for the classical case.)

Applying the wave equation differential operator to our plane wave solution

$$\left(\frac{\partial^2}{\partial x^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\vec{E}_0 e^{i(kx-\omega t)}=\left(k^2-\frac{\omega^2}{c^2}\right)\vec{E}_0 e^{i(kx-\omega t)}=0 \tag{1.2.9}$$

If the plane wave is a solution to the wave equation, this must be true for all x and t, so we must have

$$\omega = ck \tag{1.2.10}$$

Solving this equation for boundary conditions like an antenna can be quite challenging, but all we need consider at the moment is some illustration of diffraction. We take the case of a double slit experiment: if a plane wave encounters a barrier with two equal narrow parallel slit openings, the transmitted wave reaching a screen some distance further on will show a series of bright and dark stripes parallel to the slits. This pattern can be quantitatively accounted for. The two slits transmit radiation in phase with each other. At each point on the screen, the electric field vector from slit 1 must be added to the electric field vector from slit 2. At a point on the screen equidistant from the two slits, the electric field vectors will be equal. Moving away from that point in a direction perpendicular to the slits we will reach a point where the field from one slit is exactly out of phase with the field from the other slit—the screen will be dark.

In fact, the intensity if the light at any point on the screen is proportional to $|E_0|^2$.

Now consider what happens as we make the light dimmer and dimmer. How easy is it to see this diffraction pattern? Eventually we need to soup up our detection apparatus. We replace our screen and visual inspection with a series of photodetectors. Experimentally, we find that, just as in the photoelectric effect, our detectors will only detect quanta, just as if the light were made up of particles, photons. Suppose now we dim the light so that our photodetectors only detect one photon per minute coming through the slits. If we record where each photon lands, and build up a picture, we find the very same pattern of light and dark stripes that we saw with bright light.

In other words, if we send through one photon, we cannot predict where it will land, but if we send through a thousand, we will begin to discern the stripes. The best we can do for one photon is to say it will more probably land where the solution to Maxwell's wave equation gives a large $|E_0|^2$. That is to say,$|E_0(x)|^2$ is proportional to the probability of the photon being at x.

But this means each photon must have gone through both slits! The probability distribution for a single photon is given by the stripes, and the distance between the stripes depends on the distance between the slits. The photon, therefore, knows about both slits. So the bottom line is: to find where one photon will be, solve the wave equation to find the electric field everywhere on the screen. The probability of the photon landing at any particular point is proportional to $|E_0|^2$ at that point.

To illustrate how weird this really is, consider a beam of photons split into two by a half silvered mirror, the two half-beams than follow widely separated paths until they are reunited by a suitable sequence of mirrors to interfere with each other. Sending one photon at a time, we will eventually build up a diffraction pattern of some sort. So if we think of the initial photon as a "wave packet" it will split into two half "wave packets" which will finally interfere with each other. Now suppose I put 100% efficient photon detectors on both paths. If I send photons through the apparatus one at a time, I get a series of clicks from the two detectors: path 1 clicks, path 1 clicks again, path 2 clicks, etc.: a random series. I never get both clicking with one photon. (We can dim the light enough so that the photons are far apart, that is, they definitely come one at a time.) What does this tell us about the nature of the wavefunction?

You might be inclined to think that the photon goes at random, half the time it goes along one path, half the time the other. That is to say, the photon *really is* on one of the paths, we just don't know which until we detect it, and the wavefunction represents our ignorance. We *do* know that once we detect the photon on one path, there's zero probability of finding it on the other path—so that part of the wavefunction has gone! But was it really there in the first place for that particular photon? *Yes*: the other half wavepacket must have been there, because if I hadn't captured the photon with a detector in the way, the two half wavefunctions would have gone on to interfere with it to give the diffraction pattern. *So this line of thinking is wrong: we* cannot *say that the photon "really is" on one of the two paths before we detect it*.

The Nature of Matter

By the 1890's and early 1900's, most scientists believed in the existence of atoms. Not all—the distinguished German chemist Ostwald did not, for example. But nobody had a clear picture of even a hydrogen atom. The electron had just been discovered, and it was believed that the hydrogen atom had a single electron. It was suggested that maybe the electron went in circles around a central charge, but nobody believed that because Maxwell had established that accelerating charges radiate, so it was assumed that a circling electron would rapidly loose energy, spiral in to the center, and the atom would collapse. Instead, it was thought, the hydrogen atom (which was of course electrically neutral) was a ball of positively charged jelly with an electron inside, which would oscillate when heated, and emit radiation. Rough calculations, based on the accepted size of the atom, suggested that the radiation would be in the visible range, but no-one could remotely reproduce the known spectrum of hydrogen.

The big breakthrough came in 1909, when Rutherford tried to map the distribution of positive charge in a heavy atom (gold) by scattering alpha particles from it. To his amazement, he found the positive charge was all concentrated in a tiny nucleus, with a radius of order one ten-thousandth that of the atom. This meant that after all the electrons must be going in some kind of planetary orbits, and the Maxwell's equations prediction of radiation did not apply, just as it did not always apply in blackbody radiation.

The Bohr Atom

The Danish theorist Niels Bohr was visiting Manchester at the time Rutherford did this experiment, and Bohr decided that there must be certain allowed sets of electron orbits in the atom where the classical acceleration radiation did not occur: he called them "stationary states". The lowest energy stationary state would be the ground state of the atom, the others would eventually go to that state by emitting photons corresponding to energy differences between states.

But Bohr was of the opinion that looking at the very complex spectra emitted by heated atoms would never be helpful—he remarked that it would be like trying to understand fundamental biology by studying the colors of butterfly wings.

He changed his mind in February 1913, when a casual conversation with the spectroscopist H. R. Hansen revealed that one pattern *had* been discerned in the apparent chaos of spectral lines. In particular, Hansen (a colleague and former classmate of Bohr) showed him Balmer's formula for hydrogen. Balmer was a math and Latin teacher at a girls' school in Switzerland, and had found his formula in the 1880's. Balmer's formula is:

$$\frac{1}{\lambda}=R_H\left(\frac{1}{4}-\frac{1}{n^2}\right) \tag{1.2.11}$$

for the sequence of wavelengths of light emitted, with $n=3,4,5,6$ being in the visible, the lines used by Balmer in finding the formula. Hansen would doubtless have informed Bohr that the $1/4$ could be replaced by $1/m^2$, with m another integer. The constant appearing on the right hand side is called the *Rydberg constant*, R_H = 109,737 cm^{-1}. (This is the modern value—Balmer got it right to one part in 10,000, about the limit of spectral measurements at the time.)

Bohr said later: "As soon as I saw Balmer's formula, the whole thing was immediately clear to me." What he saw was that the set of allowed *frequencies* (proportional to inverse wavelengths) emitted by the hydrogen atom could all be expressed as *differences*. This immediately suggested to him a generalization of his idea of a "stationary state" lowest energy level, in which the electron did not radiate. There must be a *whole sequence* of these stationary states, with radiation only taking place as the atom jumps from one to another of lower energy, emitting a single quantum of frequency f such that

$$hf=E_n-E_m \tag{1.2.12}$$

the difference between the energies of the two states.

Evidently, from the Balmer formula and its extension to general integers m, n, these allowed non-radiating orbits, the stationary states, could be labeled 1, 2, 3, ... ,n, ... and had energies

$$E_n=-hcR_H/n^2 \tag{1.2.13}$$

using $\lambda f=c$ and the Balmer equation above.

The energies are of course negative, because these are bound states, and we take the zero of energy to be where the two particles are at rest infinitely far apart.

Bohr was very familiar with the dynamics of simple circular orbits in an inverse square field. He knew that if the energy of the orbit was $-hcR_H/n^2$, that meant the kinetic energy of the electron, $\frac{1}{2}mv^2=hcR_H/n^2$, and the potential energy would be

$$-\frac{1}{4\pi\varepsilon_0}\cdot\frac{e^2}{r_n}=-\frac{2hcR_H}{n^2} \tag{1.2.14}$$

It immediately follows that the *radius* of the n^{th} orbit is proportional to n^2, and the *speed* in that orbit is proportional to $1/n$.

*It then follows that the **angular momentum of the n^{th} orbit is just proportional to n**: and Bohr knew that Planck's constant, the basis of quantum theory had the dimensions of angular momentum!*

Evidently, then the angular momentum in the n^{th} orbit was nKh, where h is Planck's constant and K is some multiplying factor, the same for all the orbits, still to be determined.

In fact, the value of K follows from the results above. R_H, m, h, and c are all known quantities (R_H being measured experimentally by observing the lines in the Balmer series) so the above formulas immediately give the electron's speed and distance from the nucleus in the n^{th} orbit, and hence its angular momentum. Therefore, by putting in these experimentally determined quantities, we can find K.

Bohr's Semiclassical Argument to Fix the Quantum of Angular Momentum

However, Bohr found a clever *theoretical* way to determine R_H from his model: by equating his prediction of the frequency emitted when an electron goes from one orbit to another in a very large atom with the classical prediction—which would be just the orbital frequency of the electron, how many times per second it goes around, he deduced $K = 1/2\pi$ and from that the Rydberg constant that appeared before is here given in terms of h, m and e. *The rather abstract argument that the quantum predictions must match the known classical results for large slow systems actually fixes the Rydberg constant.*

His argument goes as follows: for the circular orbits

$$\frac{mv^2}{r} = \frac{1}{4\pi\varepsilon_0}\cdot\frac{e^2}{r^2} \quad so \quad mv^2 = \frac{1}{4\pi\varepsilon_0}\cdot\frac{e^2}{r}, \quad K.E. = -\frac{1}{2}P.E., \quad E = -\frac{1}{4\pi\varepsilon_0}\cdot\frac{e^2}{2r}. \tag{1.2.15}$$

With the angular momentum quantized, for the n^{th} orbit:

$$mv_n r_n = nKh \tag{1.2.16}$$

where h is Planck's constant, n an integer, K the unknown multiplying factor (ok, fixed by experiment, but we're finding it independently).

From this quantization condition we can find the radius, and hence the energy, of the n^{th} orbit:

$$-\frac{1}{4\pi\varepsilon_0}\cdot\frac{e^2}{r_n} = mv_n^2 = m\left(\frac{nKh}{mr_n}\right)^2 \tag{1.2.17}$$

Giving

$$r_n = \frac{4\pi\varepsilon_0 n^2 K^2 h^2}{me^2}, \quad E_n = -\frac{1}{4\pi\varepsilon_0}\cdot\frac{e^2}{2r_n} = -\left(\frac{1}{4\pi\varepsilon_0}\right)^2\cdot\frac{me^4}{2K^2h^2}\cdot\frac{1}{n^2} \tag{1.2.18}$$

In the large n limit,

$$E_{n+1} - E_n \cong \left(\frac{1}{4\pi\varepsilon_0}\right)^2\cdot\frac{me^4}{2K^2h^2}\cdot\frac{2}{n^3} = h\nu \tag{1.2.19}$$

so

$$\nu = \left(\frac{1}{4\pi\varepsilon_0}\right)^2\cdot\frac{me^4}{K^2h^3n^3} \tag{1.2.20}$$

where ν is the frequency of the emitted photon on jumping down one quantum number.

In the classical limit of large n, ν *must* match the orbital frequency of the electron, since Maxwell's equations will be valid. That is,

$$\nu = \frac{v_n}{2\pi r_n} = \frac{nKh}{2\pi m r_n^2} = \left(\frac{1}{4\pi\varepsilon_0}\right)^2\cdot\frac{nKh}{2\pi m}\cdot\frac{m^2e^4}{n^4K^4h^4} = \left(\frac{1}{4\pi\varepsilon_0}\right)^2\cdot\frac{me^4}{2\pi n^3K^3h^3} \tag{1.2.21}$$

Comparing the two expressions, we see that they agree if $K = 1/2\pi$

Putting $K = 1/2\pi$ into the energy level formula,

$$E_n = -\left(\frac{1}{4\pi\varepsilon_0}\right)^2\cdot\frac{me^4}{2K^2h^2}\cdot\frac{1}{n^2} = -\left(\frac{1}{4\pi\varepsilon_0}\right)^2\cdot\frac{2\pi^2me^4}{h^2}\cdot\frac{1}{n^2}. \tag{1.2.22}$$

Now the Rydberg constant is defined by

$$E_n = -hcR_H/n^2 \tag{1.2.23}$$

so the Bohr model *predicts* that

$$R_H = \left(\frac{1}{4\pi\varepsilon_0}\right)^2 \cdot \frac{2\pi^2 me^4}{ch^3} \tag{1.2.24}$$

This formula was found to be correct within the limits of experimental error in measuring the quantities on the right.

But few people believed his theory. For one thing, it soon became apparent that in the spectra of some stars (actually including the sun) there were spectral lines apparently corresponding to *half* the angular momentum quantum. How could that be?

Bohr’s response was that these lines must be from ionized helium, not hydrogen. A neutral helium atom has two electrons, a singly-ionized helium atom has just one electron, but the nucleus has a charge twice that of the hydrogen nucleus, so the factor e^4 in the Rydberg constant is replaced by $4e^4$, which leads to the observed result. But then a spectroscopist called Fowler did some very precise measurements, and found that actually the R_H for these new lines corresponded to a factor of 4.0016. How could Bohr explain *that*?

Bohr pointed out that at this level of precision, the finite mass of the nucleus must be taken into account by using a reduced mass for the electron. This gives just the right factor. This result greatly impressed Einstein, who concluded that Bohr must be on the right track.

Remark: the Bohr Atom is Still Important!

Although, as we shall see shortly, Bohr’s semiclassical analysis has long been replaced by Schrödinger’s wave function, there are recent experiments in atomic physics where the classical approach provides valuable insight. In particular, so-called Rydberg atoms, which are atoms with one electron in a spatially large orbit (large n, weakly bound), act a lot like classical systems. Such atoms can be ionized by microwave fields. For a considerable range of parameters, the onset of this ionization can be accounted for by ignoring quantum mechanics altogether, and interpreting ionization as the onset of chaotic motion in the *classical* driven system! (And, the standard perturbation theoretic methods of quantum mechanics don’t work for this system anyway, because the perturbing microwave electric field is of the same order of magnitude as the atom’s electric field at these large orbits.) We should mention that, counterintuitively, quantum mechanics *does* become important again at *very* large n (or high microwave frequency), where some tricks from condensed matter physics have been used successfully to interpret the experiments. This is a rich subject: qualitatively different phenomena occur as the ratio of microwave frequency to orbital frequency is varied.

Prince Louis de Broglie Gets His Ph.D.

The next real advance in understanding the atom came from an unlikely quarter—a student prince in Paris. Prince Louis de Broglie was a member of an illustrious family, prominent in politics and the military since the 1600’s. Louis began his university studies with history, but his elder brother Maurice studied *x*-rays in his own laboratory, and Louis became interested in physics. He worked with the very new radio telegraphy during the war.

After the war, de Broglie focused his attention on Einstein's two major achievements, the theory of special relativity and the quantization of light waves. He wondered if there could be some connection between them. Perhaps the quantum of radiation really should be thought of as a particle. It had been known for a long time that light waves carry momentum: this is famously demonstrated by the “radiometer”, a small “windmill” in a vacuum, with vanes silver on one side and blackened on the other. If the vacuum is good, the radiometer begins to rotate when exposed to light because the light bouncing off the silvered side delivers twice the momentum of the light absorbed by the blackened side. (It should be added that cheap versions of this device have poor vacua, and the heated gas near the blackened side tends to push the vanes the wrong way.)

In fact, it follows from Maxwell’s equations that the momentum density of a light beam is related to its energy density by $E = cp$. We would therefore expect this same energy-momentum relationship to be true for the photons of which the light beam is composed. Now, from special relativity we know that all particles have an energy-momentum relationship $E^2 = m_0^2c^4 + c^2p^2$, where m_0 is the rest mass of the particle. The only way this can be the same as $E = cp$ is if $m_0 = 0$, or, at least, if m_0 is so small that all our observations are on particles having kinetic energy so far in excess of their rest energy that the tiny mass is not detectable. De Broglie suspected that the photon *did* have a very tiny nonzero rest mass, so that if the speed of a sufficiently low energy quantum could be measured, it would be found to be less than c. On this point he was wrong (as far as we know!) Nevertheless, it was a very valuable conceptual breakthrough to think of the quantum of radiation as a *particle*, knowing full well that radiation is a wave. In fact, his incorrect idea that the photon (as we now call the light quantum) had a rest mass led him to analyze the relationship between particle properties and wave properties by transforming to the rest frame of the photon, and he discovered that the energy and momentum of the particle were related to the frequency and wavelength of the wave by:

$$E = hf\ ,\ p = h/\lambda \tag{1.2.25}$$

Of course, the first condition is the Planck-Einstein quantization, and the second follows trivially from it if we take $E = cp$ and $\lambda f = c$. But de Broglie showed it was more generally true—it worked even if the photon had a rest mass.

Having decided that the photon might well be a particle with a rest mass, albeit very small, it dawned on de Broglie that in other respects it might not be too different from other particles, especially the very light electron. In particular, maybe the electron also had an associated wave. The obvious objection was that if the electron was wavelike, why had no diffraction or interference effects been observed? But there was an answer. If de Broglie's relation between momentum and wavelength, $p = h/\lambda$ also held for electrons, the wavelength was sufficiently short that these effects would be easy to miss. As de Broglie himself pointed out, the wave nature of light isn't very evident in everyday life, or in ray tracing in geometrical optics. He suspected the apparently pure particle nature of electronic trajectories was analogous to the apparent straight-line propagation of rays of light, over distance scales much greater than the wavelength.

However, the wavelike properties should be important on an *atomic* scale. No progress had been made in a decade in understanding why the electronic orbits in the Bohr atom were restricted to integral values of the angular momentum in units of h. But if the electron were in some sense a wave, it would be very natural to restrict the orbits to those of standing waves, for otherwise the electron wave on going around the orbit would interfere with itself destructively.

Suppose now the electron, having momentum p, is moving in a circular orbit of radius r. Then for a standing wave, a whole number of wavelengths must fit around the circle, so for some integer n, $n\lambda = 2\pi r$. Putting this together with $p = h/\lambda$ we find:

$$2\pi r = n\lambda = nh/p \tag{1.2.26}$$

so

$$L = pr = nh/2\pi \tag{1.2.27}$$

The "standing wave" condition immediately gives Bohr's quantization of angular momentum!

This was the prince's Ph. D. thesis, presented in 1924. His thesis advisor was somewhat taken aback, and wasn't sure if this was sound work. He asked de Broglie for an extra copy of the thesis, which he sent to Einstein. Einstein wrote shortly afterwards: "*I believe it is a first feeble ray of light on this worst of our physics enigmas*". The prince got his Ph. D.

An Accident at the Phone Company Makes Everything Crystal Clear

There was an accident at the Bell Telephone Laboratories in April 1925. Clinton Davisson and L. H. Germer, looking for ways to improve vacuum tubes, were watching how electrons from an electron gun in a vacuum tube scattered off a flat nickel surface. Suddenly, while the experiment was running and the nickel target was very hot, a bottle of liquid air near the apparatus exploded, smashing one of the vacuum pipes, and air rushed into the apparatus. The hot nickel target oxidized immediately. The layer of oxide made their target useless for further investigations. They decided to clean off the oxide by heating the nickel in a hydrogen atmosphere then in vacuum. After doing this for a prolonged period, the nickel looked good, and they resumed the investigation.

To their amazement, the pattern of electron scattering from the newly cleaned nickel target was completely different from that before the accident. What had changed? On examining their newly cleaned crystal carefully, they found a clue. The original target was polycrystalline—made up of a multitude of tiny crystals, oriented randomly. During the prolonged heating of the cleaning process, the nickel had re-crystallized into a few large crystals.

To quote from their paper: "It seemed probable to us from these results that the intensity of scattering from a single crystal would exhibit a marked dependence on crystal direction, and we set about at once preparing experiments for an investigation of this dependence. We must admit that the results obtained in these experiments have proved to be quite at variance with our expectations. It seemed likely that strong beams would be found issuing from the crystal along what may be termed its transparent directions—the directions in which the atoms in the lattice are arranged along the smallest number of lines per unit area. Strong beams are indeed found issuing from the crystal, but only when the speed of bombardment lies near one or another of a series of critical values, and then in directions quite unrelated to crystal transparency.

"The most striking characteristic of these beams is a one to one correspondence ...which the strongest of them bear to the Laue beams that would be found issuing from the same crystal if the incident beam were a beam of *x*-rays. Certain others appear to be analogues ... of optical diffraction beams from plane reflection gratings—the lines of these gratings being lines or rows of atoms in the surface of the crystal. Because of these similarities ... a description ... in terms of an equivalent wave radiation ... is not only possible, but most simple and natural. This involves the association of a wavelength with the incident electron beam, and this wavelength turns out to be in acceptable agreement with the value h/mv of the undulatory mechanics, Planck's action constant divided by the momentum of the electron.

"That evidence for the wave nature of particle mechanics would be found in the reaction between a beam of electrons and a single crystal was predicted by Elsasser two years ago—shortly after the appearance of L. de Broglie's original papers on wave mechanics."

The above quotes are from *Physical Review* **30**, 705 (1927).

It should be added that the two-slit diffraction pattern is of course exhibited by a beam of electrons, has been observed experimentally many times, and has precisely the same form as that for light. Electrons and photons generate interference patterns that are identical—although the short wavelength of the electrons used presents a challenge! A double slit used by C. Jönsson in 1961 consisted of slits 0.5 microns wide 1-2 microns apart in copper foil. See D. Brandt and S Hirschi, *Am. J. Phys*. **42**, 5 (1974). (This reference from French and Taylor's *Introduction to Quantum Physics*.)

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

1.3: Wave Equations, Wavepackets and Superposition

A Challenge to Schrödinger

De Broglie’s doctoral thesis, defended at the end of 1924, created a lot of excitement in European physics circles. Shortly after it was published in the fall of 1925 Pieter Debye, a theorist in Zurich, suggested to Erwin Schrödinger that he give a seminar on de Broglie’s work. Schrödinger gave a polished presentation, but at the end Debye remarked that he considered the whole theory rather childish: why should a wave confine itself to a circle in space? It wasn’t as if the circle was a waving circular string, real waves in space diffracted and diffused, in fact they obeyed three-dimensional wave equations, and that was what was needed. This was a direct challenge to Schrödinger, who spent some weeks in the Swiss mountains working on the problem, and constructing his equation.

There is no rigorous derivation of Schrödinger’s equation from previously established theory, but it can be made very plausible by thinking about the connection between light waves and photons, and construction an analogous structure for de Broglie’s waves and electrons (and, later, other particles).

Maxwell’s Wave Equation

Let us examine what Maxwell’s equations tell us about the motion of the simplest type of electromagnetic wave—a monochromatic wave in empty space, with no currents or charges present.

As we discussed in the last lecture, Maxwell found the wave equation

$$\nabla^2 \vec{E} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = 0 \tag{1.3.1}$$

which reduces to

$$\frac{\partial^2 \vec{E}}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = 0 \tag{1.3.2}$$

for a plane wave moving in the *x*-direction, with solution

$$\vec{E}(x,t) = \vec{E}_0 e^{i(kx-\omega t)} \tag{1.3.3}$$

Applying the wave equation differential operator to this plane wave solution

$$\left(\frac{\partial^2}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\vec{E}_0 e^{i(kx-\omega t)} = \left(k^2 - \frac{\omega^2}{c^2}\right)\vec{E}_0 e^{i(kx-\omega t)} = 0 \tag{1.3.4}$$

so

$$\omega = ck \tag{1.3.5}$$

This is just the familiar statement that the wave must travel at c.

What does the Wave Equation tell us about the *Photon*?

We know from the photoelectric effect and Compton scattering that the photon energy and momentum are related to the frequency and wavelength of the light by

$$E = h\nu = \hbar\omega \tag{1.3.6}$$

$$p = \frac{h}{\lambda} = \hbar k \tag{1.3.7}$$

Notice, then, that the wave equation tells us that $\omega = ck$ and hence $E = cp$.

To put it another way, if we think of $e^{i(kx-\omega t)}$ as describing a particle (photon) it would be more natural to write the plane wave as

$$\vec{E}_0 e^{\frac{i}{\hbar}(px-Et)} \tag{1.3.8}$$

that is, in terms of the energy and momentum of the particle.

In these terms, applying the (Maxwell) wave equation operator to the plane wave yields

$$\left(\frac{\partial^2}{\partial x^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\vec{E}_0 e^{\frac{i}{\hbar}(px-Et)}=\left(p^2-\frac{E^2}{c^2}\right)\vec{E}_0 e^{\frac{i}{\hbar}(px-Et)}=0 \tag{1.3.9}$$

or

$$E^2=c^2p^2 \tag{1.3.10}$$

The wave equation operator applied to the plane wave describing the particle propagation yields the energy-momentum relationship for the particle.

Constructing a Wave Equation for a Particle with Mass

The discussion above suggests how we might extend the wave equation operator from the photon case (zero rest mass) to a particle having rest mass m_0. We need a wave equation operator that, when it operates on a plane wave, yields

$$E^2=c^2p^2+m_0^2c^4 \tag{1.3.11}$$

Writing the plane wave function

$$\varphi(x,t)=Ae^{\frac{i}{\hbar}(px-Et)} \tag{1.3.12}$$

where A is a constant, we find we can get $E^2=c^2p^2+m_0^2c^4$ by adding a constant (mass) term to the differentiation terms in the wave operator:

$$\left(\frac{\partial^2}{\partial x^2}-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}-\frac{m_0^2c^2}{\hbar^2}\right)Ae^{\frac{i}{\hbar}(px-Et)}=-\frac{1}{\hbar^2}\left(p^2-\frac{E^2}{c^2}+m_0^2c^2\right)Ae^{\frac{i}{\hbar}(px-Et)}=0 \tag{1.3.13}$$

This wave equation is called the *Klein-Gordon* equation and correctly describes the propagation of relativistic particles of mass m_0. However, it's a bit inconvenient for nonrelativistic particles, like the electron in the hydrogen atom, just as $E^2=m_0^2c^4+c^2p^2$ is less useful than $E=p^2/2m$ for this case.

A Nonrelativistic Wave Equation

Continuing along the same lines, let us assume that a nonrelativistic electron in free space (no potentials, so no forces) is described by a plane wave:

$$\psi(x,t)=Ae^{\frac{i}{\hbar}(px-Et)} \tag{1.3.14}$$

We need to construct a wave equation operator which, applied to this wave function, just gives us the ordinary nonrelativistic energy-momentum relationship, $E=p^2/2m$. The p^2 obviously comes as usual from differentiating twice with respect to x, but the only way we can get E is by having a *single* differentiation with respect to time, so this looks different from previous wave equations:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t}=-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} \tag{1.3.15}$$

This is *Schrödinger's equation* for a free particle. It is easy to check that if $\psi(x,t)$ has the plane wave form given above, the condition for it to be a solution of this wave equation is just $E=p^2/2m$.

Notice one remarkable feature of the above equation—the i on the left means that ψ *cannot* be a real function.

How Does a Varying Potential Affect a de Broglie Wave?

The effect of a potential on a de Broglie wave was considered by Sommerfeld in an attempt to generalize the rather restrictive conditions in Bohr's model of the atom. Since the electron was orbiting in an inverse square force, just like the planets around the sun, Sommerfeld couldn't understand why Bohr's atom had only circular orbits, no Kepler-like ellipses. (Recall that all the observed spectral lines of hydrogen were accounted for by energy differences between *circular* orbits.)

De Broglie's analysis of the allowed circular orbits can be formulated by assuming at some instant in time the spatial variation of the wave function on going around the orbit includes a phase term of the form $e^{ipq/\hbar}$, where here the parameter q measures distance around the orbit. Now for an acceptable wave function, the total phase change on going around the orbit must be $2n\pi$, where n is an integer. For the usual Bohr circular orbit, p is constant on going around, q changes by $2\pi r$, where r is the radius of the orbit, giving

$$\frac{1}{\hbar} p 2\pi r = 2n\pi, \ so\, pr = n\hbar, \tag{1.3.16}$$

the usual angular momentum quantization.

What Sommerfeld did was to consider a general Kepler ellipse orbit, and visualize the wave going around such an orbit. Assuming the usual relationship $p = h/\lambda$, the wavelength will vary as the particle moves around the orbit, being shortest where the particle moves fastest, at its closest approach to the nucleus. Nevertheless, the phase change on moving a short distance Δq should still be $p\Delta q/\hbar$, and requiring the wave function to link up smoothly on going once around the orbit gives

$$\oint p dq = nh \tag{1.3.17}$$

Thus only certain elliptical orbits are allowed. The mathematics is nontrivial, but it turns out that *every allowed elliptical orbit has the same energy as one of the allowed circular orbits*. That is why Bohr's theory gave all the energy levels. Actually, this whole analysis is old fashioned (it's called the "old quantum theory") but we've gone over it to introduce the idea of *a wave with variable wavelength, changing with the momentum as the particle moves through a varying potential*.

The reader may well be wondering at this point why it is at all useful to visualize a real wave going round an orbit, when we have stated that any solution of Schrödinger's equation is necessarily a complex function. As we shall see, it is often possible to find solutions, including those corresponding to Bohr's energy levels, in which the complex nature of the wave function only appears in a time varying phase factor, $e^{-iEt/\hbar}$. We should also add that if the spatial dependence is a real function, such as $\sin kx$, it represents a standing wave, not a particle circling in one direction, which would be e^{ikx}, or $e^{ipx/\hbar}$. Bearing all this in mind, it is still often instructive to sketch real wave functions, especially for one-dimensional problems.

Schrödinger's Equation for a Particle in a Potential

Let us consider first the one-dimensional situation of a particle going in the *x*-direction subject to a "roller coaster" potential. What do we expect the wave function to look like? We would expect the wavelength to be shortest where the potential is lowest, in the valleys, because that's where the particle is going fastest—maximum momentum.

With a nonzero potential present, the energy-momentum relationship for the particle becomes the energy equation

$$E = \frac{p^2}{2m} + V(x) \tag{1.3.18}$$

We need to construct a wave equation which leads naturally to this relationship. In contrast to the free particle cases discussed above, the relevant wave function here will no longer be a plane wave, since the wavelength varies with the potential. However, at a given x, the momentum is determined by the "local wavelength", that is,

$$p = -i\hbar \frac{\partial \psi}{\partial x} \tag{1.3.19}$$

It follows that the appropriate wave equation is:

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x,t)}{\partial x^2} + V(x)\psi(x,t) \tag{1.3.20}$$

This is the standard one-dimensional Schrödinger equation.

In three dimensions, the argument is precisely analogous. The only difference is that the square of the momentum is now a sum of three squared components, for the x, y and z directions, so $\frac{\partial^2}{\partial x^2}$ becomes $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \nabla^2$,

so now

$$i\hbar \frac{\partial \psi(x,y,z,t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi(x,y,z,t) + V(x,y,z)\psi(x,y,z,t) \tag{1.3.21}$$

This is the complete Schrödinger equation. So far, of course, it is based on plausibility arguments and hand-waving. Why should anyone believe that it really describes an electron wave? Schrödinger's test of his equation was the hydrogen atom. He looked for

Bohr’s “stationary states”: states in which the electron was localized somewhere near the proton, and having a definite energy. The time dependence would be the same as for a plane wave of definite energy, $e^{-iEt/\hbar}$, the spatial dependence would be a time-independent function decreasing rapidly at large distances from the proton. That is, he took

$$\psi(x,y,z,t)=e^{-iEt/\hbar}\psi(x,y,z) \tag{1.3.22}$$

He took advantage of the spherical symmetry by re-expressing the spatial wave function in spherical polar coordinates, and found his equation became a standard differential equation solved in the nineteenth century. The solution gave the shape of possible wave functions, and also allowed values of energy and angular momentum. These values were exactly the same as Bohr’s (except that the lowest allowed state in the new theory had zero angular momentum): impressive evidence that the new theory was correct.

Current Conservation

When Schrödinger published this result in 1926, he also wrote down the complex conjugate equation, and proved that taking them together it was not difficult to deduce a continuity equation:

$$\frac{\partial\rho}{\partial t}+div\vec{j}=0 \tag{1.3.23}$$

$$where\ \rho=\psi^{*}\psi=|\psi|^{2} \tag{1.3.24}$$

$$and\ \ \vec{j}=\frac{\hbar}{2mi}(\psi^{*}\vec{\nabla}\psi-\psi\vec{\nabla}\psi^{*}) \tag{1.3.25}$$

But What Do These Equations Mean?

Schrödinger believed the above continuity equations represented the conservation of electric charge, and had no further significance. He thought that after all his own equation showed the electron to be just a smooth classical wave at the deepest level. In fact, he succeeded in solving the three-dimensional equation with a Coulomb potential and he found the Bohr energy levels of the hydrogen atom. Obviously, he was on the right track! This classical revival approach, however, couldn’t deal with the unpredictability of quantum mechanics, such as where a single photon—or electron—would land in a two-slit diffraction pattern.

The truth is, Schrödinger didn’t understand his own equation. Another physicist, Max Born, published a paper a few days after Schrödinger’s in which he suggested that $|\psi(x,y,z,t)|^{2}dxdydz$ was the *relative probability* of finding the electron in a small volume $dxdydz$ at (x,y,z) at time t. This interpretation was based directly on the analogy with light waves and photons, and has turned out to be correct.

Note

ψ is called the “*amplitude*” or sometimes the “*probability amplitude*”.

Photons and Electrons

We have seen that electrons and photons behave in a very similar fashion—both exhibit diffraction effects, as in the double slit experiment, both have particle like or quantum behavior. As we have already discussed, we now have a framework for understanding photons—we first figure out how the electromagnetic wave propagates, using Maxwell’s equations, that is, we find E as a function of x,y,z,t. Having evaluated $E(x,y,z,t)$, the probability of finding a photon in a given small volume of space $dxdydz$, at time t, is proportional to $|E(x,y,z,t)|^{2}dxdydz$, the energy density.

Born assumed that Schrödinger’s wave function for the electron corresponded to the electromagnetic wave for the photon in the sense that the square of the modulus of the Schrödinger wave amplitude at a point was the relative probability density for finding the electron at that point. So the routine is the same: for given boundary conditions and a given potential, Schrödinger’s differential equation can be solved and the wave function $\psi(x,y,z,t)$ evaluated. Then, $|\psi(x,y,z,t)|^{2}dxdydz$ gives the relative probability of finding the electron at (x,y,z) at time t.

Notice, though, that this interpretation of the wave function is *not* essential in finding the allowed energy levels in a given potential, such as the Bohr orbit energies, which Schrödinger derived before the physical significance of his wave function was understood.

How Wave *Amplitude* Varies in a Roller Coaster Potential

We mentioned above that for an electron traveling along a roller coaster potential, the local wavelength is related to the momentum of the electron as it passes that point.

Perhaps slightly less obvious is that the *amplitude* of the wave varies: it will be largest at the tops of the hills (provided the particle has enough energy to get there) because that's where the particle is moving slowest, and therefore is most likely to be found.

Keeping the Wave and the Particle Together?

Suppose following de Broglie we write down the relation between the "particle properties" of the electron and its "wave properties":

$$\frac{1}{2}mv^2 = E = hf\,, \quad mv = p = h/\lambda \tag{1.3.26}$$

It would seem that we can immediately figure out the speed of the wave, just using $\lambda f = c$, say. We find:

$$\lambda f = \frac{h}{mv} \cdot \frac{\frac{1}{2}mv^2}{h} = \frac{1}{2}v \tag{1.3.27}$$

So the speed of the wave seems to be only half the speed of the electron! How could they stay together? What's wrong with this calculation?

Localizing an Electron

To answer this question, it is necessary to think a little more carefully about the wave function corresponding to an electron traveling through a cathode ray tube, say. The electron leaves the cathode, shoots through the vacuum, and impinges on the screen. At an intermediate point in this process, it is moving through the vacuum and the wave function must be nonzero over some volume, but zero in the places the electron has not possibly reached yet, and zero in the places it has definitely left.

However, if the electron has a precise energy, say exactly a thousand electron volts, it also has a precise momentum. This necessarily implies that the wave has a precise wavelength. But the only wave with a *precise* wavelength λ has the form

$$\psi(x,t) = Ae^{i(kx-\omega t)} \tag{1.3.28}$$

where $k = 2\pi/\lambda$, and $\omega = 2\pi f$. The problem is that this plane sine wave extends to infinity in both spatial directions, so cannot represent a particle whose wave function is nonzero in a limited region of space.

Therefore, to represent a localized particle, we must superpose waves having different wavelengths. Now, the waves representing electrons, *unlike the light waves representing photons*, travel at different speeds for different energies. Any intuition gained by thinking about superposing light waves of different wavelengths can be *misleading* if applied to electron waves!

Fortunately, there are many examples in nature of waves whose speed depends on wavelength. A simple example is water waves on the ocean. We consider waves having a wavelength much shorter than the depth of the ocean. What is the ω, k relationship for these waves? We know it's not $\omega = Ck$, with a constant C, because waves of different wavelengths move at different speeds. In fact, it's easy to figure out the ω, k relationship, known as the *dispersion relation*, for these waves from a simple dimensional argument. What physical parameters can the wave frequency depend on? Obviously, the wavelength λ. We will use $k = 2\pi/\lambda$ as our variable. k has dimensions L^{-1}.

These waves are driven by gravity, so g, with dimensions LT^{-2}, is relevant. Actually, that's all. For ocean waves, surface tension is certainly negligible, as is the air density, and the water's viscosity. You might think the density of the water matters, but these waves are rather like a pendulum, in that they are driven by gravity, so increasing the density would increase both force and inertial mass by the same amount.

For these deepwater waves, then, dimensional analysis immediately gives:

$$\omega^2 = Cgk \tag{1.3.29}$$

where C is some dimensionless constant we cannot fix by dimensional argument, but actually it turns out to be 1.

Wavepackets and the Principle of Superposition

To return momentarily to the electron traveling through a vacuum, it is clear physically that it must have a wave function that goes to zero far away in either direction (we'll still work in one dimension, for simplicity). A localized wave function of this type is called a "wavepacket". We shall discover that a wavepacket can be constructed by adding plane waves together. Now, the plane waves we add together will individually be solutions of the Schrödinger equation.

But does it follow that the *sum* of such solutions of the Schrödinger equation is itself a solution to the equation? The answer is yes—in other words, the Schrödinger equation

$$i\hbar\frac{\partial\psi(x,y,z,t)}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\psi(x,y,z,t)+V(x,y,z)\psi(x,y,z,t) \tag{1.3.21}$$

is a *linear* equation, that is to say, if $\psi_1(x,y,z,t)$, $\psi_2(x,y,z,t)$ are both solutions of the equation, then so is

$$\psi(x,y,z,t)=c_1\psi_1(x,y,z,t)+c_2\psi_2(x,y,z,t) \tag{1.3.30}$$

where c_1 and c_2 are arbitrary constants, as is easy to check. This is called the *Principle of Superposition*.

The essential point is that in Schrödinger's equation every term contains a factor ψ, but no term contains a factor ψ^2 (or a higher power). That's what is meant by a "linear" equation. If the equation *did* contain a constant term, or a term including ψ^2, superposition wouldn't work—the sum of two solutions to the equation would not itself be a solution to the equation.

In fact, we have been assuming this linearity all along: when we analyze interference and diffraction of waves, we just add the two wave amplitudes at each spot. For the double slit, we take it that if the wave radiating from one slit satisfies the wave equation, then adding the two waves together will give a new wave which also satisfies the equation.

The First Step in Building a Wavepacket: Adding Two Sine Waves

If we add together two sine waves with frequencies close together, we get beats. This pattern can be viewed as a string of wavepackets, and is useful for gaining an understanding of why the electron speed calculated from $\lambda f=c$ above is apparently half what it should be.

We use the trigonometric addition formula:

$$\sin((k-\Delta k)x-(\omega-\Delta\omega)t)+\sin((k+\Delta k)x-(\omega+\Delta\omega)t)=2\sin(kx-\omega t)\cos((\Delta k)x-(\Delta\omega)t) \tag{1.3.31}$$

This formula represents the phenomenon of beats between waves close in frequency. The first term, $\sin(kx-\omega t)$, oscillates at the average of the two frequencies. It is modulated by the slowly varying second term, often called the "envelope function", which oscillates once over a spatial extent of order $\pi/\Delta k$. This is the distance over which waves initially in phase at the origin become completely out of phase. Of course, going a further distance of order $\pi/\Delta k$, the waves will become synchronized again.

That is, beating two close frequencies together breaks up the continuous wave into a series of packets, the beats. To describe a single electron moving through space, we need a single packet. This can be achieved by superposing waves having a continuous distribution of wavelengths, or wave numbers within of order Δk, say, of k. In this case, the waves will be out of phase after a distance of order $\pi/\Delta k$ but since they have many different wavelengths, they will never get back in phase again.

Phase Velocity and Group Velocity

It will immediately become apparent that there are *two different velocities* in the dynamics: first, the velocity with which the *individual peaks* move to the right, and second the velocity at which the slowly varying envelope function—the beat pattern—moves. The $\lambda f=c$ individual peak velocity is determined by the term $\sin(kx-\omega t)$, it is ω/k: this is called the *phase velocity*. The speed with which the beat pattern moves, on the other hand, is determined by the term $\cos((\Delta k)x-(\Delta\omega)t)$, this speed is $\Delta\omega/\Delta k=d\omega/dk$ for close frequencies.

Going back one more time to the electron wavepacket, armed with this new insight, we can see immediately that the wave speed we calculated from $\lambda f=c$ was the *phase* velocity of the waves. The packet itself will of course move at the *group* velocity—and it is easy to check that this is just*v*.

Adding More Waves

We’ve seen how two sine waves of equal amplitude close together in frequency produce beats: if the waves are in phase at the origin, as we go along the *x*-axis they gradually fall out of phase, and cancel each other at a distance $x = \pi/2\Delta$, where 2Δ is the difference in k of the two $\sin kx$ waves. (For the moment, we are ignoring the time development of these waves: we’re just looking at *t* = 0.). If we continue along the *x*-axis to π/Δ, the two waves will be back in phase again, this is the next beat. Now, if instead of adding *two* waves, we add *many* waves, all of different k, *but* with the k’s taken from some small interval of size of order Δk, and all these waves are in phase at the origin, then, again, they will all stay more or less in phase for a distance of order $x = \pi/2\Delta$. However, as we proceed past that point, the chances of them all getting back in phase again get rapidly smaller as we increase the number of different waves.

This suggests a way to construct a wavepacket: add together a lot of waves from within a narrow frequency range, and they will *only* be in phase in a region containing the origin.

Adding waves in this way leads to a more general derivation of the formula $d\omega/dk$ for the group velocity. The standard approach is to replace the sum over plane waves by an integral, with the wavenumber k as the variable of integration, and the convention is to put a factor 2π in the denominator:

$$\psi(x,t) = \int\limits_{-\infty}^{+\infty} \frac{dk}{2\pi} e^{ikx - i\omega(k)t}\,\phi(k) \tag{1.3.32}$$

Since we are constructing a wavepacket with a fairly well-defined momentum, we will take the function $\phi(k)$ to be strongly peaked at k_0, and going rapidly to zero away from that value, so the only significant contribution to the integral is from the neighborhood of k_0. Therefore, if $\omega(k)$ is reasonably smooth (which it is) it is safe to put

$$\omega(k) = \omega(k_0) + (k - k_0)\omega'(k_0) \tag{1.3.33}$$

in the exponential.

This gives

$$\psi(x,t) = \int\limits_{-\infty}^{+\infty} \frac{dk}{2\pi} e^{ikx - i\omega(k_0)t - i(k-k_0)\omega'(k_0)t}\,\phi(k) = e^{i(k_0x - \omega(k_0)t)} \int\limits_{-\infty}^{+\infty} \frac{dk}{2\pi} e^{i(k-k_0)(x-\omega'(k_0)t)}\,\phi(k) \tag{1.3.34}$$

The first term just represents a single wave at k_0, and the peaks move at the phase velocity

$$v_{phase} = \omega/k \tag{1.3.35}$$

The second term, the integral, is the *envelope function*: here *x* only appears in the combination

$$x - \omega'(k_0)t \tag{1.3.36}$$

so the envelope, and hence the wavepacket, moves to the right at the group velocity: $v_{group} = \omega'(k_0)$. Note that if the *next* term in the Taylor expansion of $\omega(k)$ is also included, that amounts to adding wavepackets with slightly different group velocities together, and the initial (total) wavepacket will gradually widen.

The Gaussian Wavepacket

Fortunately, there is a simple explicit mathematical realization of the addition of plane waves to form a localized function: the *Gaussian wavepacket*,

$$\psi(x, t = 0) = Ae^{ik_0x}e^{-x^2/2\Delta^2} \tag{1.3.37}$$

where $p_0 = \hbar k_0$. For this wavepacket to represent one electron, with the probability of finding the electron in a small section of length dx at x equal to $|\psi|^2 dx$, and the *total* probability of finding the electron somewhere equal to one, the constant A is uniquely determined (apart from a possible phase multiplier $e^{i\delta}$, which would not affect the probability).

Using the standard result

$$\int_{-\infty}^{+\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}} \tag{1.3.38}$$

we find $|A|^2 = (\pi\Delta^2)^{-1/2}$ so

$$\psi(x, t=0) = \frac{1}{(\pi\Delta^2)^{1/4}} e^{ik_0x} e^{-x^2/2\Delta^2} \tag{1.3.39}$$

But how do we construct this particular wavepacket by superposing plane waves? That is to say, we need a representation of the form:

$$\psi(x) = \int_{-\infty}^{+\infty} \frac{dk}{2\pi} e^{ikx}\, \phi(k) \tag{1.3.40}$$

The function $\phi(k)$ represents the weighting of plane waves in the neighborhood of wavenumber k. This is a particular example of a *Fourier transform*—we will be discussing the general case in detail a little later in the course. Note that if $\phi(k)$ is a bounded function, any particular k value gives a vanishingly small contribution, the plane-wave contribution to \(y(x)\) from a range dk is $\phi(k)dk/2\pi$. In fact, $\phi(k)$ is given in terms of $\psi(x)$ by

$$\phi(k) = \int_{-\infty}^{+\infty} dx e^{-ikx}\, \psi(x) \tag{1.3.41}$$

It is perhaps worth mentioning at this point that this can be understood *qualitatively* by observing that the plane wave prefactor e^{-ikx} will interfere destructively with all plane wave components of $\psi(x)$ except that of wavenumber k, where it may at first appear that the contribution is infinite, but recall that as stated above, any particular k component has a vanishingly small weight—and, in fact, this is the right answer, as we shall show in more convincing fashion later.

In the present case, the above handwaving argument is unnecessary, because both the integrals can be carried out exactly, using the standard result:

$$\int_{-\infty}^{+\infty} e^{-ax^2+bx}\, dx = e^{b^2/4a}\sqrt{\frac{\pi}{a}} \tag{1.3.42}$$

giving

$$\phi(k) = (4\pi\Delta^2)^{\frac{1}{4}} e^{-\Delta^2(k-k_0)^2/2} \tag{1.3.43}$$

Putting this back in the integral for $\psi(x)$ shows that the integral equations are consistent.

Note the normalization integrals in *x*-space and *k*-space are:

$$\int_{-\infty}^{+\infty} |\psi|^2 dx = 1, \quad \int_{-\infty}^{+\infty} |\phi(k)|^2 \frac{dk}{2\pi} = 1 \tag{1.3.44}$$

The physical significance of the second equation above is that if the wavepacket goes through a diffraction grating so that the different *k*-components are dispersed in different directions, like the colors in white light, and a detector is arranged to register the electron if it has wavenumber between k and $k+dk$, the probability of finding it in that wavenumber range is $|\phi(k)|^2 dk/2\pi$.

Expectation Values and the Uncertainty Principle

It is clear from the expressions for $\psi(x)$ and its Fourier transform $\phi(k)$ above that the spreading of the wave function in *x*-space is inversely related to its spreading in *k*-space: the *x*-space wavefunction has spread $\sim\Delta$, the *k*-space wavefunction $\sim 1/\Delta$. This is perhaps the simplest example of Heisenberg’s famous Uncertainty Principle: in quantum mechanics, both the position and momentum of a particle cannot be known precisely at the same moment; the more exactly one is specified the less well the other is known. This is an inevitable consequence of the wave nature of the probability distribution. As we have already seen, a particle with an exact momentum has a wave of specific wavelength, and the only such wave is a plane wave extending from minus infinity to infinity, so the position of the particle is completely undetermined. A particle with precisely defined position is described by a wavepacket having all wavelengths included with equal weight—the momentum is completely undefined. We shall give more examples of the Uncertainly Principle, of efforts to evade it and of its uses in estimates, in the next lecture.

Definitions of $\Delta x, \Delta p$

The standard notation for the *expectation value* of an operator in a given quantum state is

$$\langle x\rangle = \int x|\psi(x)|^2 dx \tag{1.3.45}$$

In other words, $\langle x\rangle$ would be the statistical average outcome of making many measurements of x on identically prepared systems all in the quantum state $\psi(x)$ (ignoring the time dependence here for the sake of simplicity).

When we talk about the "uncertainty" in x, *we mean in quantum mechanics the root mean square deviation in the measurements*. This is usually written Δx (unfortunate in view of our—also standard—use of Δ in the Gaussian function above, so the reader should watch carefully!).

Therefore

$$\Delta x = \sqrt{\int (x-\langle x\rangle)^2 |\psi(x)|^2 dx} \tag{1.3.46}$$

For our wavepacket, $\langle x\rangle = 0$. It is easy to check that

$$\Delta x = \frac{\Delta}{\sqrt{2}}, \quad and\ writing\ \ p = \hbar k, \quad \Delta p = \frac{\hbar}{\Delta\sqrt{2}} \quad giving\ \ \Delta x\cdot\Delta p = \frac{\hbar}{2} \tag{1.3.47}$$

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

1.4: The Uncertainty Principle

Waves are Fuzzy

As we have shown for wavepackets, the *wave nature* of particles implies that we cannot know *both* position *and* momentum of a particle to an arbitrary degree of accuracy—if Δx represents the uncertainty in our knowledge of position, and Δp that of momentum, then

$$\Delta p \Delta x \sim h \tag{1.4.1}$$

where h is Planck's constant. In the real world, particles are three-dimensional and we should say

$$\Delta p_x \Delta x \sim h \tag{1.4.2}$$

with corresponding equations for the other two spatial directions. The fuzziness about position is related to that of momentum *in the same direction*.

Let's see how this works by trying to measure *y*-position and *y*-momentum very accurately. Suppose we have a source of electrons, say, an electron gun in a CRT (cathode ray tube, such as an old-fashioned monitor). The beam spreads out a bit, but if we interpose a sheet of metal with a slit of width w, then for particles that make it through the slit, we know y with an uncertainty $\Delta y = w$. Now, if the slit is a long way downstream from the electron gun source, we also know p_y very accurately as the electron reaches the slit, because to make it to the slit the electron's velocity would have to be aimed just right.

But does the measurement of the electron's *y* position—in other words, having it go through the slit—affect its *y* momentum? The answer is *yes*. If it didn't, then sending a stream of particles through the slit they would all hit very close to the same point on a screen placed further downstream. But we know from experiment that this is *not* what happens—a single slit diffraction pattern builds up, of angular width $\theta \sim \lambda / w$, where the electron's de Broglie wavelength λ is given by $p_x \cong h / \lambda$ (there is a negligible contribution to λ from the *y*-momentum). The consequent uncertainty in p_y is

$$\Delta p_y / p_x \sim \theta \sim \lambda / w \tag{1.4.3}$$

Putting in $p_x = h / \lambda$, we find immediately that

$$\Delta p_y \sim h / w \tag{1.4.4}$$

so the act of measuring the electron's *y* position has fuzzed out its *y* momentum by precisely the amount required by the uncertainty principle.

Trying to Beat the Uncertainty Principle

In order to understand the Uncertainty Principle better, let's try to see what goes wrong when we actually try to measure position and momentum more accurately than allowed.

For example, suppose we look at an electron through a microscope. What could we expect to see? Of course, you know that if we try to look at something *really* small through a microscope it gets blurry—a small sharp object gets diffraction patterns around its edges, indicating that we are looking at something of size comparable to the wavelength of the light being used. If we look at something much smaller than the wavelength of light—like the electron—we would expect a diffraction pattern of concentric rings with a circular blob in the middle. The size of the pattern is of order the wavelength of the light, in fact from optics it can be shown to be $\sim \lambda f / d$ where d is the diameter of the object lens of the microscope, f the focal distance (the distance from the lens to the object). We shall take $f / d \sim 1$, as it usually is. So looking at an object the size of an electron should give a diffraction pattern centered on the location of the object. That would seem to pin down its position fairly precisely.

What about the *momentum* of the electron? Here a problem arises that doesn't matter for larger objects—the light we see has, of course, bounced off the electron, and so the electron has some recoil momentum. That is, by bouncing light off the electron we have given it some momentum. Can we say how much? To make it simple, suppose we have good eyes and only need to bounce one photon off the electron to see it. We know the initial momentum of the photon (because we know the direction of the light beam we're using to illuminate the electron) and we know that after bouncing off, the photon hits the object lens and goes through the microscope, but we don't know *where* the photon hit the object lens. The whole point of a microscope is that all the light from a point, light that hits the object lens *in different places*, is all focused back to one spot, forming the image (apart from the blurriness mentioned above). So if the light has wavelength λ, its constituent photons have momentum $\sim h / \lambda$, and from our ignorance of where the photon entered the microscope we are uncertain of its *x*-direction momentum by an amount $\sim h/\lambda$. Necessarily, then, we have the same uncertainty about the electron's *x*-direction momentum, since this was imparted by the photon bouncing off.

But now we have a problem. In our attempts to minimize the uncertainty in the electron's momentum, by only using one photon to detect it, we are not going to see much of the diffraction pattern discussed above—such diffraction patterns are generated by *many* photons hitting the film, retina or whatever detecting equipment is being used. A single photon generates a single point (at best!). This point will most likely be within of order λ of the center of the pattern, but this leaves us with an uncertainty in position of order λ.

Therefore, in attempting to observe the position and momentum of a single electron using a single photon, we find an uncertainty in position $\Delta x \sim \lambda$, and in momentum $\Delta p_x \sim h/\lambda$. These results are in accordance with Heisenberg's Uncertainty Principle $\Delta x \Delta p_x \sim h$.

Of course, we could pin down the position much better if we used N photons instead of a single one. From statistical theory, it is known that the remaining uncertainty $\sim \lambda/\sqrt{N}$. But then N photons have bounced off the electron, so, since each is equally likely to have gone through any part of the object lens, the uncertainly in momentum of the electron as a consequence of these collisions goes up as $\sqrt{N}$. (The same as the average imbalance between heads and tails in a sequence of N coin flips.)

Noting that the uncertainty in the momentum of the electron arises because we don't know where the bounced-off photon passes through the object lens, it is tempting to think we could just use a smaller object lens, that would reduce Δp_x. Although this is correct, recall from above that we stated the size of the diffraction pattern was $\sim \lambda f/d$, where d is the diameter of the object lens and f its focal length. It is easy to see that the diffraction pattern, and consequently Δx, gets bigger by just the amount that Δp_x gets smaller!

Watching Electrons in the Double Slit Experiment

Suppose now that in the double slit experiment, we set out to detect which slit each electron goes through by shining a light just behind the screen and watching for reflected light from the electron immediately after it had passed through a slit. Following the discussion in Feynman's *Lectures in Physics*, Volume III, we shall now establish that if we can detect the electrons, we ruin the diffraction pattern!

Taking the distance between the two slits to be d, the dark lines in the diffraction pattern are at angles

$$(n+\frac{1}{2})\lambda_{elec} = d\sin\theta \tag{1.4.5}$$

If the light used to see which slit the electron goes through generates an uncertainty in the electron's *y* momentum $\Delta p_y{}^{\Delta p_y}$, in order not to destroy the diffraction pattern we must have

$$\Delta p_y/p < \lambda_{elec}/d \tag{1.4.6}$$

(the angular uncertainty in the electron's direction must not be enough to spread it from the diffraction pattern maxima into the minima). Here p is the electron's full momentum, $p = h/\lambda_{elec}$. Now, the uncertainty in the electron's *y* momentum, looking for it with a microscope, is $\Delta p_y \sim h/\lambda_{light}$.

Substituting these values in the inequality above we find the condition for the diffraction pattern to survive is

$$\lambda_{light} > d \tag{1.4.7}$$

the wavelength of the light used to detect which slit the electron went through must be greater than the distance between the slits. Unfortunately, the light scattered from the electron then gives one point in a diffraction pattern of size the wavelength of the light used, so *even if we see the flash this does not pin down the electron sufficiently to say which slit it went through*. Heisenberg wins again.

How the Uncertainty Principle Determines the Size of Everything

It is interesting to see how the actual physical size of the hydrogen atom is determined by the wave nature of the electron, in effect, by the Uncertainty Principle. In the ground state of the hydrogen atom, the electron minimizes its total energy. For a classical atom, the energy would be minus infinity, assuming the nucleus is a point (and very large in any case) because the electron would sit right on top of the nucleus. However, this cannot happen in quantum mechanics. Such a very localized electron would have a very large uncertainty in momentum—in other words, the kinetic energy would be large. This is most clearly seen by imagining that the electron is going in a circular orbit of radius r with angular momentum \(h/2p\). Then one wavelength of the electron's de Broglie wave just fits around the circle, $\lambda_{elec} = 2\pi r$. Clearly, as we shrink the circle's radius r, λ_{elec} goes down proportionately, and the electrons momentum

$$p = h/\lambda_{elec} = h/2\pi r \tag{1.4.8}$$

increases. Adding the electron's electrostatic potential energy we find the total energy for a circular orbit of radius r is:

$$E(r) = K.E. + P.E. = \frac{p^2}{2m} - \frac{e^2}{4\pi\varepsilon_0 r} = \frac{h^2}{8m\pi^2 r^2} - \frac{e^2}{4\pi\varepsilon_0 r} \tag{1.4.9}$$

Notice that for very large r, the potential energy dominates, the kinetic energy is negligible, and shrinking the atom lowers the total energy. However, for small enough r, the (always positive) kinetic energy term wins, and the total energy *grows* as the atom shrinks. Evidently, then, there must be a value of r for which *the total energy is a minimum*. Visualizing a graph of the total energy given by the equation above as a function of r, at the minimum point the slope of $E(r)$ is zero, $dE(r)/dr = 0$.

That is,

$$-\frac{h^2}{4m\pi^2 r} + \frac{e^2}{4\pi\varepsilon_0} = 0 \tag{1.4.10}$$

giving

$$r_{min} = \frac{\varepsilon_0 h^2}{\pi m e^2} \tag{1.4.11}$$

The total energy for this radius is the exact right answer, which is reassuring (but we don't deserve it, because we have used a naïve picture, as will become clear later.)

The point of this exercise is to see that in quantum mechanics, unlike classical mechanics, a particle cannot position itself at the exact minimum of potential energy, because that would require a very narrow wave packet and thus be expensive in kinetic energy. The ground state of a quantum particle in an attractive potential is a trade off between potential energy minimization and kinetic energy minimization. Thus the physical sizes of atoms, molecules and ultimately ourselves are determined by Planck's constant.

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

1.5: Electron in a Box

Plane Wave Solutions

The best way to gain understanding of Schrödinger's equation is to solve it for various potentials. The simplest is a one-dimensional "particle in a box" problem. The appropriate potential is $V(x) = 0$ for x between 0, L and $V(x) = \infty$ otherwise—that is to say, there are infinitely high walls at $x = 0$ and $x = L$, and the particle is trapped between them. This turns out to be quite a good approximation for electrons in a long molecule, and the three-dimensional version is a reasonable picture for electrons in metals.

Between $x = 0$ and $x = L$ we have $V = 0$, so the wave equation is just

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}. \tag{1.5.1}$$

A possible plane wave solution is

$$\psi(x,t) = Ae^{\frac{i}{\hbar}(px-Et)}. \tag{1.5.2}$$

On inserting this into the zero-potential Schrödinger equation above we find $E = p^2/2m$, as we expect.

It is very important to notice that the complex conjugate, proportional to $e^{\frac{i}{\hbar}(px-Et)}$, is *not* a solution to the Schrödinger equation! If we blindly put it into the equation we get

$$E = \frac{-p^2}{2m} \tag{1.5.1}$$

However, this is an unphysical result.

However, a wave function proportional to $e^{-\frac{i}{\hbar}(px-Et)}$ gives $E = p^2/2m$, so this plane wave *is* a solution to the equation.

Therefore, the two allowed plane-wave solutions to the zero-potential Schrödinger equation are proportional to $e^{\frac{i}{\hbar}(px-Et)}$ and $e^{-\frac{i}{\hbar}(px-Et)}$ respectively.

Note that these two solutions have the *same time dependence* $e^{-\frac{iEt}{\hbar}}$.

To decide on the appropriate solution for our problem of an electron in a box, of course we have to bring in the walls—what they mean is that $\psi = 0$ for $x < 0$ and for $x > L$ because remember $|\psi|^2$ tells us the probability of finding the particle anywhere, and, since it's in the box, it's trapped *between* the walls, so there's zero probability of finding it outside.

The condition $\psi = 0$ at $x = 0$ and $x = L$ reminds us of the vibrating string with two fixed ends—the solution of the string wave equation is standing waves of sine form. In fact, taking the difference of the two permitted plane-wave forms above gives a solution of this type:

$$\psi(x,t) = A\sin\frac{px}{\hbar}e^{-\frac{iEt}{\hbar}} \tag{1.5.3}$$

.

This wave function satisfies the Schrödinger equation between the walls, it vanishes at the $x = 0$ wall, it will also vanish at $x = L$ provided that the momentum variable satisfies:

$$\frac{pL}{\hbar} = \pi,\ 2\pi,\ 3\pi\dots \tag{1.5.4}$$

Thus the allowed values of p are $hn/2L$, where $n = 1, 2, 3\dots$, and from $E = p^2/2m$ the allowed energy levels of the particle are:

$$E=\frac{p^2}{2m}=\frac{1}{2m}\left(\frac{h}{2L}\right)^2,\ \frac{4}{2m}\left(\frac{h}{2L}\right)^2,\ \frac{9}{2m}\left(\frac{h}{2L}\right)^2,\dots \tag{1.5.5}$$

Note that these energy levels become more and more widely spaced out at high energies, in contrast to the hydrogen atom potential. (As we shall see, the harmonic oscillator potential gives equally spaced energy levels, so by studying how the spacing of energy levels varies with energy, we can learn something about the shape of the potential.)

What about the overall multiplicative constant A in the wave function? This can be real or complex. To find its value, note that at a fixed time, say $t=0$, the probability of the electron being between x and $x+dx$ is $|\psi|^2 dx$ or

$$|A|^2\sin^2\frac{px}{\hbar}dx \tag{1.5.6}$$

The *total* probability of the particle being *somewhere* between 0, L must be unity:

$$|\psi|^2 dx \tag{1.5.2}$$

or

$$\int\limits_{x=0}^{x=L}|A|^2\sin^2\frac{px}{\hbar}dx=1,\ so\ \frac{1}{2}L|A|^2=1 \tag{1.5.7}$$

Hence

$$\psi(x,t)=\sqrt{\frac{2}{L}}\sin\frac{px}{\hbar}e^{-\frac{iEt}{\hbar}} \tag{1.5.8}$$

When A is fixed in this way, by demanding that the total probability of finding the particle somewhere be unity, it is called the *normalization constant*.

Stationary States

Notice that at a later time the probability distribution for the wave function

$$\psi(x,t)=\sqrt{\frac{2}{L}}\sin\frac{px}{\hbar}e^{-\frac{iEt}{\hbar}} \tag{1.5.9}$$

is *the same*, because *time only appears as a phase factor* in this time-dependent function, and so does not affect $|\psi|^2$.

A state with a time-independent probability distribution is called a ***stationary state****.*

States with Moving Probability Distributions

Recall that the Schrödinger equation is a *linear* equation, and *the sum of any two solutions is also a solution* to the equation. That means that we can add two solutions having *different* energies, and still have a legal wave function. We shall establish that in this case, *the probability distribution varies in time*.

The simplest way to see how this must be is to look at an example. Let's add the ground state to the first excited state, and normalize the sum:

$$\psi(x,t)=\sqrt{\frac{1}{L}}\left(\sin\frac{\pi x}{L}e^{-\frac{i\pi ht}{4mL^2}}+\sin\frac{2\pi x}{L}e^{-\frac{i\pi ht}{mL^2}}\right) \tag{1.5.10}$$

$$note:\ h,\ not\ \hbar. \tag{1.5.3}$$

(You can check the normalization constant at $t=0$). For general x, the two terms in the bracket rotate in the complex plane at different rates, so their sum has a time-varying magnitude. That is to say, $|\psi(x,t)|^2$ varies in time, so the particle must be moving around—this is *not* a stationary state.

Exercise: To see this, note that at $t=0$ the wave function is:

$$\psi(x,0)=\sqrt{\frac{1}{L}}\left(\sin\frac{\pi x}{L}+\sin\frac{2\pi x}{L}\right)\tag{1.5.11}$$

and sketch this function: the particle is more likely to be found in the left-hand half of the box.

Now, suppose the time is $t=4mL^2/h$, so $e^{-\frac{i\pi ht}{4mL^2}}=-1$. At this time,

$$\psi(x,2L^2/h)=\sqrt{\frac{1}{L}}\left(-\sin\frac{\pi x}{L}+\sin\frac{2\pi x}{L}\right)\tag{1.5.11}$$

and it’s easy to see that the particle is more likely to be found in the *right*-hand half.

That is to say, this wave function, a linear sum of wave functions corresponding to different energies, has a probability distribution that sloshes back and forth in the box: and, any attempt to describe a classical-type particle motion, bouncing back and forth, *necessarily* involves adding quantum wave functions of different energies. Note that the frequency of the sloshing motion depends on the *difference* of the two energies: how constructively the two components interfere depends on the difference of the phases in the energies at the time. A single energy wave function always has a static probability distribution.

Of course, the *total* probability of finding the particle *somewhere* in the box remains unity: the normalization constant is time-independent.

The Time-Independent Schrödinger Equation: Eigenstates and Eigenvalues

The only way to prevent $|\psi(x,t)|^2$ varying in time is to have all its parts changing phase in time at the same rate. *This means they all correspond to the same energy*. If we restrict our considerations to such *stationary states*, the wave function can be factorized

$$\psi(x,t)=\psi(x)e^{-\frac{iEt}{\hbar}}\tag{1.5.12}$$

and putting this wave function into the Schrödinger equation we find

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}+V(x)\psi(x)=E\psi(x)\tag{1.5.13}$$

This is the *time-independent Schrödinger equation*, and its solutions are the spatial wave functions for stationary states, states of definite energy. These are often called *eigenstates* of the equation.

The values of energy corresponding to these eigenstates are called the *eigenvalues*.

An Important Point: What, Exactly, Happens at the Wall?

Consider again the wavefunction for the lowest energy state of a particle confined between walls at $x=0$ and $x=L$. The reader should sketch the wavefunction from some point to the left of $x=0$ over to the right of $x=L$. To the left of $x=0$, the wavefunction is exactly zero, then at $x=0$ it takes off to the right (inside the box) as a sine curve. In other words, at the origin the slope of the wavefunction ψ is zero to the left, nonzero to the right. *There is a discontinuity in the slope at the origin: this means the second derivative of* \(\psi\) *is* infinite *at the origin.* On examining the time-independent Schrödinger equation above, we see the equation can *only* be satisfied at the origin because the potential becomes infinite there—the wall is an infinite potential. (And, in fact, since ψ becomes zero on approaching the origin from inside the box, the limit must be treated carefully.)

It now becomes obvious that if the box does not have infinite walls, but merely high ones, ψ describing a confined particle cannot suddenly go to zero at the walls: the second derivative must remain finite. For non-infinite walls, ψ and its derivative must be continuous on entering the wall. This has the important physical consequence that ψ will be nonzero at least for some distance into the wall, even if classically the confined particle does not have enough energy to “climb the wall”. (Which it doesn’t, if it’s confined.) Thus, in quantum mechanics, there is a non-vanishing probability of finding the particle in a region which is “classically forbidden” in the sense that it doesn’t have enough energy to get there.

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

CHAPTER OVERVIEW

2: Some Essential Math

Topic hierarchy

2.1: Fourier Series and Integrals, the Dirac Function
2.2: Linear Algebra
2.3: Function Spaces
2.4: Complex Variable, Stationary Path Integrals

2.1: Fourier Series and Integrals, the Dirac Function

Page ID
1648

We begin with a brief review of Fourier series. Any periodic function of interest in physics can be expressed as a series in sines and cosines—we have already seen that the quantum wave function of a particle in a box is precisely of this form. The important question in practice is, for an arbitrary wave function, how good an approximation is given if we stop summing the series after N terms. We establish here that the sum after N terms, $f_N(\theta)$, can be written as a convolution of the original function with the function

$$\delta_N(x)=(1/2\pi)(\sin(N+\frac{1}{2})x)/\sin\frac{1}{2}x \tag{2.1.1}$$

that is,

$$f_N(\theta)=\int\limits_{-x}^{x}\delta_N(\theta-\theta')f(\theta')d\theta' \tag{2.1.2}$$

The structure of the function $\delta_N(x)$ (plotted below), when put together with the function $f(\theta)$, gives a good intuitive guide to how good an approximation the sum over N terms is going to be for a given function $f(\theta)$. In particular, it turns out that step discontinuities are *never* handled perfectly, no matter how many terms are included. Fortunately, *true* step discontinuities never occur in physics, but this is a warning that it is of course necessary to sum up to some N where the sines and cosines oscillate substantially more rapidly than any sudden change in the function being represented.

We go on to the Fourier transform, in which a function on the *infinite* line is expressed as an integral over a *continuum* of sines and cosines (or equivalently exponentials e^{ikx}). It turns out that arguments analogous to those that led to $\delta_N(x)$ now give a function $\delta(x)$ such that

$$f(x)=\int\limits_{-\infty}^{\infty}\delta(x-x')f(x')dx' \tag{2.1.3}$$

Confronted with this, one might well wonder what is the point of a function $\delta(x)$ which on convolution with $f(x)$ gives back the same function $f(x)$. The relevance of $\delta(x)$ will become evident later in the course, when states of a quantum particle are represented by wave functions on the infinite line, like $f(x)$, and operations on them involve integral operators similar to the convolution above. Working with operations on these functions is the continuum generalization of matrices acting on vectors in a finite-dimensional space, and $\delta(x)$ is the infinite-dimensional representation of the unit matrix. Just as in matrix algebra the eigenstates of the unit matrix are a set of vectors that span the space, and the unit matrix elements determine the set of dot products of these basis vectors, the delta function determines the generalized inner product of a *continuum* basis of states. It plays an essential role in the standard formalism for continuum states, and you need to be familiar with it!

Fourier Series

Any *reasonably smooth* real function $f(\theta)$ defined in the interval $-\pi<\theta\leq\pi$ [illegible] be expanded in a Fourier series,

$$f(\theta)=\frac{A_0}{2}+\sum_{n=1}^{\infty}(A_n\cos n\theta+B_n\sin n\theta) \tag{2.1.4}$$

where the coefficients can be found using the orthogonality condition,

$$\int\limits_{-\pi}^{\pi}\cos m\theta\cos n\theta d\theta=\pi\delta_{m,n} \tag{2.1.5}$$

and the same condition for the $\sin n\theta$'s to give:

$$A_n=\frac{1}{\pi}\int\limits_{-\pi}^{\pi}f(\theta)\cos n\theta d\theta \tag{2.1.6}$$

$$B_n=\frac{1}{\pi}\int\limits_{-\pi}^{\pi}f(\theta)\sin n\theta d\theta \tag{2.1.7}$$

Note that for an *even* function, only the A_n are nonzero, for an odd function only the B_n are nonzero.

How Smooth is "Reasonably Smooth"?

The number of terms of the series necessary to give a good approximation to a function depends on how rapidly the function changes. To get an idea of what goes wrong when a function is not "smooth", it is instructive to find the Fourier sine series for the step function

$$f(\theta)=-1\quad for\quad -\pi<\theta\leq 0,\quad f(\theta)=1\quad for\quad 0<\theta\leq\pi \tag{2.1.8}$$

Using the expression for B_n above it is easy to find:

$$f(\theta)=\frac{4}{\pi}\left(\sin\theta+\frac{\sin 3\theta}{3}+\frac{\sin 5\theta}{5}+\dots\right) \tag{2.1.9}$$

Taking the first half dozen terms in the series gives:

Gibbs' phenomenon

As we include more and more terms, the function becomes smoother but, surprisingly, *the initial overshoot at the step stays at a finite fraction of the step height*. However, the function recovers more and more rapidly, that is to say, the overshoot and "ringing" at the step take up less and less space. This overshoot is called *Gibbs' phenomenon*, and only occurs in functions with discontinuities.

How the Sum over *N* Terms is Related to the Complete Function

To get a clearer idea of how a Fourier series converges to the function it represents, it is useful to stop the series at N terms and examine how that sum, which we denote $f_N(\theta)$, tends towards $f(\theta)$.

So, substituting the values of the coefficients (Equation 2.1.6 and 2.1.7)

$$A_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\cos n\theta d\theta$$

$$B_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\sin n\theta d\theta$$

in the series (Equation 2.1.4)

$$f_N(\theta) = \frac{A_0}{2} + \sum_{n=1}^{N}(A_n\cos n\theta + B_n\sin n\theta)$$

gives

$$f_N(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta')d\theta' + \frac{1}{\pi}\sum_{n=1}^{N}\int_{-\pi}^{\pi}(\cos n\theta\cos n\theta' + \sin n\theta\sin n\theta')f(\theta')d\theta' \tag{2.1.10}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta')d\theta' + \frac{1}{\pi}\sum_{n=1}^{N}\int_{-\pi}^{\pi}\cos n(\theta-\theta')f(\theta')d\theta' \tag{2.1.11}$$

We can now use the trigonometric identity

$$\sum_{n=1}^{N}\cos nx = \frac{\sin(N+\frac{1}{2})x}{2\sin\frac{1}{2}x} - \frac{1}{2} \tag{2.1.12}$$

to find

$$f_N(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin(N+\frac{1}{2})(\theta-\theta')}{\sin\frac{1}{2}(\theta-\theta')}f(\theta')d\theta' = \int_{-\pi}^{\pi}\delta_N(\theta-\theta')f(\theta')d\theta' \tag{2.1.13}$$

where

$$\delta_N(x) = \frac{1}{2\pi}\frac{\sin(N+\frac{1}{2})x}{\sin\frac{1}{2}x} \tag{2.1.14}$$

(Note that proving the trigonometric identity is straightforward: write $z = e^{ix}$, so $\cos nx = \frac{1}{2}(z^n + z^{-n})$, and sum the geometric progressions.)

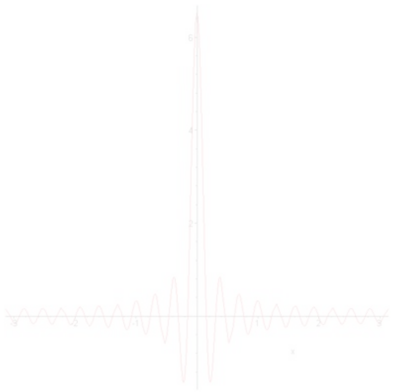
Going backwards for a moment and writing

$$\delta_N(x)=\frac{1}{\pi}\left(\sum_{n=1}^{N}\cos nx+\frac{1}{2}\right) \tag{2.1.15}$$

it is easy to check that

$$\int\limits_{-\pi}^{\pi}\delta_N(x)dx=1 \tag{2.1.16}$$

To help visualize $\delta_N(\theta)$, here is $N=20$:



We have just established that the total area under the curve = 1, and it is clear from the diagram that almost all this area is under the central peak, since the areas away from the center are almost equally positive and negative. The width of the central peak is $\pi/(N+\frac{1}{2})$, its height $(N+\frac{1}{2})/\pi$.

Exercise: for large N, approximately how far down does it dip on the first oscillation? (N/π^2)

For functions *varying slowly* compared with the oscillations, the convolution integral

$$f_N(\theta)=\int\limits_{-\pi}^{\pi}\delta_N(\theta-\theta')f(\theta')d\theta' \tag{2.1.17}$$

will give $f_N(\theta)$ close to $f(\theta)$, and for these functions $f_N(\theta)$ will tend to $f(\theta)$ [illegible] increases.

It is also clear why convoluting this curve with a step function gives an overshoot and oscillations. Suppose the function $f(\theta)$ is a step, jumping from 0 to 1 at $\theta=0$ From the convolutionary form of the integral, you should be able to convince yourself that the value of $f_N(\theta)$ at a point θ is the total area under the curve $\delta_N(\theta)$ to the *left* of that point (area below zero—that is, below the *x*-axis—of course counting negative). For $\theta=0$ this must be exactly 0.5 (since all the area under $\delta_N(\theta)$ adds to 1). *But* if we want the value of $f_N(\theta)$ at $\theta=\pi/(2N+1)$ (that is, the first point to the right of the origin where the curve cuts through the *x*-axis), we must *add* all the area to the left of $\theta=\pi/(2N+1)$, which actually adds up to a total area greater than one, since the leftover area to the right of that point is overall negative. That gives the overshoot.

A Fourier Series in Quantum Mechanics: Electron in a Box

The time-independent Schrödinger wave functions for an electron in a box (here a one-dimensional square well with infinite walls) are just the sine and cosine series determined by the boundary conditions. Therefore, any reasonably smooth initial wavefunction describing the electron can be represented as a Fourier series. The time development can then be found be multiplying each term in the series by the appropriate time-dependent phase factor.

Exercise 2.1.1

Prove that for a function $f(\theta)=\sum_{n=-\infty}^{\infty}a_ne^{in\theta}$, with the a_n in general complex,

$$\frac{1}{2\pi}\int\limits_{-\pi}^{\pi}|f(\theta)|^2d\theta=\sum_{n=-\infty}^{\infty}|a_n|^2 \tag{2.1.18}$$

The physical relevance of this result is as follows: for an electron confined to the circumference of a ring of unit radius, θ is the position of the electron. An orthonormal basis of states of the electron on this ring is the set of functions $(1/\sqrt{2\pi})e^{in\theta}$ with n an integer, a correctly normalized superposition of these states must have $\sum_{n=-\infty}^{\infty}|a_n|^2=1$, so that the total probability of finding the electron in some state is unity. But this must also mean that the total probability of finding the electron anywhere on the ring is unity—and that's the left-hand side of the above equation—the 2π's cancel.

Exponential Fourier Series

In the previous lecture, we discussed briefly how a Gaussian wave packet in *x*-space could be represented as a continuous linear superposition of plane waves that turned out to be another Gaussian wave packet, this time in *k*-space. The plan here is to demonstrate how we can arrive at that representation by carefully taking the limit of the well-defined Fourier series, going from the finite interval $(-\pi,\pi)$ to the whole line, and to outline some of the mathematical problems that arise, and how to handle them.

The first step is a trivial one: we need to generalize from real functions to complex functions, to include wave functions having nonvanishing current. A smooth *complex* function can be written in a Fourier series simply by allowing A_n and B_n to be complex, but in this case a more natural expansion would be in powers of $e^{i\theta}, e^{-i\theta}$

We write:

$$f(\theta)=\sum_{n=-\infty}^{\infty} a_n e^{in\theta} \quad with \quad a_n=\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-in\theta'} f(\theta')d\theta' \tag{2.1.19}$$

and retracing the above steps

$$f_N(\theta)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{n=-N}^{N} e^{in(\theta-\theta')} f(\theta')d\theta'=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta')d\theta'+\frac{1}{\pi}\int_{-\pi}^{\pi}\sum_{n=1}^{N}\cos n(\theta-\theta')f(\theta')d\theta' \tag{2.1.20}$$

exactly the same expression as before, therefore giving the same $\delta_N(\theta)$. This isn't surprising, because using $\cos n\theta=\frac{1}{2}(e^{in\theta}+e^{-in\theta}),\ \sin n\theta=\frac{1}{2i}(e^{in\theta}-e^{-in\theta})$ the first N terms in $A_n,\ B_n$ can simply be rearranged to a sum over $e^{in\theta}$ terms for $-N\le n\le N$.

Electron out of the Box: the Fourier Transform

To break down a *wave packet* into its plane wave components, we need to extend the range of integration from the $(-\pi,\pi)$ used above to $(-\infty,\infty)$. We do this by first rescaling from $(-\pi,\pi)$ to $(-L/2, L/2)$ and then taking the limit $L\to\infty$.

Scaling the interval from 2π to L (in the complex representation) gives:

$$f(x)=\sum_{n=-\infty}^{\infty} a_n e^{2\pi inx/L} \quad where \quad a_n=\frac{1}{L}\int_{-L/2}^{L/2} f(x)e^{-2\pi inx/L}dx \tag{2.1.21}$$

the sum in n being over all integers. This is an expression for $f(x)$ in terms of plane waves e^{ikx} where the allowed k's are $2\pi n/L$, with $n=0,\pm1,\pm2,\dots$

Retracing the steps above in the derivation of the function $\delta_N(x)$, we find the equivalent function to be

$$\delta_N^L(x)=\frac{1}{L}\left(1+2\sum_{n=1}^{N}\cos\frac{2\pi nx}{L}\right)=\frac{\sin((2N+1)\pi x/L)}{L\sin(\pi x/L)} \tag{2.1.22}$$

Studying the expression on the right, it is evident that *provided N is much greater than L*, this has the same peaked-at-the-origin behavior as the $\delta_N(x)$ we considered earlier. *But* we are interested in the limit $L\to\infty$, and there—for fixed N—this function $\delta_N^L(x)$ is low and flat.

Therefore, we must take the limit N going to infinity *before* taking L going [illegible] infinity.

This is what we do in the rest of this section.

Provided L is finite, we still have a Fourier series, representing a function of period L. Our main interest in taking L infinite is that we would like to represent a *nonperiodic* function, for example a localized wave packet, in terms of plane-wave components.

Suppose we have such a wave packet, say of length L_1, by which we mean the wave is exactly zero outside a stretch of the axis of length L_1. Why not just express it in terms of an infinite N Fourier series based on some large interval $(-L/2, L/2)$ provided the wave packet length L_1 is completely inside this interval? The point is that such an analysis would indeed accurately reproduce the wave packet inside the interval, but the *same* sum of plane waves evaluated *over all the x-axis* would reveal an *infinite string* of identical wave packets L apart! This is not what we want.

As a preliminary to taking L to infinity, let us write the exponential plane wave terms in the standard k-notation,

$$e^{2\pi inx/L}=e^{ik_nx} \tag{2.1.23}$$

So we are summing over an (infinite N) set of plane waves having wave number values

$$k_n=2\pi n/L\ ,\ n=0,\pm1,\pm2,\dots \tag{2.1.24}$$

,

a set of equally-spaced k's with separation $\Delta k=2\pi/L$.

Consider now what happens if we *double* the basic interval from $(-L/2, L/2)$ to $(-L, L)$.

The new allowed k values are $k_n=\pi n/L\ ,\ n=0,\pm1,\pm2,\dots$, so the separation is now $\Delta k=\pi/L$, *half* of what it was before. It is evident that as we increase L, the spacing between successive k_n values gets less and less.

Going back to the interval of length L, writing $La_n=a(k_n),\ k_n=2\pi n/L$ we have

$$f(x)=\sum_{n=-\infty}^{\infty} a_n e^{2\pi inx/L}=\frac{1}{L}\sum_{n=-\infty}^{\infty} a(k_n)e^{ik_nx} \tag{2.1.25}$$

Recall that the Riemann integral can be defined by

$$\int f(k)dk=\lim_{\Delta k\to0}\sum f(k_n)\Delta k \tag{2.1.26}$$

with $k_n = n\Delta k\ ,\ \ n = 0, \pm 1, \pm 2, \ldots.$

The expression on the right-hand side of the equation for $f(x)$ has the same form as the right-hand side of the Riemann integral definition, and here $\Delta k = 2\pi/L$.

That is to say,

$$2\pi f(x) = \frac{2\pi}{L}\sum_{n=-\infty}^{\infty} a(k_n)e^{ik_n x} = \sum_{n=-\infty}^{\infty} a(k_n)e^{ik_n x}\Delta k = \int_{-\infty}^{\infty} a(k)e^{ikx}dk \tag{2.1.27}$$

in the limit $\Delta k \to 0$, or equivalently $L \to \infty$. We are of course assuming here that the function $a(k_n)$, which we have only defined (for a given L) on the set of points k_n, tends to a continuous function $a(k)$ in the limit $L \to \infty$.

It follows that in the infinite L limit, we have the *Fourier transform* equations:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} a(k)e^{ikx}dk \ \ and\ \ a(k) = \lim_{L\to\infty} La_n = \lim_{L\to\infty}\int_{-L/2}^{L/2} f(x)e^{-2\pi inx/L}dx = \int_{-\infty}^{\infty} f(x)e^{-ikx}dx \tag{2.1.28}$$

Dirac's Delta Function

Now we have taken both N and L to infinity, what has happened to our function $\delta_N^L(x)$? Remember that our procedure for finding $f_N(\theta)$ in terms of $f(\theta)$ gave the equation

`f_N(\theta)=\dfrac{1}{2\pi} \int\limits_{-\pi}^{\pi} f(\theta') d\theta' +\dfrac{1}{\pi} \int\limits_{-\pi}^{\pi} \sum_{n=1}^{N} \cos n(\theta-\theta')f(\theta')d\theta' \label{2.1.10}`

and from this we found $\delta_N(\theta)$.

Following the same formal procedure with the $(L = \infty)$ Fourier transforms, we are forced to take N infinite (recall the procedure only made sense if N was taken to infinity before L), so in place of an equation for $f_N(\theta)$ in terms of $f(\theta)$, we get an equation for $f(x)$ in terms of itself! Let's write it down first and think afterwards:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(x')e^{-ikx'}dx'\right)e^{ikx}dk = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\frac{dk}{2\pi}e^{ik(x-x')}\right)f(x')dx' \tag{2.1.29}$$

In other words,

$$f(x) = \int_{-\infty}^{\infty}\delta\text{[illegible]}x')f(x')dx' \tag{2.1.30}$$

where

$$\delta(x) = \int_{-\infty}^{+\infty}\frac{dk}{2\pi}e^{ikx} \tag{2.1.31}$$

This is the *Dirac delta function*. This hand-waving approach has given a result which is not clearly defined. This integral over *x* is linearly divergent at the origin, and has finite oscillatory behavior everywhere else. To make any progress, we must provide some form of cutoff in *k*-space, then perhaps we can find a meaningful limit by placing the cutoff further and further away.

From our arguments above, we should be able to recover $\delta(x)$ as a limit of $\delta_N^L(x)$ by first taking N to infinity, then L. That is to say,

$$\delta(x) = \lim_{L\to\infty}\left(\lim_{N\to\infty}\delta_N^L(x)\right) = \lim_{L\to\infty}\left(\lim_{N\to\infty}\frac{\sin((2N+1)\pi x/L)}{L\sin(\pi x/L)}\right) \tag{2.1.32}$$

A way to understand this limit is to write $M = (2N+1)\pi/L$ and let M go to infinity before L. (This means as we take L large on its way to infinity, we're taking N far larger!)

So the numerator is just $\sin Mx$. In the limit of infinite L, for any finite x the denominator is just πx, since $\sin\theta = \theta$ in the limit of small θ.

From this,

$$\delta(x) = \lim_{M\to\infty}\frac{\sin Mx}{\pi x} \tag{2.1.33}$$

This is still a rather pathological function, in that it is oscillating more and more quickly as the infinite limit is taken. This comes about from the abrupt cutoff in the sum at the frequency N.

To see how this relates to the (also ill-defined) $\delta(x) = \int\limits_{-\infty}^{+\infty}(dk/2\pi)e^{ikx}$, recall $\delta_N^L(x)$ came from the series

$$\delta_N^L(x) = \frac{\sin(2N+1)\pi x/L}{L\sin\pi x/L} = \frac{1}{L}\left(1 + 2\sum_{n=1}^{N}\cos\frac{2\pi nx}{L}\right) \tag{2.1.34}$$

Expressing the cosine in terms of exponentials, then replacing the sum by an integral in the large N limit, in the same way we did earlier, writing $k_n = 2\pi n/L$, so the interval between successive k_n's is $\Delta k = 2\pi/L$, so $\int f(k)dk \cong (2\pi/L)\sum f(k_n)$:

$$\delta_N^L(x) = \frac{1}{L}\sum_{n=-N}^{N} e^{2\pi i n/L} \cong \int\limits_{-2\pi N/L}^{2\pi N/L} \frac{dk}{2\pi} e^{ikx} = \frac{\sin(2\pi N x/L)}{\pi x} \tag{2.1.35}$$

So it is clear that we're defining the $\delta(x)$ as a limit of the integral $\int\limits_{-2\pi N/L}^{2\pi N/L}(dk/2\pi)e^{ikx}$, which is abruptly cut off at the large values $\pm(2\pi N/L)$. In fact, this is not very physical: a much more realistic scenario for a *real* wave packet would be a *gradual diminution* in contributions from high frequency (or short wavelength) modes—that is to say, a *gentle* cutoff in the integral over k that was used to replace the sum over n. For example, a reasonable cutoff procedure would be to multiply the integrand by $exp(-\Delta^2 k^2)$, then take the limit of small Δ.

Therefore a more reasonable definition of the delta function, from a physicist's point of view, would be

$$\delta(x) = \lim_{\Delta\to 0}\int\limits_{-\infty}^{\infty}\frac{dk}{2\pi}e^{ikx}e^{-\Delta^2k^2} = \lim_{\Delta\to 0}\frac{1}{(4\pi\Delta^2)^{1/2}}e^{-x^2/4\Delta^2} = \lim_{\Delta\to 0}\delta_\Delta(x) \tag{2.1.36}$$

That is to say, the delta function can be defined as the "narrow limit" of a Gaussian wave packet with total area 1. Unlike the function $\delta_N(\theta)$, $\delta_\Delta(x)$ has no oscillating sidebands, thanks to our smoothing out of the upper k-space cutoff, so step discontinuities do not generate Gibbs' phenomenon overshoot—instead, a step will be smoothed out over a distance of order Δ.

Properties of the Delta Function

It is straightforward to verify the following properties from the definition as a limit of a Gaussian wavepacket:

$$\int \delta(x)dx = 1,\ \delta(x) = 0\ for\ x \neq 0 \tag{2.1.37}$$

$$\delta(x) = \delta(-x),\ \delta(ax) = \frac{1}{|a|}\delta(x) \tag{2.1.38}$$

$$\int \delta(a-x)\delta(x-b)dx = \delta(a-b) \tag{2.1.39}$$

Yet Another Definition, and a Connection with the Principal Value Integral

There is no unique way to define the delta function, and other cutoff procedures can give useful insights. For example, the k-space integral can be split into two and simple exponential cutoffs applied to the two halves, that is, we could take the definition to be

$$\delta(x) = \lim_{\varepsilon\to 0}\left(\int\limits_{-\infty}^{0}\frac{dk}{2\pi}e^{ikx}e^{\varepsilon k} + \int\limits_{0}^{\infty}\frac{dk}{2\pi}e^{ikx}e^{-\varepsilon k}\right) \tag{2.1.40}$$

Evaluating the integrals,

$$\delta(x) = \lim_{\varepsilon\to 0}\frac{1}{2\pi}\left(\frac{1}{ix+\varepsilon} - \frac{1}{ix-\varepsilon}\right) = \lim_{\varepsilon\to 0}\frac{1}{\pi}\left(\frac{\varepsilon}{x^2+\varepsilon^2}\right) \tag{2.1.41}$$

It is easy to check that this function is correctly normalized by making the change of variable $x = \varepsilon\tan\theta$ and integrating from $-\pi/2$ to $\pi/2$. This representation of the delta function will prove to be useful later. Note that regarded as a function of a complex variable, the delta function has two poles on the pure imaginary axis at $z = \pm i\varepsilon$.

The standard definition of the principal value integral is:

$$\int\limits_{-D}^{D} f(x)\frac{P}{x}dx = \lim_{z\to 0}\left(\int\limits_{-D}^{-z}\frac{f(x)}{x}dx + \int\limits_{z}^{D}\frac{f(x)}{x}dx\right) \tag{2.1.42}$$

It is not difficult to see that for a continuous differentiable function $f(x)$ this is equivalent to

$$\int\limits_{-D}^{D} f(x)\frac{P}{x}dx = \lim_{z\to 0}\int\limits_{-D}^{D} f(x)\frac{x}{x^2+\varepsilon^2}dx \tag{2.1.43}$$

Therefore the principal value operator can be written symbolically:

$$\frac{P}{x} = \lim_{z\to 0}\frac{x}{x^2+\varepsilon^2} = \lim_{z\to 0}\frac{1}{2}\left(\frac{1}{x+i\varepsilon} + \frac{1}{x-i\varepsilon}\right) \tag{2.1.44}$$

Putting this together with the similar representation of the delta function above, and taking the limit of $\varepsilon \to 0$ $_{\varepsilon\to 0}$ to be understood, we have the useful result:

$$\frac{1}{x\pm i\varepsilon} = \frac{P}{x} \mp i\pi\delta(x) \tag{2.1.45}$$

Exercises

1. Prove **Parseval's Theorem**:

$$If\ \ f(x)=\int_{-\infty}^{\infty}\frac{dk}{2\pi}a(k)e^{ikx},\ \ then\ \ \int_{-\infty}^{\infty}|f(x)|^2dx=\int_{-\infty}^{\infty}\frac{dk}{2\pi}|a(k)|^2. \tag{2.1.46}$$

2. Prove the rule for the Fourier Transform of a **convolution** of two functions:

$$If\ \ f(x)=\int_{-\infty}^{\infty}\frac{dk}{2\pi}a(k)e^{ikx},\ \ g(x)=\int_{-\infty}^{\infty}\frac{dk'}{2\pi}b(k')e^{ik'x},\ \ then\ \ \int_{-\infty}^{\infty}f(x-x')g(x')dx'=\int_{-\infty}^{\infty}\frac{dk}{2\pi}a(k)b(k)e^{ikx} \tag{2.1.47}$$

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

2.2: Linear Algebra

Introduction

We’ve seen that in quantum mechanics, the state of an electron in some potential is given by a wave function $\psi(\vec{x}, t)$, and physical variables are represented by operators on this wave function, such as the momentum in the *x*-direction $p_x = -i\hbar\partial/\partial x$. The Schrödinger wave equation is a *linear*equation, which means that if ψ_1 and ψ_2 are solutions, then so is $c_1\psi_1 + c_2\psi_2$, where c_1, c_2 are arbitrary complex numbers.

This linearity of the sets of possible solutions is true generally in quantum mechanics, as is the representation of physical variables by operators on the wave functions. The mathematical structure this describes, the linear set of possible states and sets of operators on those states, is in fact a *linear algebra* of operators acting on a *vector space*. From now on, this is the language we’ll be using most of the time. To clarify, we’ll give some definitions.

What is a Vector Space?

The prototypical vector space is of course the set of real vectors in ordinary three-dimensional space, these vectors can be represented by trios of real numbers (v_1, v_2, v_3) measuring the components in the *x, y* and *z* directions respectively.

The basic properties of these vectors are:

- any vector multiplied by a number is another vector in the space, $a(v_1, v_2, v_3) = (av_1, av_2, av_3)$;
- the sum of two vectors is another vector in the space, that given by just adding the corresponding components together: $(v_1 + w_1, v_2 + w_2, v_3 + w_3)$.

These two properties together are referred to as “*closure*”: adding vectors and multiplying them by numbers cannot get you out of the space.

- A further property is that there is a unique null vector $(0, 0, 0)$ and each vector has an additive inverse $(-v_1, -v_2, -v_3)$ which added to the original vector gives the null vector.

Mathematicians have generalized the definition of a vector space: a general vector space has the properties we’ve listed above for three-dimensional real vectors, but the operations of addition and multiplication by a number are generalized to more abstract operations between more general entities. The operators are, however, restricted to being commutative and associative.

Notice that the list of necessary properties for a general vector space does not include that the vectors have a magnitude—that would be an additional requirement, giving what is called a *normed* vector space. More about that later.

To go from the familiar three-dimensional vector space to the vector spaces relevant to quantum mechanics, first the real numbers (components of the vector and possible multiplying factors) are to be generalized to complex numbers, and second the three-component vector goes an *n*component vector. The consequent *n*-dimensional complex space is sufficient to describe the quantum mechanics of angular momentum, an important subject. But to describe the wave function of a particle in a box requires an *infinite* dimensional space, one dimension for each Fourier component, and to describe the wave function for a particle on an infinite line requires the set of all normalizable continuous differentiable functions on that line. Fortunately, all these generalizations are to finite or infinite sets of complex numbers, so the mathematicians’ vector space requirements of commutativity and associativity are always trivially satisfied.

We use Dirac’s notation for vectors, $|1\rangle, |2\rangle$ and call them “kets”, so, in his language, if $|1\rangle, |2\rangle$ belong to the space, so does $c_1|1\rangle + c_2|2\rangle$ for arbitrary complex constants c_1, c_2. Since our vectors are made up of complex numbers, multiplying any vector by zero gives the null vector, and the additive inverse is given by reversing the signs of all the numbers in the vector.

Clearly, the set of solutions of Schrödinger’s equation for an electron in a potential satisfies the requirements for a vector space: $\psi(\vec{x}, t)$ is just a complex number at each point in space, so only complex numbers are involved in forming $c_1\psi_1 + c_2\psi_2$, and commutativity, associativity, etc., follow at once.

Vector Space Dimensionality

The vectors $|1\rangle, |2\rangle, |3\rangle$are *linearly independent* if

$$c_1|1\rangle + c_2|2\rangle + c_3|3\rangle = 0 \tag{2.2.1}$$

implies

$$c_1 = c_2 = c_3 = 0 \tag{2.2.2}$$

A vector space is n-dimensional if the maximum number of linearly independent vectors in the space is n.

Such a space is often called $V^n(C)$, or $V^n(R)$ if only real numbers are used.

Now, vector spaces with finite dimension *n* are clearly insufficient for describing functions of a continuous variable *x*. But they are well worth reviewing here: as we've mentioned, they are fine for describing quantized angular momentum, and they serve as a natural introduction to the infinite-dimensional spaces needed to describe spatial wavefunctions.

A set of *n* linearly independent vectors in *n*-dimensional space is a *basis*—any vector can be written *in a unique way* as a sum over a basis:

$$|V\rangle = \sum v_i |i\rangle \tag{2.2.3}$$

You can check the uniqueness by taking the difference between two supposedly distinct sums: it will be a linear relation between independent vectors, a contradiction.

Since all vectors in the space can be written as linear sums over the elements of the basis, the sum of multiples of any two vectors has the form:

$$a|V\rangle + b|W\rangle = \sum (av_i + bw_i)|i\rangle \tag{2.2.4}$$

Inner Product Spaces

The vector spaces of relevance in quantum mechanics also have an operation associating a number with a pair of vectors, a generalization of the dot product of two ordinary three-dimensional vectors,

$$\vec{a}, \vec{b} = \sum a_i b_i \tag{2.2.5}$$

Following Dirac, we write the inner product of two ket vectors $|V\rangle, |W\rangle$ as $\langle W|V\rangle$. Dirac refers to this $\langle\;|\;\rangle$ form as a "bracket" made up of a "bra" and a "ket". This means that each ket vector $|V\rangle$ has an associated bra$\langle V|$. For the case of a real *n*-dimensional vector, $|V\rangle, \langle V|$ are identical—but we require for the more general case that

$$\langle W|V\rangle = \langle V|W\rangle^* \tag{2.2.6}$$

where $*$ denotes complex conjugate. This implies that for a ket $(v_1, \ldots, v_n)$ the bra will be $(v_1^*, \ldots, v_n^*)$. (Actually, bras are usually written as rows, kets as columns, so that the inner product follows the standard rules for matrix multiplication.) Evidently for the *n*-dimensional complex vector $\langle V|V\rangle$ is real and positive except for the null vector:

$$\langle V|V\rangle = \sum_1^n |v_i|^2 \tag{2.2.7}$$

For the more general inner product spaces considered later we require $\langle V|V\rangle$ to be positive, except for the null vector. (These requirements do restrict the classes of vector spaces we are considering—no Lorentz metric, for example—but they are all satisfied by the spaces relevant to nonrelativistic quantum mechanics.)

The *norm* of $|V\rangle$ is then defined by

$$|V| = \sqrt{\langle V|V\rangle} \tag{2.2.8}$$

If $|V\rangle$ is a member of $V^n(C)$, so is $a|V\rangle$, for any complex number a.

We require the inner product operation to commute with multiplication by a number, so

$$\langle W|(a|V\rangle) = a\langle W|V\rangle \tag{2.2.9}$$

The complex conjugate of the right hand side is $a^*\langle V|W\rangle$. For consistency, the bra corresponding to the ket $a|V\rangle$ must therefore be $\langle V|a^*$—in any case obvious from the definition of the bra in *n* complex dimensions given above.

It follows that if

$$|V\rangle=\sum v_i|i\rangle,\ |W\rangle=\sum w_i|i\rangle,\ then\ \langle V|W\rangle=\sum v_i^* w_j\langle i|j\rangle \tag{2.2.10}$$

Constructing an Orthonormal Basis: the Gram-Schmidt Process

To have something better resembling the standard dot product of ordinary three vectors, we need $\langle i|j\rangle=\delta_{ij}$, that is, we need to construct an *orthonormal basis* in the space. There is a straightforward procedure for doing this called the Gram-Schmidt process. We begin with a linearly independent set of basis vectors, $|1\rangle$, $|2\rangle$, $|3\rangle$,... .

We first normalize $|1\rangle$ by dividing it by its norm. Call the normalized vector $|I\rangle$. Now $|2\rangle$ cannot be parallel to $|I\rangle$, because the original basis was of linearly independent vectors, but $|2\rangle$ in general has a nonzero component parallel to $|I\rangle$, equal to $|I\rangle\langle I|2\rangle$, since $|I\rangle$ is normalized. Therefore, the vector $|2\rangle-|I\rangle\langle I|2\rangle$ is perpendicular to $|I\rangle$, as is easily verified. It is also easy to compute the norm of this vector, and divide by it to get $|II\rangle$, the second member of the orthonormal basis. Next, we take $|3\rangle$ and subtract off its components in the directions $|I\rangle$ and $|II\rangle$, normalize the remainder, and so on.

In an *n*-dimensional space, having constructed an orthonormal basis with members $|i\rangle$, any vector $|V\rangle$ can be written as a column vector,

$$|V\rangle=\sum v_i|i\rangle=\begin{pmatrix}v_1\\v_2\\.\\.\\v_n\end{pmatrix},\ where\ |1\rangle=\begin{pmatrix}1\\0\\.\\.\\0\end{pmatrix}\ and\ so\ on. \tag{2.2.11}$$

The corresponding bra is $\langle V|=\sum v_i^*\langle i|$, which we write as a row vector with the elements complex conjugated, $\langle V|=(v_1^*,v_2^*,\dots v_n^*)$. This operation, going from columns to rows and taking the complex conjugate, is called taking the *adjoint*, and can also be applied to matrices, as we shall see shortly.

The reason for representing the bra as a row is that the inner product of two vectors is then given by standard matrix multiplication:

$$\langle V|W\rangle=(v_1^*,v_2^*,\dots,v_n^*)\begin{pmatrix}w_1\\.\\.\\w_n\end{pmatrix} \tag{2.2.12}$$

(Of course, this only works with an orthonormal base.)

The Schwartz Inequality

The Schwartz inequality is the generalization to any inner product space of the result $|\vec{a},\vec{b}|^2\le|\vec{a}|^2|\vec{b}|^2$ (or $\cos^2\theta\le 1$) for ordinary three-dimensional vectors. The equality sign in that result only holds when the vectors are parallel. To generalize to higher dimensions, one might just note that two vectors are in a two-dimensional subspace, but an illuminating way of understanding the inequality is to write the vector $\vec{a}$ as a sum of two components, one parallel to $\vec{b}$ and one perpendicular to $\vec{b}$. The component parallel to $\vec{b}$ is just $\vec{b}(\vec{a}\cdot\vec{b})/|\vec{b}|^2$, so the component perpendicular to $\vec{b}$ is the vector $\vec{a}_\perp=\vec{a}-\vec{b}(\vec{a}\cdot\vec{b})/|\vec{b}|^2$. Substituting this expression into $\vec{a}_\perp\cdot\vec{a}_\perp\ge 0$, the inequality follows.

This same point can be made in a general inner product space: if $|V\rangle$, $|W\rangle$ are two vectors, then

$$|Z\rangle=|V\rangle-\frac{|W\rangle\langle W|V\rangle}{|W|^2} \tag{2.2.13}$$

is the component of $|V\rangle$ perpendicular to $|W\rangle$, as is easily checked by taking its inner product with $|W\rangle$.

Then

$$\langle Z|Z\rangle\ge 0\ \ gives\ immediately\ \ |\langle V|W\rangle|^2\le|V|^2|W|^2 \tag{2.2.14}$$

Linear Operators

A *linear operator* A takes any vector in a linear vector space to a vector in that space, $A|V\rangle=|V'\rangle$ and satisfies

$$A(c_1|V_1\rangle+c_2|V_2\rangle)=c_1A|V_1\rangle+c_2A|V_2\rangle \tag{2.2.15}$$

with c_1, c_2 arbitrary complex constants.

The *identity operator* I is (obviously!) defined by:

$$I|V\rangle = |V\rangle \quad for\ all\ |V\rangle \tag{2.2.16}$$

For an n-dimensional vector space with an orthonormal basis $|1\rangle, \dots, |n\rangle$, since any vector in the space can be expressed as a sum $|V\rangle = \sum v_i|i\rangle$, the linear operator is completely determined by its action on the basis vectors—this is all we need to know. It's easy to find an expression for the identity operator in terms of bras and kets.

Taking the inner product of both sides of the equation $|V\rangle = \sum v_i|i\rangle$ with the bra $\langle i|$ gives $\langle i|V\rangle = v_i$, so

$$|V\rangle = \sum v_i|i\rangle = \sum |i\rangle\langle i|V\rangle \tag{2.2.17}$$

Since this is true for any vector in the space, it follows that that the identity operator is just

$$I = \sum_1^n |i\rangle\langle i| \tag{2.2.18}$$

This is an important result: it will reappear in many disguises.

To analyze the action of a general linear operator A, we just need to know how it acts on each basis vector. Beginning with $A|1\rangle$, this must be some sum over the basis vectors, and since they are orthonormal, the component in the $|i\rangle$ direction must be just $\langle i|A|1\rangle$.

That is,

$$A|1\rangle = \sum_1^n |i\rangle\langle i|A|1\rangle = \sum_1^n A_{i1}|i\rangle\,,\ writing\ \langle i|A|1\rangle = A_{i1} \tag{2.2.19}$$

So if the linear operator A acting on $|V\rangle = \sum v_i|i\rangle$ gives $|V'\rangle = \sum v_i'|i\rangle$, that is, $A|V\rangle = |V'\rangle$, the linearity tells us that

$$\sum v_i'|i\rangle = |V'\rangle = A|V\rangle = \sum v_j A|j\rangle = \sum_{i,j} v_j|i\rangle\langle i|A|j\rangle = \sum_{i,j} v_j A_{ij}|i\rangle \tag{2.2.20}$$

where in the fourth step we just inserted the identity operator.

Since the $|i\rangle$'s are all orthogonal, the coefficient of a particular $|i\rangle$ on the left-hand side of the equation must be identical with the coefficient of the same $|i\rangle$ on the right-hand side. That is, $v_i' = A_{ij}v_j$.

Therefore the operator A is simply equivalent to *matrix multiplication*:

$$\begin{pmatrix} v_1' \\ v_2' \\ . \\ . \\ v_n' \end{pmatrix} = \begin{pmatrix} \langle 1|A|1\rangle & \langle 1|A|2\rangle & . & . & \langle 1|A|n\rangle \\ \langle 2|A|1\rangle & \langle 2|A|2\rangle & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \langle n|A|1\rangle & \langle n|A|2\rangle & . & . & \langle n|A|n\rangle \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ . \\ . \\ v_n \end{pmatrix} \tag{2.2.21}$$

Evidently, then, applying two linear operators one after the other is equivalent to successive matrix multiplication—and, therefore, *since matrices do not in general commute, nor do linear operators*. (Of course, if we hope to represent quantum variables as linear operators on a vector space, this has to be true—the momentum operator $p = -i\hbar d/dx$ certainly doesn't commute with x!)

Projection Operators

It is important to note that a linear operator applied successively to the members of an orthonormal basis might give a new set of vectors which *no longer span the entire space*. To give an example, the linear operator $|1\rangle\langle 1|$ applied to any vector in the space picks out the vector's component in the $|1\rangle$ direction. It's called a *projection operator*. The operator $(|1\rangle\langle 1| + |2\rangle\langle 2|)$ projects a vector into its components in the subspace spanned by the vectors $|1\rangle$ and $|2\rangle$, and so on—if we extend the sum to be over the whole basis, we recover the identity operator.

Exercise: prove that the $n\times n$ matrix representation of the projection operator $(|1\rangle\langle 1| + |2\rangle\langle 2|)$ has all elements zero except the first two diagonal elements, which are equal to one.

There can be no *inverse* operator to a nontrivial projection operator, since the information about components of the vector perpendicular to the projected subspace is lost.

The Adjoint Operator and Hermitian Matrices

As we've discussed, if a ket $|V\rangle$ in the *n*-dimensional space is written as a column vector with n (complex) components, the corresponding bra is a *row* vector having as elements the complex conjugates of the ket elements. $\langle W|V\rangle=\langle V|W\rangle^*$ then follows automatically from standard matrix multiplication rules, and on multiplying $|V\rangle$ by a complex number a to get $a|V\rangle$ (meaning that each element in the column of numbers is multiplied by a) the corresponding bra goes to $\langle V|a^*=a^*\langle V|$.

But suppose that instead of multiplying a ket by a number, we operate on it with a linear operator. What generates the parallel transformation among the bras? In other words, if $A|V\rangle=|V'\rangle$, what operator sends the bra $\langle V|$ to $\langle V'|$? It must be a linear operator, because A is linear, that is, if under A $|V_1\rangle\to|V_1'\rangle$, $|V_2\rangle\to|V_2'\rangle$ and $|V_3\rangle=|V_1\rangle+|V_2\rangle$, then under A $|V_3\rangle$ is required to got to $|V_3'\rangle=|V_1'\rangle+|V_2'\rangle$. Consequently, under the parallel *bra* transformation we must have $\langle V_1|\to\langle V_1'|$, $\langle V_2|\to\langle V_2'|$ and $\langle V_3|\to\langle V_3'|$,—the bra transformation is necessarily also linear. Recalling that the bra is an *n*-element *row* vector, the most general linear transformation sending it to another bra is an $n\times n$ matrix operating on the bra from the *right*.

This bra operator is called the *adjoint* of A, written$A^\dagger$. That is, the ket $A|V\rangle$ has corresponding bra $\langle V|A^\dagger$. In an orthonormal basis, using the notation $\langle Ai|$ to denote the bra $\langle i|A^\dagger$ corresponding to the ket $A|i\rangle=|Ai\rangle$, say,

$$(A^\dagger)_{ij}=\langle i|A^\dagger|j\rangle=\langle Ai|j\rangle=\langle j|Ai\rangle^*=A^*_{ji} \tag{2.2..22}$$

So *the adjoint operator is the transpose complex conjugate*.

Important: for a *product* of two operators (prove this!),

$$(AB)^\dagger=B^\dagger A^\dagger \tag{2.2..23}$$

An operator equal to its adjoint $A=A^\dagger$ is called *Hermitian*. As we shall find in the next lecture, Hermitian operators are of central importance in quantum mechanics. An operator equal to *minus* its adjoint, $A=-A^\dagger$, is *anti* Hermitian (sometimes termed skew Hermitian). These two operator types are essentially generalizations of real and imaginary number: any operator can be expressed as a sum of a Hermitian operator and an anti Hermitian operator,

$$A=\frac{1}{2}(A+A^\dagger)+\frac{1}{2}(A-A^\dagger) \tag{2.2.24}$$

The definition of adjoint naturally extends to *vectors and numbers*: the adjoint of a ket is the corresponding bra, the adjoint of a number is its complex conjugate. This is useful to bear in mind when taking the adjoint of an operator which may be partially constructed of vectors and numbers, such as projection-type operators. The adjoint of a product of matrices, vectors and numbers is the product of the adjoints *in reverse order*. (Of course, for numbers the order doesn't matter.)

Unitary Operators

An operator is *unitary* if $U^\dagger U=1$. This implies first that U operating on any vector gives a vector having the same norm, since the new norm $\langle V|U^\dagger U|V\rangle=\langle V|V\rangle$. Furthermore, inner products are preserved, $\langle W|U^\dagger U|V\rangle=\langle W|V\rangle$. *Therefore, under a unitary transformation the original orthonormal basis in the space must go to another orthonormal basis.*

Conversely, any transformation that takes one orthonormal basis into another one is a unitary transformation. To see this, suppose that a linear transformation A sends the members of the orthonormal basis $(|1\rangle_1,|2\rangle_1,\ldots,|n\rangle_1)$ to the different orthonormal set $(|1\rangle_2,|2\rangle_2,\ldots,|n\rangle_2)$ so $A|1\rangle_1=|1\rangle_2$, etc. Then the vector $|V\rangle=\sum v_i|i\rangle_1$ will go to $|V'\rangle=A|V\rangle=\sum v_i|i\rangle_2$, having the same norm, $\langle V'|V'\rangle=\langle V|V\rangle=\sum|v_i|^2$. A matrix element $\langle W'|V'\rangle=\langle W|V\rangle=\sum w_i^*v_i$, but also $\langle W'|V'\rangle=\langle W|A^\dagger A|V\rangle$. That is, $\langle W|V\rangle=\langle W|A^\dagger A|V\rangle$ for *arbitrary* kets $|V\rangle$, $|W\rangle$. This is only possible if $A^\dagger A=1$, so A is unitary.

A unitary operation amounts to a rotation (possibly combined with a reflection) in the space. Evidently, since $U^\dagger U=1$, the adjoint $U^\dagger$ rotates the basis back—it is the inverse operation, and so $UU^\dagger=1$ also, that is, U and $U^\dagger$ commute.

Determinants

We review in this section the *determinant* of a matrix, a function closely related to the operator properties of the matrix.

Let's start with 2×2 matrices:

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{2.2.25}$$

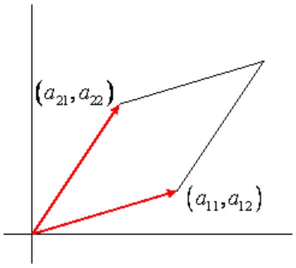
The determinant of this matrix is defined by:

$$\det A=|A|=a_{11}a_{22}-a_{12}a_{21} \tag{2.2.26}$$

Writing the two rows of the matrix as vectors:

$$\begin{aligned} \vec{a}_1^R &= (a_{11},a_{12}) \\ \vec{a}_2^R &= (a_{21},a_{22}) \end{aligned} \tag{2.2.27}$$

(R denotes row), $\det A=\vec{a}_1^R\times\vec{a}_2^R$ is just the *area* (with appropriate sign) of the parallelogram having the two row vectors as adjacent sides:



This is zero if the two vectors are parallel (linearly dependent) and is not changed by adding any multiple of $\vec{a}_2^R$ to $\vec{a}_2^R$ (because the new parallelogram has the same base and the same height as the original—check this by drawing).

Let’s go on to the more interesting case of 3×3 matrices:

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{2.2.28}$$

The determinant of A is defined as

$$\det A=\varepsilon_{ijk}a_{1i}a_{2j}a_{3k} \tag{2.2.29}$$

where $\varepsilon_{ijk}=0$ if any two are equal, +1 if $ijk=123,\ 231\ or\ 312$ (that is to say, an even permutation of 123) and –1 if ijk is an odd permutation of 123. Repeated suffixes, of course, imply summation here.

Writing this out explicitly,

$$\det A=a_{11}a_{22}a_{33}+a_{21}a_{32}a_{13}+a_{31}a_{12}a_{23}-a_{11}a_{32}a_{23}-a_{21}a_{12}a_{33}-a_{31}a_{22}a_{13} \tag{2.2.30}$$

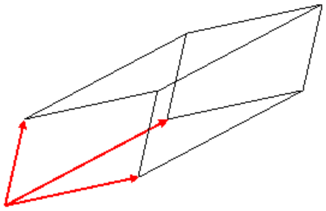
Just as in two dimensions, it’s worth looking at this expression in terms of vectors representing the rows of the matrix

$$\begin{aligned} \vec{a}_1^R &= (a_{11},a_{12},a_{13}) \\ \vec{a}_2^R &= (a_{21},a_{22},a_{23}) \\ \vec{a}_3^R &= (a_{31},a_{32},a_{33}) \end{aligned} \tag{2.2.31}$$

so

$$A=\begin{pmatrix} \vec{a}_1^R \\ \vec{a}_2^R \\ \vec{a}_3^R \end{pmatrix},\ \textit{and we see that}\ \det A=(\vec{a}_1^R\times\vec{a}_2^R)\cdot\vec{a}_3^R \tag{2.2.32}$$

This is the volume of the parallelopiped formed by the three vectors being adjacent sides (meeting at one corner, the origin).



This parallelepiped volume will of course be zero if the three vectors lie in a plane, and it is not changed if a multiple of one of the vectors is added to another of the vectors. That is to say, *the determinant of a matrix is not changed if a multiple of one row is added to another row*. This is because the determinant is linear in the elements of a single row,

$$\det\begin{pmatrix}\vec{a}_1^R+\lambda\vec{a}_2^R\\ \vec{a}_2^R\\ \vec{a}_3^R\end{pmatrix}=\det\begin{pmatrix}\vec{a}_1^R\\ \vec{a}_2^R\\ \vec{a}_3^R\end{pmatrix}+\lambda\det\begin{pmatrix}\vec{a}_2^R\\ \vec{a}_2^R\\ \vec{a}_2^R\end{pmatrix}\tag{2.2.33}$$

and the last term is zero because two rows are identical—so the triple vector product vanishes.

A more general way of stating this, applicable to larger determinants, is that for a determinant with two identical rows, the symmetry of the two rows, together with the *antisymmetry* of ε_{ijk}, ensures that the terms in the sum all cancel in pairs.

Since the determinant is not altered by adding some multiple of one row to another, if the rows are linearly dependent, one row could be made identically zero by adding the right multiples of the other rows. Since every term in the expression for the determinant has one element from each row, the determinant would then be identically zero. For the three-dimensional case, the linear dependence of the rows means the corresponding vectors lie in a plane, and the parallelepiped is flat.

The algebraic argument generalizes easily to $n\times n$ determinants: they are *identically zero if the rows are linearly dependent*.

The generalization from 3×3 to $n\times n$ 3×3 to n×n determinants is that $\det A=\varepsilon_{ijk}a_{1i}a_{2j}a_{3k}$ becomes:

$$\det A=\varepsilon_{ijk\dots p}a_{1i}a_{2j}a_{3k}\dots a_{np}\tag{2.2.34}$$

where $ijk\dots p$ is summed over all permutations of $132\dots n$, and the ε symbol is zero if any two of its suffixes are equal, +1 for an even permutation and -1 for an odd permutation. (*Note*: any permutation can be written as a product of swaps of neighbors. Such a representation is in general not unique, but for a given permutation, all such representations will have either an odd number of elements or an even number.)

An important theorem is that for a product of two matrices A, B the determinant of the product is the product of the determinants, $\det AB=\det A\times\det B$. This can be verified by brute force for 2×2 matrices, and a proof in the general case can be found in any book on mathematical physics (for example, Byron and Fuller).

It can also be proved that if the *rows* are linearly independent, the determinant cannot be zero.

(*Here's a proof*: take an $n\times n$ matrix with the n row vectors linearly independent. Now consider the components of those vectors in the $n-1$ dimensional subspace perpendicular to $(1,0,\dots,0)$. These n vectors, each with only $n-1$ components, must be linearly dependent, since there are more of them than the dimension of the space. So we can take some combination of the rows below the first row and subtract it from the first row to leave the first row $(a,0,0,\dots,0)$, and a cannot be zero since we have a matrix with n linearly independent rows. We can then subtract multiples of this first row from the other rows to get a determinant having zeros in the first column below the first row. Now look at the $n-1$ by $n-1$ determinant to be multiplied by a.

Its rows must be linearly independent since those of the original matrix were. Now proceed by induction.)

To return to three dimensions, it is clear from the form of

$$\det A=a_{11}a_{22}a_{33}+a_{21}a_{32}a_{13}+a_{31}a_{12}a_{23}-a_{11}a_{32}a_{23}-a_{21}a_{12}a_{33}-a_{31}a_{22}a_{13}\tag{2.2.30}$$

that we could equally have taken the *columns* of A as three vectors, $A=(\vec{a}_1^C,\vec{a}_2^C,\vec{a}_3^C)$ in an obvious notation, $\det A=(\vec{a}_1^C\times\vec{a}_2^C)\cdot\vec{a}_3^C$, and linear dependence among the *columns* will also ensure the vanishing of the determinant—so, in fact, linear dependence of the columns ensures linear dependence of the rows.

This, too, generalizes to $n \times n$: in the definition of determinant $\det A = \varepsilon_{ijk\dots p} a_{1i} a_{2j} a_{3k} \dots a_{np}$, the row suffix is fixed and the column suffix goes over all permissible permutations, with the appropriate sign—but the same terms would be generated by having the *column* suffixes kept in numerical order and allowing the row suffix to undergo the permutations.

An Aside: Reciprocal Lattice Vectors

It is perhaps worth mentioning how the inverse of a 3×3 matrix operator can be understood in terms of vectors. For a set of linearly independent vectors $(\vec{a}_1, \vec{a}_2, \vec{a}_3)$, a reciprocal set $(\vec{b}_1, \vec{b}_2, \vec{b}_3)$ can be defined by

$$\vec{b}_1 = \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \times \vec{a}_2 \cdot \vec{a}_3} \tag{2.2.35}$$

and the obvious cyclic definitions for the other two reciprocal vectors. We see immediately that

$$\vec{a}_i \cdot \vec{b}_j = \delta_{ij} \tag{2.2.36}$$

from which it follows that the inverse matrix to

$$A = \begin{pmatrix} \vec{a}_1^R \\ \vec{a}_2^R \\ \vec{a}_3^R \end{pmatrix} \; is \; B = \begin{pmatrix} \vec{b}_1^C & \vec{b}_2^C & \vec{b}_3^C \end{pmatrix} \tag{2.2.37}$$

(These reciprocal vectors are important in *x*-ray crystallography, for example. If a crystalline lattice has certain atoms at positions $n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$, where n_1, n_2, n_3 are integers, the reciprocal vectors are the set of normals to possible planes of the atoms, and these planes of atoms are the important elements in the diffractive *x*-ray scattering.)

Eigenkets and Eigenvalues

If an operator A operating on a ket $|V\rangle$ gives a multiple of the same ket,

$$A|V\rangle = \lambda|V\rangle \tag{2.2.38}$$

then $|V\rangle$ is said to be an *eigenket* (or, just as often, *eigenvector*, or *eigenstate*!) of A with *eigenvalue* λ.

Eigenkets and eigenvalues are of central importance in quantum mechanics: dynamical variables are operators, a physical measurement of a dynamical variable yields an eigenvalue of the operator, and forces the system into an eigenket.

In this section, we shall show how to find the eigenvalues and corresponding eigenkets for an operator A. We'll use the notation $A|a_i\rangle = a_i|a_i\rangle$ for the set of eigenkets $|a_i\rangle$ with corresponding eigenvalues a_i. (Obviously, in the eigenvalue equation here the suffix i is not summed over.)

The first step in solving $A|V\rangle = \lambda|V\rangle$ is to find the allowed eigenvalues a_i.

Writing the equation in matrix form:

$$\begin{pmatrix} A_{11}-\lambda & A_{12} & . & . & A_{1n} \\ A_{21} & A_{22}-\lambda & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ A_{n1} & . & . & . & A_{nn}-\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ . \\ . \\ v_n \end{pmatrix} = 0 \tag{2.2.39}$$

This equation is actually telling us that the *columns* of the matrix $A - \lambda I$ are linearly dependent! To see this, write the matrix as a row vector each element of which is one of its columns, and the equation becomes

$$(\vec{M}_1^C, \vec{M}_2^C, \dots, \vec{M}_n^C) \begin{pmatrix} v_1 \\ . \\ . \\ . \\ v_n \end{pmatrix} = 0 \tag{2.2.40}$$

which is to say

$$v_1\vec{M}_1^C + v_2\vec{M}_2^C + \ldots + v_n\vec{M}_n^C = 0 \tag{2.2.41}$$

the columns of the matrix are indeed a *linearly dependent* set.

We know that means the determinant of the matrix $A-\lambda I$ is zero,

$$\begin{vmatrix} A_{11}-\lambda & A_{12} & . & . & A_{1n} \\ A_{21} & A_{22}-\lambda & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ A_{n1} & . & . & . & A_{nn}-\lambda \end{vmatrix} = 0 \tag{2.2.42}$$

Evaluating the determinant using $\det A = \varepsilon_{ijk\ldots p} a_{1i} a_{2j} a_{3k} \ldots a_{np}$ gives an n^{th} order polynomial in λ sometimes called the *characteristic polynomial*. Any polynomial can be written in terms of its roots:

$$C(\lambda - a_1)(\lambda - a_2) \ldots (\lambda - a_n) = 0 \tag{2.2.43}$$

where the a_i's, the roots of the polynomial, and C is an overall constant, which from inspection of the determinant we can see to be $(-1)^n$. (It's the coefficient of λ^n.) The polynomial roots (which we don't yet know) are in fact the eigenvalues. For example, putting $\lambda = a_1$ in the matrix, $\det(A - a_1 I) = 0$, which means that $(A - a_1 I)|V\rangle = 0$ has a nontrivial solution $|V\rangle$, and this is our eigenvector $|a_1\rangle$.

Notice that the diagonal term in the determinant $(A_{11}-\lambda)(A_{22}-\lambda)\ldots(A_{nn}-\lambda)$ generates the leading two orders in the polynomial $(-1)^n(\lambda^n - (A_{11} + \ldots + A_{nn})\lambda^{n-1})$, (and some lower order terms too). Equating the coefficient of λ^{n-1} λ^{n-1} here with that in $(-1)^n(\lambda - a_1)(\lambda - a_2)\ldots(\lambda - a_n)$,

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} A_{ii} = TrA \tag{2.2.44}$$

Putting $\lambda = 0$ in both the determinantal and the polynomial representations (in other words, equating the λ-independent terms),

$$\prod_{i=1}^{n} a_i = \det A \tag{2.2.45}$$

So we can find both the sum and the product of the eigenvalues directly from the determinant, and for a 2×2 matrix this is enough to solve the problem.

For anything bigger, the method is to solve the polynomial equation $\det(A - \lambda I) = 0$ to find the set of eigenvalues, then use them to calculate the corresponding eigenvectors. This is done one at a time.

Labeling the first eigenvalue found as a_1, the corresponding equation for the components v_i v_i of the eigenvector $|a_1\rangle$ is

$$\begin{pmatrix} A_{11}-a_1 & A_{12} & . & . & A_{1n} \\ A_{21} & A_{22}-a_1 & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ A_{n1} & . & . & . & A_{nn}-a_1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ . \\ . \\ v_n \end{pmatrix} = 0 \tag{2.2.46}$$

This looks like n equations for the n numbers v_i, but it isn't: remember the rows are linearly dependent, so there are only $n-1$ independent equations. However, that's enough to determine

the ratios of the vector components $v_1, \ldots, v_n$, then finally the eigenvector is normalized. The process is then repeated for each eignevalue. (Extra care is needed if the polynomial has coincident roots—we'll discuss that case later.)

Eigenvalues and Eigenstates of Hermitian Matrices

For a *Hermitian* matrix, it is easy to establish that the eigenvalues are always *real*. (*Note*: A basic postulate of Quantum Mechanics, discussed in the next lecture, is that physical observables are represented by Hermitian operators.) Taking (in this section) A to be hermitian,$A = A^\dagger$, and labeling the eigenkets by the eigenvalue, that is,

$$A|a_1\rangle = a_1|a_1\rangle \tag{2.2.47}$$

the inner product with the bra $\langle a_1|$ gives $\langle a_1|A|a_1\rangle = a_1\langle a_1|a_1\rangle$. But the inner product of the *adjoint* equation (remembering $A = A^\dagger$)

$$\langle a_1|A = a_1^*\langle a_1| \tag{2.2.48}$$

with $|a_1\rangle$ gives $\langle a_1|A|a_1\rangle = a_1^*\langle a_1|a_1\rangle$, so $a_1 = a_1^*$, and all the eigenvalues must be real.

They certainly don't have to all be different—for example, the unit matrix I is Hermitian, and all its eigenvalues are of course 1. But let's first consider the case where they *are* all different.

It's easy to show that the eigenkets belonging to *different* eigenvalues are *orthogonal*.

If

$$\begin{aligned} A|a_1\rangle &= a_1|a_1\rangle \\ A|a_2\rangle &= a_2|a_2\rangle \end{aligned} \tag{2.2.49}$$

take the adjoint of the first equation and then the inner product with $|a_2\rangle$, and compare it with the inner product of the second equation with $\langle a_1|$:

$$\langle a_1|A|a_2\rangle = a_1\langle a_1|a_2\rangle = a_2\langle a_1|a_2\rangle \tag{2.2.50}$$

so $\langle a_1|a_2\rangle = 0$ unless the eigenvalues are equal. (If they *are* equal, they are referred to as *degenerate* eigenvalues.)

Let's first consider the nondegenerate case: A has all eigenvalues distinct. The eigenkets of A, appropriately normalized, form an orthonormal basis in the space.

Write

$$|a_1\rangle = \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix}, \textit{ and consider the matrix } V = \begin{pmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \dots & v_{nn} \end{pmatrix} = \begin{pmatrix} |a_1\rangle & |a_2\rangle & \dots & |a_n\rangle \end{pmatrix} \tag{2.2.51}$$

Now

$$AV = A\begin{pmatrix} |a_1\rangle & |a_2\rangle & \dots & |a_n\rangle \end{pmatrix} = \begin{pmatrix} a_1|a_1\rangle & a_2|a_2\rangle & \dots & a_n|a_n\rangle \end{pmatrix} \tag{2.2.52}$$

so

$$V^\dagger AV = \begin{pmatrix} \langle a_1| \\ \langle a_2| \\ \vdots \\ \langle a_n| \end{pmatrix} \begin{pmatrix} a_1|a_1\rangle & a_2|a_2\rangle & \dots & a_n|a_n\rangle \end{pmatrix} = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{pmatrix} \tag{2.2.53}$$

Note also that, obviously, V is unitary:

$$V^\dagger V = \begin{pmatrix} \langle a_1| \\ \langle a_2| \\ \vdots \\ \langle a_n| \end{pmatrix} \begin{pmatrix} |a_1\rangle & |a_2\rangle & \dots & |a_n\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \tag{2.2.54}$$

We have established, then, that for a Hermitian matrix with distinct eigenvalues (nondegenerate case), the unitary matrix V having columns identical to the normalized eigenkets of A *diagonalizes* A, that is, $V^\dagger AV$ is diagonal. Furthermore, its (diagonal) elements equal the corresponding eigenvalues of A.

Another way of saying this is that the unitary matrix V is the transformation from the original orthonormal basis in ths space to the basis formed of the normalized eigenkets of A.

Proof that the Eigenvectors of a Hermitian Matrix Span the Space

We'll now move on to the general case: what if some of the eigenvalues of A are the same? In this case, any linear combination of them is also an eigenvector with the same eigenvalue. *Assuming they form a basis in the subspace*, the Gram Schmidt procedure can be used to make it orthonormal, and so part of an orthonormal basis of the whole space.

However, we have not actually established that the eigenvectors *do* form a basis in a degenerate subspace. Could it be that (to take the simplest case) the two eigenvectors for the single eigenvalue turn out to be parallel? This is actually the case for some 2×2 matrices—for example, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, we need to prove it is *not* true for Hermitian matrices, and nor are the analogous statements for higher-dimensional degenerate subspaces.

A clear presentation is given in Byron and Fuller, section 4.7. We follow it here. The procedure is by induction from the 2×2 case. The general 2×2 Hermitian matrix has the form

$$\begin{pmatrix} a & b \\ b^* & c \end{pmatrix} \tag{2.2.55}$$

where a, c are real. It is easy to check that if the eigenvalues are degenerate, this matrix becomes a real multiple of the identity, and so trivially has two orthonormal eigenvectors. Since we already know that if the eigenvalues of a 2×2 Hermitian matrix are distinct it can be diagonalized by the unitary transformation formed from its orthonormal eigenvectors, we have established that *any* 2×2 Hermitian matrix can be so diagonalized.

To carry out the induction process, we now assume any $(n-1)\times(n-1)$ Hermitian matrix can be diagonalized by a unitary transformation. We need to prove this means it's also true for an $n\times n$ Hermitian matrix A. (Recall a unitary transformation takes one complete orthonormal basis to another. If it diagonalizes a Hermitian matrix, the new basis is necessarily the set of orthonormalized eigenvectors. Hence, if the matrix can be diagonalized, the eigenvectors do span the *n*-dimensional space.)

Choose an eigenvalue a_1 of A, with normalized eigenvector $|a_1\rangle = (v_{11}, v_{21}, \dots, v_{n1})^T$. (We put in T for transpose, to save the awkwardness of filling the page with a few column vectors.) We construct a unitary operator V by making this the first column, then filling in with $n-1$ other normalized vectors to construct, with $|a_1\rangle$, an *n*-dimensional orthonormal basis.

Now, since $A|a_1\rangle = a_1|a_1\rangle$, the first column of the matrix AV will just be $a_1|a_1\rangle$, and the *rows* of the matrix $V^\dagger = V^{-1}$ will be $\langle a_1|$ followed by $n-1$ normalized vectors orthogonal to it, so the first column of the matrix $V^\dagger AV$ $V^\dagger AV$ will be a_1 followed by zeros. It is easy to check that $V^\dagger AV$ is Hermitian, since A is, so its first row is also zero beyond the first diagonal term.

This establishes that for an $n\times n$ Hermitian matrix, a unitary transformation exists to put it in the form:

$$V^\dagger AV = \begin{pmatrix} a_1 & 0 & . & . & 0 \\ 0 & M_{22} & . & . & M_{2n} \\ 0 & . & . & . & . \\ 0 & . & . & . & . \\ 0 & M_{n2} & . & . & M_{nn} \end{pmatrix} \tag{2.2.56}$$

But we can now perform a second unitary transformation in the $(n-1)\times(n-1)$ subspace orthogonal to $|a_1\rangle$ (this of course leaves $|a_1\rangle$ invariant), to complete the full diagonalization—that is to say, the existence of the $(n-1)\times(n-1)$ diagonalization, plus the argument above, guarantees the existence of the $n\times n$ diagonalization: the induction is complete.

Diagonalizing a Hermitian Matrix

As discussed above, a Hermitian matrix is diagonal in the orthonormal basis of its set of eigenvectors: $|a_1\rangle, |a_2\rangle, \dots, |a_n\rangle$, since

$$\langle a_i|A|a_j\rangle = \langle a_i|a_j|a_j\rangle = a_j\langle a_i|a_j\rangle = a_j\delta_{ij} \tag{2.2.57}$$

If we are given the matrix elements of A in some other orthonormal basis, to diagonalize it we need to rotate from the initial orthonormal basis to one made up of the eigenkets of A.

Denoting the initial orthonormal basis in the standard fashion

$$|1\rangle=\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix},\ |2\rangle=\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix},\ |i\rangle=\begin{pmatrix}0\\\vdots\\1\\\vdots\\0\end{pmatrix}\ldots\ (1\,in\,i^{th}\,place\,down),\ |n\rangle=\begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix} \tag{2.2.58}$$

the elements of the matrix are $A_{ij}=\langle i|A|j\rangle$.

A transformation from one orthonormal basis to another is a *unitary* transformation, as discussed above, so we write it

$$|V\rangle\rightarrow|V'\rangle=U|V\rangle \tag{2.2.59}$$

Under this transformation, the matrix element

$$\langle W|A|V\rangle\rightarrow\langle W'|A|V'\rangle=\langle W|U^{\dagger}AU|V\rangle \tag{2.2.60}$$

So we can find the appropriate transformation matrix U by requiring that $U^{\dagger}AU$ $U^{\dagger}AU$ be diagonal with respect to the *original* set of basis vectors. (Transforming the operator in this way, leaving the vector space alone, is equivalent to rotating the vector space and leaving the operator alone. Of course, in a system with more than one operator, the same transformation would have to be applied to all the operators).

In fact, just as we discussed for the nondegenerate (distinct eigenvalues) case, the unitary matrix U we need is just composed of the normalized eigenkets of the operator A,

$$U=(|a_1\rangle,|a_2\rangle,\ldots,|a_n\rangle) \tag{2.2.61}$$

And it follows as before that

$$(U^{\dagger}AU)_{ij}=\langle a_i|a_j|a_j\rangle=\delta_{ij}a_j,\ a\,diagonal\,matrix. \tag{2.2.62}$$

(The repeated suffixes here are of course *not* summed over.)

If some of the eigenvalues are the same, the Gram Schmidt procedure may be needed to generate an orthogonal set, as mentioned earlier.

Functions of Matrices

The same unitary operator U that diagonalizes an Hermitian matrix A will also diagonalize A^2, because

$$U^{-1}A^2U=U^{-1}AAU=U^{-1}AUU^{-1}AU \tag{2.2.63}$$

so

$$U^{\dagger}A^2U=\begin{pmatrix}a_1^2&0&0&.&0\\0&a_2^2&0&.&0\\0&0&a_3^2&.&0\\.&.&.&.&.\\0&.&.&.&a_n^2\end{pmatrix} \tag{2.2.64}$$

Evidently, this same process works for any power of A, and formally for any function of A expressible as a power series, but of course convergence properties need to be considered, and this becomes trickier on going from finite matrices to operators on infinite spaces.

Commuting Hermitian Matrices

From the above, the set of powers of an Hermitian matrix all commute with each other, and have a common set of eigenvectors (but not the same eigenvalues, obviously). In fact it is not difficult to show that any two Hermitian matrices that commute with each other have the same set of eigenvectors (after possible Gram Schmidt rearrangements in degenerate subspaces).

If two $n\times n$ Hermitian matrices A, B commute, that is, $AB=BA$, and A has a *nondegenerate* set of eigenvectors $A|a_i\rangle=a_i|a_i\rangle$, then $AB|a_i\rangle=BA|a_i\rangle=Ba_i|a_i\rangle=a_iB|a_i\rangle$, that is, $B|a_i\rangle$ is an eigenvector of A with eigenvalue a_i. Since A

is nondegenerate, $B|a_i\rangle$ must be some multiple of $|a_i\rangle$, and we conclude that A, B have the same set of eigenvectors.

Now suppose A *is* degenerate, and consider the $m \times m$ $m\times m$subspace S_{a_i} spanned by the eigenvectors $|a_i, 1\rangle$, $|a_i, 2\rangle, \dots$ of A having eigenvalue a_i. Applying the argument in the paragraph above, $B|a_i, 1\rangle$, $B|a_i, 2\rangle, \dots$ must also lie in this subspace. Therefore, if we transform B with the same unitary transformation that diagonalized A, B will not in general be diagonal in the subspace S_{a_i}, but it will be what is termed *block diagonal*, in that if B operates on any vector in S_{a_i} it gives a vector in S_{a_i}.

B can be written as two diagonal blocks: one $m \times m$, one $(n-m)\times(n-m)$, with zeroes outside these diagonal blocks, for example, for $m = 2,\ n = 5$:

$$\begin{pmatrix} b_{11} & b_{12} & 0 & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 & 0 \\ 0 & 0 & b_{33} & b_{34} & b_{35} \\ 0 & 0 & b_{43} & b_{44} & b_{45} \\ 0 & 0 & b_{53} & b_{54} & b_{55} \end{pmatrix} \tag{2.2.65}$$

And, in fact, if there is only *one* degenerate eigenvalue that second block will only have nonzero terms on the diagonal:

$$\begin{pmatrix} b_{11} & b_{12} & 0 & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 & 0 \\ 0 & 0 & b_{3} & 0 & 0 \\ 0 & 0 & 0 & b_{4} & 0 \\ 0 & 0 & 0 & 0 & b_{5} \end{pmatrix} \tag{2.2.65}$$

B therefore operates on two subspaces, one *m*-dimensional, one (*n*-*m*)-dimensional, *independently*—a vector entirely in one subspace stays there.

This means we can complete the diagonalization of B with a unitary operator that *only* operates on the $m \times m$ block S_{a_i}. Such an operator will also affect the eigenvectors of A, but that doesn't matter, because all vectors in this subspace are eigenvectors of A with the *same* eigenvalue, so as far as A is concerned, we can choose any orthonormal basis we like—the basis vectors will still be eigenvectors.

This establishes that any two *commuting* Hermitian matrices can be diagonalized at the same time. Obviously, this can never be true of *noncommuting* matrices, since all diagonal matrices commute.

Diagonalizing a Unitary Matrix

Any unitary matrix can be diagonalized by a unitary transformation. To see this, recall that *any* matrix M can be written as a sum of a Hermitian matrix and an anti Hermitian matrix,

$$M = \frac{M + M^\dagger}{2} + \frac{M - M^\dagger}{2} = A + iB \tag{2.2.66}$$

where both A, B are Hermitian. This is the matrix analogue of writing an arbitrary complex number as a sum of real and imaginary parts.

If $A,\ B$ commute, they can be simultaneously diagonalized (see the previous section), and therefore M can be diagonalized. Now, if a *unitary* matrix is expressed in this form $U = A + iB$ with $A,\ B$ Hermitian, it easily follows from $UU^\dagger = U^\dagger U = 1$ that $A,\ B$ commute, so *any unitary matrix* U *can be diagonalized by a unitary transformation*. More generally, if a matrix M commutes with its adjoint $M^\dagger$, it can be diagonalized.

(*Note*: it is *not* possible to diagonalize M unless both $A,\ B$ are simultaneously diagonalized. This follows from $U^\dagger AU,\ U^\dagger iBU$ being Hermitian and antiHermitian for any unitary operator U, so their off-diagonal elements cannot cancel each other, they must all be zero if *M* has been diagonalized by U, in which case the two transformed matrices $U^\dagger AU,\ U^\dagger iBU$ are diagonal, therefore commute, and so do the original matrices $A,\ B$.)

It is worthwhile looking at a specific example, a simple rotation of one orthonormal basis into another in three dimensions. Obviously, the axis through the origin about which the basis is rotated is an eigenvector of the transformation. It's less clear what the other two eigenvectors might be—or, equivalently, what are the eigenvectors corresponding to a two dimensional rotation of basis in a plane? The way to find out is to write down the matrix and diagonalize it.

The matrix

$$U(\theta)=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix} \tag{2.2.67}$$

Note that the determinant is equal to unity. The eigenvalues are given by solving

$$\begin{vmatrix}\cos\theta-\lambda & \sin\theta\\ -\sin\theta & \cos\theta-\lambda\end{vmatrix}=0\ to\,give\ \lambda=e^{\pm i\theta} \tag{2.2.68}$$

The corresponding eigenvectors satisfy

$$\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}u_1^{\pm}\\ u_2^{\pm}\end{pmatrix}=e^{\pm i\theta}\begin{pmatrix}u_1^{\pm}\\ u_2^{\pm}\end{pmatrix} \tag{2.2.69}$$

The eigenvectors, normalized, are:

$$\begin{pmatrix}u_1^{\pm}\\ u_2^{\pm}\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ \pm i\end{pmatrix} \tag{2.2.70}$$

Note that, in contrast to a Hermitian matrix, the eigenvalues of a unitary matrix do not have to be real. In fact, from $U^{\dagger}U=1$, sandwiched between the bra and ket of an eigenvector, we see that any eigenvalue of a unitary matrix must have unit modulus—it's a complex number on the unit circle. With hindsight, we should have realized that one eigenvalue of a two-dimensional rotation had to be $e^{i\theta}$, the product of two two-dimensional rotations is given be adding the angles of rotation, and a rotation through π changes all signs, so has eigenvalue -1. Note that the eigenvector itself is independent of the angle of rotation—the rotations all commute, so they must have common eigenvectors. Successive rotation operators applied to the plus eigenvector add their angles, when applied to the minus eigenvector, all angles are subtracted.

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

2.3: Function Spaces

Linear Algebra in Infinite Dimensions

The motivation for our review of linear algebra was the observation that the set of solutions to Schrödinger's equation satisfies some of the basic requirements of a vector space, in that linear combinations of solutions give another solution to the equation. Furthermore, Schrödinger's equation itself, as a differential operator acting on a function, suggests that the concept of a matrix operator acting on vectors in an *n*-dimensional vector space can be extended to more general operators, such as differential operators, acting on functions in an *infinite*-dimensional space.

Our analysis of linear vector spaces began by defining an *inner product*, which was used to establish an *orthonormal basis* for the space. Constructing a well-defined basis for the space of all functions on the real axis sounds impossible, and probably is. Fortunately, we don't need to be so all-encompassing. For one thing, we are not interested in functions with discontinuities, because in quantum mechanics that would be a wavefunction corresponding to infinite energy. (We *can* allow discontinuities in slope, although, as discussed in the Electron in a Box lecture, that occurs only where the potential is infinite. Infinite potentials are of course unphysical, but are convenient approximations in some cases, so we'll keep that option open.) Another important restriction arises from the requirement that the wavefunction describe a single particle—it must be *normalizable*, that is to say the *norm*

$$\int_{-\infty}^{\infty} \psi^*(x)\psi(x)dx < \infty \tag{2.3.1}$$

and in fact ψ must be scaled so that this integral is equal to unity for actual computation of probabilities. Note that $\psi(x)$ means $\psi(x, t=0)$, but the norm turns out to be time-independent, as it must be, for the case of a single particle.

Building on the analogy with *n*-dimensional vector spaces, the requirement of finite norm suggests a definition for the inner product in function space:

$$\langle f|g\rangle = \int_{-\infty}^{\infty} f^*(x)g(x)dx \tag{2.3.2}$$

This definition satisfies Dirac's requirement that $\langle f|g\rangle^* = \langle g|f\rangle$, gives a positive norm, and is linear in f, g. The space of functions with this inner product, and with finite norm $\sqrt{\langle f|f\rangle}$, is written $L_2(-\infty, \infty)$ or just L_2. The functions are said to be "square integrable".

Notice that this inner product resembles the linear algebraic bra-ket product if we imagine every point on the line as an independent basis vector—mathematically meaningless, of course, but a hint of where we're going.

Electron in a Box Again

As a preliminary to discussing functions on the infinite line, it is worth considering those restricted to the finite interval $(0, L)$ and vanishing at the two ends. These are precisely the conditions satisfied by the electron-in-a-box wavefunctions (see the earlier lecture):

$$|n\rangle = \psi_n(x, t=0) = \sqrt{\frac{2}{L}} \sin\frac{n\pi x}{L} \tag{2.3.3}$$

Recall from the Fourier Series lecture that any function without discontinuities can be represented as a sum over Fourier components. For the present case of functions equal to zero at the two ends (as any physical wavefunction in a box must be) the sine kets $|n\rangle$ above form a complete set, that is to say, at $t=0$, any $\psi(x)$ satisfying the boundary condition can be written:

$$|\psi(x)\rangle = \sum_{n=1}^{\infty} a_n|n\rangle \tag{2.3.4}$$

where, from the orthonormality of the basis set $|n\rangle$, the Fourier coefficients $a_n = \langle n|\psi\rangle$,so (making explicit that $\psi(x)$is in fact a ket in this vector space)

$$|\psi\rangle = \sum_{n=1}^{\infty} |n\rangle\langle n|\psi\rangle \tag{2.3.5}$$

giving an identity operator in the space of continuous functions vanishing at 0 and L:

$$I=\sum_{n=1}^{\infty}|n\rangle\langle n| \tag{2.3.6}$$

exactly analogous to that in finite-dimensional vector spaces. The inner product of two functions

$$\psi(x)=\sum_{n=1}^{\infty}a_n|n\rangle,\ \phi(x)=\sum_{n=1}^{\infty}b_n|n\rangle \tag{2.3.7}$$

defined as in the preceding section by

$$\langle\phi|\psi\rangle=\int_{-\infty}^{\infty}\phi^*(x)\psi(x)dx \tag{2.3.8}$$

is equivalently, in terms of Fourier coefficients,

$$\langle\phi|\psi\rangle=\sum_{n=1}^{\infty}b_n^*a_n \tag{2.3.9}$$

and the normalization

$$\langle\psi|\psi\rangle=\int_{-\infty}^{\infty}\psi^*(x)\psi(x)dx=\sum_{n=1}^{\infty}|a_n|^2=1 \tag{2.3.10}$$

So for the electron-in-a-box wavefunctions, the orthonormal basis of sine functions gives a well-defined infinite-dimensional vector space.

We have previously stated that the standard interpretation of the wavefunction $\psi(x)$ is that $|\psi(x)|^2dx$ is the probability of finding the particle in a small interval dx near x, and on integrating over all x the total probability of finding the particle is one. But we could also look for the particle in a particular *state*, rather than in a particular small interval dx. In this case, $|a_n|^2$ is the probability of finding the particle in the n^{th} state. This is consistent with the previous interpretation, and is parallel to our earlier analysis of the probability of a particle having a particular momentum. The state coefficient a_n is called the *amplitude*, or sometimes the *probability amplitude*.

You might be wondering how we would measure that a particle is in a particular state. The answer is to wait for it to jump out. If an atom is excited (for example by a short burst of radiation) it will be excited to a state which is a linear superposition of different energy eigenstates, $\sum a_n|E_n\rangle$, rather than to a single eigenstate. It will usually return to the ground state by emitting one or a series of photons, and the frequency of an emitted photon reveals the energy difference between the atomic states involved. For a collection of atoms excited in the same way, the relative intensities of different spectral lines give the relative probabilities of different states. Of course, a long almost monochromatic wave packet of incoming radiation will tend to put all the excited atoms into the same state.

Exercise 2.3.1

Write out the identity operator for the electron in a box $I=\sum_{n=1}^{\infty}|n\rangle\langle n|$ using the explicit form $|n\rangle=\sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}$. Prove this is equivalent to the delta function when operating on other functions *within the box*. What is the behavior of this function *outside* the box?

Functions on the Infinite Line

What happens if we take the analysis of the previous section and let L go to infinity? This is parallel to the analysis (two lectures back) of going from Fourier series to the Fourier transform, the sum over a *series* of plane waves satisfying a boundary condition becoming an integral over the *continuum* of *all* plane waves. In that lecture, we saw that as L went to infinity, the amplitude of the normalized eigenstates $|n\rangle$ went to zero as $1/\sqrt{L}$, and therefore so did the individual coefficients $a_n=\langle n|\psi\rangle$. However, the density of these eigenstates in momentum space increased as L, so overall the factors of L cancelled and the sum tended to a finite integral, specifically

$$\psi(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}a(k)e^{ikx}dk \tag{2.3.11}$$

with

$$a(k)=\int_{-\infty}^{\infty}\psi(x)e^{-ikx}dx \tag{2.3.12}$$

For the electron in a box (Fourier series) above we wrote the corresponding equation in Dirac notation as

$$|\psi(x)\rangle=\sum_{n=1}^{\infty}a_n|n\rangle \tag{2.3.13}$$

with

$$a_n=\langle n|\psi\rangle,\ so\ I=\sum_{n=1}^{\infty}|n\rangle\langle n| \tag{2.3.14}$$

It's tempting to write down the analogous equations for the infinite line case, by translating the Fourier transform equations into Dirac notation, and blindly writing $e^{ikx}=|k\rangle$:

$$|\psi(x)\rangle=\int_{-\infty}^{\infty}\frac{dk}{2\pi}a(k)|k\rangle,\ a(k)=\int_{-\infty}^{\infty}\psi(x)e^{-ikx}dx=\langle k|\psi(x)\rangle,\ I=\int_{-\infty}^{\infty}\frac{dk}{2\pi}|k\rangle\langle k| \tag{2.3.15}$$

This looks good, but has a problem—in contrast to the Fourier series basis functions $|n\rangle$, these Fourier transform "basis states" $|k\rangle$ are infinitely long plane wave states e^{ikx} and therefore not normalizable in the sense we've used that term so far. They're not even in the space we're supposed to be working in!

Furthermore, $|\langle k|\psi(x)\rangle|^2$ is *not* the probability that a measurement of the momentum of the electron will yield precisely the value $p=\hbar k$. *The correct probabilistic interpretation for a continuum of k-values is exactly parallel to the continuum of x-values in ordinary space*: $|\langle k|\psi(x)\rangle|^2 dk$ is the probability that a measurement of momentum would find the k-value to be in a small interval of width dk near k. The probability goes to zero with the width of the interval, and so is *vanishingly small* if we demand an *exact* value of k.

But we never measure k with infinite precision anyway—that would take an infinitely large apparatus. The *physically* significant quantity *is* the probability of finding k in a small interval dk—in practice, with real detectors, we are always integrating over some (small) *range* in k.

This means we might be ok with this continuum basis of states: we don't want them to be normalized in the traditional fashion $\langle k|k\rangle=1$, because that would correspond to a finite probability of the particle having a mathematically precise value of k, which makes no physical sense—in fact it's nonsense. The normalization we need is one that makes sense in the context of an integral over a small interval in k—but still of course over a continuous infinity of basis states!

From our earlier definition of the delta function, we can express orthogonality of these $|k\rangle$ states:

$$\langle k'|k\rangle=\int_{-\infty}^{\infty}e^{i(k-k')x}dx=2\pi\delta(k-k') \tag{2.3.16}$$

and, since the δ-function is normalized in the sense that it has total weight one in an integral, *we take this equation as the definition of the normalization of the functions* $|k\rangle$. That is to say, we take the state $|k\rangle$ to have wave function Ae^{ikx} with $A=1$.

Now the delta function is only meaningful inside an integral, therefore so is our normalization, and the formalism, a continuum basis of plane wave states with delta function orthogonality, although perhaps leaving something to be desired from a strict mathematical perspective, turns out to be a consistent and reliable way of formulating quantum mechanics.

Exercise 2.3.1

From the expression for the identity operator above:

$|\psi\rangle=\int_{-\infty}^{\infty}\frac{dk}{2\pi}|k\rangle\langle k|\psi\rangle$

substitute $|k\rangle=e^{ikx}$ and check that this makes sense.

Note: some authors prefer to define the normalized plane wave states by $|k\rangle=\sqrt{1/2\pi}e^{ikx}$, in which case $\langle k'|k\rangle=\delta(k'-k)$, and the $dk/2\pi$ appearing in the above integral for the identity operator becomes simply dk. With our convention, dk *always* appears with a 2π in the denominator.

Further Note: some prefer to go to a huge, but not infinite box, so the basis momentum eigenstates wave functions are the discrete set $|k\rangle=\sqrt{\frac{1}{L}}e^{ikx}$, or in three dimensions $|k\rangle=\sqrt{\frac{1}{V}}e^{i\vec{k}\cdot\vec{x}}$, V being the volume. For this huge box, it is safe to replace the sum over discrete momentum states by an integral, bearing in mind that the density of states in phase space being proportional to L gives $\sum_n\equiv\int L\frac{dk}{2\pi}$ or $\int V\frac{d^3k}{(2\pi)^3}$ in three dimensions. The L or V factors finally cancel in computations, as we shall discover later.

Schrödinger's Equation as an Operator on a Vector Space

As we recounted at the beginning of this course, when Schrödinger was challenged to find a wave equation for the electron wave, he constructed one parallel to the electromagnetic "photon wave equation", that is to say, he took the energy-momentum equation and wrote

$$E=i\hbar\frac{\partial}{\partial t},\ p_x=-i\hbar\frac{\partial}{\partial x}\tag{2.3.17}$$

He discovered that the three-dimensional version of the differential equation constructed in this way could be solved by standard analytic methods for an electron in an inverse-square force field—the hydrogen atom. The standing wave solutions yielded the right set of energy levels—those Bohr had found earlier with his simplistic model. This confirmed that indeed the wave equation describing the propagation of the electron waves had been discovered, and it was

$$\left(E-\frac{p^2}{2m}-V(x)\right)\psi(x,t)=0\tag{2.3.18}$$

with $E,\ p$ the differential operators given above. Since the operator in brackets is *linear*, the solutions $\psi(x,t)$ form a *linear vector space*.

Differential Operators: the Momentum Operator on L_2

Our task now is to recast this old approach of differential operators acting on wave functions in the equivalent Dirac language. Let's begin with the simplest, the momentum operator. First, we need to show that it is Hermitian. The trick is to integrate by parts:

$$\langle\phi|p_x|\psi\rangle=-i\hbar\int_{-\infty}^{\infty}dx\phi^*(x)\frac{d\psi(x)}{dx}=i\hbar\int_{-\infty}^{\infty}dx\psi(x)\frac{d\phi^*(x)}{dx}-[\phi^*(x)\psi(x)]_{-\infty}^{\infty}\tag{2.3.19}$$

The last term, the contribution from the infinite endpoints of the integration, *must be zero* because square-integrable functions must go to zero at infinity, so

$$\langle\phi|p_x|\psi\rangle=i\hbar\int_{-\infty}^{\infty}dx\psi(x)\frac{d\phi^*(x)}{dx}\tag{2.3.20}$$

Now $p_x|\phi\rangle=-i\hbar d\phi/dx=|p_x\phi\rangle$, so $\langle p_x\phi|=i\hbar d\phi^*/dx=\langle\phi|p_x^\dagger$, and

$$\langle\phi|p_x|\psi\rangle=i\hbar\int_{-\infty}^{\infty}dx\psi(x)\frac{d\phi^*(x)}{dx}=\langle\phi|p_x^\dagger|\psi\rangle\tag{2.3.21}$$

we have established that $p_x=p_x^\dagger$ between any two states in the space: so this is an operator identity, and $p_x=-i\hbar d/dx$ is Hermitian. (The i is important: the differential operator d/dx alone is not Hermitian, it's *anti* Hermitian in L_2!

So p_x is a Hermitian operator, and therefore has *real* eigenvalues, which it *must have* since momentum is a physical quantity. But what are its eigenvectors? We already know, of course, that they are the plane wave states—this is the whole reason this particular operator was chosen in building the wave equation in the first place. Strictly speaking, though, as we've already discussed, these plane wave states are not in L_2. Nevertheless, any smooth function in L_2 can be expressed as an integral over these states, so they *do* form a complete basis for the functions relevant to physics.

(It is true that later, in scattering theory and some other places, we may talk about plane waves without always doing an integral: such loose talk should be understood as referring to a very long but finite wave packet, well approximated by a plane wave during the scattering event.)

The Position Operator and Its Eigenstates

The "position" is just the co-ordinate x, manifestly always real, and a Hermitian operator.

Proof:

$$\langle\varphi|x|\psi\rangle = \int \varphi^*(x)x\psi(x)dx = (\int \psi^*(x)x\varphi(x)dx)^* = \langle\psi|x|\varphi\rangle^* \tag{2.3.22}$$

We shall make clear that in this context we regard x as an operator by writing it with a little hat, $\hat{x}$. It is equally clear that the eigenstates of $\hat{x}$, states in which the particle has probability equal to one of being at a particular position, must be delta functions corresponding to that position: that's the only function with zero probability of finding the particle anywhere else. So if $|a\rangle$ is an eigenstate of x with eigenvalue a,

$$|a\rangle = C\delta(x-a) \tag{2.3.23}$$

where C is a constant. But, whatever value we choose for C, this wave function, like the momentum eigenstate, isn't normalizable —so, in fact, $|a\rangle$ itself could never be the wave function of a particle!

Example 2.3.1

Take your favorite definition of the delta function, and prove that it isn't normalizable, as defined in L_2.

Solution

(It wouldn't be physically reasonable anyway—to localize a particle to a point would take infinite energy.) But the set of all $|a\rangle$'s is certainly *complete*, and therein lies its value: it *is* a basis for the space. The convention is to "normalize" these kets, or rather to construct an "orthonormal set", by analogy with the orthonormalization convention for the plane-wave momentum states, that is, to take

$$\langle a|b\rangle = \delta(a-b) \tag{2.3.24}$$

From the earlier result

$$\int \delta(a-x)\delta(x-b)dx = \delta(a-b) \tag{2.3.25}$$

it follows immediately that $C=1$.

Therefore,

$$\langle x|\psi\rangle = \int_{-\infty}^{\infty} dx'\delta(x-x')\psi(x') = \psi(x) \tag{2.3.26}$$

Taking the inner product of $|\psi\rangle$ with the bra $\langle x|$ just gives the value of $\psi\Psi$ at the point x. Consequently any function $\psi(x)$ in L_2 can be written:

$$|\psi\rangle = \int_{-\infty}^{\infty} dx|x\rangle\langle x|\psi\rangle \tag{2.3.27}$$

Exercise 2.3.1

Check that Equation 2.3.27is true by finding $\langle x'|\psi\rangle$.

It follows from the Equation 2.3.27that the identity operator in L_2 can be written in terms of the eigenstates of $\hat{x}$:

$$I = \int_{-\infty}^{\infty} dx|x\rangle\langle x| \tag{2.3.28}$$

From this, $|k\rangle$ can be written

$$|k\rangle = \int_{-\infty}^{\infty} dx|x\rangle\langle x|k\rangle = \int_{-\infty}^{\infty} dxe^{ikx}|x\rangle \tag{2.3.29}$$

and

$$|x\rangle=\int_{-\infty}^{\infty}\frac{dk}{2\pi}|k\rangle\langle k|x\rangle=\int_{-\infty}^{\infty}\frac{dk}{2\pi}e^{-ikx}|k\rangle \tag{2.3.30}$$

These are possibly the least rigorous equations in this section—we're expressing one set of states outside of L_2 in terms of another such set, using both sets as bases in L_2! Obviously, this is only meaningful with an $|x\rangle$ state defined as a zero-width limit of narrowing Gaussians (say) and a $|k\rangle$ state as a limit of longer and longer wavepackets, tending to a single *k*-value. Yet despite the lack of rigor in the above presentation, these states, used with care, are in fact reliable and efficient tools for analyzing quantum mechanical problems. We shall use them often.

Exercise: show these equations are consistent by substituting $|k\rangle$ from the first into the right-hand side of the second, to give $|x\rangle=|x\rangle$.

The Hamiltonian Operator

The Hamiltonian operator gives the time development of the wavefunction. It corresponds to the total energy. If the wavefunction corresponds to a definite energy, the time dependence can be factored out, and the spatial wave function is a solution of Schrödinger's time independent equation:

$$H\psi(x)=\left(\frac{p^2}{2m}+V(x)\right)\psi(x)=-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x)}{\partial x^2}+V(x)\psi(x)=E\psi(x) \tag{2.3.31}$$

Since we only consider the space L_2 of wave functions on which both p and x are Hermitian, H must be Hermitian, and therefore has real eigenvalues.

The Basic Rules of Quantum Mechanics

Any quantum mechanical wave function must be normalizable, because the norm represents the total probability of finding the particle (or, more generally, the system) somewhere in its phase space, so

First Basic Rule: *any state of the particle is a ket $|\psi\rangle$, symbolizing a function $\psi(x)$ in L_2.*

> Mathematicians use the term *Hilbert space* to refer to inner-product spaces of normalizable functions such that any convergent sequence in the space has a limit in the space (a property that, for example, the rational numbers don't have, but the real numbers do) . Our functions above for the electron in the box do form such a space, with the sine waves an orthonormal basis. However, on going to the infinite line, although we still have normalizable wave functions, the two bases we have discussed above, the plane waves (momentum basis) and the delta functions (position basis)*are not themselves in the space*—by which we mean they are not normalized as defined in L_2.
>
> But these bases are both *complete*, meaning any wave function can be expressed in terms of a (continuous) sum over the elements of either of them.
>
> Constructing these complete but not conventionally normalized bases was Dirac's doing, and is extremely convenient in describing quantum mechanics. But it upset the mathematicians. Fortunately, they later justified it by inventing the theory of *distributions*, which are generalized functions, and include delta functions.
>
> *Bottom Line*: we shall follow the other physicists in using the term "Hilbert space" *more loosely* than mathematicians do, to refer to L_2, *extended* to include these non-normalizable bases.

Next Basic Rule: *A physical variable, or observable, corresponds to a Hermitian operator A acting on L_2.*

> We shall *assume* that the eigenkets of any such variable span the space: this is always true for a finite dimensional space, as previously discussed, but not for a general Hermitian operator in a Hilbert space, so this is a nontrivial assumption.
>
> For an operator with a *discrete* set of eigenvalues, $A|n\rangle=\lambda_n|n\rangle$, any wave function can be written
>
> $$|\psi\rangle=\sum c_n|n\rangle,\ with\ c_n=\langle n|\psi\rangle \tag{2.3.32}$$

Rule for Relating Operators to Experiments: *any measurement of the value of the physical variable A will yield one of the eigenvalues λ_n of the operator A, and the probability of finding the particular value λ_n is equal to $|c_n|^2=|\langle n|\psi\rangle|^2$.*

> The ***expectation value*** of an observable A is the average value of a series of measurements on identical quantum systems,

$$\langle A\rangle=\langle\psi|A|\psi\rangle=\sum|c_n|^2\lambda_n \tag{2.3.33}$$

> It is important to note that two measurements of the same observable A *on the same system*, one measurement being made *immediately* after the other, must yield the same result. That is to say, if the first measurement reads λ_n, the second must be λ_n with 100% probability. But this can only happen if the wave function after the first measurement is $|n\rangle$, which in general it wasn't before the first measurement. The jargon description of this is that the act of measurement "collapses the wave function" into one of the eigenstates of the variable being measured.

Measuring a Continuum Variable: For variables like position and momentum having continuum sets of eigenvectors, the statistical interpretation is in terms of finding the particle within some small range—the probability of finding it between x and $x+dx$— is

$$\langle\psi|\int_x^{x+dx} dx|x\rangle\langle x|\psi\rangle=\int_x^{x+dx}|\psi(x)|^2dx \tag{2.3.34}$$

and the expectation value of x is

$$\langle\psi|x|\psi\rangle=\langle\psi|\hat{x}\int_{-\infty}^{\infty} dx|x\rangle\langle x|\psi\rangle=\int_{-\infty}^{\infty} x|\psi(x)|^2dx \tag{2.3.35}$$

where we've put a little hat on the x to remind us that it is an *operator*, with eigenkets $|x\rangle$.

2.4: Complex Variable, Stationary Path Integrals

Analytic Functions

Suppose we have a complex function $f=u+iv$ of a complex variable $z=x+iy$, defined in some region of the complex plane, where u, v, x, y are real. That is to say,

$$f(z)=u(x,y)+iv(x,y), \tag{2.4.1}$$

with $u(x,y)$ and $v(x,y)$ real functions in the plane.

We now assume that in this region $f(z)$ is differentiable, that is to say,

$$\frac{df(z)}{dz}=\lim_{\Delta z\to 0}\frac{f(z+\Delta z)-f(z)}{\Delta z} \tag{2.4.2}$$

is well-defined. What does this tell us about the functions $u(x,y)$ and $v(x,y)$, the real and imaginary parts of $f(z)$?

In fact, the property of differentiability for a function of a complex variable tells us a lot! It does *not* just mean that the function is reasonably smooth. The crucial difference from a function of a real variable is that Δz can approach zero *from any direction* in the complex plane, and *the limit in these different directions must be the same*. Of course, there are only two independent directions, so what we are really saying is

$$\frac{\partial f(x+iy)}{\partial x}=\frac{\partial f(x+iy)}{\partial(iy)}, \tag{2.4.3}$$

which we can write in terms of $u,\ v$:

$$\frac{\partial u(x,y)}{\partial x}+i\frac{\partial v(x,y)}{\partial x}=\frac{\partial u(x,y)}{\partial(iy)}+i\frac{\partial v(x,y)}{\partial(iy)}. \tag{2.4.4}$$

Equating real and imaginary parts of this equation we find:

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y},\quad \frac{\partial v}{\partial x}=-\frac{\partial u}{\partial y}. \tag{2.4.5}$$

These are called the *Cauchy-Riemann equations*.

It immediately follows that both $u(x,y)$ and $v(x,y)$ must satisfy the two-dimensional Laplacian equation,

$$\frac{\partial^2 u(x,y)}{\partial x^2}+\frac{\partial^2 u(x,y)}{\partial y^2}=0, \tag{2.4.6}$$

that is,

$$\nabla^2 u=0\ \ and\ \ \nabla^2 v=0. \tag{2.4.7}$$

Notice that this implies (just as for an electrostatic potential) that $u(x,y)$ cannot have an absolute minimum or maximum inside the region of analyticity. If $df(z)/dz=0$, but the second-order partial derivatives are nonzero, then they must have opposite sign, signaling a saddlepoint. In the general case, a two-dimensional version of Gauss' theorem can be used to show there is no local extremum.

Furthermore,

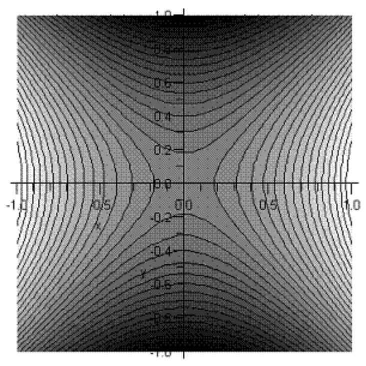
$$\nabla u\cdot\nabla v=\left(\frac{\partial u}{\partial x},\frac{\partial u}{\partial y}\right)\cdot\left(\frac{\partial v}{\partial x},\frac{\partial v}{\partial y}\right)=0. \tag{2.4.8}$$

That is to say, the *contour lines of constant* $u(x,y)$ *are everywhere orthogonal to the contour lines of constant* $v(x,y)$. (The gradient being orthogonal to the contour lines everywhere.) The important point is that *just requiring differentiability* of a function of a complex variable imposes a *strong* constraint on its real and imaginary parts, the functions $u(x,y)$ and $v(x,y)$.

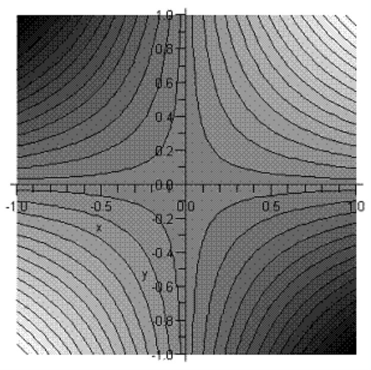
Example 2.4.1

A Simple Example: $f(z) = z^2$.

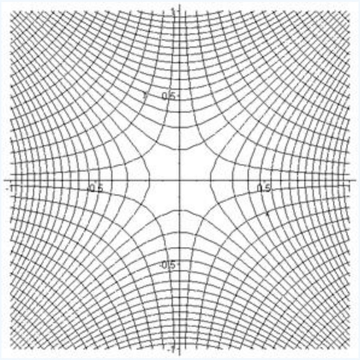
It is worthwhile building a clear picture of the real and imaginary parts of the function z^2. The real part is $x^2 - y^2$, and its contour lines in the square -1 to 1 are shown below. **The darker shades are the lower ground**. At the origin, there is a saddlepoint with higher ground in both directions of the real axis, lower ground in the pure imaginary directions. The lines $x = y$, $x = -y$ (not shown) are contours at the same level (zero) as the origin.



What about the imaginary part? $Im\ z^2 = 2xy$ has contours:



Putting the two sets of contour lines on the same diagram it is clear that they always cut each other orthogonally:



(Incidentally, this picture has a physical realization. It represents the field lines and equipotentials of a quadrupole magnet, used for focusing beams of charged particles.)

Example 2.4.2

$f(z) = 1/z$

The definition of differentiation above can be used to show that

$$\frac{d}{dz}\frac{1}{z}=-\frac{1}{z^2} \tag{2.4.9}$$

just as for a real variable, so the function can be differentiated everywhere in the complex plane except at the origin. The singularity at the origin is termed a “pole”, for obvious reasons.

Contour Integration: Cauchy’s Theorem

Cauchy’s theorem states that *the integral of a function of a complex variable around a closed contour in the complex plane is zero if the function is analytic in the region enclosed by the contour*.

This theorem can be proved at various levels of rigor, we shall give a basic physicist’s proof using Stokes theorem, that the integral of a vector function around a contour (now in ordinary, not complex, space) is equal to the integral of the curl of that function over an area spanning the contour, provided of course the curl is well-defined everywhere on the area,

$$\oint \vec{P}\cdot\overrightarrow{ds}=\int curl\,\vec{P}\cdot d\vec{A}. \tag{2.4.10}$$

Taking the special case where the contour and the area are confined to the x, y plane, and writing $\vec{P}=(P,Q)$, $curl\,\vec{P}=(\partial Q/\partial x-\partial P/\partial y)$, and Stokes’ theorem becomes:

$$\oint(Pdx+Qdy)=\iint\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dxdy, \tag{2.4.11}$$

known in this form as Green’s theorem (and easy to prove: the two terms are separately equal,

$$\oint Qdy=\iint(\partial Q/\partial x)dxdy. \tag{2.4.12}$$

This can be established by dividing the area into strips of infinitesimal width dy parallel to the x-axis, integrating with respect to x within a strip, to give $\int_{strip}(\partial Q/\partial x)dx=Q(x_2,y_1)-Q(x_1,y_1)$, the co-ordinates of the points on the contour at the ends of the strip, then adding the contributions from all the parallel strips just gives the integral around the contour.)

Now back to the complex plane: write as before

$$z=x+iy,\quad f(z)=u(x,y)+iv(x,y) \tag{2.4.13}$$

from which

$$\oint f(z)dz=\oint(u+iv)(dx+idy)=\oint(udx-vdy)+i\oint(vdx+udy). \tag{2.4.14}$$

Now apply Green’s theorem to the two integrals on the right, replacing $P,\ Q$ with first $u,\ -v$ then with $v,\ u$. This gives:

$$\iint\left(-\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)dxdy+i\int\int\left(\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y}\right)dxdy \tag{2.4.15}$$

and both these integrals are identically zero from the Cauchy-Riemann equations.

So the integral of a function of a complex variable around a closed contour in the complex plane can only depend on nonanalytic behavior inside the contour. Consider for example a pole, take the simplest case of integrating $1/z$ around the unit circle, the conventional direction is counterclockwise. Then $1/z=e^{-i\theta}$, $dz=ie^{i\theta}d\theta$,

$$\oint\frac{dz}{z}=\oint id\theta=2\pi i. \tag{2.4.16}$$

This is called the *residue* at the pole.

Moving the Contour of Integration

Cauchy’s theorem has a very important consequence: for an integral from, say, z_a to z_b in the complex plane, moving the contour in a region where the function is analytic will not affect the result, because the difference between the integral over the original contour and that over the shifted contour is an integral around a closed circuit, and therefore zero, provided the function is analytic in the region enclosed.

For an integral around a closed contour, if the only singularities enclosed by the contour are poles, the contour may be shrunk and broken to become a sum of separate small contours, one around each pole, then the integral around the original contour is the sum of the residues at the poles.

Other Singularities: Cuts, Sheets, etc.

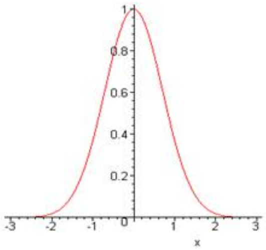
Poles are of course not the only possible singularities. For example, $\log z$ has a singularity at the origin. Now, $\log z = \log re^{i\theta} = \log r + i\theta$. The singularity at the origin is from the $\log r$ term, but notice that if we go around the unit circle, θ increases by 2π, and if we go around again it increases by a further 2π. This means that the value of $\log z$ is not uniquely defined: any given point in the complex plane has values differing by $2n\pi i$, n any integer. This is handled by replacing the single complex plane with a pile of sheets, and a cut going out from the origin. To find $\log z$, you need to know not only z, but also which sheet you're on: going up one sheet means $\log z$ has increased by $2\pi i$. When you cross the cut, you go to the next sheet, like a multilevel parking garage. The cut can go out from the origin in any direction, the standard arrangement is along the real axis, either positive or negative.

The square root function similarly has a cut, but only two sheets.

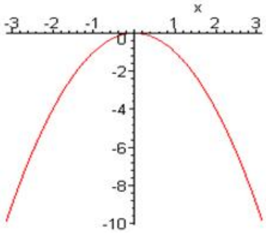
Evaluating Rapidly Oscillating Integrals by Steepest Descent

How to evaluate $\int_{-\infty}^{\infty} e^{iax^2} dx$ *in an unambiguous fashion: an introduction to moving the contour of integration and the Method of Steepest Descent.*

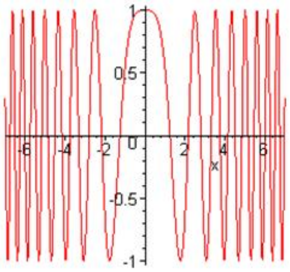
The familiar Gaussian integral $\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$ is easy to understand. Plotting the integrand, (here for $a = 1$) there is a peak of height 1 and width of order $1/\sqrt{a}$.



But what about the result $\int_{-\infty}^{\infty} e^{-iax^2} dx = \sqrt{\frac{\pi}{ia}}$? This (correct) result is far less obvious! Here the integrand e^{-iax^2} is always on the unit circle in the complex plane, and equal to 1 at $x = 0$. It is instructive to plot the *phase* of the integrand $\varphi(x) = -ax^2$ as a function of x (taking $a = 1$ in the graph below).



The phase is stationary at the origin, so contributions from that region add coherently. To help visualize the integrand better, here's a plot of the real part:



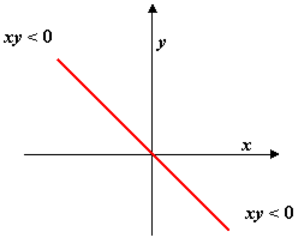
It is evident that almost all the contribution to the integral comes from the central region where the phase is stationary, the increasingly rapid oscillations away from the origin ensuring that very little comes from elsewhere.

So how do we actually evaluate the integral? In the complex plane $z=x+iy$, we can write

$$I=\int_{-\infty}^{\infty} e^{-iaz^2}dz \ \ (along\,real\,axis) \ = \int_{-\infty}^{\infty} e^{-ia(x^2-y^2+2ixy)}dz. \tag{2.4.17}$$

Notice that the *amplitude* (or *modulus*) of the integrand $e^{2axy}=1$ on the real axis, so it does not go to zero at infinity, although there are essentially no contributions from large x because of the rapid oscillations.

A cleaner way to see what's going on is to rotate the contour of integration around the origin to the 45 degree line $x=-y$. It's safe to do this because the amplitude of the integrand decreases on going from the real axis into the region $xy<0$, and in fact tends to zero on going to infinity in that region.



(To give a more precise argument, suppose we replace the infinite integral by one from $-L$ to L, so we will be taking the limit of L going to infinity at the end. Then the distorted contour has first a vertical part, from $(-L,0)$ to $(-L,L)$ then the diagonal contour from $(-L,L)$ to $(L,-L)$, finally another vertical leg from $(L,-L)$ to $(L,0)$. Now, on the first vertical part, the integral is clearly less than the integral of the modulus, that is,

$$\int_{(-L,0)}^{(-L,L)} e^{-ia(x^2-y^2+2ixy)}dz<\int_0^L e^{-2aLy}dy<\int_0^{\infty} e^{-2aLy}dy=1/2aL \tag{2.4.18}$$

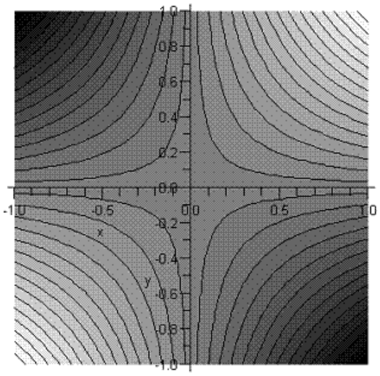
so in the limit of the original integral being over the whole real axis, the contributions from the vertical parts of the contour vanish.

The integral becomes \(\int_{-\infty}^{\infty} e^{-2ax^2}dz\) with \(dz=\sqrt{2}e^{-i\pi/4}dx\), so \[I=\sqrt{2}e^{-i\pi/4}\int_{-\infty}^{\infty} e^{-2ax^2}dx=\sqrt{2}e^{-i\pi/4}\sqrt{\frac{\pi}{2a}}=\sqrt{\frac{\pi}{ia}}, \label{2.4.18}\]

the required result.

General Steepest Descent Method

In fact, the contour rotation trick used above to make the integral easier to evaluate is a particular case of a method having wide applicability in evaluating contour integrals of the form $\int e^{iaf(z)}dz$. The basic strategy is to distort the contour of integration in the complex *z*-plane so that the amplitude of the integrand is as small as possible over as much of the contour of integration as possible. Actually, that is exactly what we did in the example above. To see this, it is helpful to plot the contour lines (lines of constant value) of the *modulus* of the integrand, $|e^{-iax^2}|=e^{2axy}$.



The convention here is white for the high ground, **black for the valleys**.

We want to keep the contour of integration as low as possible for as long as possible. The map above is of a “saddlepoint”: hills rise to the northeast and the southwest of the origin, valleys fall away to the northwest and the southeast. The strategy is to stay in the valleys (small integrand) as much as possible—however, to get from $-\infty$ to $+\infty$ we have to get from one valley to the other, and that means going over the saddlepoint at the origin. Obviously, to get the integrand as small as possible at all stages in the integration we must go down from the saddlepoint in both directions by the steepest possible route, and it is evident that this is right down the center of the valley, just the contour we chose above. *Note that this steepest descent path is also one of stationary phase*. This is because for any analytic function of a complex variable $f(z)$, the lines of constant $Re\, f(z)$ are perpendicular to those of constant $Im\, f(z)$. For a function $e^{f(z)}$, the steepest descent line is perpendicular to the lines of constant $Re\, f(z)$, and is therefore a line of constant $Im\, f(z)$, that is, constant phase of $e^{f(z)}$.

Saddlepoints of Analytic Functions

Suppose we have a function $f(z)$ analytic in some region R of the complex plane, and at some point z_0 inside R the derivative $\frac{df(z)}{dz}=0$. Then in the neighborhood of z_0,

$$f(z)=f(z_o)+\frac{1}{2}f''(z_o)(z-z_o)^2+\ldots \tag{2.4.19}$$

Close enough to z_0 we can neglect the higher order terms, and for the case of$f''(z_o)$ real, the contour lines of the real and imaginary parts of $f(z)$ will then be exactly those we have plotted for z^2 above. For $f''(z_o)$ complex, the plots will be rotated by an angle equal to the phase of $f''(z_o)$. That is to say, for any *analytic function*, near any point where $df(z)/dz=0$, the real and imaginary parts of the function have *saddlepoints* with contour maps rotated versions of those above.

Integrating Through a Saddlepoint

We consider now integrals of the form

$$\int_C e^{f(z)}dz \tag{2.4.20}$$

where C is some path in a region where $f(z)$ is analytic. This means the value of the integral will not be affected by distorting the path, provided it stays in the region of analyticity. (The path of integration is usually called the *contour* of integration—we’ll call it path here, to avoid confusion with *our* contours, which have the standard geographic meaning, joining points having the same value of some parameter.)

Note that with the exponential form of the integrand, the *real* part of $f(z)$ determines the *magnitude* of the integrand, the *imaginary* part of $f(z)$ determines its *phase*.

The strategy is to arrange the path of integration so that as much as possible of it is in the valleys, where the integrand is small, then to go over the saddlepoint by the steepest possible route, which would be staying on the imaginary axis in the case of z^2 plotted above. It is important to note that this “steepest descent” route is also a path along which *the imaginary part of $f(z)$ remains constant*, so the contributions along this path are *all in phase*, that is to say, they add coherently.

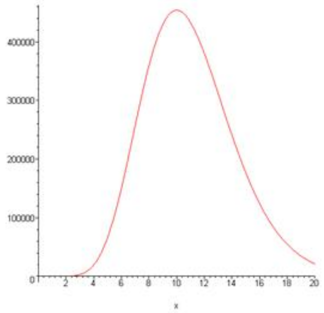
The bottom line is that by directing the path of integration through the saddlepoint along the steepest route for the magnitude of the integrand, the biggest contributions to the integral are all in phase. Along this path, the integral has standard Gaussian form. If the function *f*(*z*) is sufficiently large, it may be that the contribution of the integral away from the saddlepoint can be neglected. This method is therefore often valuable in cases where some parameter becomes large: we give a number of examples to clarify this point.

Saddlepoint Estimation of $n!$

We use the identity

$$n!=\int_0^\infty t^n e^{-t}dt=\int_0^\infty e^{f(t)}dt \;\; with \;\; f(t)=n\ln t-t. \tag{2.4.21}$$

To picture $t^n e^{-t}$, here it is for $n=10$:



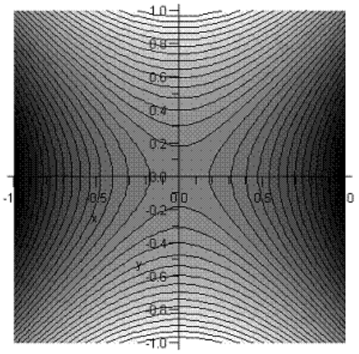
Note that

$$f'(t)=\frac{n}{t}-1, \quad f'=0 \; for \; t=n, \quad f''(t)=-\frac{n}{t^2}. \tag{2.4.22}$$

Therefore, in the neighborhood of the maximum value of $f(t)$ at $t=n$,

$$f(t)=n\ln n-n-\frac{1}{2n}(t-n)^2+higher\,order\,terms. \tag{2.4.23}$$

For integer n, the function is analytic in any finite region of the complex plane. Taking $n=10$, as in the real-axis graph above, and plotting the contours of $Re\, t^n e^{-t}$ in the neighborhood of $t=10$, we find:



It is clear that the integral along the real axis is in fact a steepest descent path. The reason we look at this straightforward case is to gain some experience about when it is reasonable to throw away all the contribution to the integral except that near the saddlepoint. If we simply take

$$f(t)=f(n)-\frac{1}{2n}(t-n)^2=n\ln n-n-\frac{1}{2n}(t-n)^2 \tag{2.4.24}$$

and take the t integration to be over the whole real axis, not just positive t, it is a Gaussian integral and

$$n!=\int_0^\infty e^{f(t)}dt\cong e^{f(n)}\int_0^\infty e^{-(1/2n)(t-n)^2}dt=\sqrt{2\pi n}n^n e^{-n}. \tag{2.4.25}$$

More precise, and considerably more complicated, methods give the leading correction to this expression. It is down by a factor of $1/12n$, so the naïve Gaussian saddlepoint result is accurate within 1% for $n=10$, and improves as n increases.

The Delta Function

Recall that the delta function can be defined by the limit of a Gaussian integral

$$\delta(x)=\lim_{\Delta\to 0}\frac{1}{(4\pi\Delta^2)^{1/2}}e^{-x^2/4\Delta^2}. \tag{2.4.26}$$

It is easy to see how this leads to

$$\int f(x)\delta(x)dx=f(0) \tag{2.4.27}$$

for an integral along the real axis with a function $f(x)$ reasonably well-behaved near the origin. Shankar mentions that the definition also works even if Δ^2 is replaced by $i\Delta^2$. In that case, the absolute value of the function is the same everywhere on the real axis, and increases as Δ^{-1} on taking Δ small. The reason it still works is that the phase oscillations are so rapid everywhere except at the origin, where the phase is momentarily stationary, so all the contribution comes from there.

However, it is easier to believe

$$\delta(x)=\lim_{\Delta\to 0}\frac{1}{(4\pi i\Delta^2)^{1/2}}e^{-x^2/4i\Delta^2} \tag{2.4.28}$$

on going into the complex plane. If we change variables from x to x, where $x^2=ix^2$, the integral again becomes a simple real Gaussian. But, regarding x as a complex variable, transforming to x is just equivalent to rotating the axes by $\pi/4$, or multiplying by the square root of i. The steepest descent route through the origin is now along the line at $\pi/4$ to the real axis. So this is a perfectly good definition of the δ-function provided we can distort the path of integration from the real axis to the line $x=y$. (Strictly speaking, the path would now include two octants of a circle at very large R—their contribution vanishes in the limit of R going to infinity.)

CHAPTER OVERVIEW

3: Mostly 1-D Quantum Mechanics

Topic hierarchy

3.1: 1-D Schrödinger Equation - Example Systems
3.2: General Uncertainty Principal
3.3: Energy-Time Uncertainty Principle
3.4: The Simple Harmonic Oscillator
3.5: Propagators and Representations
3.6: Coherent States
3.7: Path Integrals
3.8: Path Integrals for the SHO
3.9: Appendix- Some Exponential Operator Algebra

Thumbnail: A particle in a 1D infinite potential well of dimension L.

3.1: 1-D Schrödinger Equation - Example Systems

Curvature of Wavefunctions

Schrödinger's equation in the form

$$\frac{d^2\psi(x)}{dx^2} = \frac{2m(V(x)-E)}{\hbar^2}\psi(x) \tag{3.1.1}$$

can be interpreted by saying that the left-hand side, the rate of change of slope, is the *curvature* – so the curvature of the function is proportional to $(V(x)-E)\psi(x)$. This means that if $E > V(x)$, for $\psi(x)$ positive $\psi(x)$ is curving negatively, for $\psi(x)$ negative $\psi(x)$ is curving positively. *In both cases,* $\psi(x)$ *is always curving towards the x-axis* -- so, for $E > V(x)$, $\psi(x)$ has a kind of stability: its curvature is always bringing it back towards the axis, and so generating oscillations. The simplest example is that of a constant potential $V(x) = V_0 < E$, for which the wavefunction is

$$\psi(x) = A\sin(kx+\delta) \tag{3.1.2}$$

with δ a constant and

$$k = \sqrt{(2m/\hbar^2)(E-V_0)}. \tag{3.1.3}$$

On the other hand, for $V(x) > E$, the curvature is always *away* from the axis. This means that $\psi(x)$ tends to diverge to infinity. Only with *exactly* the right initial conditions will the curvature be just right to bring the wavefunction to zero as x goes to infinity. (This is possible because as $\psi(x)$ tends to zero, the curvature tends to zero, too.)

For a constant potential $V_0 > E$, the wavefunction is

$$\psi(x) = Ae^{\alpha x} + Be^{-\alpha x} \tag{3.1.4}$$

with

$$\alpha = \sqrt{(2m/\hbar^2)(V_0-E)}. \tag{3.1.5}$$

Of course, this wavefunction will diverge in at least one direction! However, as we shall see below, there are situations with spatially varying potentials where this wavefunction is only relevant for positive x, and the coefficients A, B are functions of the energy -- for certain energies it turns out that $A = 0$, and the wavefunction converges.

One Dimensional Infinite Depth Square Well

In an earlier lecture, we considered in some detail the allowed wavefunctions and energies for a particle trapped in an infinitely deep square well, that is, between infinitely high walls a distance L apart. For that case, the potential between the walls is identically zero so the wavefunction has the form (Equation 3.1.2). The wavefunction $\psi(x)$ necessarily goes to zero right at the walls, since it cannot have a discontinuity, and must be zero just inside the wall. Even a quantum particle cannot penetrate an infinite wall!

An immediate consequence is that the lowest state cannot have zero energy, since $k = 0$ gives a constant $\psi(x)$. Rather, the lowest energy state must have the minimal amount of bending of the wavefunction necessary for it to be zero at *both* walls but nonzero in between -- this corresponds to half a period of a sine or cosine (depending on the choice of origin), these functions being the solutions of Schrödinger's equation in the zero potential region between the walls. The allowed wavefunctions (eigenstates) found as the energy increases have successively 0, 1, 2, … zeros (nodes) in the well.

Parity of a wavefunction

Notice that the allowed wave eigenfunctions of the Hamiltonian for the infinite well are symmetrical or antisymmetrical about the center:

$$\psi(x) = \pm\psi(-x). \tag{3.1.6}$$

We call the operator that reflects a function in the origin the *parity* operator P, so these eigenstates of the Hamiltonian are *also* eigenstates of the parity operator, with eigenvalues ± 1. This is because the Hamiltonian is itself symmetric: d^2/dx^2 is even in

> x, and so is $V(x)$, so $[H, P] = 0$, and the two operators can be simultaneously diagonalized, that is, a **common set of eigenstates** can be constructed.

Finite Depth Square Well

If the potential at the walls is not infinite, the parity operator P will continue to commute with the Hamiltonian H as long as the potential is symmetric, $V(x) = V(-x)$.

We take

$$\begin{aligned} V(x) &= V_0, & x &< -L/2, \\ V(x) &= 0, & -L/2 \le x &\le L/2, \\ V(x) &= V_0, & L/2 &< x. \end{aligned} \tag{3.1.7}$$

We only need look for solutions symmetric or antisymmetric about the origin. This is important from a practical point of view, because it allows us to integrate Schrödinger's equation numerically out from the origin in the positive direction: $\psi(x)$ in the negative direction is fixed by symmetry (or antisymmetry). Since it's a second-order equation, we need two boundary conditions to get going, for symmetric states, we can take at x=0, for antisymmetric states, $\psi(x) = 0$, $\psi'(x) = 1$ at $x = 0$. (Of course, we will have to normalize $\psi(x)$ correctly eventually.)

The numerical strategy is to pick a value for the energy E, choose one of the boundary conditions above and integrate $\psi(x)$ numerically to a large positive value of x. For almost all values of E, the wavefunction will be exponentially increasing with x. For the particular values corresponding to bound states, it will be exponentially decreasing.

It is well worth while building up an intuition for this by playing with the spreadsheet accompanying this lecture: the spreadsheet does the numerical integration for any E and well depth, and has a macro to locate the nearest bound state.

Joining the wavefunctions Inside and Outside of the Well

The numerical method mentioned above works for any symmetric potential. Fortunately, for the square well, an analytic/graphical method is very effective, and provides more insight.

Let us begin by considering how the lowest energy state wavefunction is affected by having finite instead of infinite walls. Inside the well, where $V = 0$, the solution to Schrödinger's equation is still of cosine form (for a symmetric state). However, Schrödinger's equation now has a nonzero solution *inside* the wall $(x > L/2)$, where $V = V_0$:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{d\,x^2} + V_0\psi(x) = E\psi(x), \tag{3.1.8}$$

has two exponential solutions one increasing with x, the other decreasing,

$$e^{\alpha x} \; and \; e^{-\alpha x} \tag{3.1.9}$$

where $\alpha = \sqrt{2m(V_0 - E)/\hbar^2}$. We are assuming here that $E < V_0$, so the particle is bound to the well. We shall find the lowest energy state is always bound in a finite square well, however weak the potential.

Now, Schrödinger's equation must be valid everywhere, including the point $x = L/2$. Since the potential is finite, *the wavefunction $\psi(x)$ and its first derivative must be continuous at $x = L/2$.*

Suppose, then, we choose a particular energy E. Then the wavefunction inside the well (taking the symmetric case) is proportional to coskx, where $k = \sqrt{2mE/\hbar^2}$. The wavefunction (*and* its derivative!) must match a sum of exponential terms at $x = L/2$, so

$$\begin{aligned} \cos(kL/2) &= Ae^{\alpha L/2} + Be^{-\alpha L/2} \\ -k\sin(kL/2) &= \alpha Ae^{\alpha L/2} - \alpha Be^{-\alpha L/2} \end{aligned} \tag{3.1.10}$$

Solving these equations for the coefficients A, B in the usual way, we find that in general the cosine solution inside the well goes smoothly into a linear combination of exponentially increasing and decreasing terms in the wall. *However, this cannot in general represent a bound state in the well*. The increasing solution increases *without limit* as x goes to infinity, so since the square of the wavefunction is proportional to the probability of finding the particle at any point, the particle is infinitely more likely to be found at infinity than anywhere else. It got away! This clearly makes no sense -- we're trying to find wavefunctions for particles that stay in, or at least close to, the well. We are forced to conclude that the *only* exponential wavefunction that makes sense is the one for which A *is exactly zero*, so that there is only a *decreasing* wave in the wall.

Finding the Bound State Energies Graphically

If we demand that the wavefunction decrease exponentially as x goes to infinity, or, in other words, *require A to be zero*, k must satisfy the condition given be dividing one of the boundary equations above by the other:

$$\tan(kL/2) = \alpha/k. \tag{3.1.11}$$

This equation cannot be solved analytically, but is easy to solve graphically by plotting the two sides as functions of k (recall $\alpha = \sqrt{2m(V_0 - E)/\hbar^2}$, and $k = \sqrt{2mE/\hbar^2}$) and finding where the curves intersect.

From

$$\alpha^2 = \frac{2mV_0}{\hbar^2} - k^2 \;\; we\ find \;\; \frac{\alpha}{k} = \sqrt{\frac{2mV_0}{\hbar^2 k^2} - 1} \tag{3.1.12}$$

note that this is real only for

$$k < \frac{\sqrt{2mV_0}}{\hbar}. \tag{3.1.13}$$

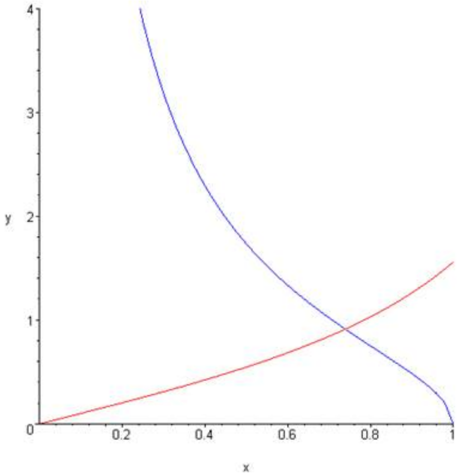
(Because if this inequality is *not* satisfied, the particle has enough kinetic energy to get out of the well!)

Now the condition for a bound state can be written

$$\tan(kL/2) = \frac{\alpha}{k} = \sqrt{\frac{2mV_0}{\hbar^2 k^2} - 1} \tag{3.1.14}$$

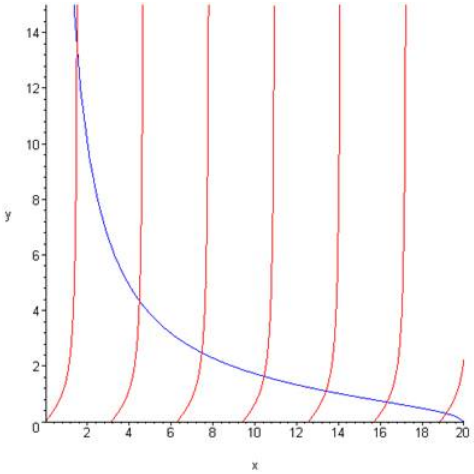
Cleaning up the appearance of the equation by choosing variables $x = \frac{kL}{2}$, $a^2 = \frac{mV_0L^2}{2\hbar^2}$, and plotting $\tan x$ and $\sqrt{(\frac{a}{x})^2 - 1}$ from $x = 0$ to $x = a$ allowed bound state k-values correspond to the points of intersection of the two curves. The bound state energies are then given by $E = \hbar^2 k^2/2m$.

The variable a is a measure of the attractive strength of the well. Here are the two curves for a shallow well ($a = 1$):



It is interesting to note that however small a is, the curve $y = \sqrt{(\frac{a}{x})^2 - 1}$ goes to infinity as x goes to zero, so will always intersect $y = \tan x$: there will always be a bound state.

A deeper well, $a = 20$, gives several bound states:



For the lower energies at least, the allowed k–values are approximately linearly spaced, at about $\pi/2, 3\pi/2, 5\pi/2, \ldots$ so the bound state energies are not far off the 1, 9, 25,… pattern of the infinitely deep square well -- remembering that we are only looking at the *even* parity (cosine) solutions!

Exercise 3.1.1

Use the spreadsheet with *D* = 50, *W* = 4 and find all the even bound state energies. How well do they fit this pattern? Can you account for the deviation? Examine the wavefunctions for the different eigenenergies: note how far it penetrates the wall, and how much that changes the boundary condition at the wall from that for an infinite wall. Which one of the bound state energies is most affected by this, and how is it affected? Would you expect that from the graphical solution?

The odd parity solutions, sine waves inside the well, can be found by an exactly similar analysis. One difference is that an arbitrarily weak well will not bind an odd parity state. The point is that for a weak potential to bind an even state, it only has to curve the wavefunction slightly to get from one exponentially decaying to the left to one exponentially decaying to the right. These curves decay very slowly for a weak potential, and give a bound state in which the particle is most likely to be found outside the well. On the other hand, in an odd solution the wavefunction within the well has to have enough total curvature to fit together two decaying wavefunctions which have opposite sign. This takes much more bending, and cannot be achieved with a very weak potential.

Exercise 3.1.2

Check this last statement, by considering what fraction of a wavelength of the oscillating wavefunction inside the well is necessary to make a connection between the decaying wavefunctions in the walls to the left and right.

The Delta Function Potential

One limiting case of a square well is a very narrow deep well, which can be approximated by a delta function when the range of variation of the wavefunction is much greater than the range of the potential, so Schrödinger’s equation becomes

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}+\lambda\delta(x)\psi(x)=E\psi(x) \tag{3.1.15}$$

with λ negative for an attractive potential.

The infinity of the δ-function cannot be balanced by the finite right hand side, so the wavefunction must have a discontinuity in slope at the origin.

To find the ground state energy, note first that as a one-dimensional attractive potential there will be a bound state: any change in slope is sufficient to connect an exponentially increasing function coming in from $-\infty$ to a decreasing one going to $+\infty$, since the rates of increase and decrease can be arbitrarily slow.

Away from the origin, then, we can take the wavefunction to be

$$\psi(x) = Ae^{-\alpha|x|}, \tag{3.1.16}$$

the energy of the state being $E = -\hbar^2\alpha^2/2m$.

The discontinuity in slope at the origin is just

$$\frac{d\psi(x=+\varepsilon)}{dx} - \frac{d\psi(x=-\varepsilon)}{dx} = -2A\alpha. \tag{3.1.17}$$

To match this with the δ-function singularity, we integrate the Schrödinger equation term by term from $-\varepsilon$ to $+\varepsilon$ in the limit of ε going to zero:

$$\int_{-\varepsilon}^{\varepsilon} -\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}dx + \int_{-\varepsilon}^{\varepsilon} \lambda\delta(x)\psi(x)dx = \int_{-\varepsilon}^{\varepsilon} E\psi(x)dx. \tag{3.1.18}$$

Note first that the right-hand side, having a finite integrand, must go to zero in the limit of ε going to zero.

The δ-function term must integrate to $\lambda\psi(0) = \lambda A$.

The first term just gives the discontinuity in slope,

$$\int_{-\varepsilon}^{\varepsilon} -\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}dx = -\frac{\hbar^2}{2m}\int_{-\varepsilon}^{\varepsilon}\frac{d}{dx}\left(\frac{d\psi}{dx}\right)dx = -\frac{\hbar^2}{2m}\left[\frac{d\psi}{dx}\right]_{-\varepsilon}^{\varepsilon} = \frac{\hbar^2}{2m}2A\alpha. \tag{3.1.19}$$

Schrödinger's equation is therefore satisfied if $\hbar^2\alpha/m + \lambda = 0$ (remembering λ is negative for an attractive potential).

The energy of the bound state is

$$E = -\frac{\hbar^2\alpha^2}{2m} = -\frac{m\lambda^2}{2\hbar^2}. \tag{3.1.20}$$

Exercise 3.1.1

Rederive this result by taking the limit of a narrow deep well, tending to a δ-function, with a cosine wavefunction inside.

A Potential Step

Our analysis so far has been limited to real-valued solutions of the time-independent Schrödinger equation. This is fine for analyzing bound states in a potential, or standing waves in general, but cannot be used, for example, to represent an electron traveling through space after being emitted by an electron gun, such as in an old fashioned TV tube. The reason is that a *real*-valued wavefunction $\psi(x)$, in an energetically allowed region, is made up of terms locally like $\cos kx$ and $\sin kx$, multiplied in the full wavefunction by the time dependent phase factor $e^{-iEt/\hbar}$, giving equal amplitudes of right moving waves $e^{i(px-Et)/\hbar}$ and left moving waves $e^{-i(px+Et)/\hbar}$. So, for an electron definitely moving to the right, even the time-independent part of the wavefunction must necessarily be complex.

Consider an electron of energy E moving in one dimension through a region of zero potential from large negative x and encountering an upward step potential of height V_0 ($V_0 < E$) at the origin $x = 0$. Of course, strictly speaking, the electron should be represented by a wave packet, and hence could not have a precisely defined energy E, but we assume here that it is a very long wave packet, very close to a plane wave, so we take it that the wavefunction is:

$$\psi(x,t) = Ae^{ikx}e^{-iEt/\hbar} \quad for \quad x < 0 \tag{3.1.21}$$

(A more precise analysis, in which an incoming wave packet is used, can be done by solving for the plane-wave components individually. In the limit of a wavepacket long compared to the de Broglie wavelength, the result is the same as that found here.)

Visualizing the classical picture of a particle approaching a hill (smoothing off the corners a bit) that it definitely has enough energy to surmount, we would perhaps expect that the wavefunction continues beyond $x = 0$ in the form

$$\psi(x,t) = Be^{ik_1x}e^{-iEt/\hbar} \quad for \quad x > 0, \tag{3.1.22}$$

where k_1 corresponds to the slower speed the particle will have after climbing the hill.

Schrödinger's equation requires that the wavefunction have no discontinuities and no kinks (discontinuities in slope) so the $x<0$ and $x>0$ wavefunctions must match smoothly at the origin. For them to have the same value, we see from above that $A=B$. For them to have the same slope we must have $kA=k_1B$. Unfortunately, the only way to satisfy both these equations with our above wavefunctions is to take $k=k_1$ -- which means there is no step potential at all!

Example 3.1.1

What is wrong with the above reasoning?

Solution?

The answer is that we have been led astray by the depiction of the particles as little balls rolling along in a potential, with enough energy to get up the hill, etc. Schrödinger's equation is a *wave equation*. Building intuition about solutions should rely on experience with waves. We should be thinking about a light wave going from air into glass, for example. If we do, we realize that at *any* interface *some of the light gets reflected*. This means that our expression for the wavefunction for $x<0$ is incomplete, we need to add a *reflected* wave, giving

$$\begin{aligned}\psi(x,t)=Ae^{ikx}e^{-iEt/\hbar}+Ce^{-ikx}e^{-iEt/\hbar} \quad & for \quad x<0,\\ \psi(x,t)=Be^{ik_1x}e^{-iEt/\hbar} \quad & for \quad x>0.\end{aligned} \tag{3.1.23}$$

Now matching the wavefunction and its derivative at the origin,

$$\begin{gathered}A+C=B\\ k(A-C)=k_1B.\end{gathered} \tag{3.1.24}$$

The fraction of the wave that is reflected

$$R=\frac{C^2}{A^2}=\left(\frac{k-k_1}{k+k_1}\right)^2. \tag{3.1.25}$$

Evidently, the fraction of the wave transmitted

$$T=1-R=\frac{4kk_1}{(k+k_1)^2}. \tag{3.1.26}$$

Example 3.1.2

Isn't the amount transmitted just given by B^2/A^2?

Solution?

The answer is no. The ratio B^2/A^2 gives the relative probability of finding a particle in some small region in the transmitted stream relative to that in the incoming stream, but the particles in the transmitted stream are moving more slowly, by a factor k_1/k. That means that just comparing the densities of particles in the transmitted and incoming streams is not enough. The physically significant quantity is the *probability current* flowing past a given point, and this is the product of the density *and* the speed. Therefore, the transmission coefficient is B^2k_1/A^2k.

Exercise 3.1.1

Prove that even a step *down* gives rise to some reflection.

Tunneling through a Square Barrier

If a plane wave coming in from the left encounters a step at the origin of height $V_0>E$, the incoming energy, there will be total reflection, but with an exponentially decaying wave penetrating some distance into the step. This, by the way, is a general wave phenomenon, not confined to quantum mechanics. If a light wave traveling through a piece of glass is totally internally at the surface, there will be an exponentially decaying electromagnetic field in the air outside the surface. If another piece of glass with a parallel (flat) surface is brought close, some light will "tunnel through" the air gap into the second piece of glass. We are considering here the quantum analogue of this classical behavior.

Suppose then we replace the step with a barrier,

$$\begin{array}{cccc} V=0 & for & x<0, & call\ this\ region\ I \\ V=V_0 & for & 0<x<L, & this\ is\ region\ II \\ V=0 & for & L<x, & region\ III. \end{array} \tag{3.1.27}$$

In this situation, the wavefunction will still decay exponentially into the barrier (assuming the barrier is thick compared to the exponential decay length), but on reaching the far end at $x=L$, a plane wave solution is again allowed, so there is a nonzero probability of finding the particle beyond the barrier, moving with its original speed. This phenomenon is called *tunneling*, since in the classical (particle) picture the particle doesn't have enough energy to get over the top of the barrier.

The way to solve the problem is to solve the Schrödinger equation in the three regions, then apply the boundary conditions. Since we are interested in the probability of a particle getting through the barrier, we do not need to worry about normalizing the wavefunction, so for simplicity we take an incoming wave of unit amplitude. In region *I*, there will also be a reflected wave, so

$$\psi_I(x)=e^{ikx}+Be^{-ikx}. \tag{3.1.28}$$

In region II, there will in general be both exponentially decreasing and exponentially growing solutions, so we take

$$\psi_{II}(x)=Ce^{\alpha x}+De^{-\alpha x}. \tag{3.1.29}$$

Recall $\alpha=\sqrt{(2m/\hbar^2)(V_0-E)}$

In region III, there is only the outgoing wave, to make the equations easy we absorb a phase factor in the coefficient, and write:

$$\psi_{III}(x)=Se^{ik(x-L)}. \tag{3.1.30}$$

We now require $\psi(x)$ and $\psi'(x)$ be continuous at $x=0,\ L$. Elementary computations lead to

$$\begin{aligned} 2ik &= (ik+\alpha)C+(ik-\alpha)D \\ 2\alpha Ce^{\alpha L} &= (ik+\alpha)S \\ 2\alpha De^{-\alpha L} &= -(ik-\alpha)S. \end{aligned} \tag{3.1.31}$$

Solving these equations gives

$$S(k)=\frac{2k\alpha}{2k\alpha\cosh\alpha L+i(\alpha^2-k^2)\sinh\alpha L}. \tag{3.1.32}$$

The probability of tunneling is:

$$|S(k)|^2=\frac{4k^2\alpha^2}{(k^2+\alpha^2)^2\sinh^2\alpha L+4k^2\alpha^2}. \tag{3.1.33}$$

An important limit is that of a barrier thick compared with the decay length, $\alpha L\gg 1$.

In this limit, $\sinh^2\alpha L\approx\frac{1}{4}e^{2\alpha L}$, and using $\alpha=\sqrt{(2m/\hbar^2)(V_0-E)},\quad E=\hbar^2k^2/2m$, we find

$$|S(k)|^2\cong 16\left(\frac{E}{V_0}\right)\left(1-\frac{E}{V_0}\right)e^{-2\alpha L}. \tag{3.1.34}$$

In typical tunneling problems, the far and away dominant term is the $e^{-2\alpha L}$, which may differ from unity by many orders of magnitude.

The Spherically Symmetric Three-Dimensional Problem

The methods developed above for the one-dimensional system are almost immediately applicable to a very important three-dimensional case: a particle in a spherically symmetric potential. A more detailed treatment will be given later -- we restrict ourselves here to *spherically symmetric* solutions of Schrödinger's equation $\psi(r,\theta,\phi)=\psi(r)$, a subspace of the space of all possible solutions that always includes the ground state.

The kinetic energy operator on states in this subspace (where $d/d\theta\equiv 0,\ d/d\phi\equiv 0$) is

$$\frac{\vec{p}^2}{2m}=-\frac{\hbar^2\vec{\nabla}^2}{2m}=-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{d}{dr}(r^2\frac{d}{dr}). \tag{3.1.35}$$

It is easy to check that if we write the wavefunction

$$\psi(r)=\frac{u(r)}{r} \tag{3.1.36}$$

the function $u(r)$ obeys the one-dimensional equation

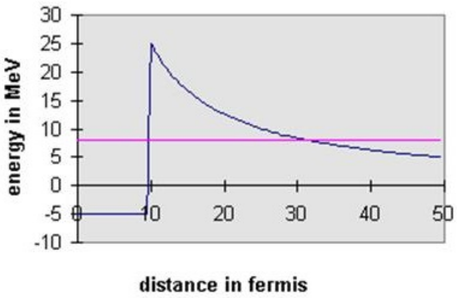
$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2}+V(r)u(r)=Eu(r) \tag{3.1.37}$$

exactly like a particle in one dimension, except that here r is only positive, and $u(r)$ must go to zero at the origin. (If $u(r)$ does *not* go to zero, $\psi(r)$ will be at best of order $1/r$ near the origin, and, going back momentarily to three dimensions, $\vec{\nabla}^2\psi\sim\delta(\vec{r})$, so Schrödinger's equation will not be satisfied with any realizable potential.)

Exercise: for a spherical square well, $V=0$ if $r<r_0$, $V=V_0$ for $r\geq r_0$, find the minimum value of V_0 for which a bound state exists for given r_0 and particle mass m. Sketch the wavefunction.

Alpha Decay

A good example of tunneling, and one which historically helped establish the validity of quantum ideas at the nuclear level, is α-decay. Certain large unstable nuclei decay radioactively by emitting an α-particle, a tightly bound state of two protons and two neutrons. It is thought that α-particles may exist, at least as long lived resonances, inside the nucleus. For such a particle, the strong but short ranged nuclear force creates a spherical finite depth well having a steep wall more or less coinciding with the surface of the nucleus. However, we must also include the electrostatic repulsion between the α-particle and the rest of the nucleus, a potential $(1/4\pi\varepsilon_0)(Z-2)2e^2/r$ outside the nucleus. This means that, as seen from inside the nucleus, the wall at the surface may not be a step but a barrier, in the sense we used the word above, a step up followed by a slide down the electrostatic curve:



Therefore, an α-particle bouncing around inside the nucleus may have enough energy to tunnel through to the outside world.

The α-particles are emitted with spherical symmetry, so the wavefunction can be written $\psi(r)=u(r)/r$, as discussed above, and Schrödinger's equation is

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2}+V(r)u(r)=Eu(r),\quad r>0 \tag{3.1.38}$$

It is evident that the more energetic the α-particle is, the thinner the barrier it faces. Since the wavefunction decays exponentially in the barrier, this can make a huge difference in tunneling rates. It is not difficult to find the energy with which the α-particle hits the nuclear wall, because this will be the same energy with which it escapes. Therefore, if we measure the energy of an emitted α, since we think we know the shape of the barrier pretty well, we should be able, at least numerically, to predict the tunneling rate. The only other thing we need to know is how many times per second α's bounce off the wall. The size of the nucleus is of order 10^{-14} meters (10 fermis), if we assume an α moves at, say, 10^7 meters per second, it will bang into the wall 10^{21} times per second. This is a bit handwaving, but all α-radioactive nuclei are pretty much the same size, so perhaps it's safe to assume this will be about the same for all of them.

This is a huge number -- the probability of transmission is evidently very tiny! In other words, the decay length of the wavefunction inside the barrier is extremely short (except for the very last bit as it emerges into the outside world). It's so short, in fact, that we can get results in good agreement with experiment by dividing the barrier into a sequence of square barriers and using the above $\alpha L \gg 1$ formula for each of them,

$$\psi(r+\Delta r) \approx \psi(r) e^{-\alpha(r)\delta r} \tag{3.1.39}$$

(neglecting the slowly varying and completely unimportant $1/r$ term) so the total attenuation of the wavefunction in this approximation is the product of terms like this,

$$\begin{aligned} \psi(exit) = \psi(nucl) \prod_i e^{-\alpha(r_i)\Delta r} \\ \approx \psi(nucl) e^{-\int \alpha(r) dr} \\ = \psi(nucl) e^{-\int \sqrt{(2m/\hbar^2)V(r)-E}\, dr} \end{aligned} \tag{3.1.40}$$

Since we know the actual shape of the barrier, this integral can be done with elementary functions, and the result for the transmission rate is

$$|S(k)|^2 \approx A e^{-B/\sqrt{E}}. \tag{3.1.41}$$

Here $B = (\pi\sqrt{2m}/\hbar)(Z-2)2e^2$ is *not* an adjustable parameter: and plotting $\ln|S(k)|^2$ against $1/\sqrt{E}$ for Polonium^{212} (which emits α's with energy 8.95MeV, and lasts 3×10^{-7} seconds) Thorium^{232} (4.05MeV α's, 1.4×10^{10} years), and several intermediate lifetime nuclei gives a straight line plot with the correct slope within a few percent! These elements can all be understood in terms of essentially the same barrier being tunneled through at the different heights corresponding to the α energy.

The treatment here is a slightly simplified version of the WKB method, to be discussed in detail later. Further refinements make little difference to the final result in this case.

Source: many of the topics covered in this lecture are elementary, and treated in any quantum textbook. For some of them I've followed (more or less) the excellent book by French and Taylor, *An Introduction to Quantum Physics*, Norton, 1978.

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

3.2: General Uncertainty Principal

Uncertainty and Non-Commutation

As we discussed in the Linear Algebra lecture, if two physical variables correspond to *commuting* Hermitian operators, they can be diagonalized simultaneously -- that is, they have a common set of eigenstates. In these eigenstates *both* variables have precise values at the same time, there is no “Uncertainty Principle” requiring that as we know one of them more accurately, we increasingly lose track of the other. For example, the energy and momentum of a free particle can both be specified exactly. More interesting examples will appear in the sections on angular momentum and spin.

But if two operators do *not* commute, in general one *cannot* specify both values precisely. Of course, such operators could still have *some* common eigenvectors, but the interesting case arises in attempting to measure A and B simultaneously for a state $|\psi\rangle$ in which the commutator $[A,B]$ has a nonzero expectation value, $\langle\psi|[A,B]|\psi\rangle\neq 0$.

A Quantitative Measure of “Uncertainty”

Our task here is to give a *quantitative* analysis of how accurately noncommuting variables can be measured together. We found earlier using a semi-quantitative argument that for a free particle, $\Delta p\cdot\Delta x\sim\hbar$ at best. To improve on that result, we need to be precise about the uncertainty $\Delta\, A$ in a state $|\psi\rangle$.

We *define* $\Delta\, A$ as the root mean square deviation:

$$(\Delta\, A)^2=\langle\psi|(A-\langle A\rangle)^2|\psi\rangle \tag{3.2.1}$$

where

$$\langle A\rangle=\langle\psi|A|\psi\rangle. \tag{3.2.2}$$

To make the equations more compact, we define $\hat{a}$ by

$$A=\langle A\rangle+\hat{a}. \tag{3.2.3}$$

(We’ll put a caret (a hat) on the $\hat{a}$ to remind ourselves it’s an operator -- and, of course, it’s a *Hermitian* operator, like A.) We also drop the ψ bra and ket, on the understanding that this whole argument is for a particular state. Now

$$(\Delta\, A)^2=\langle(A-\langle A\rangle)^2\rangle=\langle\hat{a}^2\rangle. \tag{3.2.4}$$

Introduce an operator B in exactly similar fashion, $B=\langle B\rangle+\hat{b}$, having the property that $\langle\psi|[A,B]|\psi\rangle\neq 0$.

The Generalized Uncertainty Principle

The quantitative measure of how the combined “uncertainty” of measuring two variables relates to their lack of commutativity is most simply presented as a

Theorem

$$(\Delta\, A)^2(\Delta\, B)^2\ \geq\ \frac{1}{4}\langle i[A,B]\rangle^2. \tag{3.2.5}$$

Remember that for A,B to be Hermitian, then $[A,B]$ is *anti*-Hermitian: so $\langle i[A,B]\rangle$ is real! To make this clear, take adjoints: if H is Hermitian then

$$(\langle\psi|H|\psi\rangle)^*=\langle\psi|H^\dagger|\psi\rangle=\langle\psi|H|\psi\rangle \tag{3.2.6}$$

so $\langle\psi|H|\psi\rangle$ is real.

If K is *anti*-Hermitian, $K^\dagger=-K$, then

$$(\langle\psi|K|\psi\rangle)^*=\langle\psi|K^\dagger|\psi\rangle=-\langle\psi|K|\psi\rangle \tag{3.2.7}$$

from which $\langle\psi|K|\psi\rangle$ is *pure imaginary*.)

Proof of Theorem

Define

$$|\psi_a\rangle = \hat{a}|\psi\rangle, \quad |\psi_b\rangle = \hat{b}|\psi\rangle. \tag{3.2.8}$$

Then

$$(\Delta A)^2(\Delta B)^2 = \langle\psi|\hat{a}^2|\psi\rangle\langle\psi|\hat{b}^2|\psi\rangle = \langle\psi_a|\psi_a\rangle\langle\psi_b|\psi_b\rangle \tag{3.2.9}$$

Using Schwartz's inequality

$$\langle\psi_a|\psi_a\rangle\langle\psi_b|\psi_b\rangle \geq |\langle\psi_a|\psi_b\rangle|^2 \tag{3.2.10}$$

gives immediately

$$(\Delta A)^2(\Delta B)^2 \geq |\langle\psi_a|\psi_b\rangle|^2 = |\langle\psi|\hat{a}\hat{b}|\psi\rangle|^2. \tag{3.2.11}$$

The operator $\hat{a}\hat{b}$ is neither Hermitian nor antiHermitian. To evaluate the mod squared of its expectation value, we break the amplitude into real and imaginary parts:

$$\langle\psi|\hat{a}\hat{b}|\psi\rangle = \langle\psi|\frac{1}{2}(\hat{a}\hat{b}+\hat{b}\hat{a})|\psi\rangle + \langle\psi|\frac{1}{2}[\hat{a},\hat{b}]|\psi\rangle. \tag{3.2.12}$$

(The first term on the right-hand side is the expectation value of a Hermitian matrix, and so is real, the second term is the expectation value of an antiHermitian matrix, so is pure imaginary.)

It follows immediately that

$$|\langle\psi|\hat{a}\hat{b}|\psi\rangle|^2 \geq |\langle\psi|\frac{1}{2}[\hat{a},\hat{b}]|\psi\rangle|^2. \tag{3.2.13}$$

But $[A,B] = [\hat{a},\hat{b}]$, so

$$(\Delta\ A)^2(\Delta\ B)^2 \geq \frac{1}{4}\langle i[A,B]\rangle^2. \tag{3.2.14}$$

□

Minimizing the Uncertainty

For a particle in one dimension denote

$$A = x \tag{3.2.15}$$

$$B = p = -i\hbar\frac{d}{dx} \tag{3.2.16}$$

so

$$[A,B] = -i\hbar\left(x\frac{d}{dx} - \frac{x}{dx}x\right) = i\hbar. \tag{3.2.17}$$

(It important in that last step to understand that the operator $\frac{d}{dx}$ operates on everything to its right, and, as we are always finding matrix elements of operators, there will be a following ket it operates on, so $\frac{d}{dx}x = 1 + x\frac{d}{dx}$.)

We conclude that

$$(\Delta x)^2(\Delta p)^2 \geq \frac{1}{4}\hbar^2. \tag{3.2.18}$$

Example 3.2.1

Is there a wavefunction for which the inequality in Equation 3.2.18 becomes an equality?

Solution

That would require

$$|\langle\psi_a|\psi_b\rangle|^2 = \langle\psi_a|\psi_a\rangle\langle\psi_b|\psi_b\rangle \tag{3.2.19}$$

which can only be true if the two vectors are parallel,

$$|\psi_b\rangle = \lambda|\psi_a\rangle \tag{3.2.20}$$

or, written explicitly,

$$\left(-i\hbar\frac{d}{dx} - \langle p\rangle\right)\psi(x) = \lambda(x - \langle x\rangle)\psi(x). \tag{3.2.21}$$

Actually , that's not enough: we *also* need $\langle\psi|\frac{1}{2}(\hat{a}\hat{b} + \hat{b}\hat{a})|\psi\rangle$ to be zero. (Look at the equation above giving $\langle\psi|\hat{a}\hat{b}|\psi\rangle$ in terms of its real and imaginary parts, and how we used it to establish the inequality.)

Writing $|\psi_b\rangle = \lambda|\psi_a\rangle$ as $\hat{b}|\psi\rangle = \lambda\hat{a}|\psi\rangle$ and $\langle\psi|\hat{b} = \lambda^*\langle\psi|\hat{a}$ we find

$$\langle\psi|\frac{1}{2}(\hat{a}\hat{b} + \hat{b}\hat{a})|\psi\rangle = (\lambda + \lambda^*)\langle\psi|\hat{a}^2|\psi\rangle, \tag{3.2.22}$$

so this will be zero *if and only if* λ *is pure imaginary*.

Turning to the differential equation, we first take the simplest case where $\langle x\rangle$ and $\langle p\rangle$ are both zero. The first requirement just sets the origin, but the second stipulates that our wavefunction has no net momentum.

For this simple case, $|\psi_b\rangle = \lambda|\psi_a\rangle$ becomes

$$\begin{gathered} -i\hbar\frac{d\psi(x)}{dx} = \lambda x\psi(x) \\ \frac{d\psi}{\psi} = \frac{i\lambda}{\hbar}xdx \\ \psi = Ce^{i\lambda x^2/2\hbar} \end{gathered} \tag{3.2.23}$$

and recalling that λ is pure imaginary, this is a Gaussian wave packet! It is straightforward to check that the solution with $\langle x\rangle$ and $\langle p\rangle$ nonzero is

$$\psi(x) = Ce^{ix/\hbar}e^{-\alpha(x-<x>)^2/2\hbar} \tag{3.2.24}$$

where $\alpha = -i\lambda$ is real, and C is the usual Gaussian normalization constant.

Exercise 3.2.1

Confirm Equation 3.2.24.

The conclusion is then that the Gaussian wave packet gives the optimum case for minimizing the joint uncertainties in position and momentum.

Note that the condition $\hat{b}|\psi\rangle = \lambda\hat{a}|\psi\rangle$ does *not* mean that $|\psi\rangle$ is an eigenstate of either $\hat{a}$ or $\hat{b}$, but it *is* an eigenstate of the *non-Hermitian* operator $\hat{b} - \lambda\hat{a} = \hat{b} + i\alpha\hat{a}$, with eigenvalue zero. We shall soon see that this non-Hermitian operator and its adjoint play important roles in the quantum mechanics of the simple harmonic oscillator.

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

3.3: Energy-Time Uncertainty Principle

Model of a Decaying State

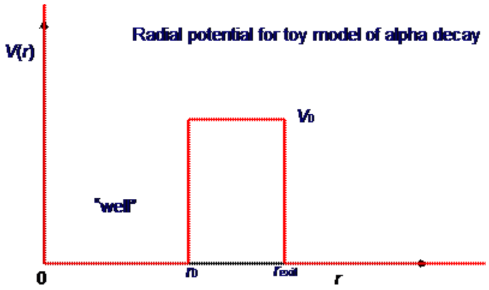
The momentum-position uncertainty principle

$$\Delta p \cdot \Delta x \geq \hbar \tag{3.3.1}$$

has an energy-time analog,

$$\Delta E \cdot \Delta t \geq \hbar \tag{3.3.2}$$

Evidently, though, this must be a different kind of relationship to the momentum-position one, because t is not a dynamical variable, so this cannot have anything to do with non-commutation. To illustrate the meaning of Equation 3.3.2, let us reconsider α -decay, but with a slightly simplified potential to clarify what's going on. Specifically, we replace the combined nuclear force/electrostatic repulsion barrier with a square barrier, high enough and thick enough that there is a small probability per unit time of the particle tunneling out of the well.



If the barrier thickness were increased to infinity (keeping r_0 fixed) there would be a true bound state having energy E_0, and for E_0 well below V_0, having approximately an integral number of half wavelengths in the well. For a barrier of finite thickness, there is of course some nonzero probability of the particle escaping -- so no longer a true bound state, but for a thick barrier the difference may be hard to detect.

As with the α-decay analysis, we'll look at this semi-classically, picturing the particle as bouncing off the walls backwards and forwards inside, time τ between hits, and at each hit probability of penetration some small number ε. Therefore, the probability that the particle is still in the well after a time $t = n\tau$ is

$$P(n\tau) = (1-\varepsilon)^n. \tag{3.3.3}$$

Since ε really is *very* small for α -decay (less than 10^{-12}), we can conveniently write P as a function of time by using the formula

$$e^{-x} = \lim_{\varepsilon \to 0} (1-\varepsilon)^{x/\varepsilon}. \tag{3.3.4}$$

From this, the probability of the particle being in the well after n bounces, time $t = n\tau$,

$$P(t = n\,\tau) = (1-\varepsilon)^n = (1-\varepsilon)^{\varepsilon n/\varepsilon} \cong e^{-\varepsilon n}. \tag{3.3.5}$$

The standard notation is to introduce a variable $\Gamma = \hbar\varepsilon/\tau$ having the dimensions of *energy* (ε being dimensionless), in terms of which

$$P(t) \cong e^{-n\varepsilon} = e^{-n\Gamma\tau/\hbar} = e^{-\Gamma t/\hbar}. \tag{3.3.6}$$

Evidently $\Gamma/\hbar$ measures the *decay rate*. In other words, $\Gamma/\hbar$ is an inverse lifetime.

The exponential decay law for radioactive elements is completely confirmed experimentally, it is the basis of the "half-life" rule: for any given amount of a radioactive nucleus, half of it will decay in a time period -- the half-life -- fixed for that nucleus.

Obviously, if the modulus of the wave function of the particle in the well is decreasing with time, the time dependence of ψ can no longer be just the $e^{-iE_0t/\hbar}$ of the original "bound state". The wave function inside the well will stay the same shape, but gradually decrease in amplitude:

$$|\psi(t)|2 = |\psi(0)|^2 e^{-\Gamma t/\hbar}. \tag{3.3.7}$$

This is a far slower time dependence than that of the $e^{-iE_0t/\hbar}$ term, so it is an excellent approximation to put the time dependences together in one exponential factor:

$$\psi(t)=\psi(0)e^{-i(E_0-i\Gamma/2)t/\hbar}. \tag{3.3.8}$$

At this point, the analogy with $\Delta p\cdot\Delta x\geq\hbar$ emerges. Recall we introduced the p, x uncertainty principle by finding what spread Δp in the Fourier components of a wave packet were necessary in order for the wave packet to die away from its center over a distance of order Δx. A true localized wave packet has a continuum of p components, but the right expression for the spread in momentum space Δp turns out to be given by taking just two waves, $e^{ipx/\hbar}$ and $e^{i(p+\Delta p)x/\hbar}$, and noticing that they fall out of phase in a distance $\Delta x\sim\hbar/\Delta p$. The more precise derivation based on Gaussian wave packets reaches essentially the same conclusion.

Now, in the present situation the wave function decays in *time* rather than space, but the argument is very similar. To construct the decaying wave function we must add together "plane waves in time" $e^{-iEt/\hbar}$ corresponding to different energies. The required spread in energy ΔE follows from an argument just like the one above for space: if the wave function is to be die away in a time of order Δt, in other words to be "localized in time", it must be constructed of waves having a range ΔE in energies such that $e^{-iEt/\hbar}$ and $e^{-i(E+\Delta E)t/\hbar}$ get out of sync in a time Δt. This gives immediately $\Delta E\cdot\Delta t\geq\hbar$.

It's worth looking a little further into just what superposition of energy "plane waves" $e^{-iEt/\hbar}$ gives the required exponential-in-time behavior of the wave function for α-decay. Writing

$$\psi(t)=\psi(0)\int c(E)e^{-iEt/\hbar}dE/2\pi \tag{3.3.9}$$

the Fourier coefficients are given by

$$c(E)=\int_0^\infty e^{i(E-E_0+i\Gamma/2)t/\hbar}dt=\frac{i\hbar}{E-E_0+i\Gamma/2}. \tag{3.3.10}$$

This tells us that the energy with which the α -particle emerges has a probability distribution which is easily normalized to give

$$|c(E)|^2=\frac{1}{\pi}\frac{\Gamma/2}{(E-E_0)^2+(\Gamma/2)^2}. \tag{3.3.11}$$

This distribution (called *Lorentzian*) has a narrow peak of width of order Γ and height of order $1/\Gamma$ centered at E_0. (Strictly speaking, this is an approximation in that $c(E)$ must of course be zero for E negative -- we ignore that tiny correction here.)

In fact, for the α-decaying nucleus, this energy spread is undetectably small, but that is certainly *not* the case for other decaying states, where this same analysis applies. In particular, some of the resonant states created in collisions of elementary particles have masses of order 1,000 Mev, and lifetimes of order 10^{-23} seconds -- corresponding to a width Γ of the order of 10% of E_0! Obviously, these transient bound states are far from eigenstates of the Hamiltonian— but you will find them listed in particle tables.

Resonances

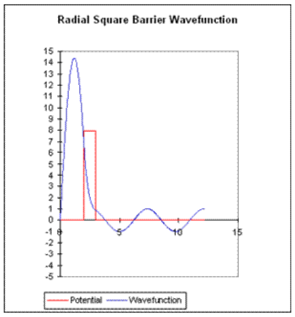
The *time-reversed* wave function for α -decay is also a perfectly good solution to Schrödinger's equation. In principle, if we could arrange for an α -particle to have a spherical ingoing wave function within the narrow energy range corresponding to the quasi bound state, ψ would become very large inside the nucleus, meaning that the α -particle would spend a very long time there compared with the time spent at any other point on the way in: this is called *resonance*. This particular experiment is never likely to take place, but precisely analogous experiments involving particles scattering off each other, and electrons scattering resonantly from atoms, are common.

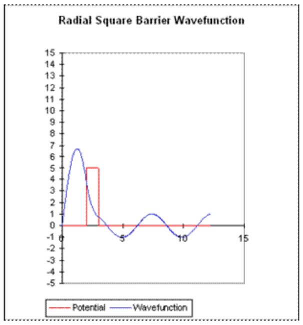
One can interpret the wave function for a *true* bound state as a standing wave, radially containing a whole number of half wavelengths, so that when the wave is reflected at the walls it has just the right phase to interfere constructively with itself. A resonance will have the same wavelength requirement (but the reflection at the walls is of course no longer perfect).

We shall see later that in the case of nonzero angular momentum, the centrifugal barrier provides an effective repulsive term, which together with an attractive force can create a barrier configuration having similar effect to that in the toy model.

It is instructive to use the spreadsheet to discover how the maximum value of $|\psi|^2$ varies as the barrier height is increased, and to vary the energy to explore how Γ varies with barrier height.

The pictures below are for the energy right at resonance for two different barrier heights. Evidently, if the height of the barrier is increased, $|\psi|^2$ inside at resonance increases as well. This is easy to understand -- the higher barrier is more difficult to penetrate, so the particle spends even longer inside. The lifetime Δt is clearly proportional to $|\psi|^2$ inside.





Exercise 3.3.1

Open the spreadsheet, find the resonant energy for these cases, and then detune the energy away from resonance. As the energy moves away from resonance, which ψ drops to half its maximum value first?

This energy spread of the resonance is the Γ above, and is related to the “lifetime” of the resonance by $\Delta E \cdot \Delta t \geq \hbar$. How does the ratio of the two Γ’s for these two barriers relate to the ratio of the two maximum values of $|\psi|^2$? Give a physical explanation of your findings.

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

3.4: The Simple Harmonic Oscillator

Einstein’s Solution of the Specific Heat Puzzle

The simple harmonic oscillator, a nonrelativistic particle in a potential $\frac{1}{2}kx^2$, is an excellent model for a wide range of systems in nature. In fact, not long after Planck’s discovery that the black body radiation spectrum could be explained by assuming energy to be exchanged in quanta, Einstein applied the same principle to the simple harmonic oscillator, thereby solving a long-standing puzzle in solid state physics—the mysterious *drop in specific heat* of all solids at low temperatures. Classical thermodynamics, a very successful theory in many ways, predicted no such drop -- with the standard equipartition of energy, kT in each mode (potential plus kinetic), the specific heat should remain more or less constant as the temperature was lowered (assuming no phase change).

To explain the anomalous low temperature behavior, Einstein assumed each atom to be an independent (quantum) simple harmonic oscillator, and, just as for black body radiation, he assumed the oscillators could only absorb or emit energy in *quanta*. Consequently, at low enough temperatures there is rarely sufficient energy in the ambient thermal excitations to excite the oscillators, and they freeze out, just as blue oscillators do in low temperature black body radiation. Einstein’s picture was later somewhat refined -- the basic set of oscillators was taken to be standing sound wave oscillations in the solid rather than individual atoms (making the picture even more like black body radiation in a cavity) but the main conclusion -- the drop off in specific heat at low temperatures -- was not affected.

The Classical Simple Harmonic Oscillator

The classical equation of motion for a one-dimensional simple harmonic oscillator with a particle of mass m attached to a spring having spring constant k is

$$m\frac{d^2x}{dt^2} = -kx. \tag{3.4.1}$$

The solution is

$$x = x_0\sin(\omega t+\delta), \quad \omega = \sqrt{\frac{k}{m}}, \tag{3.4.2}$$

and the momentum $p = mv$ has time dependence

$$p = mx_0\omega\cos(\omega t+\delta). \tag{3.4.3}$$

The total energy

$$(1/2m)(p^2+m^2\omega^2x^2) = E \tag{3.4.4}$$

is clearly constant in time.

It is often useful to picture the time-development of a system in ***phase space***, in this case a two-dimensional plot with position on the x -axis, momentum on the y -axis. Actually, to have (x, y) coordinates with the same dimensions, we use $(m\omega x, p)$.

It is evident from the above expression for the total energy that in these variables the point representing the system in phase space moves clockwise around a circle of radius $\sqrt{2mE}$ centered at the origin.

Note that in the *classical* problem we could choose any point $(m\omega x, p)$, place the system there and it would then move in a circle about the origin. In the *quantum* problem, on the other hand, we cannot specify the initial coordinates $(m\omega x, p)$ precisely, because of the uncertainly principle. The best we can do is to place the system initially in a small cell in phase space, of size $\Delta x\cdot\Delta p = \hbar/2$. In fact, we shall find that in quantum mechanics phase space is always divided into cells of essentially this size for each pair of variables.

Schrödinger’s Equation and the Ground State Wavefunction

From the classical expression for total energy given above, the Schrödinger equation for the quantum oscillator follows in standard fashion:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}+\frac{1}{2}m\omega^2x^2\psi(x)=E\psi(x). \tag{3.4.5}$$

What will the solutions to this Schrödinger equation look like? Since the potential $\frac{1}{2}m\omega^2x^2$ increases without limit on going away from $x=0$, it follows that no matter how much kinetic energy the particle has, for sufficiently large x the potential energy dominates, and the (bound state) wavefunction decays with increasing rapidity for further increase in x. (Obviously, for a real physical oscillator there is a limit on the height of the potential—we will assume that limit is much greater than the energies of interest in our problem.)

We know that when a particle penetrates a barrier of constant height V_0 (greater than the particle's kinetic energy) the wavefunction decreases exponentially into the barrier, as $e^{-\alpha x}$, where

$$\alpha=\sqrt{2m(V_0-E)/\hbar^2}. \tag{3.4.6}$$

In contrast to this constant height barrier, the "height" of the simple harmonic oscillator potential *continues to increase* as the particle penetrates to larger x. Obviously, in this situation the decay will be faster than exponential. If we (rather naïvely) assume it is more or less *locally* exponential, but with a local α varying with V_0, neglecting E relative to V_0 in the expression for α suggests that \alpha itself is proportional to x (since the potential is proportional to x^2, and $\alpha\propto\sqrt{V}$) so maybe the wavefunction decays as $e^{-(constant)x^2}$?

To check this idea, we insert $\psi(x)=e^{-\ x^2/2b^2}$ in the Schrödinger equation, using

$$\frac{d^2\psi}{dx^2}=-\frac{1}{b^2}\psi+\frac{x^2}{b^4}\psi \tag{3.4.7}$$

to find

$$-\frac{\hbar^2}{2m}\left(-\frac{1}{b^2}+\frac{x^2}{b^4}\right)\psi(x)+\frac{1}{2}m\omega^2x^2\psi(x)=E\psi(x). \tag{3.4.8}$$

The $\psi(x)$ is just a factor here, and it is never zero, so can be cancelled out. This leaves a quadratic expression which must have the same coefficients of x^0, x^2 on the two sides, that is, the coefficient of x^2 on the left hand side must be zero:

$$\frac{\hbar^2}{2mb^4}=\frac{m\omega^2}{2}, sob=\sqrt{\frac{\hbar}{m\omega}}. \tag{3.4.9}$$

This fixes the wavefunction. Equating the constant terms fixes the energy:

$$E=\frac{\hbar^2}{2mb^2}=\frac{1}{2}\hbar\omega. \tag{3.4.10}$$

So the conjectured form for the wavefunction is in fact the *exact* solution for the lowest energy state! (It's the lowest state because it has no nodes.)

Also note that even in this ground state the energy is *nonzero*, just as it was for the square well. The central part of the wavefunction must have some curvature to join together the decreasing wavefunction on the left to that on the right. This "zero point energy" is sufficient in one physical case to melt the lattice -- helium is liquid even down to absolute zero temperature (checked down to microkelvins!) because the wavefunction spread destabilizes the solid lattice that will form with sufficient external pressure.

Higher Energy States

It is clear from the above discussion of the ground state that $b=\sqrt{\frac{\hbar}{m\omega}}$ is the natural unit of length in this problem, and $\hbar\omega$ that of energy, so to investigate higher energy states we reformulate in dimensionless variables,

$$\xi=\frac{x}{b}=x\sqrt{\frac{m\omega}{\hbar}},\quad \varepsilon=\frac{E\hbar}{\omega}. \tag{3.4.11}$$

Schrödinger's equation becomes

$$\frac{d^2\psi(\xi)}{d\xi^2} = (\xi^2 - 2\varepsilon)\psi(\xi). \tag{3.4.12}$$

Deep in the barrier, the ε term will become negligible, and just as for the ground state wavefunction, higher bound state wavefunctions will have $e^{-\xi^2/2}$ behavior, multiplied by some more slowly varying factor (it turns out to be a polynomial).

Exercise: find the relative contributions to the second derivative from the two terms in $x^n e^{-x^2/2}$. For given n, when do the contributions involving the first term become small? Define “small”.

The standard approach to solving the general problem is to factor out the $e^{-\xi^2/2}$ term,

$$\psi(\xi) = h(\xi)e^{-\xi^2/2} \tag{3.4.13}$$

giving a differential equation for $h(\xi)$:

$$\frac{d^2h}{d\xi^2} - 2\xi\frac{dh}{d\xi} + (2\varepsilon - 1)h = 0 \tag{3.4.14}$$

We try solving this with a power series in ξ:

$$h(\xi) = h_0 + h_1\xi + h_2\xi^2 = \dots \tag{3.4.15}$$

Inserting this in the differential equation, and requiring that the coefficient of each power ξ^n vanish identically, leads to a recurrence formula for the coefficients h_n:

$$h_{n+2} = \frac{(2n+1-2\varepsilon)}{(n+1)(n+2)}h_n. \tag{3.4.16}$$

Evidently, the series of odd powers and that of even powers are independent solutions to Schrödinger’s equation. (Actually this isn’t surprising: the potential is even in x, so the parity operator P commutes with the Hamiltonian. Therefore, unless states are degenerate in energy, the wavefunctions will be even or odd in x.) For large n, the recurrence relation simplifies to

$$h_{n+2} \approx \frac{2}{n}h_n, \quad n \gg \varepsilon. \tag{3.4.17}$$

The series therefore tends to

$$\sum \frac{2n\xi^{2n}}{(2n-2)(2n-4)\dots 2} = 2\xi^2 \sum \frac{\xi^{2(n-1)}}{(n-1)!} = e^{\xi^2}. \tag{3.4.18}$$

Multiply this by the $e^{-\xi^2/2}$ factor to recover the full wavefunction, we find ψ diverges for large ξ as $e^{+\xi^2/2}$.

Actually we should have expected this -- for a general value of the energy, the Schrödinger equation has the solution $\approx Ae^{+\xi^2/2} + Be^{-\xi^2/2}$ at large distances, and only at certain energies does the coefficient A vanish to give a normalizable bound state wavefunction.

So how do we find the *nondiverging* solutions? It is clear that the infinite power series must be stopped! The key is in the recurrence relation.

If the energy satisfies

$$2\varepsilon = 2n+1, \quad n \textit{ an integer}, \tag{3.4.19}$$

then h_{n+2} *and all higher coefficients vanish.*

This requirement in fact *completely determines the polynomial* (except for an overall constant) because with $2\varepsilon = 2n+1$ the coefficients hm for $m < n$ are determined by

$$h_{m+2} = \frac{(2m+1-2\varepsilon)}{(m+1)(m+2)}h_m = \frac{(2m+1-(2n+1))}{(m+1)(m+2)}h_m. \tag{3.4.20}$$

This n^{th} order polynomial is called a *Hermite polynomial* and written $H_n(\xi)$. The standard normalization of the Hermite polynomials $H_n(\xi)$ is to take the coefficient of the highest power ξ^n to be 2^n. The other coefficients then follow using the recurrence relation above, giving:

$$H_0(\xi)=1,\ \ H_1(\xi)=2\xi,\ \ H_2(\xi)=4\xi^2-2,\ \ H_3(\xi)=8\xi^3-\frac{1}{2}\xi,\ \ etc. \tag{3.4.21}$$

So the bottom line is that the wavefunction for the nth excited state, having energy $\varepsilon=n+\frac{1}{2}$, is $\psi_n(\xi)=C_nH_n(\xi)e^{-\xi^2/2}$, where C_n is a normalization constant to be determined in the next section.

It can be shown (see exercises at the end of this lecture) that $H_n{}'(\xi)=2nH_{n-1}(\xi)$. Using this, beginning with the ground state, one can easily convince oneself that the successive energy eigenstates each have one more node -- the n^{th} state has n nodes. This is also evident from numerical solution using the spreadsheet, watching how the wavefunction behaves at large x as the energy is cranked up.

The spreadsheet can also be used to plot the wavefunction for large n, say $n=200$. It is instructive to compare the probability distribution with that for a *classical* pendulum, one oscillating with fixed amplitude and observed many times at random intervals. For the pendulum, the probability peaks at the end of the swing, where the pendulum is slowest and therefore spends most time. The $n=200$ distribution amplitude follows this pattern, but of course oscillates. However, in the large n limit these oscillations take place over undetectably small intervals.

The *classical* pendulum when not at rest clearly has a time-dependent probability distribution -- it swings backwards and forwards. This means it *cannot* be in an eigenstate of the energy. In fact, the quantum state most like the classical is a *coherent state* built up of neighboring energy eigenstates. We shall discuss coherent states later in the course.

Operator Approach to the Simple Harmonic Oscillator (Ladder Operators)

Having scaled the position coordinate x to the dimensionless ξ defined as

$$\xi=x/b=x\sqrt{m\omega/\hbar} \tag{3.4.22}$$

and let us also scale the momentum from p to $\pi=-i\ d/d\xi$, so

$$\pi=\frac{bp}{\hbar}=p/\sqrt{\hbar m\omega}. \tag{3.4.23}$$

The Hamiltonian is then

$$H=\frac{p^2+m^2\omega^2x^2}{2m}=\frac{\hbar\omega}{2}(\pi^2+\xi^2). \tag{3.4.24}$$

Dirac had the brilliant idea of factorizing this expression: the obvious thought $(\xi^2+\pi^2)=(\xi+i\pi)(\xi-i\pi)$ isn't quite right, because it fails to take account of the noncommutativity of the operators, but the symmetrical version

$$H=\frac{\hbar\omega}{4}[(\xi+i\pi)(\xi-i\pi)+(\xi-i\pi)(\xi+i\pi)] \tag{3.4.25}$$

is fine, and we shall soon see that it leads to a very easy way of finding the eigenvalues and operator matrix elements for the oscillator, far simpler than using the wavefunctions we found above. Interestingly, Dirac's factorization here of a second-order differential operator into a product of first-order operators is close to the idea that led to his most famous achievement, the Dirac equation, the basis of the relativistic theory of electrons, protons, etc.

To continue, we define new operators $a, a^\dagger$ by

$$a=\xi+i\pi2\surd=\frac{1}{\sqrt{2\hbar m\omega}}(m\omega x+ip),\ \ a^\dagger=\frac{\xi-i\pi}{\sqrt{2}}=\frac{1}{\sqrt{2\hbar m\omega}}(m\omega x-ip). \tag{3.4.26}$$

(We've expressed a in terms of the original variables x, p for later use.)

From the commutation relation $[i\pi,\xi]=1$ it follows that

$$[a,a^\dagger]=1. \tag{3.4.27}$$

Therefore the Hamiltonian can be written:

$$H=\hbar\omega(a^\dagger a+\frac{1}{2})=\hbar\omega(N+\frac{1}{2}),\ \ where\ \ N=a^\dagger a. \tag{3.4.28}$$

Note that the operator N can only have *non-negative* eigenvalues, since

$$\langle\psi|N|\psi\rangle = \langle\psi|a^\dagger a|\psi\rangle = \langle\psi_a|\psi_a\rangle \geq 0. \tag{3.4.29}$$

Now

$$[N, a^\dagger] = a^\dagger a a^\dagger - a^\dagger a^\dagger a \tag{3.4.30}$$

$$= a^\dagger[a, a^\dagger] \tag{3.4.31}$$

$$= a^\dagger \tag{3.4.32}$$

Suppose N has an eigenfunction $|\nu\rangle$ with eigenvalue ν,

$$N|\nu\rangle = \nu|\nu\rangle. \tag{3.4.33}$$

From the two equations above

$$Na^\dagger|\nu\rangle = a^\dagger N|\nu\rangle + a^\dagger|\nu\rangle = (\nu+1)a^\dagger|\nu\rangle \tag{3.4.34}$$

so $a^\dagger|\nu\rangle$ is an eigenfunction of N with eigenvalue $\nu+1$. Operating with $a^\dagger$ again and again, we climb an infinite ladder of eigenstates equally spaced in energy.

$a^\dagger$ is often termed a ***creation operator***, since the quantum of energy $\hbar\omega$ added each time it operates is equivalent to an added photon in black body radiation (electromagnetic oscillations in a cavity).

It is easy to check that the state $a|\nu\rangle$ is an eigenstate with eigenvalue $\nu-1$, provided it is nonzero, so the operator a takes us *down* the ladder. However, this cannot go on indefinitely -- we have established that N cannot have negative eigenvalues. We must eventually reach a state $|\nu\rangle$ for which $a|\nu\rangle = 0$, the operator a *annihilates* the state. (At each step down, a annihilates one quantum of energy -- so a is often called an *annihilation* or *destruction* operator.)

Since the norm squared of $a|\nu\rangle$, $|a|\nu\rangle|^2 = \langle\nu|a^\dagger a|\nu\rangle = \langle\nu|N|\nu\rangle = \nu\langle\nu|\nu\rangle$, and since $\langle\nu|\nu\rangle > 0$ for any nonvanishing state, it must be that *the lowest eigenstate* (the $|\nu\rangle$ for which $a|\nu\rangle = 0$) *has* $\nu = 0$. It follows that the ν 's on the ladder are *the positive integers*, so from this point on we relabel the eigenstates with n in place of ν.

That is to say, we have proved that the only possible eigenvalues of N are zero and the positive integers: 0, 1, 2, 3… .

N is called the ***number operator***: it measures the number of quanta of energy in the oscillator above the irreducible ground state energy (that is, above the “zero-point energy” arising from the wave-like nature of the particle).

Since from above the Hamiltonian

$$H = \hbar\omega(a^\dagger a + \frac{1}{2}) = \hbar\omega(N + \frac{1}{2}) \tag{3.4.35}$$

the energy eigenvalues are

$$H|n\rangle = (n + \frac{1}{2})\hbar\omega|n\rangle. \tag{3.4.36}$$

It is important to appreciate that Dirac’s factorization trick and very little effort has given us *all* the eigenvalues of the Hamiltonian

$$H = \frac{\hbar\omega}{2}(\pi^2 + \xi^2). \tag{3.4.37}$$

Contrast the work needed in this section with that in the standard Schrödinger approach. We have also established that the lowest energy state $|0\rangle$, having energy $\frac{1}{2}\hbar\omega$, must satisfy the first-order differential equation $a|0\rangle = 0$, that is,

$$(\xi + i\pi) \mid 0 > = (\xi + \frac{d}{d\xi})\psi_0(\xi) = 0. \tag{3.4.38}$$

The solution, unnormalized, is

$$\psi_0(\xi) = Ce^{-\xi^2/2}. \tag{3.4.39}$$

(In fact, we've seen this equation and its solution before: this was the condition for the "least uncertain" wavefunction in the discussion of the Generalized Uncertainty Principle.)

We denote the *normalized* set of eigenstates $|0\rangle, |1\rangle, |2\rangle, \ldots |n\rangle \ldots$ with $\langle n|n\rangle = 1$. Now $a^\dagger|n\rangle = C_n|n+1\rangle$, and C_n is easily found:

$$\left|C_n\right|^2 = |\, Cn\, |^2 \, \langle n+1|n+1\rangle = \langle n|aa^\dagger|n\rangle = (n+1), \tag{3.4.40}$$

and

$$a^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle. \tag{3.4.41}$$

Therefore, if we take the set of orthonormal states $|0\rangle, |1\rangle, |2\rangle, \ldots |n\rangle \ldots$ as the basis in the Hilbert space, the *only* nonzero matrix elements of $a^\dagger$ are $\langle n+1|a^\dagger|n\rangle = \sqrt{n+1}$. That is to say,

$$a^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & \ldots \\ \sqrt{1} & 0 & 0 & 0 & \ldots \\ 0 & \sqrt{2} & 0 & 0 & \ldots \\ 0 & 0 & \sqrt{3} & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.4.42}$$

(The column vectors in the space this matrix operates on have an infinite number of elements: the lowest energy, the ground state component, is the entry at the *top* of the infinite vector -- so up the energy ladder is down the vector!)

The adjoint

$$a = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \ldots \\ 0 & 0 & \sqrt{2} & 0 & \ldots \\ 0 & 0 & 0 & \sqrt{3} & \ldots \\ 0 & 0 & 0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.4.43}$$

So

$$a|n\rangle = \sqrt{n}\,|n-1\rangle. \tag{3.4.44}$$

For practical computations, we need to find the matrix elements of the position and momentum variables between the normalized eigenstates. Now

$$x = \sqrt{\hbar/2m\omega}\,(a^\dagger + a), \quad p = i\sqrt{m\omega\hbar/2}\,(a^\dagger - a) \tag{3.4.45}$$

so

$$x = \sqrt{\hbar/2m\omega}\begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \ldots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \ldots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \ldots \\ 0 & 0 & \sqrt{3} & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad p = i\sqrt{m\omega\hbar/2}\begin{pmatrix} 0 & -\sqrt{1} & 0 & 0 & \ldots \\ \sqrt{1} & 0 & -\sqrt{2} & 0 & \ldots \\ 0 & \sqrt{2} & 0 & -\sqrt{3} & \ldots \\ 0 & 0 & \sqrt{3} & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.4.46}$$

These matrices are, of course, Hermitian (not forgetting the i factor in p.)

To find the matrix elements between eigenstates of any product of x 's and p 's, express all the x 's and p 's in terms of a 's and $a^\dagger$'s, to give a sum of products of a 's and $a^\dagger$'s. Each product in this sum can be evaluated sequentially from the right, because each a or $a^\dagger$ has only one nonzero matrix element when the product operates on one eigenstate.

Normalizing the Eigenstates in *x*-space

The normalized ground state wavefunction is

$$\psi_0(\xi)=Ce^{-\xi^2/2}=\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}e^{-m\omega x^2/2\hbar}, \tag{3.4.47}$$

where we have gone back to the x variable, and normalized using $\int_{-\infty}^{\infty}e^{-ax^2}dx=\sqrt{\pi/a}$.

To find the normalized wavefunctions for the higher states, they are first constructed formally by applying the creation operator $a^\dagger$ repeatedly on the ground state $|0\rangle$. Next, the result is translated into x -space (actually $\xi=x/b$) by writing $a^\dagger$ as a differential operator, acting on $\psi_0(\xi)$.

Using $\langle n|a^\dagger|n-1\rangle=\sqrt{n}$,

$$|n\rangle=\frac{a^\dagger}{\sqrt{n}}|n-1\rangle=\dots=\frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle. \tag{3.4.48}$$

Now

$$a^\dagger=(1/\sqrt{2})(\xi-i\pi)=(1/\sqrt{2})(\xi-d/d\xi), \tag{3.4.49}$$

so

$$\psi_n(\xi)=\frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle=\frac{1}{\sqrt{n!}}\left(\frac{1}{\sqrt{2}}(\xi-\frac{d}{d\xi})\right)^n\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}e^{-\xi^2/2}. \tag{3.4.50}$$

We need to check that this expression is indeed the same as the Hermite polynomial wavefunction derived earlier, and to do that we need some further properties of the Hermite polynomials.

Some Properties of Hermite Polynomials

The mathematicians *define* the Hermite polynomials by:

$$H_n(\xi)=(-)^n e^{\xi^2}\frac{d^n}{d\xi^n}e^{-\xi^2} \tag{3.4.51}$$

so

$$H_0(\xi)=1,\quad H_1(\xi)=2\xi,\quad H_2(\xi)=4\xi^2-2,\quad H_3(\xi)=8\xi^3-\frac{1}{2}\xi,\quad etc. \tag{3.4.52}$$

It follows immediately from the definition that the coefficient of the leading power is 2^n.

It is a straightforward exercise to check that H_n is a solution of the differential equation

$$\left(\frac{d^2}{d\xi^2}-2\xi\frac{d}{d\xi}+2n\right)H_n(\xi)=0, \tag{3.4.53}$$

so these are indeed the same polynomials we found by the series solution of Schrödinger’s equation earlier (recall the equation for the polynomial component of the wavefunction was

$$\frac{d^2h}{d\xi^2}-2\xi\frac{dh}{d\xi}+(2\varepsilon-1)h=0, \tag{3.4.54}$$

with

$$2\varepsilon=2n+1. \tag{3.4.55}$$

We have found $\psi_n(\xi)$ in the form

$$\psi_n(\xi)=\frac{1}{\sqrt{n!}}(\frac{1}{\sqrt{2}}\left(\xi-\frac{d}{d\xi})\right)^n\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}e^{-\xi^2/2}. \tag{3.4.56}$$

We shall now prove that the polynomial component is exactly equivalent to the Hermite polynomial as defined at the beginning of this section.

We begin with the operator identity:

$$(\xi-\frac{d}{d\xi})=-e^{\xi^2/2}\frac{d}{d\xi}e^{-\xi^2/2} \tag{3.4.57}$$

Both sides of this expression are to be regarded as *operators*, that is, it is assumed that both are operating on some function $f(\xi)$.

Now take the n^{th} power of both sides: on the right, we find, for example,

$$(-e^{\xi^2/2}\frac{d}{d\xi}e^{-\xi^2/2})^3=(-)^3e^{\xi^2/2}\frac{d}{d\xi}e^{-\xi^2/2}e^{\xi^2/2}\frac{d}{d\xi}e^{-\xi^2/2}e^{\xi^2/2}\frac{d}{d\xi}e^{-\xi^2/2}=(-)^3e^{\xi^2/2}\frac{d^3}{d\xi^3}e^{-\xi^2/2} \tag{3.4.58}$$

since the intermediate exponential terms cancel against each other.

So

$$(\xi-\frac{d}{d\xi})n=(-)^ne^{\xi^2/2}\frac{d^n}{d\xi^n}e^{-\xi^2/2} \tag{3.4.59}$$

and substituting this into the expression for $\psi_n(\xi)$ above,

$$\begin{aligned}\psi_n(\xi)=&\tfrac{1}{\sqrt{2^nn!}}(-)^n(e^{\xi^2/2}\tfrac{d^n}{d\xi^n}e^{-\xi^2/2})(\tfrac{m\omega}{\pi\hbar})^{1/4}e^{-\xi^2/2}\\ &=\tfrac{1}{\sqrt{2^nn!}}(-)^n(\tfrac{m\omega}{\pi\hbar})^{1/4}e^{-\xi^2/2}(e^{\xi^2}\tfrac{d^n}{d\xi^n}e^{-\xi^2})\\ =&\tfrac{1}{\sqrt{2^nn!}}(\tfrac{m\omega}{\pi\hbar})^{1/4}H_n(\xi)e^{-\xi^2/2},\ \ with\ \ \xi=\sqrt{\tfrac{m\omega}{\hbar}}x.\end{aligned} \tag{3.4.60}$$

This established the equivalence of the two approaches to Schrödinger’s equation for the simple harmonic oscillator, and provides us with the overall normalization constants without doing integrals. (The expression for $\psi_n(\xi)$ above satisfies $\int|\psi_n|^2dx=1$.)

Exercise 3.4.1

Exercises:

Use $H_n(\xi)=(-)^ne^{\xi^2}\frac{d^n}{d\xi^n}e^{-\xi^2}$ to prove:

a. the coefficient of ξ_n is 2^n.
b. $H_n{}'(\xi)=2nH_{n-1}(\xi)$
c. $H_{n+1}(\xi)=2\xi H_n(\xi)-2nH_{n-1}(\xi)$
d. $\int_{-\infty}^{\infty}e^{-\xi^2}H_n^2(\xi)\,d\xi=2^nn!\sqrt{\pi}$ (Hint: rewrite as $\int_{-\infty}^{\infty}H_n(\xi)(-)^n\frac{d^n}{d\xi^n}e^{-\xi^2}\,d\xi$, then integrate by parts n times, and use (a).)
e. $\int_{-\infty}^{\infty}e^{-\xi^2}H_n(\xi)\,H_m(\xi)d\xi=0$, for $m\neq n$.

It’s worth doing these exercises to become more familiar with the Hermite polynomials, but in evaluating matrix elements (and indeed in establishing some of these results) it is almost always far simpler to work with the creation and annihilation operators.

Exercise 3.4.2

Use the creation and annihilation operators to find $\langle n|x^4|n\rangle$. This matrix element is useful in estimating the energy change arising on adding a small nonharmonic potential energy term to a harmonic oscillator.

Time-Dependent Wavefunctions

The set of normalized eigenstates $|0\rangle,|1\rangle,|2\rangle,\dots|n\rangle\dots$ discussed above are of course solutions to the time-*independent* Schrödinger equation, or in ket notation eigenstates of the Hamiltonian $H|n\rangle=(n+\frac{1}{2})\hbar\omega|n\rangle$. Putting in the time-dependence explicitly,

$$|n,t\rangle=e^{-iHt/\hbar}|n,t=0\rangle=e^{-i(n+\frac{1}{2})\omega t}|n\rangle. \tag{3.4.61}$$

It is necessary to include the time dependence when dealing with a state which is a superposition of states of different energies, such as $(1/\sqrt{2})(|0\rangle+|1\rangle)$, which then becomes

$$(1/\sqrt{2})(e^{-i\omega t/2}|0\rangle+e^{-3i\omega t/2}|1\rangle). \tag{3.4.62}$$

Expectation values of combinations of position and/or momentum operators in such states are best evaluated by expressing everything in terms of annihilation and creation operators.

Solving Schrödinger's Equation in Momentum Space

In the lecture on Function Spaces, we established that the basis of $|x\rangle$ states (eigenstates of the position operator) and that of $|k\rangle$ states (eigenstates of the momentum operator) were both complete bases in Hilbert space (physicist's definition) so we could work equally well with either from a formal point of view. Why then do we almost always work in x -space? Well, probably because we live in x -space, but there's another reason. The momentum operator in the x -space representation is $p=-i\hbar d/dx$, so Schrödinger's equation, written $(p^2/2m+V(x))\psi(x)=E\psi(x)$, with p in operator form, is a second-order differential equation. Now consider what happens to Schrödinger's equation if we work in p -space. Since the operator identity $[x,p]=i\hbar$ is true regardless of representation, we must have $x=i\hbar d/dp$. So for a particle in a potential $V(x)$, writing Schrödinger's equation in p -space we are confronted with the nasty looking operator $V(i\hbar d/dp)$! This will produce a differential equation in general a lot harder to solve than the standard x -space equation -- so we stay in x -space.

But there are two potentials that can be handled in momentum space: first, for a *linear* potential $V(x)=-Fx$, the momentum space analysis is actually easier -- it's just a first-order equation. Second, for a particle in a *quadratic* potential -- a simple harmonic oscillator -- the two approaches yield the *same* differential equation. That means that the eigenfunctions in momentum space (scaled appropriately) must be *identical* to those in position space -- the simple harmonic eigenfunctions are their own Fourier transforms!

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

3.5: Propagators and Representations

We've spent most of the course so far concentrating on the eigenstates of the Hamiltonian, states whose time-dependence is merely a changing phase. We did mention much earlier a superposition of two different energy states in an infinite well, resulting in a wavefunction sloshing backwards and forwards. It's now time to cast the analysis of time dependent states into the language of bras, kets and operators. We'll take a time-independent Hamiltonian H, with a complete set of orthonormalized eigenstates, and as usual

$$i\hbar\frac{\partial}{\partial t}\psi(x,t)=-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t)+V(x)\psi(x,t), \tag{3.5.1}$$

Or, as we would now write it

$$i\hbar\frac{\partial}{\partial t}|\psi(x,t)\rangle=H|\psi(x,t)\rangle. \tag{3.5.2}$$

Since H is itself time independent, this is very easy to integrate!

$$|\psi(x,t)\rangle=e^{-iH(t-t_0)/\hbar}|\psi(x,t_0)\rangle. \tag{3.5.3}$$

The exponential operator that generates the time-dependence is called the *propagator*, because it describes how the wave propagates from its initial configuration, and is usually denoted by U:

$$|\psi(x,t)\rangle=U(t-t_0)|\psi(x,t_0)\rangle. \tag{3.5.4}$$

It's appropriate to call the propagator U, because it's a unitary operator:

$$U(t-t_0)=e^{-iH}(t-t_0) \tag{3.5.5}$$

so

$$U^{\dagger}(t-t_0)=e^{iH^{\dagger}}(t-t_0)=e^{iH}(t-t_0)=U-1(t-t_0). \tag{3.5.6}$$

Since H is Hermitian, U is unitary. It immediately follows that

$$\langle\psi(x,t)|\psi(x,t)\rangle=\langle\psi(x,t_0)|U^{\dagger}U(t-t_0)|\psi(x,t_0)\rangle=\langle\psi(x,t_0)|\psi(x,t_0)\rangle \tag{3.5.7}$$

the norm of the ket vector is conserved, or, translating to wavefunction language, a wavefunction correctly normalized to give a total probability of one stays that way. (This can also be proved from the Schrödinger equation, of course, but this is quicker.)

This is all very succinct, but unfortunately the exponential of a second-order differential operator doesn't sound too easy to work with. Recall, though, that any function of a Hermitian operator has the same set of eigenstates as the original operator. This means that the eigenstates of $e^{-iH(t-t_0)/\hbar}$ are the same as the eigenstates of H, and if $H|\psi_n\rangle=E_n|\psi_n\rangle$, then

$$e^{-iH(t-t_0)/\hbar}|\psi_n\rangle=e^{-iEn(t-t_0)/\hbar}|\psi_n\rangle. \tag{3.5.8}$$

This is of course nothing but the time dependent phase factor for the eigenstates we found before — and, as before, to find the time dependence of any general state we must express it as a superposition of these eigenkets, each having its own time dependence. But how do we do that in the operator language? Easy: we simply insert an identity operator, the one constructed from the complete set of eigenkets, thus:

$$|\psi(t)\rangle=e^{-iH(t-t_0)/\hbar}\sum_{n=1}^{\infty}|\psi_n\rangle\langle\psi_n|\psi(t_0)\rangle=\sum_{n=1}^{\infty}e^{-iE_n(t-t_0)/\hbar}|\psi_n\rangle\langle\psi_n|\psi(t_0)\rangle. \tag{3.5.9}$$

Staring at this, we see that it's just what we had before: at the initial time $t=t_0$, the wavefunction can be written as a sum over the eigenkets:

$$|\psi(t_0)\rangle=\sum|\psi_n(t_0)\rangle\langle\psi_n(t_0)|\psi(t_0)\rangle=\sum c_n|\psi_n(t_0)\rangle \tag{3.5.10}$$

with

$$c_n=\langle\psi_n|\psi\rangle \tag{3.5.11}$$

and

$$\sum|c_n|^2=1 \tag{3.5.12}$$

and the usual generalization for continuum eigenvalues, and the time development is just given by inserting the phases:

$$|\psi(t)\rangle=\sum c_n e^{-iE_n(t-t_0)/\hbar}|\psi_n(t_0)\rangle. \tag{3.5.13}$$

The expectation value of the energy E in $|\psi\rangle$,

$$\langle E\rangle = \langle\psi|H|\psi\rangle = \sum |c_n|^2 E_n \tag{3.5.14}$$

and is (of course) time independent.

The expectation value of the particle position x is

$$\langle\psi(t)|x|\psi(t)\rangle = \sum_{n,m} c_n^* c_m e^{i(E_n-E_m)(t-t_0)/\hbar}\langle\psi_n(t_0)|x|\psi_m(t_0)\rangle \tag{3.5.15}$$

and is not in general time-independent. (It is real, of course, on adding the n,m term to the m, n term.)

This analysis is only valid for a *time-independent Hamiltonian*. The important extension to a system in a time-dependent external field, such as an atom in a light beam, will be given later in the course.

The Free Particle Propagator

To gain some insight into what the propagator U looks like, we'll first analyze the case of a particle in one dimension with no potential at all. We'll also take $t_0 = 0$ to make the equations less cumbersome. For a free particle in one dimension

$$E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m} \tag{3.5.16}$$

the energy eigenstates are also momentum eigenstates, we label them $|k\rangle$, so

$$U(t) = e^{-iHt/\hbar} = \in_{-\infty}^{\infty} e^{-iHt/\hbar dk^2\pi}|k\rangle\langle k| = \int_{-\infty}^{\infty} e^{-i\hbar}k^2t/2mdk^2\pi|k\rangle\langle k|. \tag{3.5.17}$$

a particle is at x_0: $\psi(x,t=0) = \delta(x-x_0) = |x_0\rangle$: what is the probability amplitude for finding it at x at a later time t? (This would be just its wavefunction at the later time.)

$$\langle x|U(t,0)|x_0\rangle = \int_{-\infty}^{\infty} e^{-i\hbar k^2t/2m}dk^2\pi\langle x|k\rangle\langle k|x0\rangle = \int_{-\infty}^{\infty} e^{-i\hbar k^2t/2m}dk2\pi e^{-ik(x_0-x)} = m^2\pi\hbar it ----\sqrt{} e^{im(x_0-x)2}2\hbar t, \tag{3.5.18}$$

Exercise 3.5.1

On examining Equation 3.5.18, though, it turns out to be nonsense! Noting that the term in the exponent is pure imaginary, $|\psi(x,t)|^2 = \frac{m}{2\pi\hbar t}$ independent of x! This particle apparently instantaneously fills all of space, but then its probability dies away as 1/t…

Question: Where did we go wrong?

Answer: Notice first that $|\psi(x,t)|^2$ is constant throughout space. This means that the normalization, ∫|ψ(x,t)|2dx=∞ ! And, as we've seen above, the normalization stays constant in time — the propagator is unitary. Therefore, our initial wavefunction must have had infinite norm. That's exactly right — we took the initial wavefunction

\][ψ(x,t=0)=δ(x−x_0)=| x_0 ⟩.\]

Think of the δ-function as a limit of a function equal to 1/Δ over an interval of length Δ , with Δ going to zero, and it's clear the normalization goes to infinity as 1/Δ . This is not a meaningful wavefunction for a particle. Recall that continuum kets like |x0⟩ are normalized by ⟨x|x′⟩ =δ(x−x′) , they do not represent wavefunctions individually normalizable in the usual sense. The only meaningful wavefunctions are integrals over a range of such kets, such as ∫dxψ(x)|x⟩ . In an integral like this, notice that states |x⟩ within some tiny x -interval of length δx, say, have total weight ψ(x)δx , which goes to zero as δx is made smaller, but by writing ψ(x,t=0)=δ(x−x0)=| x0 ⟩ we took a single such state and gave it a finite weight. This we can't do.

Of course, we do want to know how a wavefunction initially localized near a point develops. To find out, we must apply the propagator to a legitimate wavefunction — one that is normalizable to begin with. The simplest "localized particle" wavefunction from a practical point of view is a Gaussian wave packet,

$$\psi(x',0) = e^{ip_0x'/\hbar}e^{-x'2/2d^2(\pi d^2)}1/4. \tag{3.5.19}$$

(I've used d in place of Shankar's Δ here to try to minimize confusion with Δx, etc.)

The wavefunction at a later time is then given by the operation of the propagator on this initial wavefunction:

$$\psi(x,t) = \int U(x,t;x',0)eip0x'/\hbar e - x'2/2d2(\pi d2)1/4dx' = m2\pi\hbar it ----\sqrt{}\int eim(x-x')2/2\hbar teip0x'/\hbar e - x'2/2d2(\pi d2)1/4dx'. \tag{3.5.20}$$

The integral over x′ is just another Gaussian integral, so we use the same result,

$$\int -\infty\infty dx'e-ax'2+bx'=\pi a--\sqrt{}eb2/4a. \tag{3.5.21}$$

Looking at the expression above, we can see that

$$b=-im\hbar t\cdot(x-p0tm),a=12d2-im2\hbar t. \tag{3.5.22}$$

This gives

$$\psi(x,t)=\pi-1/4d(1+i\hbar tmd2)\sqrt{}exp(imx22\hbar t)exp(-im\hbar t(x-p0tm)22(1+i\hbar tmd2)) \tag{3.5.23}$$

where the second exponential is the term eb2/4a . As written, the small t limit is not very apparent, but some algebraic rearrangement yields:

$$\psi(x,t)=\pi-1/4d(1+i\hbar t/md2)\sqrt{}exp(-(x-p0t/m)22d2(1+i\hbar t/md2))exp(ip0\hbar(x-p0t/2m)). \tag{3.5.24}$$

It is clear that this expression goes to the initial wave packet as t goes to zero. Although the phase has contributions from all three terms here, the main phase oscillation is in the third term, and one can see the phase velocity is one-half the group velocity, as discussed earlier.

The resulting probability density:

$$|\psi(x,t)|2=1\pi(d2+\hbar 2t2/m2d2)\sqrt{}\cdot exp-(x-p0t/m)2(d2+\hbar 2t2/m2d2). \tag{3.5.25}$$

This is a Gaussian wave packet, having a width which goes as ℏt/md for large times, where d is the width of the initial packet in x -space — so ℏ/md is the spread in velocities ⟨Δv⟩ within the packet, hence the gradual spreading ⟨Δv⟩t in x -space.

It's amusing to look at the limit of this as the width d of the initial Gaussian packet goes to zero, and see how that relates to our δ -function result. Suppose we are at distance x from the origin, and there is initially a Gaussian wave packet centered at the origin, width d≪x. At time t∼mxd/ℏ , the wave packet has spread to x and has |ψ(x,t)|2 of order 1/x at x. Thereafter, it continues to spread at a linear rate in time, so locally |ψ(x,t)|2 must decrease as 1/t to conserve probability. In the δ -function limit d→0 , the wavefunction instantly spreads through a huge volume, but then goes as 1/t as it spreads into an even huger volume. Or something.

Schrödinger and Heisenberg Representations

Assuming a Hamiltonian with no explicit time dependence, the time-dependent Schrödinger equation has the form

$$i\hbar\frac{\partial}{\partial t}|\psi(x,t)\rangle=H|\psi(x,t)\rangle \tag{3.5.26}$$

and as discussed above, the formal solution can be expressed as:

$$|\psi(x,t)\rangle=e^{-iHt/\hbar}|\psi(x,t=0)\rangle. \tag{3.5.27}$$

Now, any measurement on a system amounts to measuring a matrix element of an operator between two states (or, more generally, a function of such matrix elements).

In other words, the physically significant time dependent quantities are of the form

$$\langle\varphi(t)|A|\psi(t)\rangle=\langle\varphi(0)|e^{iHt/\hbar}Ae^{-iHt/\hbar}|\psi(0)\rangle \tag{3.5.28}$$

where A is an operator, which we are assuming has no explicit time dependence.

So in this Schrödinger picture, the time dependence of the measured value of an operator like x or p comes about because we measure the matrix element of an unchanging operator between bras and kets that are changing in time.

Heisenberg took a different approach: he assumed that the ket describing a quantum system did not change in time, it remained at |ψ(0)⟩, but the operators evolved according to:

$$AH(t)=e^{iHt/\hbar}AH(0)e^{-iHt/\hbar}. \tag{3.5.29}$$

Clearly, this leads to the same physics as before. The equation of motion of the operator is:

$$i\hbar\frac{d}{dt}AH(t)=[AH(t),H]. \tag{3.5.30}$$

The Hamiltonian itself does not change in time — energy is conserved, or, to put it another way, H commutes with e−iHt/ℏ. But for a nontrivial Hamiltonian, say for a particle in one dimension in a potential,

$$H=\frac{p^2}{2m}+V(x) \tag{3.5.31}$$

the separate components will have time-dependence, parallel to the classical case: the kinetic energy of a swinging pendulum varies with time. (For a particle in a potential in an energy eigenstate the expectation value of the kinetic energy is constant, but this is not the case for any other state, that is, for a superposition of different eigenstates.) Nevertheless, the commutator of x, and p will be time-independent:

$$[xH(t),pH(t)] = e^{iHt/\hbar}[xH(0),pH(0)]e^{-iHt/\hbar} = e^{iHt/\hbar}i\hbar e^{-iHt/\hbar} = i\hbar. \tag{3.5.32}$$

(The Heisenberg operators are identical to the Schrödinger operators at t=0.)

Applying the general commutator result [A,BC]=[A,B]C+B[A,C] ,

$$[xH(t),\ p2H(t)2m] = i\hbar pH(t)m \tag{3.5.33}$$

so

$$dxH(t)dt = pH(t)m \tag{3.5.34}$$

and since

$$[xH(t),pH(t)] = i\hbar,\ \ pH(t) = -i\hbar d/dxH(t), \tag{3.5.35}$$

$$dpH(t)dt = 1i\hbar[pH(t),V(xH(t))] = -\nabla V(xH(t)). \tag{3.5.36}$$

This result could also be derived by writing V(x) as an expansion in powers of x, then taking the commutator with p.

Exercise 3.5.1

Exercise: check this.

Notice from the above equations that the operators in the Heisenberg Representation obey the classical laws of motion! Ehrenfest's Theorem, that the expectation values of operators in a quantum state follow the classical laws of motion, follows immediately, by taking the expectation value of both sides of the operator equation of motion in a quantum state.

Simple Harmonic Oscillator in the Heisenberg Representation

For the simple harmonic oscillator, the equations are easily integrated to give:

$$xH(t) = xH(0)cos\omega t + (pH(0)/m\omega)sin\omega tpH(t) = pH(0)cos\omega t - m\omega xH(0)sin\omega t. \tag{3.5.37}$$

We have put in the H subscript to emphasize that these are operators. It is usually clear from the context that the Heisenberg representation is being used, and this subscript may be safely omitted.

The time-dependence of the annihilation operator a is:

$$a(t) = eiHt/\hbar a(0)e - iHt/\hbar \tag{3.5.38}$$

with

$$H = \hbar\omega(a\dagger(t)a(t) + 12/). \tag{3.5.39}$$

Note again that although H is itself time-independent, it is necessary to include the time-dependence of individual operators within H.

$$i\hbar ddta(t) = [a(t),H] = \hbar\omega[a(t),a\dagger(t)a(t)] = \hbar\omega[a(t),a\dagger(t)]a(t) = \hbar\omega\ a(t) \tag{3.5.40}$$

so

$$a(t) = a(0)e - i\omega t. \tag{3.5.41}$$

Actually, we could have seen this as follows: if |n⟩ are the energy eigenstates of the simple harmonic oscillator,

$$e - iHt/\hbar|n\rangle = e - in\ \hbar\omega t/\hbar|n\rangle = e - in\omega t|n\rangle. \tag{3.5.42}$$

Now the only nonzero matrix elements of the annihilation operator aˆ between energy eigenstates are of the form

$$\langle n-1|a(t)|n\rangle = \langle n-1|eiHt/\hbar a(0)e - iHt/\hbar|n\rangle = ei\omega(n-1)t\langle n-1|a(0)|n\rangle e - i\omega nt = \langle n-1|a(0)|n\rangle e - i\omega t. \tag{3.5.43}$$

Since this time-dependence is true of all energy matrix elements (trivially so for most of them, since they're identically zero), and the eigenstates of the Hamiltonian span the space, it is true as an operator equation.

Evidently, the expectation value of the operator a(t) in any state goes clockwise in a circle centered at the origin in the complex plane. That this is indeed the classical motion of the simple harmonic oscillator is confirmed by recalling the definition a=ξ+iπ2√ =12ℏmω√(mωx+ip) , so the complex plane corresponds to the (mωx,p) phase space discussed near the beginning of the lecture on the Simple Harmonic Oscillator. We'll discuss this in much more detail in the next lecture, on Coherent States.

The time-dependence of the creation operator is just the adjoint equation:

$$a\dagger(t)=a\dagger(0)ei\omega t. \tag{3.5.44}$$

Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

3.6: Coherent States

What is the wavefunction of a Swinging Pendulum?

Consider a macroscopic simple harmonic oscillator, and to keep things simple assume there are no interactions with the rest of the universe. We know how to describe the motion using classical mechanics: for a given initial position and momentum, classical mechanics correctly predicts the future path, as confirmed by experiments with real (admittedly not perfect) systems. But from the Hamiltonian we could also write down Schrödinger's equation, and from that predict the future behavior of the system. Since we already know the answer from classical mechanics and experiment, quantum mechanics must give us the same result in the limiting case of a large system.

It is a worthwhile exercise to see just how this happens. Evidently, we cannot simply follow the classical method of specifying the initial position and momentum -- the uncertainty principle won't allow it. What we can do, though, is to take an initial state in which the position and momentum are specified *as precisely as possible*. Such a state is called a *minimum uncertainty* state (the details can be found in my earlier lecture on the Generalized Uncertainty Principle).

In fact, the *ground state* of a simple harmonic oscillator *is* a minimum uncertainty state. This is not too surprising -- it's just a localized wave packet centered at the origin. The system is as close to rest as possible, having only zero-point motion. What*is* surprising is that there are excited states of the pendulum in which this ground state wave packet swings backwards and forwards indefinitely, a quantum realization of the classical system, and the wave packet is always one of minimum uncertainty. Recall that this *doesn't* happen for a *free* particle on a line—in that case, an initial minimal uncertainty wave packet spreads out because the different momentum components move at different speeds. But for the oscillator, the potential somehow keeps the wave packet together, a minimum uncertainty wave packet at all times. These remarkable quasi-classical states are called *coherent states*, and were discovered by Schrödinger himself. They are important in many quasi-classical contexts, including laser radiation.

Our task here is to construct and analyze these coherent states and to find how they relate to the usual energy eigenstates of the oscillator.

Classical Mechanics of the Simple Harmonic Oscillator

To define the notation, let us briefly recap the dynamics of the *classical* oscillator: the constant energy is

$$E=\frac{p^2}{2m}+\frac{1}{2}kx^2 \tag{3.6.1}$$

or

$$p^2+(m\omega x)^2=2mE,\quad \omega=\sqrt{k/m}. \tag{3.6.2}$$

The classical motion is most simply described in *phase space*, a two-dimensional plot in the variables $(m\omega x,p)$. In this space, the point $(m\omega x,p)$ corresponding to the position and momentum of the oscillator at an instant of time moves as time progresses at constant angular speed \omega in a clockwise direction around the circle of radius $\sqrt{2mE}$ centered at the origin.

(*Note*: phase space is usually defined in terms of the variables (x,p), but in describing the simple harmonic oscillator, the variables $(m\omega x,p)$ are more convenient, they have the same dimensions.)

This motion is elegantly described by regarding the two-dimensional phase space as a complex plane, and defining the dimensionless complex variable

$$z=\frac{m\omega x+ip}{\sqrt{2\hbar m\omega}}. \tag{3.6.3}$$

The time evolution in phase space is simply

$$z(t)=z_0e^{-i\omega t}. \tag{3.6.4}$$

The particular choice of (quantum!) scaling factor in defining z amounts to defining the unit of energy as $\hbar\omega$, the natural quantum unit for the oscillator: it is easy to check that if the classical energy $E=(n+\frac{1}{2})\hbar\omega$ then the dimensionless $|z|^2$ is simply the number $n+\frac{1}{2}$ (which is of course very large, so the $\frac{1}{2}$ is insignificant).

Minimum Uncertainty Wavepackets

We established in the lecture on the Generalized Uncertainty Principle that any minimum uncertainty one-dimensional wavefunction (so $\Delta p\cdot\Delta\ x=\hbar/2$) for a particle must satisfy the linear differential equation (here $\hat{p}=-i\hbar d/dx$)

$$(\hat{p}-\langle p\rangle)\psi(x)=\lambda(\hat{x}-\langle x\rangle)\psi(x) \tag{3.6.5}$$

where $\langle x\rangle$, $\langle p\rangle$, λ are constants, and λ is *pure imaginary*. The equation is easy to solve: any minimum uncertainly one-dimensional wavefunction is a Gaussian wave packet, having expectation value of momentum $\langle p\rangle$, centered at $\langle x\rangle$ and having width $(\Delta x)^2=-\hbar/2i\lambda$. (Δx is defined for a state $|\psi\rangle$ by $(\Delta x)^2=\langle\psi|(x-\langle x\rangle)^2|\psi\rangle$.)

That is to say, the minimum uncertainly solution is:

$$\psi(x)=Ce^{i\langle p\rangle x/\hbar}e^{i\lambda(x-\langle x\rangle)^2/2\hbar}=Ce^{i\langle p\rangle x/\hbar}e^{-(x-\langle x\rangle)^2/4(\Delta x)^2} \tag{3.6.6}$$

with C the normalization constant.

In fact, the simple harmonic oscillator *ground state* $\psi_0(x)=\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}e-m\omega x^2/2\hbar$ is just such a minimum uncertainty state, with

$$\lambda=im\omega,\quad \langle x\rangle=\langle p\rangle=0\ ;\quad (\Delta x)^2=\frac{\hbar}{2m\omega},\quad (\Delta p)^2=\frac{\hbar m\omega}{2},\quad \Delta p\cdot\Delta x=\frac{\hbar}{2}. \tag{3.6.7}$$

Furthermore, it is easy to see that the *displaced* ground state $\psi_0(x-x_0)=Ce^{-m\omega(x-x_0)^2/2\hbar}$, with $\langle x\rangle=x_0$, and writing the normalization constant $(m\omega/\pi\hbar)^{1/4}=C$, must also be a minimum uncertainty state, *with the same* \(\lambda=im\omega\). (It satisfies the necessary differential equation.) Of course, in contrast to the ground state, this displaced state is no longer an eigenstate of the Hamiltonian, and will therefore change with time.

(Both these states, $\langle x\rangle=x_0$ and $\langle x\rangle=0$, have the same spread in x -space $(\Delta x)^2=\hbar/2m\omega$, and the same spread in p -space, the only difference in the p direction being a phase factor $e^{ip\langle x_0\rangle/\hbar}$ for the displaced state.)

What about the higher eigenstates of the oscillator Hamiltonian? They are *not* minimally uncertain states -- for the n^{th} state, $\Delta p\cdot\Delta x=n\hbar/2$, as is easily checked using $\frac{1}{2}(\Delta p)^2/2m=\frac{1}{2}k(\Delta x)^2\sim\frac{1}{2}n\hbar\omega$. So, if we construct a minimally uncertain higher energy state, it will *not* be an eigenstate of the Hamiltonian.

Exercise 3.6.1

Exercise: prove $\Delta p\cdot\Delta x=n\hbar/2$ for the n^{th} energy eigenstate. (Hint: use creation and annihilation operators.)

Eigenstates of the Annihilation Operator are Minimum Uncertainty States

Notation: We'll write

$$\langle x(t=0)\rangle=x_0,\quad \langle p(t=0)\rangle=p_0. \tag{3.6.8}$$

We restrict our attention here to *those minimum uncertainty states having the same spatial width as the oscillator ground state*-- these are what we need, and these are the ones we'll show to be eigenstates of the annihilation operator. (Actually, more general minimum uncertainty states, known as squeezed states, are also interesting and important, but we'll not consider them here.)

Suppose that at $t=0$ the oscillator wavefunction is the minimum uncertainty state

$$\psi(x,t=0)=Ce^{ip_0x/\hbar}ei^{\lambda(x-x_0)^2/2\hbar}=Ce^{ip_0x/\hbar}e^{-m\omega(x-x_0)^2/2\hbar} \tag{3.6.9}$$

centered at $(p_0,\ m\omega x_0)$ in phase space (as defined above for the classical oscillator), and with $\lambda=im\omega$ to give it the same spatial extent as the ground state.

From the preceding section, this $\psi(x,0)$ satisfies the minimum uncertainty equation

$$(\hat{p}-p_0)\psi(x,0)=im\omega(\hat{x}-x_0)\psi(x,0). \tag{3.6.10}$$

Rearranging this equation (and multiplying by $-i$) shows it in a different light:

$$(m\omega\hat{x}+i\hat{p})\psi(x,0)=(m\omega x_0+ip_0)\psi(x,0). \tag{3.6.11}$$

This is an eigenvalue equation! The wave packet $\psi(x,0)$ is an eigenstate of the operator $(m\omega\hat{x}+i\hat{p})$ with eigenvalue $(m\omega x_0+ip_0)$. It is *not*, of course, an eigenstates of either $\hat{p}$ or $\hat{x}$ taken individually.

Furthermore, the operator $(m\omega\hat{x}+i\hat{p})$ is just a constant times the annihilation operator $\hat{a}$ -- recall

$$\hat{a}=\frac{1}{\sqrt{2\hbar m\omega}}(m\omega\hat{x}+i\hat{p}). \tag{3.6.12}$$

Therefore, this minimally uncertain initial wave packet $\psi(x,0)$ is an eigenstate of the annihilation operator $\hat{a}$, with eigenvalue $(m\omega x_0+ip_0)/\sqrt{2\hbar m\omega}$. (By the way, it's ok for $\hat{a}$ to have complex eigenvalues, because $\hat{a}$ is not a Hermitian operator.)

We can now make the connection with the complex plane representation of the classical operator: the eigenvalue $(m\omega x_0+ip_0)/\sqrt{2\hbar m\omega}$ is precisely the parameter z_0 labeling the position of the classical operator in phase space in natural dimensionless units!

That is to say, a minimum uncertainty oscillator wave packet

$$\psi(x,t=0)=Ce^{ip_0x/\hbar}e^{-m\omega(x-x_0)^2/2\hbar} \tag{3.6.13}$$

centered at $(m\omega x_0,p_0)$ in phase space and having the same spatial extent as the ground state, is an eigenstate of the annihilation operator

$$\hat{a}\psi(x,t=0)=z_0\psi(x,t=0). \tag{3.6.14}$$

with eigenvalue the position of its center in phase space, that is,

$$z_0=\frac{m\omega x_0+ip_0}{\sqrt{2\hbar m\omega}}. \tag{3.6.15}$$

Time Development of the Minimal Wave Packet

Turning now to the *time development* of the state, it is convenient to use the ket notation

$$|\psi(x,t=0)\rangle=|x_0,p_0\rangle \tag{3.6.16}$$

with $|x,p\rangle$ denoting a minimum uncertainly wave packet (with the same spatial width as the ground state) having those expectation values of position and momentum.

The time development of the ket, as usual, is given by

$$|\psi(x,t)\rangle=e^{-iHt/\hbar}|x_0,p_0\rangle. \tag{3.6.17}$$

We shall show that $|\psi(x,t)\rangle$ *remains* an eigenstate of the annihilation operator *for all times* t: it therefore continues to be a minimum uncertainty wave packet! (And, of course, with constant spatial extent.)

The key point in establishing this is that the annihilation operator itself has a simple time development in the Heisenberg representation,

$$\hat{a}(t)=e^{iHt/\hbar}\hat{a}e^{-iHt/\hbar}=\hat{a}e^{-i\omega t}. \tag{3.6.18}$$

To prove this, consider the matrix elements of $\hat{a}(t)$ between any two eigenstates $|n\rangle$ of the Hamiltonian

$$H|n\rangle=(n+\frac{1}{2})\hbar\omega|n\rangle \tag{3.6.19}$$

so

$$\langle m|\hat{a}(t)|n\rangle=e^{i(m+\frac{1}{2})\hbar\omega t/\hbar}\langle m|\hat{a}|n\rangle e^{-i(n+\frac{1}{2})\hbar\omega t/\hbar}=\langle n-1|\hat{a}|n\rangle e^{-i\omega t}. \tag{3.6.20}$$

Since the only nonzero matrix elements of the annihilation operator $\langle m|\hat{a}|n\rangle$ are for $m=n-1$, and the energy eigenstates form a complete set, this simple time dependence is true as an *operator* equation

$$\hat{a}(t)=e^{iHt/\hbar}\hat{a}e^{-iHt/\hbar}=\hat{a}e^{-i\omega t}. \tag{3.6.21}$$

It is now easy to prove that

$$|\psi(x,t)\rangle = e^{-iHt/\hbar}|x_0,p_0\rangle \tag{3.6.22}$$

is always an eigenstate of $\hat{a}$:

$$\begin{aligned} \hat{a}|\psi(x,t)\rangle &= \hat{a}e^{-iHt/\hbar}|x_0,p_0\rangle \\ &= e^{-iHt/\hbar}(e^{iHt/\hbar}\hat{a}e^{-iHt/\hbar})|x_0,p_0\rangle \\ &= e^{-iHt/\hbar}e^{-i\omega t}\hat{a}|x_0,p_0\rangle \\ &= e^{-iHt/\hbar}e^{-i\omega t}(m\omega x_0+ip_0)/\sqrt{2\hbar m\omega}|x_0,p_0\rangle \\ &= \left(e^{-i\omega t}(m\omega x_0+ip_0)/\sqrt{2\hbar m\omega}\right)|\psi(x,t)\rangle. \end{aligned} \tag{3.6.23}$$

Therefore the annihilation operator, which at $t=0$ had the eigenvalue

$$z_0 = (m\omega x_0+ip_0)/\sqrt{2\hbar m\omega}, \tag{3.6.24}$$

corresponding to a minimal wave packet centered at $(m\omega x_0, p_0)$ in phase space, evolves in time t to another minimal packet (because it's still an eigenstate of the annihilation operator), and writing

$$|\langle x(t)\rangle, \langle p(t)\rangle\rangle = e^{-iHt/\hbar}|x_0,p_0\rangle, \tag{3.6.25}$$

the new eigenvalue of $\hat{a}$

$$z(t) = \frac{(m\omega\langle x(t)\rangle + i\langle p(t)\rangle)}{\sqrt{2\hbar m\omega}} = \frac{(m\omega x_0+ip_0)}{\sqrt{2\hbar m\omega}}e^{-i\omega t} = z(0)e^{-i\omega t}. \tag{3.6.26}$$

Therefore, the center of the wave packet in phase space follows the classical path in time. This is made explicit by equating real and imaginary parts:

$$\langle x(t)\rangle = x_0\cos\omega t + (p_0/m\omega)\sin\omega t,\ \langle p(t)\rangle = p_0\cos\omega t - m\omega x_0\sin\omega t. \tag{3.6.27}$$

So we've found Schrödinger's "best possible" quantum description of a classical oscillator.

A Remark on Notation

We have chosen to work with the original position and momentum variables, and the complex parameter expressed as a function of those variables, throughout. We could have used the dimensionless variables introduced in the lecture on the simple harmonic oscillator,

$$\xi = x/b = x\sqrt{m\omega/\hbar},\quad \pi = bp/\hbar = p/\sqrt{\hbar m\omega},\quad \hat{a} = (\hat{\xi}+i\hat{\pi})/\sqrt{2}. \tag{3.6.28}$$

This would of course also give $z = (\xi+i\pi)/\sqrt{2}$, a more compact representation, but one more thing to remember.

It's also common to denote the eigenstates of $\hat{a}$ by α, $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$, very elegant, but we've used z to keep reminding ourselves that this eigenvalue, unlike most of those encountered in quantum mechanics, is a complex number. Finally, some use the dimensionless variables $X = \sqrt{2\hbar/m\omega}x$, $P = \sqrt{1/(2m\omega\hbar)}p$, differing from ξ, π by a factor of $\sqrt{2}$. The eigenvalue equation for the annihilation operator is very neat in this notation: $\hat{a}|z\rangle = (X+iP)|z\rangle$. We've avoided it, though, because our recommended textbook, Shankar, uses X, P for the ordinary position and momentum operators.

The Translation Operator

It's worth repeating the exercise for the simple case of the oscillator initially at rest a distance x_0 from the center. This gives a neat tie-in with the *translation* operator (defined below).

Let us then take the initial state to be

$$\psi(x,0) = Ce^{-m\omega(x-x_0)^2/2\hbar} = \psi_0(x-x_0) \tag{3.6.29}$$

where $\psi_0(x)$ is the ground state wavefunction -- so we've moved the packet to the right by x_0.

Now do a Taylor series expansion (taking x_0 to be the variable!):

$$\psi_0(x-x_0) = \psi_0(x) - x_0\frac{d}{dx}\psi_0(x) + \frac{x_0^2}{2!}\frac{d^2}{dx^2}\psi_0(x) - \cdots = e^{-x_0\frac{d}{dx}}\psi_0(x). \tag{3.6.30}$$

It’s clear from this that the *translation operator* $e^{-x_0\frac{d}{dx}}$ shifts the wavefunction a distance x_0 to the right.

Since $\hat{p}=-i\hbar d/dx$, the translation operator can also be written as $e^{-ix_0\hat{p}/\hbar}$, and from this it can be expressed in terms of $\hat{a},\hat{a}^\dagger$, since

$$\hat{a}=\frac{1}{\sqrt{2\hbar m\omega}}(m\omega\hat{x}+i\hat{p}),\quad \hat{a}^\dagger=\frac{1}{\sqrt{2\hbar m\omega}}(m\omega\hat{x}-i\hat{p}),\tag{3.6.31}$$

($\hat{p}$, $\hat{x}$ being Hermitian) so

$$\hat{p}=i\sqrt{\hbar m\omega}2(\hat{a}^\dagger-\hat{a}).\tag{3.6.32}$$

Therefore the displaced ground state wavefunction can be written

$$\psi_0(x-x_0)=e^{-ix_0\hat{p}/\hbar}\psi_0(x)=e^{x_0\sqrt{m\omega/2\hbar}(\hat{a}^\dagger-\hat{a})}\psi_0(x)=e^{z_0(\hat{a}^\dagger-\hat{a})}\psi_0(x)\tag{3.6.33}$$

for *real* $z_0=x_0\sqrt{m\omega/2\hbar}$, since p_0 is zero for this initial state (the wavefunction is real).

In the ket notation, we have established that the minimal uncertainty state centered at x_0, and having zero expectation value for the momentum, is

$$|x_0,0\rangle=e^{z_0(\hat{a}^\dagger-\hat{a})}|0,0\rangle.\tag{3.6.34}$$

But it’s not exactly obvious that this is an eigenstate of $\hat{a}$ with eigenvalue z_0! (As it must be.)

It’s worth seeing how to prove that just from the properties of the operators -- but to do that, we need a couple of theorems concerning exponentials of operators given in the Appendix.

First, if the commutator $[A,B]$ commutes with A and B, then $e^{A+B}=e^Ae^Be-\frac{1}{2}[A,B]$. This result simplifies the right hand side of the above equation, for

$$\begin{aligned}e^{z_0(\hat{a}^\dagger-\hat{a})}|0,0\rangle&=e^{z_0\hat{a}^\dagger}e^{-z_0\hat{a}}e^{-z_0^2[\hat{a}^\dagger,\hat{a}]/2}|0,0\rangle\\&=e^{-z_0^2/2}e^{z_0\hat{a}^\dagger}|0,0\rangle\end{aligned}\tag{3.6.35}$$

where we have used $e^{-z_0\hat{a}}|0,0\rangle=|0,0\rangle$.

This is simpler, but it’s still not obvious that we have an eigenstate of $\hat{a}$: we need the commutator

$$[\hat{a},e^{z_0\hat{a}^\dagger}].\tag{3.6.36}$$

The second theorem we need is: if the commutator of two operators $[A,B]=c$ itself

commutes with A and B, then

$$[A,e^{\lambda B}]=\lambda ce^{\lambda B}.\tag{3.6.37}$$

(This is easily proved by expanding the exponential—see the Appendix.)

Applying this to our case,

$$[\hat{a},e^{z_0\hat{a}^\dagger}]=z_0e^{z_0\hat{a}^\dagger}.\tag{3.6.38}$$

It follows immediately that $e^{-z_0^2/2}e^{z_0\hat{a}^\dagger}|0,0\rangle$ is indeed an eigenstate of $\hat{a}$ with eigenvalue $z_0=x_0\sqrt{m\omega/2\hbar}$. (It must also be correctly normalized because the translation $|x_0,0\rangle=e^{z_0(\hat{a}^\dagger-\hat{a})}|0,0\rangle$ is a unitary operation for real z_0.)

How do we generalize this translation operator to an arbitrary state, with nonzero $\langle x\rangle$, $\langle p\rangle$? Thinking in terms of the complex parameter space z, we need to be able to move in both the x and the p directions, using both $\hat{p}=-i\hbar d/dx$ and $\hat{x}=i\hbar d/dp$. This is slightly tricky since these operators do not commute, but their commutator is just a number, so (using the theorem proved in the Appendix) this will only affect the overall normalization.

Furthermore, both $\hat{p}$ and $\hat{x}$ are combinations of $\hat{a},\hat{a}^\dagger$, so for the generalization of $e^{-i\langle x_0\rangle\hat{p}/\hbar}$ from real x_0 to complex z to be unitary, it must have an *antihermitian* combination of $\hat{a},\hat{a}^\dagger$ in the exponent -- a unitary operator has the form $U=e^{iH}$, where H is Hermitian, so iH is antihermitian.

We are led to the conclusion that

$$|\langle p\rangle, \langle x\rangle\rangle = e^{(z\hat{a}^\dagger - z^*\hat{a})}|0\rangle = |z\rangle, \tag{3.6.39}$$

conveniently labeling the coherent state using the complex parameter z of its center in phase space. Since this generalized translation operator is unitary, the new state is automatically correctly normalized.

How Do These States Relate to the Energy Eigenstates?

The equation above suggests the possibility of representing the displaced state $|z\rangle$ in the standard energy basis $|n\rangle$. We can simplify with the same trick used for the spatial displacement case in the last section, that is, the theorem $e^{A+B} = e^A e^B e - \frac{1}{2}[A, B]$ where now $A = z\hat{a}^\dagger$, $B = -z^*\hat{a}$:

$$|z\rangle = e^{z\hat{a}^\dagger - z^*\hat{a}}|0\rangle = e^{-|z|^2/2} e^{z\hat{a}^\dagger} e^{-z^*\hat{a}}|0\rangle = e^{-|z|^2/2} e^{z\hat{a}^\dagger}|0\rangle \tag{3.6.40}$$

using $e^{-z^*\hat{a}}|0\rangle = |0\rangle$ since $\hat{a}|0\rangle = 0$.

It is now straightforward to expand the exponential:

$$|z\rangle = e^{-|z|^2/2} e^{z\hat{a}^\dagger}|0\rangle = e^{-|z|^2/2}(1 + za^\dagger + (za^\dagger)^2 2! + \ldots)|0\rangle \tag{3.6.41}$$

and recalling that the normalized energy eigenstates are

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle \tag{3.6.42}$$

we find

$$|z\rangle = e^{-|z|^2/2}(|0\rangle + z|1\rangle + \frac{z^2}{\sqrt{2!}}|2\rangle + \frac{z^3}{\sqrt{3!}}|3\rangle + \ldots). \tag{3.6.43}$$

Exercise: Check that this state is correctly normalized, and is an eigenstate of $\hat{a}$.

Time Development of an Eigenstate of *a* Using the Energy Basis

Now that we have expressed the eigenstate $|z\rangle$ as a sum over the eigenstates $|n\rangle$ of the Hamiltonian, finding its time development in this representation is straightforward.

Since $|n(t)\rangle = e^{-in\omega t}|n\rangle$,

$$|z(t)\rangle = e^{-|z_0|^2/2}(|0\rangle + z_0 e^{-i\omega t}|1\rangle + \frac{z_0^2 e^{-2i\omega t}}{\sqrt{2!}}|2\rangle + \sqrt{z_0 3 e^{-3i\omega t}}\sqrt{3!}|3\rangle + \ldots) \tag{3.6.44}$$

which can be written

$$|z(t)\rangle = e^{-|z_0|^2/2} e^{z_0 e^{-i\omega t}\hat{a}^\dagger}|0\rangle, \tag{3.6.45}$$

equivalent to the result $z(t) = z_0 e^{-i\omega t}$ derived earlier.

Some Properties of the Set of Eigenstates of $\hat{a}$

In quantum mechanics, any physical variable is represented by a Hermitian operator. The eigenvalues are real, the eigenstates are orthogonal (or can be chosen to be so for degenerate states) and the eigenstates for a complete set, spanning the space, so any vector in the space can be represented in a unique way as a sum over these states.

The operator $\hat{a}$ is not Hermitian. Its eigenvalues are *all the numbers in the complex plane*. The eigenstates belonging to different eigenvalues are never orthogonal, as is immediately obvious on considering the ground state and a displaced ground state. The overlap does of course decrease rapidly for states far away in phase space.

The state overlap can be computed using $|z\rangle = e^{-|z|^2/2} e^{z\hat{a}^\dagger}|0\rangle$:

$$\langle w|z\rangle = \langle 0|e^{w^*\hat{a}} e^{-|w|^2/2} e^{-|z|^2/2} e^{z\hat{a}^\dagger}|0\rangle \tag{3.6.46}$$

and we can then switch the operators $e^{-w^*\hat{a}}, e^{z\hat{a}^\dagger}$ using the theorem from the Appendix $e^Be^A = e^Ae^Be-[A,B]$, then since $\langle 0|\hat{a}^\dagger = \hat{a}|0\rangle = 0$, we're left with

$$\langle w|z\rangle = \langle 0|e^{w^*z}e^{-|w|^2/2}e^{-|z|^2/2}|0\rangle, \tag{3.6.47}$$

from which

$$|\langle w|z\rangle|^2 = e^{-|w-z|^2}. \tag{3.6.48}$$

Finally, using $|z\rangle = e^{-|z|^2/2}\left(|0\rangle + z|1\rangle + \frac{z^2}{\sqrt{2!}}|2\rangle + \frac{z^3}{\sqrt{3!}}|3\rangle + \ldots\right)$, we can construct a unit operator using the $|z\rangle$,

$$I = \iint \frac{dxdy}{\pi}|z\rangle\langle z| \tag{3.6.49}$$

where the integral is over the whole complex plane $z = x + iy$ (this x is not, of course, the original position x, recall for the wavefunction just displaced along the axis $z_0 = x_0\sqrt{m\omega/2\hbar}$). Therefore, the $|z\rangle$ span the whole space.

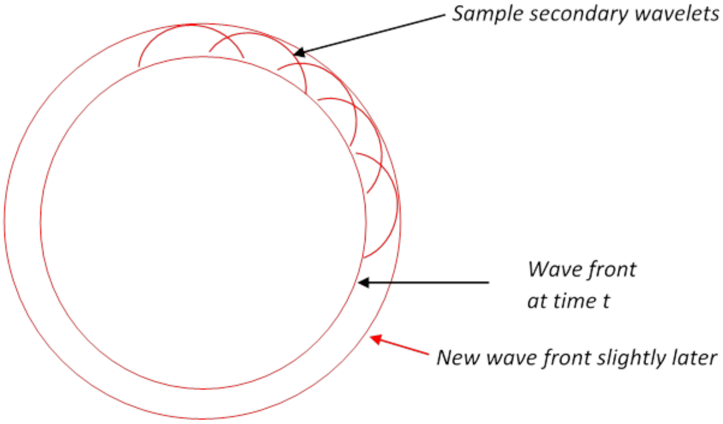
Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

3.7: Path Integrals

Huygen’s Picture of Wave Propagation

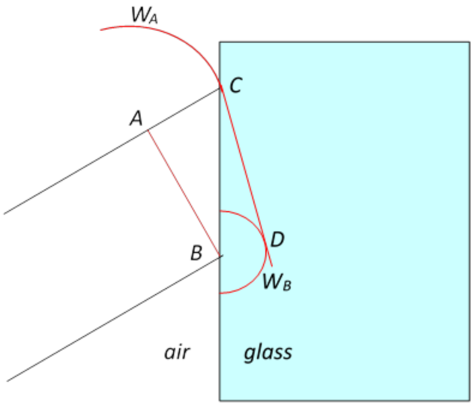
If a point source of light is switched on, the wavefront is an expanding sphere centered at the source. Huygens suggested that this could be understood if at any instant in time each point on the wavefront was regarded as a source of secondary wavelets, and the new wavefront a moment later was to be regarded as built up from the sum of these wavelets. For a light shining continuously, this process just keeps repeating.



Huygens’ picture of how a spherical wave propagates: each point on the wave front is a source of secondary wavelets that generate the new wave front.

What use is this idea? For one thing, it explains refraction—the change in direction of a wavefront on entering a different medium, such as a ray of light going from air into glass.

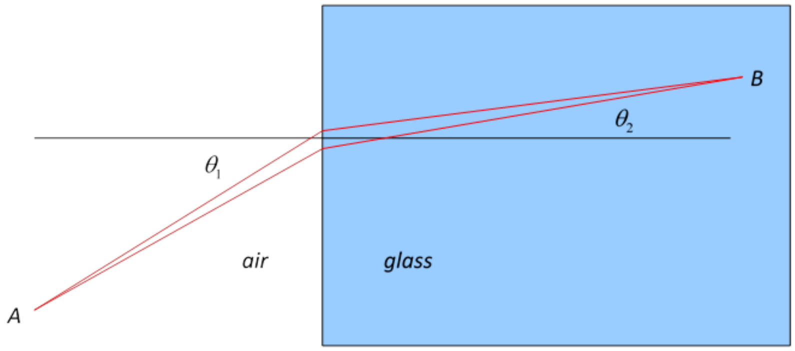
If the light moves more slowly in the glass, velocity v instead of c, with $v < c$, then Huygen’s picture explains Snell’s Law, that the ratio of the sines of the angles to the normal of incident and transmitted beams is constant, and in fact is the ratio c/v. This is evident from the diagram below: in the time the wavelet centered at A has propagated to C, that from B has reached D, the ratio of lengths AC/BD being c/v. But the angles in Snell’s Law are in fact the angles ABC, BCD, and those right-angled triangles have a common hypotenuse BC, from which the Law follows.



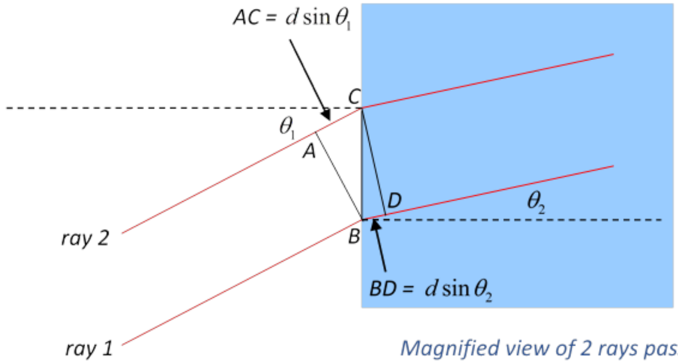
Huygens’ explanation of refraction: showing two wavelets from the wavefront AB. W_B is slowed down compared with W_A, since it is propagating in glass. This turns the wave front through an angle.

Fermat's Principle of Least Time

We will now temporarily forget about the wave nature of light, and consider a narrow ray or beam of light shining from point A to point B, where we suppose A to be in air, B in glass. Fermat showed that the path of such a beam is given by the *Principle of Least Time*: a ray of light going from A to B by any other path would take longer. How can we see that? It's obvious that any deviation from a straight line path in air or in the glass is going to add to the time taken, but what about moving slightly the point at which the beam enters the glass?



Where the air meets the glass, the two rays, separated by a small distance $CD = d$ along that interface, will look parallel:



Magnified view of 2 rays passing through interface: ray 1 is the minimum time path. Rays encounter interface distance CB = d apart.

(Feynman gives a nice illustration: a lifeguard on a beach spots a swimmer in trouble some distance away, in a diagonal direction. He can run three times faster than he can swim. What is the quickest path to the swimmer?)

Moving the point of entry up a small distance d, the light has to travel an extra $d\sin\theta_1$ in air, but a distance less by $d\sin\theta_2$ in the glass, giving an extra travel time $\Delta t = d\sin\theta_1/c - d\sin\theta_2/v$. For the *classical* path, Snell's Law gives $\sin\theta_1/\sin\theta_2 = n = c/v$, so $\Delta t = 0$ to first order. But if we look at a series of possible paths, each a small distance d away from the next at the point of crossing from air into glass, Δt becomes of order d/c away from the classical path.

Suppose now we imagine that the light actually travels along all these paths with about equal amplitude. What will be the total contribution of all the paths at B? Since the times along the paths are different, the signals along the different paths will arrive at B with different phases, and to get the total wave amplitude we must add a series of unit $2D$ vectors, one from each path. (Representing the amplitude and phase of the wave by a complex number for convenience—for a real wave, we can take the real part at the end.)

When we map out these unit $2D$ vectors, we find that in the neighborhood of the classical path, the phase varies little, but as we go away from it the phase spirals more and more rapidly, so those paths interfere amongst themselves destructively. To formulate this a little more precisely, let us assume that some close by path has a phase difference φ from the least time path, and goes from air to glass a distance x away from the least time path: then for these close by paths, $\varphi = ax^2$, where a depends on the geometric arrangement and the wavelength. From this, the sum over the close by paths is an integral of the form $\int e^{iax^2}dx$. (We are assuming the wavelength of light is far less than the size of the equipment.) This is a standard integral, its value is $\sqrt{\pi/ia}$, all its weight is concentrated in a central area of width $1/\sqrt{a}$, exactly as for the real function e^{-ax^2}.

This is the explanation of Fermat's Principle—only near the path of least time do paths stay approx_imately in phase with each other and add constructively. So this classical path rule has an underlying wave-phase explanation. In fact, the central role of phase in this analysis is sometimes emphasized by saying the light beam follows *the path of stationary phase*.

Of course, we're not summing over *all* paths here—we assume that the path in air from the source to the point of entry into the glass is a straight line, clearly the subpath of stationary phase.

Classical Mechanics: The Principle of Least Action

Confining our attention for the moment to the mechanics of a single nonrelativistic particle in a potential, with Lagrangian $L = T - V$, the action S is defined by

$$S = \int_{t_1}^{t_2} L(x, \dot{x})dt. \tag{3.7.1}$$

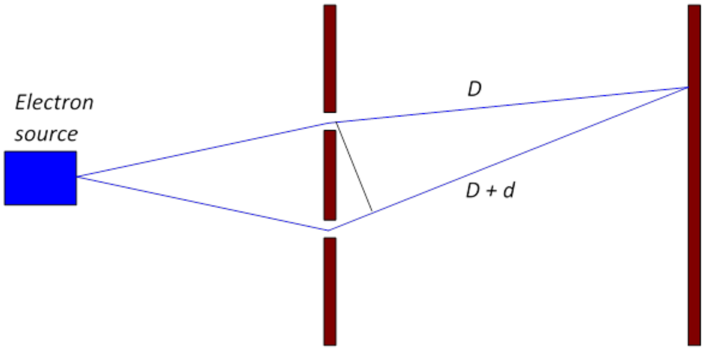
Newton's Laws of Motion can be shown to be equivalent to the statement that a particle moving in the potential from A at t_1 to B at t_2 travels along the path that minimizes the action. This is called the *Principle of Least Action*: for example, the parabolic path followed by a ball thrown through the air minimizes the integral along the path of the action $T - V$ where T is the ball's kinetic energy, V its gravitational potential energy (neglecting air resistance, of course). Note here that the initial and final times are fixed, so since we'll be summing over paths with different lengths, necessarily the particles speed will be different along the different paths. In other words, it will have different energies along the different paths.

With the advent of quantum mechanics, and the realization that any particle, including a thrown ball, has wave like properties, the rather mysterious Principle of Least Action looks a lot like Fermat's Principle of Least Time. Recall that Fermat's Principle works because the total phase along a path is the integrated time elapsed along the path, and for a path where that integral is stationary for small path variations, neighboring paths add constructively, and no other sets of paths do. If the Principle of Least Action has a similar explanation, then the wave amplitude for a particle going along a path from A to B must have a phase equal to some constant times the action along that path. If this is the case, then the observed path followed will be just that of least action, or, more generally, of *stationary* action, for only near that path will the amplitudes add constructively, just as in Fermat's analysis of light rays.

Going from Classical Mechanics to Quantum Mechanics

Of course, if we write a phase factor for a path e^{icS} where S is the action for the path and c is some constant, c must necessarily have the dimensions of inverse action. Fortunately, there is a natural candidate for the constant c. The wave nature of matter arises from quantum mechanics, and the fundamental constant of quantum mechanics, Planck's constant, is in fact a unit of action. (Recall action has the same dimensions as Et, and therefore the same as px, manifestly the same as angular momentum.) It turns out that the appropriate path phase factor is $e^{iS/\hbar}$

That the phase factor is $e^{iS/\hbar}$, rather than $e^{iS/h}$, say, can be established by considering the double slit experiment for electrons (Peskin page 277).



This is analogous to the light waves going from a source in air to a point in glass, except now we have vacuum throughout (electrons don't get far in glass), and we close down all but two of the paths.

Suppose electrons from the top slit, Path I, go a distance D to the detector, those from the bottom slit, Path II, go $D+d$, with $d \ll D$. Then if the electrons have wavelength λ we know the phase difference at the detector is $2\pi d/\lambda$. To see this from our formula for summing over paths, on Path I the action $S=Et=\frac{1}{2}mv_1^2t$, and $v_1=D/t$, so

$$S_1=\frac{1}{2}mD^2/t. \tag{3.7.2}$$

For Path II, we must take $v_2=(D+d)/t$. Keeping only terms of leading order in d/D, the action difference between the two paths

$$S_2-S_1=mDd/t \tag{3.7.3}$$

so the phase difference

$$\frac{S_2-S_1}{\hbar}=\frac{mvd}{\hbar}=\frac{2\pi pd}{h}=\frac{2\pi d}{\lambda}. \tag{3.7.4}$$

This is the known correct result, and this fixes the constant multiplying the action/*h* in the expression for the path phase.

In quantum mechanics, such as the motion of an electron in an atom, we know that the particle does not follow a well-defined path, in contrast to classical mechanics. Where does the crossover to a well-defined path take place? Taking the simplest possible case of a free particle (no potential) of mass m moving at speed v, the action along a straight line path taking time t from A to B is $\frac{1}{2}mv^2t$. If this action is of order Planck's constant h, then the phase factor will not oscillate violently on moving to different paths, and a range of paths will contribute. In other words, quantum rather than classical behavior dominates when $\frac{1}{2}mv^2t$ is of order h. But vt is the path length L, and mv/h is the wavelength λ, so we conclude that we must use quantum mechanics when the wavelength h/p is significant compared with the path length. Interference sets in when the *difference* in path actions is of order h, so in the atomic regime many paths must be included.

Feynman (in Feynman and Hibbs) gives a nice picture to help think about summing over paths. He begins with the double slit experiment for an electron. We suppose the electron is emitted from some source A on the left, and we look for it at a point B on a screen to the right. In the middle is a thin opaque barrier with the familiar two slits. Evidently, to find the amplitude for the electron to reach B we sum over two paths. Now suppose we add *another* two-slit barrier. We have to sum over four paths. Now add another. Next, replace the two slits in each barrier by several slits. We must sum over a multitude of paths! Finally, increase the number of barriers to some large number N, and at the same time increase the number of slits to the point that there are no barriers left. We are left with a sum over all possible paths through space from A to B, multiplying each path by the appropriate action phase factor. This is reminiscent of the original wave propagation picture of Huygens: if one pictures it at successive time intervals of picoseconds, say, from each point on the wavefront waves go out 3 mm in all directions, then in the next time interval each of those sprouts more waves in all directions. One could write this as a sum over all zigzag paths with random 3 mm steps.

In fact, the sum over paths is even more daunting than Feynman's picture suggests. All the paths going through these many slitted barriers are progressing in a forward direction, from A towards B. Actually, if we're summing over *all* paths, we should be including the possibility of paths zigzagging backwards and forwards as well, eventually arriving at B. We shall soon see how to deal systematically with all possible paths.

Review: Standard Definition of the Free Electron Propagator

As a warm up exercise, consider an electron confined to one dimension, with no potential present, moving from x' at time 0 to x at time T. We'll follow Feynman in using T for the final time, so we can keep t for the continuous (albeit sometimes discretized) time variable over the interval 0 to T.

(As explained previously, when we write that the electron is initially at x', we mean its wave function is a normalizable state, such as a very narrow Gaussian, centered at x'. The propagator then represents the probability amplitude, that is, the wave function, at point x after the given time T.) The propagator is given by

$$|\psi(x,t=T)\rangle=U(T)|\psi(x,t=0)\rangle, \tag{3.7.5}$$

or, in Schrödinger wave function notation,

$$\psi(x,T)=\int U(x,T;x',0)\psi(x',0)\;dx'. \tag{3.7.6}$$

It is clear that for this to make sense, as $T\to 0, U(x,T;x',0)\to\Delta(x-x')$.

In the lecture on propagators, we found

$$\langle x|U(T,0)|x'\rangle=\int_{-\infty}^{\infty}e^{-i\hbar k^2T/2m}\frac{dk}{2\pi}\langle x|k\rangle\langle k|x'\rangle=\int_{-\infty}^{\infty}e^{-i\hbar k^2T/2m}\frac{dk}{2\pi}e^{-ik(x-x')}=\sqrt{\frac{m}{2\pi\hbar iT}}e^{im(x-x')2/2\hbar T}. \tag{3.7.7}$$

Summing over Paths

Let us formulate the sum over paths for this simplest one-dimensional case, the free electron, more precisely. Each path is a continuous function of time $x(t)$ in the time interval $0\le t\le T$, with boundary conditions $x(0)=x'$, $x(T)=x$. Each path contributes a term $e^{iS/\hbar}$, where

$$S[x(t)]=\int_0^T L(x(t),\dot{x}(t))dt=\int_0^T\frac{1}{2}m\dot{x}^2(t)dt \tag{3.7.8}$$

(for the free electron case) evaluated along that path.

The integral over all paths is written:

$$\langle x|U(T,0)|x'\rangle=\int D[x(t)]\;e^{iS[x(t)]/\hbar} \tag{3.7.9}$$

This rather formal statement begs the question of how, exactly, we perform the sum over paths: what is the appropriate measure in the space of paths?

A natural approach is to measure the paths in terms of their deviation from the classical path, since we know that path dominates in the classical limit. The classical path for the free electron is just the straight line from x' to x, traversed at constant velocity, since there are no forces acting on the electron.

We write

$$x(t)=x_{cl}(t)+y(t) \tag{3.7.10}$$

where

$$x_{cl}(0)=x',\quad x_{cl}(T)=x \tag{3.7.11}$$

and therefore

$$y(0)=0,\qquad y(T)=0. \tag{3.7.12}$$

Then

$$\begin{gathered}\langle x|U(T,0)|x'\rangle=\int D[y(t)]\;e^{iS[x_{cl}(t)+y(t)]/\hbar},\\ S[x_{cl}(t)+y(t)]=\int_0^T\tfrac{1}{2}m(\dot{x}_{cl}(t)+\dot{y}(t))^2dt\\ =S[x_{cl}(t)]+\int_0^T m\dot{x}_{cl}(t)\dot{y}(t)dt+\int_0^T\tfrac{1}{2}m\dot{y}^2(t)dt.\end{gathered} \tag{3.7.13}$$

The middle term on the bottom line is zero, as it has to be since it is a linear term in the deviation from the minimum path. To see this explicitly, one can integrate by parts: the end terms are zero, from the boundary condition on y, and the other term is the acceleration of the particle along the classical path, which is zero.

Therefore

$$\langle x|U(T,0)|x'\rangle=e^{iS[x_{cl}(t)]/\hbar}\int D[y(t)]\;e^{iS[y(t)]/\hbar} \tag{3.7.14}$$

The y- paths, being the deviation from the classical path from x' to x, necessarily begin and end at the y- origin, since *all* paths summed over go from x' to x.

The classical path, motion from x' to x at a constant speed $v=(x'-x)/T$, has action Et, with E the classical energy $\frac{1}{2}mv^2$, so

$$U(x,T;x',0)=A(T)e^{im(x-x')2/2\hbar T}. \tag{3.7.15}$$

This gives the correct exponential term. The prefactor A, representing the sum over the deviation paths y(t), cannot depend on x or x', and is fixed by the requirement that as t goes to zero, U must approach a δ- function, giving the prefactor found previously.

Proving that the Sum-Over-Paths Definition of the Propagator is Equivalent to the Sum-Over-Eigenfunctions Definition

The first step is to construct a practical method of summing over paths. Let us begin with a particle in one dimension going from x' at time 0 to x at time T. The paths can be enumerated in a crude way, reminiscent of Riemann integration: divide the time interval 0 to T into N equal intervals each of duration \varepsilon, so $t_0=0,\; t_1=t_0+\varepsilon,\; t_2=t_0+2\varepsilon,\ldots,\; t_N=T$.

Next, define a particular path from x to x' by specifying the position of the particle at each of the intermediate times, that is to say, it is at x_1 at time t_1, x_2 at time t_2 and so on. Then, simplify the path by putting in straight line bits connecting x_0 to x_1, x_1 to x_2, etc. The justification is that in the limit of ε going to zero, taken at the end, this becomes a true representation of the path.

The next step is to sum over all possible paths with a factor $e^{iS/\hbar}$ for each one. The sum is accomplished by integrating over all possible values of the intermediate positions $x_1,x_2,\ldots,x_{N-1}$, and then taking N to infinity.

The action on the zigzag path is

$$S=\int_0^T dt(\frac{1}{2}m\dot{x}^2-V(x))\rightarrow\sum_i\left[\frac{m(x_i+1-x_i)^2}{2\varepsilon}-\varepsilon\,V(\frac{x_i+1+x_i}{2})\right] \tag{3.7.16}$$

We define the “integral over paths” written $\int D[x(t)]$ by

$$\lim_{\substack{\varepsilon\to 0\\ N\to\infty}}\frac{1}{B(\varepsilon)}\int_{-\infty}^{\infty}\int\ldots\int\frac{dx_1}{B(\varepsilon)}\ldots\frac{dx_{N-1}}{B(\varepsilon)} \tag{3.7.17}$$

where we haven’t yet figured out what the overall weighting factor $B(\varepsilon)$ is going to be. (It is standard convention to have that extra $B(\varepsilon)$ outside.)

To summarize: the propagator $U(x,T;x',0)$ is the contribution to the wave function at x at time

$t=T$ from that at x' at the earlier time t=0.

Consequently, $U(x,T;x',0)$ regarded as a function of x,T is, in fact, nothing but the Schrödinger wave function $\psi(x,T)$, and therefore must *satisfy Schrödinger’s equation*

$$i\hbar\frac{\partial}{\partial T}U(x,T;x',0)=\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V(x)\right)U(x,T;x',0). \tag{3.7.18}$$

We shall now show that defining $U(x,T;x',0)$ as a sum over paths, it does in fact satisfy Schrödinger’s equation, and furthermore goes to a δ- function as time goes to zero.

$$U(x,T;x',0)=\int D[x(t)]e^{iS[x(t)]/\hbar}=\lim_{\substack{\varepsilon\to 0\\ N\to\infty}}\frac{1}{B(\varepsilon)}\int_{-\infty}^{\infty}\int\ldots\int\frac{dx_1}{B(\varepsilon)}\ldots\frac{dx_{N-1}}{B(\varepsilon)}e^{iS(x_1,\;\ldots\;,x_{N-1})/\hbar}. \tag{3.7.19}$$

We shall establish this equivalence by proving that it satisfies the same differential equation. It clearly has the same initial value—as $t\,'$ and t coincide, it goes to $\delta(x-x\,')$ in both representations.

To differentiate $U(x,T;x',0)$ with respect to t, we isolate the integral over the last path variable, x_{N-1}:

$$U(x,T;x',0)=\int\frac{dx_{N-1}}{B(\varepsilon)}e^{\left[\frac{im(x-x_{N-1})^2}{2\hbar\varepsilon}-\frac{i}{\hbar}\varepsilon V(\frac{x+x_{N-1}}{2})\right]}U(x_{N-1},T-\varepsilon;x',0) \tag{3.7.20}$$

Now in the limit ε going to zero, almost all the contribution to this integral must come from close to the point of stationary phase, that is, $x_{N-1}=x$. In that limit, we can take $U(x_{N-1},t-\varepsilon;x',t\,')$ to be a slowly varying function of x_{N-1}, and replace it by the leading terms in a Taylor expansion about x, so

$$U(x,T;x',0) = \int \frac{dx_{N-1}}{B(\varepsilon)} e^{\frac{im(x-x_{N-1})^2}{2\hbar\varepsilon}} \left(1-\frac{i}{\hbar}\varepsilon V\left(\frac{x+x_{N-1}}{2}\right)\right)\left(U(x,T-\varepsilon)+(x_{N-1}-x)\frac{\partial U}{\partial x}+\frac{(x_{N-1}-x)^2}{2}\frac{\partial^2 U}{\partial x^2}\right) \tag{3.7.21}$$

The x_{N-1} dependence in the potential V can be neglected in leading order—that leaves standard Gaussian integrals, and

$$U(x,T;x',0) = \frac{1}{B(\varepsilon)}\sqrt{\frac{2\pi\hbar\varepsilon}{-im}}\left(1-\frac{i\varepsilon}{\hbar}V(x)+\frac{i\varepsilon\hbar}{2m}\frac{\partial^2}{\partial x^2}\right)U(x,T-\varepsilon;x',0). \tag{3.7.22}$$

Taking the limit of ε going to zero fixes our unknown normalizing factor,

$$B(\varepsilon) = \sqrt{\frac{2\pi\hbar\varepsilon}{-im}} \tag{3.7.23}$$

giving

$$i\hbar\frac{\partial}{\partial T}U(x,T;x',0) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V(x)\right)U(x,T;x',0), \tag{3.7.24}$$

thus establishing that the propagator derived from the sum over paths obeys Schrödinger’s equation, and consequently gives the same physics as the conventional approach.

Explicit Evaluation of the Path Integral for the Free Particle Case

The required correspondence to the Schrödinger equation result fixes the unknown normalizing factor, as we’ve just established. This means we are now in a position to evaluate the sum over paths explicitly, at least in the free particle case, and confirm the somewhat hand-waving result given above.

The sum over paths is

$$U(x,T;x',0) = \int D[x(t)]e^{iS[x(t)]/\hbar} = \lim_{\substack{\varepsilon\to 0 \\ N\to\infty}} 1B(\varepsilon)\int_{-\infty}^{\infty}\int\ldots\int\frac{dx_1}{B(\varepsilon)}\ldots\frac{dx_{N-1}}{B(\varepsilon)}e^{i\sum_i\frac{im(x_i+1-x_i)^2}{2\hbar\varepsilon}}. \tag{3.7.25}$$

Let us consider the sum for small but finite ε. In particular, we’ll divide up the interval first into halves, then quarters, and so on, into 2^n small intervals. The reason for this choice will become clear.

Now, we’ll integrate over half the paths: those for i odd, leaving the even x_i values fixed for the moment. The integrals are of the form

$$\begin{aligned} &\int_{-\infty}^{\infty} dye^{(ia/2)[(x-y)^2+(y-z)^2]} = e^{(ia/2)(x^2+z^2)}\int_{-\infty}^{\infty} dye^{iay^2-iay(x+z)} \\ &= e^{(ia/2)(x^2+z^2)}\sqrt{\frac{\pi}{-ia}}e^{(-ia/4)(x+z)^2} = \sqrt{\frac{\pi}{-ia}}e^{(ia/4)(x-z)^2} \end{aligned} \tag{3.7.26}$$

using the standard result $\int_{-\infty}^{\infty} dxe^{-ax^2+bx} = \sqrt{\frac{\pi}{a}}e^{b^2/4a}$.

Now put in the value $a = m/\hbar\varepsilon$: the factor $\sqrt{\frac{\pi}{-ia}} = \sqrt{\frac{\pi\hbar\varepsilon}{-im}}$ cancels the normalization factor $B(\varepsilon) = \sqrt{\frac{2\pi\hbar\varepsilon}{-im}}$ except for the factor of 2 inside the square root. But we need that factor of 2, because we’re left with an integral—over the remaining even numbered paths—exactly like the one before except that the time interval has doubled, both in the normalization factor and in the exponent, $\varepsilon\to 2\varepsilon$.

So we’re back where we started. We can now repeat the process, halving the number of paths again, then again, until finally we have the same expression but with only the fixed endpoints appearing.

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

3.8: Path Integrals for the SHO

Hand-waving Argument

Recall that the free particle propagator has the form

U(x,T;x′,0)=\sqrt{\frac{m}{2\pi \hbar iT}}\exp \left{ \frac{im(x-x′)^2}{2\hbar T} \right}. \tag{3.8.1}

From a classical mechanical evaluation of the action, we can show the simple harmonic oscillator propagator has the form

$$U(x,T;x',0)=A(T)expim\omega 2\hbar sin\omega T[(x2+x'2)cos\omega T-2xx'] \tag{3.8.1}$$

where A(T) is a so far unknown function of time. For sufficiently small T, though, much less than the period of the oscillator, the potential will have no significant effect, so the propagator must tend to the free particle propagator. That is, for T tending to zero, A(T)=m/2\pi \hbar iT--------√. But we also know that at time t=2\pi /\omega , all simple harmonic wavefunctions return to their t=0 values, so a particle localized at x' will be localized at x' again after one period. In the exponent we do indeed see this cyclic behavior. In the prefactor, we must replace T by (sin\omega T)/\omega , that is,

$$A(T)=\sqrt{\frac{m\omega}{2\pi\hbar i\sin\omega T}}. \tag{3.8.3}$$

Note this expression is also consistent with the free particle propagator in the limit \omega going to zero, in other words, the vanishing of the simple harmonic oscillator potential.

Approximating Integrals by Stationary Phase Techniques

Actually, this result can be derived from the integral over the fluctuations about the classical path. The argument is closely analogous to that for the free particle, and the following equation is a straightforward generalization of that case (discussed in the previous lecture):

$$\langle x|U(T,0)|x'\rangle=\int D[y(t)]\ eiS[x_{cl}(t)+y(t)]/\hbar,\quad S[x_{cl}(t)+y(t)]= \tag{3.8.2}$$

$$\int_0^T 12m[(\dot{x_{cl}}(t)+\dot{y}(t))2-\omega^2(x_{cl}(t)+y(t))2]dt=S[x_{cl}(t)]+\int_0^T m(\dot{x_{cl}}(t)\dot{y}(t)-\omega^2x_{cl}(t)y(t))dt+$$

$$\int_0^T 12m(\dot{y}2(t)-\omega^2y2(t))dt.$$

Just as for the free particle, the middle term in that last line is zero, as it has to be since it’s the linear term in the path deviation from the stationary point. As before, the y-path (the deviation from the classical least action path) is zero at the two endpoints, so we can integrate by parts to get

$$\int_0^T m(\dot{x_{cl}}(t)\dot{y}(t)-\omega^2x_{cl}(t)y(t))dt=\int_0^T m(-x\ddot{}cl(t)-\omega^2x_{cl}(t))y(t)dt \tag{3.8.3}$$

identically zero from the classical equation of motion (after all, this is how it’s derived from least action!)

The hard part is evaluating the integral over paths in the third term. To see how to do this,

it’s worth briefly reviewing a stationary phase integral in ordinary space, that is,

$$I=\int_{-\infty}^{\infty}dx1\int_{-\infty}^{\infty}dx2...\int_{-\infty}^{\infty}dxNexp\lceil\lfloor\{\ |\ |\ i\ f(x1,\ldots,xN)\rceil\rfloor\}\ |\ | \tag{3.8.4}$$

where we assume that the real function f has an absolute minimum at the point (x01,…,x0N) and take variables yi=xi-x0i. Then to leading order near the stationary point, f(x)=f(x0)+12yiAijyj. Taking only this term, the integral becomes

$$I=e^{if(x^0)}\int_{-\infty}^{\infty}dy1\int_{-\infty}^{\infty}dy2...\int_{-\infty}^{\infty}dyNexp\lceil\lfloor\{\ |\ |\ 12iyiAijyj\rceil\rfloor\}\ |\ |. \tag{3.8.5}$$

The matrix A is positive definite and symmetric, it has real orthogonal eigenvectors, and we can choose our coordinate axes in the directions of those eigenvectors. In this case, Ais a diagonal matrix, and its diagonal elements are just its eigenvalues \lambda i.

So I becomes:

I=e^{if(x^0)}\int_{-\infty}^{\infty}dy1\int_{-\infty}^{\infty}dy2...\int_{-\infty}^{\infty}dyNexp⌈⌊{||12i\lambda i y2i⌉⌋}||=e^{if(x^0)}∏1N2\pi i\lambda i---√=e^{if(x^0)}(2\pi i)N/2(DetA)-1/2.

Recall that Det A is invariant under an orthogonal transformation, and so is unaffected by transforming back from the eigenvector basis to the original coordinates.

Applying Stationary Phase to the Integral Over Paths

We're now ready to generalize this result from an integral over an N-dimensional space of yi 's to an integral over the infinite dimensional space of paths y(t) with boundary conditions y(0)=0 and y(t)=0. Note that for the particular case at hand, the Simple Harmonic Oscillator, the leading order quadratic term is in fact the whole story, and so gives the exact result. This is not, of course, true for general quantum systems.

As a preliminary step, we integrate by parts to write

$$\int_0^T (dydt)2dt = -\int_0^T y(t)\frac{d^2}{dt^2}y(t). \tag{3.8.6}$$

The path integral then becomes:

$$A(T) = \int 00exp\lceil\lfloor| i\hbar \int_0^T y(t)[12m(-\frac{d^2}{dt^2}-\omega^2)]y(t)\; dt\rceil\rfloor|\; D[y(t)]. \tag{3.8.7}$$

Comparing this with the finite-dimensional version, we see that the N- component vector (yi) is replaced by the continuous function y(t), and the matrix A by a differential operator acting on the space of path functions y(t) equal to zero at t=0 and t=T.

The operator is well-defined and Hermitian, with eigenstates and eigenvalues:

$$(-\frac{d^2}{dt^2}-\omega^2)sin(\frac{n\pi t}{T}) = (\frac{n^2\pi^2}{T^2}-\omega^2)sin(\frac{n\pi t}{T}). \tag{3.8.8}$$

Any reasonable path satisfying the boundary conditions can be written as a sum over these eigenstates with Fourier coefficients y_n .

$$y(t) = \sum n = 1\infty y_n 12-\sqrt{}sin(\frac{n\pi t}{T}). \tag{3.8.9}$$

The functional integral over all paths now becomes an infinite product of ordinary integrals over the variables y_n . This is precisely analogous to the finite integral above, except that there is no maximum value N. The operator is diagonal with respect to this base formed of its eigenstates, and therefore can be represented as an infinite diagonal matrix with the diagonal elements equal to the eigenvalues—and, just as for the finite matrix case, the gaussian integrals can be carried out in succession, each one giving a term

$$\int_{-\infty}^{\infty} dy_n expim2\hbar\lambda_n y2n = 2\hbar\pi im\lambda_n -----\sqrt{}\;\; , where\lambda_n = (\frac{n^2\pi^2}{T^2}-\omega^2). \tag{3.8.10}$$

When we take the infinite product of all these terms, we find in the denominator the square root of the product of all the eigenvalues—and that product is the determinant of the operator. That is to say,

$$[Det(-\frac{d^2}{dt^2}-\omega^2)]-1/2 = \prod_{n=1}^{\infty}(\frac{n^2\pi^2}{T^2}-\omega^2)-1/2 = K(T)\prod_{n=1}^{\infty}(1-\omega^2T2n2\pi2)-1/2 = K(T)(sin\omega T\omega T)-1/2. \tag{3.8.11}$$

In the last step, we used the infinite-product representation for the sine function:

$$sinx = x\prod_{n=1}^{\infty}(1-x2n2\pi2). \tag{3.8.12}$$

The rather disconcerting factor K(T), which does not depend on \omega and hence has nothing to do with the dy_n amics, can be thrown in with similar time-dependent and constant factors involved in the measure of the integral over paths, and their overall contribution can be nailed down by the simple observation that for \omega =0, we must recover the known free-particle propagator. So, we can be quite careless about overall multiplying factors!

Therefore, matching to the free-particle propagator in the limit \omega =0, we find A(T)=m\omega /2\pi \hbar isin\omega T-------------√ for the simple harmonic oscillator.

I used an excellent book, Solitons and Instantons, by R. Rajaraman, for this section. The original work was all by Feynman.

3.9: Appendix- Some Exponential Operator Algebra

Suppose that the commutator of two operators A,B

$$[A,B]=c, \tag{3.9.1}$$

where c commutes with A and B, usually it’s just a number, for instance 1 or $i\hbar$.

Then

$$[A,e^{\lambda B}] \;=\left[A,1+\lambda B+\left(\frac{\lambda^2}{2!}\right)B^2+\left(\frac{\lambda^3}{3!}\right)B^3+\ldots\right] \tag{3.9.2}$$

$$=\lambda c+\left(\frac{\lambda^2}{2!}\right)2Bc+\left(\frac{\lambda^3}{3!}\right)3B^2c+\ldots \tag{3.9.3}$$

$$=\lambda ce^{\lambda B}. \tag{3.9.4}$$

That is to say, the commutator of A with $e^{\lambda B}$ is proportional to $e^{\lambda B}$ itself. That is reminiscent of the simple harmonic oscillator commutation relation $[H,a^\dagger]=\hbar\omega a^\dagger$ which led directly to the ladder of eigenvalues of H separated by $\hbar\omega$. Will there be a similar “ladder” of eigenstates of A in general?

Assuming A (which is a general operator) has an eigenstate $|a\rangle$ with eigenvalue a,

$$A|a\rangle=a|a\rangle. \tag{3.9.5}$$

Applying $[A,e^{\lambda B}]=\lambda ce^{\lambda B}$ to the eigenstate $|a\rangle$:

$$Ae^{\lambda B}|a\rangle=e^{\lambda B}A|a\rangle+\lambda ce^{\lambda B}|a\rangle=(a+\lambda c)|a\rangle. \tag{3.9.6}$$

Therefore, unless it is identically zero, $e^{\lambda B}|a\rangle$ is *also* an eigenstate of A, with eigenvalue $a+\lambda c$. We conclude that instead of a *ladder* of eigenstates, we can apparently generate a whole *continuum* of eigenstates, since λ can be set arbitrarily!

To find more operator identities, premultiply $[A,e^{\lambda B}]=\lambda ce^{\lambda B}$ by $e^{-\lambda B}$ to find:

$$\begin{aligned} e^{-\lambda B}Ae^{\lambda B} &=A+\lambda[A,B] \\ &=A+\lambda c. \end{aligned}$$

This identity is *only* true for operators A,B whose commutator c is a number. (Well, c *could* be an operator, provided it still commutes with both A and B).

Our next task is to establish the following very handy identity, which is also only true if $[A,B]$ commutes with A and B:

$$e^{A+B}=e^{A}e^{B}e-\frac{1}{2}[A,B]. \tag{3.9.7}$$

The proof is as follows:

Proof

Take $f(x)=e^{Ax}e^{Bx}$,

$$\begin{aligned} \frac{df}{dx} &=Ae^{Ax}e^{Bx}+e^{Ax}e^{Bx}B \\ &=f(x)(e^{-Bx}Ae^{Bx}+B) \\ &=f(x)(A+x[A,B]+B). \end{aligned}$$

It is easy to check that the solution to this first-order differential equation equal to one at $x=0$ is

$$f(x)=e^{x(A+B)}e^{\frac{1}{2}x^2[A,B]}$$

so taking $x=1$ gives the required identity,

$$e^{A+B}=e^{A}e^{B}e^{-\frac{1}{2}[A,B]}.$$

□

It also follows that $e^B e^A = e^A e^B e^{-[A,B]}$ provided—as always—that $[A, B]$ commutes with A and B.

Contributor

- Michael Fowler (Beams Professor, Department of Physics, University of Virginia)

CHAPTER OVERVIEW

4: Angular Momentum, Spin and the Hydrogen Atom

4.1: Angular Momentum Operator Algebra
4.2: Orbital Eigenfunctions- 2-D Case
4.3: Note on Curvilinear Coordinates
4.4: Orbital Eigenfunctions in 3-D
4.5: Spin
4.6: The Hydrogen Atom
4.7: Adding Angular Momenta
4.8: Tensor Operators

Thumbnail: Hydrogen atom. (Public Domain; Bensaccount).

4.1: Angular Momentum Operator Algebra

Preliminaries: Translation and Rotation Operators

As a warm up to analyzing how a wave function transforms under rotation, we review the effect of *linear translation* on a single particle wave function $\psi(x)$. We have already seen an example of this: the coherent states of a simple harmonic oscillator discussed earlier were (at $t=0$) identical to the ground state except that they were centered at some point displaced from the origin. In fact, the operator creating such a state from the ground state is a translation operator.

The *translation operator* $T(a)$ is *defined* at that operator which when it acts on a wave function ket $|\psi(x)\rangle$ gives the ket corresponding to that wave function moved over by a, that is,

$$T(a)|\psi(x)\rangle = |\psi(x-a)\rangle, \tag{4.1.1}$$

so, for example, if $\psi(x)$ is a wave function centered at the origin, $T(a)$ moves it to be centered at the point a.

We have written the wave function as a ket here to emphasize the parallels between this operation and some later ones, but it is simpler at this point to just work with the wave function as a function, so we will drop the ket bracket for now. The form of $T(a)$ as an operator on a function is made evident by rewriting the Taylor series in operator form:

$$\psi(x-a) = \psi(x) - a\frac{d}{dx}\psi(x) + \frac{a^2}{2!}\frac{d^2}{dx^2}\psi(x) - \cdots = e^{-a\dfrac{d}{dx}}\psi(x) = T(a)\psi(x). \tag{4.1.2}$$

Now for the quantum connection: the differential operator appearing in the exponential is in quantum mechanics proportional to the momentum operator ($\hat{p} = -i\hbar d/dx$) so the translation operator

$$T(a) = e^{-ia\hat{p}/\hbar}. \tag{4.1.3}$$

An important special case is that of an infinitesimal translation,

$$T(\varepsilon) = e^{-i\varepsilon\hat{p}/\hbar} = 1 - i\varepsilon\hat{p}/\hbar. \tag{4.1.4}$$

The momentum operator $\hat{p}$ is said to be the *generator* of the translation.

> Shankar writes (page 281) $T(\varepsilon)|x\rangle = |x+\varepsilon\rangle$. Here $|x\rangle$ denotes a delta-function type wave function centered at x. It might be better if he had written $T(\varepsilon)|x_0\rangle = |x_0+\varepsilon\rangle$, then we would see right away that this translates into the wave function transformation $T(\varepsilon)\delta(x-x_0) = \delta(x-x_0-\varepsilon)$, the sign of ε now obviously consistent with our usage above.)

It is important to be clear about whether the *system* is being translated by a, as we have done above or whether, alternately, the *coordinate axes* are being translated by a, that latter would result in the *opposite* change in the wave function. Translating the coordinate axes, along with the apparatus and any external fields by $-a$ relative to the wave function would of course give the same physics as translating the wave function by $+a$. In fact, these two equivalent operations are analogous to the time development of a wave function being described either by a Schrödinger picture, in which the bras and kets change in time, but not the operators, and the Heisenberg picture in which the operators develop but the bras and kets do not change. To pursue this analogy a little further, in the "Heisenberg" case

$$\hat{x} \to T^{-1}(\varepsilon)\hat{x}T(\varepsilon) = e^{i\varepsilon\hat{p}/\hbar}\hat{x}e^{-i\varepsilon\hat{p}/\hbar} = \hat{x} + i\varepsilon[\hat{p},\hat{x}]/\hbar = \hat{x} + \varepsilon \tag{4.1.5}$$

and $\hat{p}$ is unchanged since it commutes with the operator. So there are two possible ways to deal with translations: transform the bras and kets, *or* transform the operators. We shall almost always leave the operators alone, and transform the bras and kets.

We have established that *the momentum operator is the generator of spatial translations* (the generalization to three dimensions is trivial). We know from earlier work that the Hamiltonian is the generator of *time* translations, by which we mean

$$\psi(t+a) = e^{-iHa/\hbar}\psi(t). \tag{4.1.6}$$

It is tempting to conclude that the *angular momentum* must be the operator generating *rotations* of the system, and, in fact, it is easy to check that this is correct. Let us consider an infinitesimal rotation $\delta\vec{\theta}$ about some axis through the origin (the infinitesimal vector being in the direction of the axis). A wavefunction $\psi(\vec{r})$initially localized at $\overrightarrow{r_0}$ will shift to be localized at $\overrightarrow{r_0} + \delta\overrightarrow{r_0}$, where $\delta\overrightarrow{r_0} = \delta\vec{\theta} \times \overrightarrow{r_0}$. So, how does a wave function transform under this small rotation? Just as for the translation case, $\psi(\vec{r}) \to \psi(\vec{r} - \delta\vec{r})$. If you don't understand the minus sign, reread the discussion on translations and the sign of ε.

Thus

$$\psi(\vec{r}) \to \psi(\vec{r}) - \frac{i}{\hbar}\delta\vec{r}.\hat{\vec{p}}\psi(\vec{r}) \tag{4.1.7}$$

to first order in the infinitesimal quantity, so the rotation operator

$$R(\delta\vec{\theta})\psi(\vec{r}) = (1 - \frac{i}{\hbar}\delta\vec{\theta} \times \vec{r}.\hat{\vec{p}})\psi(\vec{r}) = (1 - \frac{i}{\hbar}\delta\vec{\theta}.\vec{r} \times \hat{\vec{p}})\psi(\vec{r}) = (1 - \frac{i}{\hbar}\delta\vec{\theta}.\hat{\vec{L}})\psi(\vec{r}). \tag{4.1.8}$$

If we write this as

$$R(\delta\vec{\theta})\psi(\vec{r}) = e^{-\frac{i}{\hbar}\delta\vec{\theta}.\hat{\vec{L}}}\psi(\vec{r}) \tag{4.1.9}$$

it is clear that a finite rotation is given by multiplying together a large number of these operators, which just amounts to replacing $\delta\vec{\theta}$ by $\vec{\theta}$ in the exponential. Another way of going from the infinitesimal rotation to a full rotation is to use the identity

$$\lim_{N\to\infty}(1 + \frac{A\theta}{N})^N = e^{A\theta} \tag{4.1.10}$$

which is clearly valid even if A is an operator.

We have therefore established that the orbital angular momentum operator $\hat{\vec{L}}$ is the generator of spatial rotations, by which we mean that if we rotate our apparatus, and the wave function with it, the appropriately transformed wave function is generated by the action of $R(\vec{\theta})$ on the original wave function. It is perhaps worth giving an explicit example: suppose we rotate the system, and therefore the wave function, through an infinitesimal angle $\delta\theta_z$ about the z -axis. Denote the rotated wave function by $\psi_{rot}(x, y)$. Then

$$\begin{aligned} \psi_{rot}(x,y) = (1 - \frac{i}{\hbar}(\delta\theta_z)\hat{L}_z)\psi(x,y) = (1 - \frac{i}{\hbar}(\delta\theta_z)(-i\hbar(x\frac{d}{dy} - y\frac{d}{dx})))\psi(x,y) = (1 - (\delta\theta_z)(x\frac{d}{dy} - y\frac{d}{dx}))\psi(x,y) \\ = \psi(x + (\delta\theta_z)y,\ y - (\delta\theta_z)x). \end{aligned} \tag{4.1.11}$$

That is to say, the value of the new wave function at (x, y) is the value of the old wave function at the point which was rotated into (x, y).

Quantum Generalization of the Rotation Operator

However, it has long been known that in quantum mechanics, orbital angular momentum is *not* the whole story. Particles like the electron are found experimentally to have an internal angular momentum, called spin. In contrast to the spin of an ordinary macroscopic object like a spinning top, the electron's spin is *not* just the sum of orbital angular momenta of internal parts, and any attempt to understand it in that way leads to contradictions.

To take account of this new kind of angular momentum, we generalize the orbital angular momentum $\hat{\vec{L}}$ to an operator $\hat{\vec{J}}$ which is *defined* as the generator of rotations on *any* wave function, including possible spin components, so

$$R(\delta\vec{\theta})\psi(\vec{r}) = e^{-\frac{i}{\hbar}\delta\vec{\theta}.\hat{\vec{J}}}\psi(\vec{r}). \tag{4.1.12}$$

This is of course identical to the equation we found for $\hat{\vec{L}}$, but there we derived if from the quantum angular momentum operator including the momentum components written as differentials. But up to this point $\psi(\vec{r})$ has just been a complex valued function of position. From now on, the wave function at a point can have *several components*, so it is in some vector space, and the rotation operator will operate in this space as well as being a differential operator with respect to position. For example, the wave function could be a vector at each point, so rotation of the system could rotate this vector as well as moving it to a different $\vec{r}$.

To summarize: $\psi(\vec{r})$ is in general an n -component function at each point in space, $R(\delta\vec{\theta})$ is an $n \times n$ matrix in the component space, and the above equation is the *definition* of $\hat{\vec{J}}$. Starting from this definition, we will find $\hat{\vec{J}}$'s properties.

The first point to make is that in contrast to translations, rotations do not commute even for a classical system. Rotating a book through $\pi/2$ first about the z -axis then about the x -axis leaves it in a different orientation from that obtained by rotating from the same starting position first $\pi/2$ about the x -axis then $\pi/2$ about the z -axis. Even small rotations do not commute, although the commutator is second order. Since the R-operators are representations of rotations, they will reflect this commutativity structure, and we can see just how they do that by considering ordinary classical rotations of a real vector in three-dimensional space.

The matrices rotating a vector by θ about the x,y and z axes are

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \quad R_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \quad R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.1.13}$$

In the limit of rotations about infinitesimal angles (ignoring higher order terms),

$$R_x(\varepsilon)=1+\varepsilon\begin{pmatrix}0&0&0\\0&0&-1\\0&1&0\end{pmatrix},\quad R_y(\varepsilon)=1+\varepsilon\begin{pmatrix}0&0&1\\0&0&0\\-1&0&0\end{pmatrix},\quad R_z(\varepsilon)=1+\varepsilon\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}. \tag{4.1.14}$$

It is easy to check that

$$[R_x(\varepsilon),R_y(\varepsilon)]=\varepsilon^2\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}=R_z(\varepsilon^2)-1. \tag{4.1.15}$$

The rotation operators on quantum mechanical kets must, like all rotations, follow this same pattern, that is, we must have

$$((1-\frac{i}{\hbar}\varepsilon J_x)(1-\frac{i}{\hbar}\varepsilon J_y)-(1-\frac{i}{\hbar}\varepsilon J_y)(1-\frac{i}{\hbar}\varepsilon J_x)+\frac{i}{\hbar}\varepsilon^2 J_z)|\psi\rangle=0 \tag{4.1.16}$$

where we have used the definition of the infinitesimal rotation operator on kets, $R(\delta\vec{\theta})\psi(\vec{r})=e^{-\frac{i}{\hbar}\delta\vec{\theta}.\hat{\vec{J}}}\psi(\vec{r})$. The zeroth and first-order terms in ε all cancel, the second-order term gives $[J_x,J_y]=i\hbar J_z$. The general statement is:

$$[J_i,J_j]=i\hbar\varepsilon_{ijk}J_k \tag{4.1.17}$$

This is one of the most important formulas in quantum mechanics.

Consequences of the Commutation Relations

The commutation formula $[J_i,J_j]=i\hbar\varepsilon_{ijk}J_k$, which is, after all, a straightforward extension of the result for ordinary classical rotations, has surprisingly far-reaching consequences: it leads directly to the directional quantization of spin and angular momentum observed in atoms subject to a magnetic field.

It is by now very clear that in quantum mechanical systems such as atoms the total angular momentum, and also the component of angular momentum in a given direction, can only take certain values. Let us try to construct a basis set of angular momentum states for a given system: a complete set of kets corresponding to all allowed values of the angular momentum. Now, angular momentum is a *vector* quantity: it has magnitude and direction. Let's begin with the magnitude, the natural parameter is the length squared:

$$J^2=J_x^2+J_y^2+J_z^2. \tag{4.1.18}$$

Now we must specify direction -- but here we run into a problem. J_x,J_y and J_z are all mutually non-commuting, so we cannot construct a set of common eigenkets of any two of them, which we would need for a precise specification of direction. They *do* all commute with J^2, since it is spherically symmetric and therefore cannot be affected by any rotation (and, it's easy to check this commutation explicitly).

The bottom line, then, is that in attempting to construct eigenkets describing the different possible angular momentum states of a quantum system, the best we can do is to find the common eigenkets of J^2 and *one* direction, say J_z. The commutation relations do not allow us to be more precise about direction, analogous to the Uncertainty Principle for position and momentum, which also comes from noncommutativity of the relevant operators.

We conclude that the appropriate angular momentum basis is the set of common eigenkets of the commuting Hermitian matrices J^2,J_z :

$$J^2|a,b\rangle=a|a,b\rangle J_z|a,b\rangle=b|a,b\rangle. \tag{4.1.19}$$

Our next task is to find the allowed values of a and b.

> *The Uncertainty Principle limits our knowledge about the direction of angular momenta with the best we can do is to find the common eigenkets of J^2 and one direction, say J_z.*

4.2: Orbital Eigenfunctions- 2-D Case

Previously, we established that the operators $\vec{J}^2$, J_z have a common set of eigenkets ($|j,m\rangle$ with:

$$\vec{J}^2|j,m\rangle = j(j+1)\hbar^2|j,m\rangle \tag{4.2.1}$$

and

$$J_z|j,m\rangle = m\hbar|j,m\rangle \tag{4.2.2}$$

where j, m are integers or half odd integers, and we found the matrix elements of J_+, J_- (and hence those of J_x, J_y) between these eigenkets. This purely formal structure, therefore, nails down the allowed values of total angular momentum and of any measured component. But there are other things we need to know: for example, how is an electron in a particular angular momentum state in an atom affected by an external field? To compute that, we need to know the wave function ψ.

If a system has spherical symmetry, such as an electron in the Coulomb field of a hydrogen nucleus, then the Hamiltonian H and the operators $\vec{J}^2$, J_z have a common set of eigenkets $|E,j,m\rangle$. The spherically symmetric Hamiltonian is unchanged by rotation, so must commute with any rotation operator, $[H,\vec{J}^2]=0$ and $[H,J_z]=0$. Recall that commuting Hermitian operators can be diagonalized simultaneously -- and therefore have a common set of eigenkets.

Fortunately, many systems of interest do have spherical symmetry, at least to a good approximation, the basic example of course being the hydrogen atom, so the natural set of basis states is the common eigenkets of energy and angular momentum. It turns out that even when the spherical symmetry is broken, the angular momentum eigenkets may still be a useful starting point, with the symmetry breaking treated using perturbation theory.

> *Commuting Hermitian operators can be diagonalized simultaneously and therefore have a common set of eigenkets.*

Two-Dimensional Models

As a warm-up exercise for the complications of the three-dimensional spherically symmetric model, it is worth analyzing a two-dimensional *circularly* symmetric model, that is,

$$H\psi(x,y) = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\psi(x,y)+V\left(\sqrt{x^2+y^2}\right)\psi(x,y)=E\psi(x,y). \tag{4.2.3}$$

(In this section, we'll denote the particle mass by M, to avoid confusion with the angular momentum quantum number m -- but be warned you are often going to find m used for both in the same discussion!)

The two-dimensional angular momentum operator is

$$L=\vec{r}\times\vec{p}=xp_y-yp_x=-i\hbar(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}). \tag{4.2.4}$$

It is a straightforward exercise to check that for the circularly-symmetric Hamiltonian above,

$$[H,L]=0. \tag{4.2.5}$$

Confirm that Equation 4.2.5 is accurate.

To take advantage of the circular symmetry, we switch to polar variables (r,ϕ), where

$$r=\sqrt{x^2+y^2},\quad \phi=\tan^{-1}(y/x),\quad \textit{so}\ \ x=r\cos\phi,\ \ y=r\sin\phi. \tag{4.2.6}$$

Transforming the Hamiltonian and angular momentum into (r,ϕ) coordinates,

$$H\psi(r,\phi)=-\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\phi^2}\right)\psi(r,\phi)+V(r)\psi(r,\phi)=E\psi(r,\phi) \tag{4.2.7}$$

and

$$L=-i\hbar\frac{\partial}{\partial\phi}. \tag{4.2.8}$$

Confirm Equation 4.2.8 is correct.

The angular momentum eigenfunctions $\psi_m(r,\phi)$ satisfy

$$L\psi_m(r,\phi)=-i\hbar\frac{\partial}{\partial\phi}\psi_m(r,\phi)=m\hbar\psi_m(r,\phi) \tag{4.2.9}$$

equivalent to $L|m\rangle=m\hbar|m\rangle$. So $\psi_m(r,\phi)=R(r)e^{im\phi}$, and for this to be a single-valued wave function, m must be an integer. (This also ensures the hermiticity of the operator -- the integration-by-parts check has canceling contributions from $\phi=0$ and $\phi=2\pi$.)

Notice this means that *any* function of r multiplied by $e^{im\phi}$ is an eigenfunction of angular momentum with eigenvalue $m\hbar$, and in fact any eigenfunction of L with eigenvalue $m\hbar$ must be of this form. So we can factor out the r -dependence, and write a complete set of orthonormal eigenfunctions of L, normalized by integrating around the circle:

$$\Phi_m(\phi)=\frac{1}{\sqrt{2\pi}}e^{im\phi} \tag{4.2.10}$$

where ϕ is an integer.

It is interesting to note that this would be a complete set of wave functions for a particle confined to a ring -- rather like the original Bohr orbits. In fact, nanotech rings in which electrons have wave functions like this can be manufactured. Note also that in such rings one can also have *real* wave functions $\sqrt{1/\pi}\sin m\phi$, $\sqrt{1/\pi}\cos m\phi$, which are still energy eigenstates, but *not* angular momentum eigenstates, since they are standing waves, linear superpositions of waves going around the ring in opposite directions.

The common eigenstates of the Hamiltonian and the angular momentum evidently have the form

$$|E,m\rangle=\psi_{E,m}(r,\phi)=R_{E,m}(r)\Phi_m(\phi). \tag{4.2.11}$$

We should emphasize that although the angular part of the wave function does not depend on the radial potential, the radial component $R_{E,m}(r)$ *does* depend on the angular momentum m. This becomes obvious on putting this $\psi_{E,m}(r,\phi)$ into the (r,ϕ) version of Schrödinger’s equation,

$$-\frac{\hbar^2}{2M}(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\phi^2})R_{E,m}(r)\Phi_m(\phi)+V(r)R_{E,m}(r)\Phi_m(\phi)=ER_{E,m}(r)\Phi_m(\phi), \tag{4.2.12}$$

noting that $\partial^2/\partial\phi^2\equiv-m^2$, and canceling out the common factor $\Phi_m(\phi)$ to give

$$-\frac{\hbar^2}{2M}(\frac{d^2}{dr^2}+\frac{1}{r}\frac{d}{dr}-\frac{m^2}{r^2})R_{E,m}(r)+V(r)R_{E,m}(r)=ER_{E,m}(r). \tag{4.2.13}$$

In this one-dimensional equation for the radial wave function $R_{E,m}(r)$, the angular momentum term $\hbar^2m^2/2Mr^2=L^2/2Mr^2$ evidently is equivalent to a repulsive potential. It’s called the “centrifugal barrier” and is easy to understand from classical mechanics. To see this, consider a classical particle bound (in two dimensions) by an attractive central force $V(r)$. Split the momentum into a radial component p_r and a component in the direction perpendicular to the radius, $p_\perp$. The angular momentum $L=rp_\perp$ and is constant (since the force is central). The energy

$$E=\frac{p_r^2}{2M}+\frac{p_\perp^2}{2M}+V(r)=\frac{p_r^2}{2M}+\frac{L^2}{2Mr^2}+V(r) \tag{4.2.14}$$

substituting $p_\perp=L/r$. Since $L=m\hbar$, the angular part is exactly equivalent to the above Schrödinger equation.

But what about the radial part? Why isn’t p_r just equal to $-i\hbar\partial/\partial r$, and p_r^2 equal to $\hbar^2\partial^2/\partial r^2$? We know the more complicated differentiation with respect to r in the Schrödinger equation above must be correct, because it came from $\partial^2/\partial x^2+\partial^2/\partial y^2$ and $r=\sqrt{x^2+y^2}$, $\phi=\tan-1(y/x)$.

To see why p_r equal to $-i\hbar\partial/\partial r$ is incorrect, even though it satisfies $[r,p_r]=i\hbar$, recall what happens in x -space. We argued there that $p_x=-i\hbar\partial/\partial x$ for a plane wave because from the photon analogy, acting on the plane wave state $Ce^{ip_xx/\hbar}$ this operator gives the rate of change of phase and therefore the momentum. But a *radial* wave is a little different: picture a photon wave coming out of a single narrow slit, that is, a slit having width far smaller than the photon wavelength. The photon wave will radiate outwards with equal amplitude in all directions (180°) but the *wave amplitude will decrease* with distance from the slit to conserve probability. For a long (narrow) slit, this is essentially a two-dimensional problem, so the wave function will be $\psi(r) \cong Ce^{ip_r r/\hbar}/\sqrt{r}$. We know that if we measure the momentum of photons at different distances from the slit we'll get the same result. The wavelength determines the photon's momentum, and it isn't changing. The color stays the same. However, $-i\hbar\partial/\partial r$ operating on $\psi(r)$ doesn't just give p_r: it picks up *an extra term* from differentiating the $\sqrt{r}$, so it is obviously *not* giving us the right momentum. Fortunately, this is easy to fix: we define the operator

$$\hat{p_r} = -i\hbar(\frac{\partial}{\partial r} + \frac{1}{2r}) \tag{4.2.15}$$

which eliminates the extra term, *and still satisfies* $[r, p_r] = i\hbar$.

However, there is still a small problem. If we substitute this $\hat{p_r}$ in the classical expression for the energy, following the procedure we used successfully to find Schrödinger's equation in Cartesian coordinates, we find

$$\begin{aligned} H &= \frac{p_r^2}{2M} + \frac{L^2}{2Mr^2} + V(r) \\ &= \frac{-\hbar^2\left(\frac{\partial}{\partial r} + \frac{1}{2r}\right)^2}{2M} + \frac{L^2}{2Mr^2} + V(r) \\ &= \frac{-\hbar^2\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\cdot\frac{\partial}{\partial r} - \frac{1}{4r^2}\right)}{2M} + \frac{L^2}{2Mr^2} + V(r) \end{aligned} \tag{4.2.16}$$

This is almost -- but not quite -- the same as the equation we found by transforming from Cartesian coordinates. The difference is the term $\hbar^2/8Mr^2$. So which is right? Actually our first one was right -- this second one, derived directly from the classical Hamiltonian, does give the same result in the classical limit, because the difference between them vanishes for $\hbar \to 0$. We conclude that beginning with the classical Hamiltonian, and replacing dynamical variables with the appropriate quantum operators, cannot guarantee that we get the correct quantum Hamiltonian: it might be off by some term of order $\hbar$. This would become evident in predicting properties of truly quantum systems, such as atomic energy levels. Problems of this kind are common in constructing quantum theories starting from a classical theory: essentially, in a classical theory, the order of variables in an expression is irrelevant, but in the quantum theory there can only be *one* correct order of noncommuting variables such as $\partial/\partial r$ and r in any expression.

What can we say about the radial wave function $R_{E,m}(r)$? If both the energy and the potential at the origin are finite, then for small r:

$$R_{E,m}(r) \approx Ar^m \text{ or } Ar^{-m} \tag{4.2.17}$$

However, the wave function cannot be discontinuous, so $R_{E,m}(r) \approx Ar^{|m|}$. To make further progress in finding the wave function, we need to know the potential. Specific examples will be analyzed in due course. It is interesting to note that the allowed wave functions, proportional to $rme^{im\phi}$, $rme^{-im\phi}, m > 0$, are the complex functions z^m, $(z^*)^m$ if the two-dimensional space is mapped into the complex plane. Representing many-electron wave functions in the plane in this way was a key to understanding the quantum Hall effect.

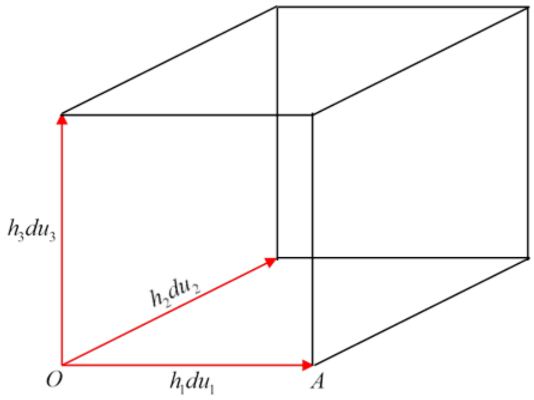
4.3: Note on Curvilinear Coordinates

Div, Grad and Curl in Orthogonal Curvilinear Coordinates

Problems with a particular symmetry, such as cylindrical or spherical, are best attacked using coordinate systems that take full advantage of that symmetry. For example, the Schrödinger equation for the hydrogen atom is best solved using spherical polar coordinates. For this and other differential equation problems, then, we need to find the expressions for differential operators in terms of the appropriate coordinates.

We only look at *orthogonal* coordinate systems, so that locally the three axes (such as r, θ, φ) are a mutually perpendicular set. We denote the curvilinear coordinates by (u_1, u_2, u_3). The standard Cartesian coordinates for the same space are as usual (x, y, z).

Suppose now we take an infinitesimally small cube with edges parallel to the local curvilinear coordinate directions, and therefore with faces satisfying $u_i = constant, \; i = 1, 2, 3$ for the three pairs of faces.



The lengths of the edges are then $h_1du_1, \; h_2du_2$ and h_3du_3, where $h_1, \; h_2, \; h_3$ are in general functions of $u_1, \; u_2, u_3$. That is to say, the distance across the cube from one corner to the opposite corner

$$ds^2 = h_1^2du_1^2 + h_2^2du_2^2 + h_3^2du_3^2 = dx^2 + dy^2 + dz^2 \tag{4.2.1}$$

It is clear that the gradient of a function ψ in the u_1 direction is

$$(\nabla\psi)_1 = \lim_{du_1 \to 0} \frac{\psi(A) - \psi(0)}{h_1du_1} = \frac{1}{h_1}\frac{\partial\psi}{\partial u_1} \tag{4.3.2}$$

The divergence of a vector field $\vec{V}$ in curvilinear coordinates is found using Gauss' theorem, that the total vector flux through the six sides of the cube equals the divergence multiplied by the volume of the cube, in the limit of a small cube.

The area of the face bracketed by h_2du_2 and h_3du_3 is $h_2du_2h_3du_3$. For that face, the component of the vector field contributing to the flow from the cube is $-V_1$, so the flow across the face is $-V_1h_2h_3du_2du_3$. To find the flow across the opposite (parallel) face of the cube, corresponding to an increase in u_1 of du_1, we must bear in mind that $h_2, \; h_3$ and V_1 all vary with u_1, so the flow will be:

$$V_1h_2h_3du_2du_3 + \frac{\partial}{\partial u_1}(h_2h_3V_1)du_1du_2du_3 \tag{4.3.3}$$

The first term here of course cancels the contribution from the other face. The remaining term, plus the terms with 123 replaced with 231 and 312 from the two other pairs of opposite faces, must, applying Gauss' theorem, add to give

$$\vec{\nabla}\cdot\vec{V} \times volume = \vec{\nabla}\cdot\vec{V} \; h_1h_2h_3du_1du_2du_3. \tag{4.3.4}$$

This gives:

$$\vec{\nabla}\cdot\vec{V} = \frac{1}{h_1h_2h_3}\{\frac{\partial}{\partial u_1}(h_2h_3V_1) + \frac{\partial}{\partial u_2}(h_3h_1V_2) + \frac{\partial}{\partial u_3}(h_1h_2V_3)\} \tag{4.3.5}$$

.

Putting this together with the expression for the gradient gives immediately the expression for the Laplacian operator in curvilinear coordinates:

$$\nabla^2\psi = \frac{1}{h_1h_2h_3}\{\frac{\partial}{\partial u_1}(\frac{h_2h_3}{h_1}\frac{\partial\psi}{\partial u_1}) + \frac{\partial}{\partial u_2}(\frac{h_3h_1}{h_2}\frac{\partial\psi}{\partial u_2}) + \frac{\partial}{\partial u_3}(\frac{h_1h_2}{h_3}\frac{\partial\psi}{\partial u_3})\} \tag{4.3.6}$$

The curl of a vector field $\vec{A}$ is found by integrating around one of the square faces. Thus, the 1-component of $\vec{\nabla}\times\vec{A}$ is given by integrating $\vec{A}\cdot\overrightarrow{ds}$ around the (23) square with two of its sides h_2du_2 and h_3du_3. The integral must equal $(\vec{\nabla}\times\vec{A})_1$ multiplied by the area $h_2du_2h_3du_3$. This gives

$$(\vec{\nabla}\times\vec{A})_1 = \frac{1}{h_2h_3}\{\frac{\partial}{\partial u_2}(A_3h_3) - \frac{\partial}{\partial u_3}(A_2h_2)\}. \tag{4.3.7}$$

Cylindrical Coordinates

$$\begin{gathered} x = r\cos\varphi \\ y = r\sin\varphi \\ z = z \\ ds^2 = dr^2 + r^2d\varphi^2 + dz^2 \end{gathered} \tag{4.3.8}$$

Here $(u_1, u_2, u_3) = (r, \varphi, z)$, and $(h_1, h_2, h_3) = (1, r, 1)$.

Therefore, for example,

$$\nabla^2\psi = \frac{1}{r}\frac{\partial}{\partial r}(r\frac{\partial\psi}{\partial r}) + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\varphi^2} + \frac{\partial^2\psi}{\partial z^2}. \tag{4.3.9}$$

Spherical Polar Coordinates

$$\begin{gathered} x = r\sin\theta\cos\varphi \\ y = r\sin\theta\sin\varphi \\ z = r\cos\theta \\ ds^2 = dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\varphi^2 \end{gathered} \tag{4.3.10}$$

So $(u_1, u_2, u_3) = (r, \theta, \varphi)$ and $(h_1, h_2, h_3) = (1, r, r\sin\theta)$.

Here

$$\nabla^2\psi = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial\psi}{\partial r}) + \frac{1}{r^2 sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\frac{\partial\psi}{\partial\theta}) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2}. \tag{4.3.11}$$

4.4: Orbital Eigenfunctions in 3-D

The Angular Momentum Operators in Spherical Polar Coordinates

The angular momentum operator

$$\vec{L}=\vec{r}\times\vec{p}=-i\hbar\vec{r}\times\vec{\nabla}. \tag{4.4.1}$$

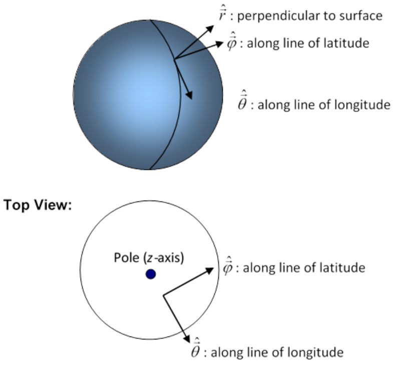
In spherical polar coordinates,

$$\begin{gathered}x=r\sin\theta\cos\phi\\ y=r\sin\theta\sin\phi\\ z=r\cos\theta\\ ds^2=dr^2+r^2d\theta^2+r^2\sin^2\theta d\phi^2\end{gathered} \tag{4.4.2}$$

the gradient operator is

$$\vec{\nabla}=\hat{r}\frac{\partial}{\partial r}+\hat{\theta}\frac{1}{r}\frac{\partial}{\partial\theta}+\hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi} \tag{4.4.3}$$

where now *the little hats denote unit vectors*: $\hat{r}$ is radially outwards, $\hat{\theta}$ points along a line of longitude away from the north pole (and therefore in the direction of increasing θ) and $\hat{\phi}$ points along a line of latitude in an anticlockwise direction as seen looking down on the north pole (that is, in the direction of increasing ϕ).



The three unit vectors in the spherical polar system

Here $\hat{r},\hat{\theta},\hat{\phi}$ form an orthonormal local basis, and

$$\hat{r}\times\hat{\theta}=\hat{\phi},\quad \hat{r}\times\hat{\phi}=-\hat{\theta} \tag{4.4.4}$$

as should be clear from the diagram.

So

$$\vec{r}\times\vec{\nabla}=\hat{\phi}\frac{\partial}{\partial\theta}-\hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi} \tag{4.4.5}$$

(Explicitly, $\hat{\phi}=(-\sin\phi,\cos\phi,0)$ and $\hat{\theta}=(\cos\theta\cos\phi,\cos\theta\sin\phi,-\sin\theta)$.)

The vector $\hat{\phi}$ has zero component in the *z*-direction, the vector $\hat{\theta}$ has component $-\sin\theta$ in the *z*-direction, so we can immediately conclude that

$$L_z=(\vec{r}\times\vec{p})_z=-(i\hbar\vec{r}\times\vec{\nabla})_z=-i\hbar\frac{\partial}{\partial\phi} \tag{4.4.6}$$

just as in the two-dimensional case.

The operator

$$L^2=-\hbar^2\left(\hat{\phi}\frac{\partial}{\partial\theta}-\hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\right)\cdot\left(\hat{\phi}\frac{\partial}{\partial\theta}-\hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\right). \tag{4.4.7}$$

To evaluate this expression, we use $\hat{\phi}^2=1,\ \hat{\theta}^2=1,\ \hat{\phi}\cdot\hat{\theta}=0$ *but* we must also check the effects of the differential operators in the first expression on the variables in the second, *including the unit vectors.*

From the explicit coordinate expressions for the unit vectors, or by staring at the diagram, you should be able to establish the following: $\partial\hat{\phi}/\partial\theta=0,\ \partial\hat{\theta}/\partial\theta$ is in the *r*-direction, $\partial\hat{\phi}/\partial\phi$ is a horizontal unit vector pointing inwards perpendicular to $\hat{\phi}$, and having component $-\cos\theta$ in the $\hat{\theta}$-direction, $\partial\hat{\theta}/\partial\phi=\hat{\phi}\cos\theta$.

Therefore, the *only* "differentiation of a unit vector" term that contributes to L^2 is

$$\hbar^2\hat{\theta}\frac{1}{\sin\theta}\cdot\frac{\partial\hat{\phi}}{\partial\phi}\frac{\partial}{\partial\theta}=-\hbar^2\cot\theta\frac{\partial}{\partial\theta}. \tag{4.4.8}$$

The $\hat{\phi}\frac{\partial}{\partial\theta}$ acting on the $\sin\theta$ in $-\hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}$ contributes nothing because $\hat{\phi}\cdot\hat{\theta}=0$.

Hence

$$\begin{aligned}L^2&=-\hbar^2\left(\frac{\partial^2}{\partial\theta^2}+\cot\theta\frac{\partial}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)\\ &=-\hbar^2\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)\end{aligned} \tag{4.4.9}$$

Now, we know that L^2 and L_z have a common set of eigenkets (since they commute) and we've already established that those of L_z are $\Phi_m(\phi)=e^{im\phi}/\sqrt{2\pi}$, with m an integer, so the eigenkets of L^2 must have this same f dependence, so they must be of the form $\Theta_l^m(\theta)\Phi(\phi)$, where $\Theta_l^m(\theta)$ is a (suitably normalized) solution of the equation

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\sin\theta\frac{d\Theta_l^m(\theta)}{d\theta}-\frac{m^2}{\sin^2\theta}\Theta_l^m(\theta)=-l(l+1)\Theta_l^m(\theta) \tag{4.4.10}$$

more conveniently written

$$\sin\theta\frac{d}{d\theta}\sin\theta\frac{d\Theta_l^m(\theta)}{d\theta}+\left(l(l+1)\sin^2\theta-m^2\right)\Theta_l^m(\theta)=0. \tag{4.4.11}$$

To summarize: the solutions to this differential equation, with integer $l,\ m,\ |m|\le l$, will (together with $\Phi_m(\phi)$) give the complete set of eigenstates of L^2, L_z in the coordinate representation.

Finding the *m* = *l* Eigenket of L^2, L_z

Recall now that for the simple harmonic oscillator, the easiest wave function to find was that of the ground state, the solution of the simple *linear* equation $\hat{a}\Psi_0=0$ (as well as being a solution of the quadratic Schrödinger equation, of course). The other state wave functions could then be found by applying the creation operator in differential form the necessary number of times.

A similar strategy works here: we can easily find the *highest* state on the l ladder, $m=l$, the state $|l,l\rangle$, since it satisfies the linear equation $L_+|l,l\rangle=0$, where $L_+=L_x+iL_y$. We just need to cast this equation in coordinate form. In Cartesian coordinates, $L_+=-i\hbar(\hat{r}\times\vec{\nabla})_+$, and we've already shown that

$$\vec{r}\times\vec{\nabla}=\hat{\phi}\frac{\partial}{\partial\theta}-\hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}. \tag{4.4.12}$$

Therefore

$$(\vec{r}\times\vec{\nabla})_+=\hat{\phi}_+\frac{\partial}{\partial\theta}-\hat{\theta}_+\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}, \tag{4.4.13}$$

and using $\hat{\phi}=(-\sin\phi,\cos\phi,0)$, $\hat{\theta}=(\cos\theta\cos\phi,\cos\theta\sin\phi,-\sin\theta$ we see that $\hat{\theta}_+$, the component of $\hat{\theta}$ in the + direction, is $\phi_+=\phi_x+i\phi_y=ie^{i\phi}$, and similarly $\hat{\theta}_+=\cos\theta e^{i\phi}$.

So

$$\begin{aligned}L_+&=\hbar e^{i\phi}\left(\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\phi}\right)\\ L_-&=-\hbar e^{-i\phi}\left(\frac{\partial}{\partial\theta}-i\cot\theta\frac{\partial}{\partial\phi}\right)\end{aligned} \tag{4.4.14}$$

and $L_+|l,l\rangle=0$ becomes

$$\left(\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\phi}\right)\Theta_l^l(\theta)e^{il\phi}=0 \tag{4.4.15}$$

That is,

$$\left(\frac{d}{d\theta}-l\cot\theta\right)\Theta_l^l(\theta)=0 \tag{4.4.16}$$

The solution to this equation is

$$\Theta_l^l(\theta)=N(\sin\theta)^l \tag{4.4.17}$$

where N is the normalization constant. The $m\neq l$ wave functions are generated by applying the lowering operator L_-.

Normalizing the *m* = *l* Eigenket

The standard notation for the normalized eigenkets $|l,m\rangle$ is $Y_l^m(\theta,\phi)=\Theta_l^m(\theta)\Phi_m(\phi)$. These functions, being eigenkets of Hermitian operators with different eigenvalues, must satisfy

$$\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi}Y_{l'}^{m'*}(\theta,\phi)Y_l^m(\theta,\phi)\sin\theta d\theta d\phi=\int Y_{l'}^{m'*}(\theta,\phi)Y_l^m(\theta,\phi)d\Omega=\delta_{l'l}\delta_{m'm}\ . \tag{4.4.18}$$

So, our first job is to normalize $\Theta_l^l(\theta)=N(\sin\theta)^l$ (taking $\Phi_l(\phi)=e^{il\phi}/\sqrt{2\pi}$ already normalized)

$$|N|^2\int_0^{\pi}(\sin\theta)^{2l+1}d\theta=1 \tag{4.4.19}$$

The integral can be evaluated using the substitution $\mu=\cos\theta$ to give $\int_{-1}^{1}(1-\mu^2)^l d\mu$, then making the further substitution $u=\frac{1}{2}(1-\mu)$ to give $2^{2l+1}\int_0^1 u^l(1-u)^l du$, which can be integrated by parts to give

$$|N|^2 2^{2l+1}(l!)^2/(2l+1)!=1. \tag{4.4.20}$$

Therefore

$$Y_l^l(\theta,\phi)=(-1)^l\left(\frac{(2l+1)!}{4\pi}\right)^{1/2}\frac{1}{2^l l!}(\sin\theta)^l e^{il\phi}=c_l(\sin\theta)^l e^{il\phi} \tag{4.4.21}$$

where we have fixed the sign in accord with the standard convention, and we will denote the rather cumbersome normalization constant by c_l.

Notice that for large values of *l*, this function is heavily weighted around the equator, as we would expect -- for a given total angular momentum one gets a maximum component in the *z*-direction when the motion is concentrated in the *x*, *y* plane. This looks like a Bohr orbit.

Finding the Rest of the Eigenkets: the Details

Now that $|l,l\rangle$ is normalized, we can automatically produce correctly normalized $|l,m\rangle$'s, since we know the matrix element of the lowering operator between normalized states. We don't have to do any more integrals.

For example, $L_-|l,l\rangle=\hbar\sqrt{2l}|l,l-1\rangle$, equivalently (the $\hbar$'s of course cancel)

$$Y_l^{l-1}(\theta,\phi)=\frac{(-1)}{\sqrt{2l}}e^{-i\phi}\left(\frac{\partial}{\partial\theta}-i\cot\theta\frac{\partial}{\partial\phi}\right)Y_l^l. \tag{4.4.22}$$

That is,

$$\begin{aligned}Y_l^{l-1}(\theta,\phi)&=c_l e^{-i\phi}\left(\frac{\partial}{\partial\theta}-i\cot\theta\frac{\partial}{\partial\phi}\right)\sin^l\theta\cdot e^{il\phi}\\ &=-c_l e^{i(l-1)\phi}\sqrt{2l}\sin^{l-1}\theta\cos\theta\end{aligned} \tag{4.4.23}$$

(both terms giving equal contributions).

Note that this function is actually *zero* on the equator, but for large l it peaks close to the equator (on both sides).

In principle, we can reapply this differential operator over and over to generate all the $|l,m\rangle$ states, but this gets very messy. However, there is a neat theorem concerning the lowering operator that makes it all straightforward:

$$L_-e^{im\phi}f(\theta)=e^{i(m-1)\phi}\left(\sin^{1-m}\theta\frac{d}{d(\cos\theta)}\sin^m\theta\right)f(\theta) \tag{4.4.24}$$

Exercise: prove this.

So

$$L_-e^{il\phi}\sin^l\theta=e^{i(l-1)\phi}\left(\sin^{1-l}\theta\frac{d}{d(\cos\theta)}\sin^l\theta\right)\sin^l(\theta) \tag{4.4.25}$$

and applying the operator again,

`\begin{matrix} (L_-)^2e^{il\phi}\sin^l\theta=L_-e^{i(l-1)\phi}\left( \sin^{1-l}\theta\dfrac{d}{d(\cos\theta)}\sin^l\theta \right) \sin^l(\theta) \\ =e^{i(l-2)\phi}\left( \sin^{2-l}\theta\dfrac{d}{d(\cos\theta)}\sin^{l-1}\theta \righ`

So the point of introducing this odd-looking representation of the lowering operator is that the $\sin^{1-l}\theta$ term in the middle is exactly *canceled* when the operator is applies twice, and similar cancellations occur on repeating the operation, giving the (relatively) simple representation:

$$Y_l^{\,m}(\theta,\phi)=c_l\sqrt{\frac{(l+m)!}{(2l)!(l-m)!}}e^{im\phi}\sin^{-m}\theta\frac{d^{l-m}}{d(\cos\theta)^{l-m}}\sin^{2l}\theta \tag{4.4.26}$$

(Where did all those factorials come from? They're the product of all the inverse square root factors in

$$|l,m-1\rangle=\frac{1}{\sqrt{(l+m)(l-m+1)}}L_-|l,m\rangle \tag{4.4.27}$$

for the number of lowerings necessary.)

Note that for $m=0$ the function is

$$Y_l^{\,0}(\theta,\phi)=c_l\sqrt{\frac{1}{(2l)!}}\frac{d^l}{d(\cos\theta)^l}\sin^{2l}\theta, \tag{4.4.28}$$

and in fact not a function of ϕ at all. This isn't surprising, since it has zero angular momentum about the *z*-direction, the appropriate $\Phi(\phi)$ is just constant.

For $m=-l$, the differentiation becomes trivial, because, writing $\cos\theta=\mu$, the differentiation becomes $\dfrac{d^{2l}}{d\mu^{2l}}(1-\mu^2)^l$ and only the μ^{2l} term survives, giving

$$Y_l^{-l}(\theta,\phi)=(-1)^lc_le^{-il\phi}\sin^l\theta. \tag{4.4.29}$$

Of course, this could also have been found from the linear equation $L_-|l,-l\rangle=0$, and we could have equally generated all the states by applying L_+ to this state. In fact, this gives a different -- but of course equivalent -- expression for the $Y_l^{\,m}(\theta,\phi)$:

$$Y_l^{\,m}(\theta,\phi)=(-1)^mc_l\sqrt{\frac{(l-m)!}{(2l)!(l+m)!}}e^{im\phi}\sin^m\theta\frac{d^{l+m}}{d(\cos\theta)^{l+m}}\sin^{2l}\theta \tag{4.4.30}$$

(from Messiah, page 522).

Relating the $Y_l^{\,m}$'s to the Legendre Functions

The *Legendre polynomials* $P_n(\cos\theta)$ are defined by:

$$\begin{aligned}P_n(\cos\theta)&=\frac{1}{2^nn!}\frac{d^n}{d(\cos\theta)^n}\sin^{2n}\theta,\ \textit{or}\\ P_n(\mu)&=\frac{1}{2^nn!}\frac{d^n}{d\mu^n}(1-\mu^2)^n\end{aligned} \tag{4.4.31}$$

where $\mu=\cos\theta$, so $d\mu=-\sin\theta d\theta$. From this form, it is easy to show that $P_n(1)=1$ (all n differentiations must take out a $(1-\mu^2)$ factor to give a nonzero contribution), and $P_n(\mu)$ must have n zeros in the interval (-1, 1). $P_n(\mu)$ alternates between an even function and an odd function.

The normalization of the $P_n(\mu)$'s is

$$\begin{aligned}\textstyle\int_{-1}^1(P_n(\mu))^2d\mu&=\left(\frac{1}{2^nn!}\right)^2\textstyle\int_{-1}^1\frac{d^n}{d\mu^n}(\mu^2-1)^n\frac{d^n}{d\mu^n}(\mu^2-1)^nd\mu\\ &=(-1)^n\left(\frac{1}{2^nn!}\right)^2\textstyle\int_{-1}^1(\mu^2-1)^n\frac{d^{2n}}{d\mu^{2n}}(\mu^2-1)^nd\mu\\ &=(2n)!\left(\frac{1}{2^nn!}\right)^2\textstyle\int_{-1}^1(\mu^2-1)^nd\mu\\ &=\frac{2}{2n+1}\end{aligned} \tag{4.4.32}$$

where in that last line we used the result for the integral obtained earlier in this lecture for normalizing $Y_l^{\,l}$.

Doing the same repeated integration by parts for two different Legendre polynomials proves they are orthogonal,

$$\int_{-1}^1P_m(\mu)P_n(\mu)d\mu=0,\ m\neq n \tag{4.4.33}$$

The *associated Legendre functions* are defined (for *n* and *m* zero or positive integers, $n\geq m$ $^{n\geq m}$) by:

$$\begin{aligned}P_n^m(\mu)&=(1-\mu^2)^{m/2}\frac{d^m}{d\mu^m}P_n(\mu)\\ &=(-1)^n\frac{(1-\mu^2)^{m/2}}{2^nn!}\frac{d^{n+m}}{d\mu^{n+m}}(1-\mu^2)^n.\end{aligned} \tag{4.4.34}$$

Following Messiah in requiring $Y_l^{\,0}(0,0)$ be real and positive, we find

$$Y_l^{\,0}(\theta,\phi)=\sqrt{\frac{2l+1}{4\pi}}P_l(\cos\theta) \tag{4.4.35}$$

where the coefficient just reflects the differing normalization conventions. Similarly, the spherical harmonics with nonzero *m* are proportional to the associated Legendre functions (the odd ones are *not* polynomials in $\cos\theta$, despite Shankar p. 337, since they include odd powers of $\sin\theta$),

$$Y_l^m(\theta,\phi)=\sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}(-1)^m e^{im\phi}P_l^m(\cos\theta). \tag{4.4.36}$$

The Spherical Harmonics as a Basis

We have found explicit expressions for the spherical harmonics: an orthonormal set of eigenfunctions of L^2 and L_z defined on the surface of a sphere,

`\int^{\pi}_{\theta=0}\int^{2\pi}_{\phi=0} Y^{m'^{*}}_{l'}(\theta,\phi)Y^m_l(\theta,\phi)\sin\theta d\theta d\phi=\int Y^{m'^*}_{l'}(\theta,\phi)Y^m_l(\theta,\phi)d\Omega=\delta_{l'l}\delta_{m'm} \; . \label{4.4.16}`

They form a *complete* set:

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l}|l,m\rangle\langle l,m|=I \tag{4.4.37}$$

or

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l}Y_l^{m^*}(\theta,\phi)Y_l^m(\theta',\phi')=\delta(\cos\theta-\cos\theta')\delta(\phi-\phi')=\delta(\Omega-\Omega') \tag{4.4.38}$$

in the notation of Messiah, where W refers to a point on the spherical surface.

(Formal proof of the completeness is given in Byron and Fuller, *Mathematics of Classical and Quantum Physics*.)

The above equation could also be written

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\langle\theta,\phi|l,m\rangle\langle l,m|\theta',\phi'\rangle=\langle\theta,\phi|\theta',\phi'\rangle=\delta(\cos\theta-\cos\theta')\delta(\phi-\phi') \tag{4.4.39}$$

where the ket $|\theta',\phi'\rangle$ is to be understood as a localized ket, the spherical-surface version of $|x\rangle$, normalized by its delta function inner product with the bra $\langle\theta,\phi|$, exactly analogous to $\langle x|x'\rangle=\delta(x-x')$, bearing in mind that the infinitesimal area element is $-d(\cos\theta)d\phi$, (a positive quantity in the relevant interval, 0 to π).

This completeness means that any reasonable function on the surface of the sphere can be expressed as a sum over spherical harmonics with appropriate coefficients, in other words the spherical generalization of a Fourier series.

In fact, L^2 is equivalent to ∇^2 on the spherical surface, so the Y_l^m are the eigenfunctions of the operator ∇^2. Just as in one dimension the eigenfunctions of d^2/dx^2 have the spatial dependence of the eigenmodes of a vibrating string, the spherical harmonics have the spatial dependence of the eigenmodes of a vibrating spherical balloon. Of course, to describe the displacement of the balloon skin (which must be real!) with these eigenfunctions, we can no longer use the eigenfunctions of the *z*-component of angular momentum, since they are complex except in the trivial zero case. We must rearrange the eigenfunctions of L^2, for example replacing the pair $e^{i\phi},\ e^{-i\phi}$ with $\cos\phi,\sin\phi$. These real solutions, essentially

$$\frac{1}{\sqrt{2}}(|l,l\rangle\pm|l,-l\rangle) \tag{4.4.40}$$

, have l nodal lines (zeroes) of longitude. Moving down one notch in $|m|$, the (real) state with $|m|=l-1$ has $l-1$ longitudinal nodes, but has added a *latitudinal* node: the equator. Then $|m|=l-2$ has $l-2$ longitudinal nodes, 2 latitudinal nodal lines -- *there are always l nodal lines total*.

Some of these modes of vibration have been observed in the sun after a sunspot storm. The spherical harmonics are also used in analyzing the cosmic background radiation.

Some Low Order Spherical Harmonics

Let's look in more detail at the lowest order spherical harmonics. For the first few, the normalization of the highest state $|l,l\rangle$ is pretty easy to do from scratch: factoring out the ϕ dependence as before, $Y_l^m(\theta,\phi)=\Theta_l^m(\theta)\Phi_m(\phi)$, and taking the normalized $\Phi_m(\phi)=e^{im\phi}/\sqrt{2\pi}$, the θ normalization for $|l,l\rangle$ is just

`|N|^2\int^{\pi}_0 (\sin\theta)^{2l+1}d\theta=1 \label{4.4.17}`

This is easily accomplished for $l=0,1,2$.

All we then need is

$$L_\pm=\pm\hbar e^{\pm i\phi}\left(\frac{\partial}{\partial\theta}\pm i\cot\theta\frac{\partial}{\partial\phi}\right) \tag{4.4.41}$$

,

$$L_-|l,m\rangle=\hbar\sqrt{l(l+1)-m(m-1)}\,|l,m-1\rangle \tag{4.4.42}$$

, and finally the *sign convention* that $Y_l^0(0,0)$ be real and positive.

With a few elementary steps, it can be established that:

$$\begin{gathered}Y_0^0=\frac{1}{\sqrt{4\pi}}\\ Y_1^1=-\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\phi}\\ Y_1^0=\sqrt{\frac{3}{4\pi}}\cos\theta\\ Y_1^{-1}=\sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\phi}\end{gathered} \tag{4.4.43}$$

$$\begin{gathered}Y_2^2=\sqrt{\frac{15}{32\pi}}\sin^2\theta e^{2i\phi},\ Y_2^1=-\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{i\phi},\ Y_2^0=\sqrt{\frac{5}{16\pi}}(3\cos^2\theta-1)\\ Y_2^{-2}=\sqrt{\frac{15}{32\pi}}\sin^2\theta e^{-2i\phi},\ Y_2^{-1}=\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{-i\phi},\end{gathered} \tag{4.4.44}$$

<

It is often useful to write the Y_l^m in terms of Cartesian coordinates,

$$(x,y,z)=(r\sin\theta\cos\phi,\ r\sin\theta\sin\phi,\ r\cos\theta) \tag{4.4.45}$$

so

$$Y_1^1(x,y,z) = -\sqrt{\frac{3}{8\pi}} \cdot \frac{x+iy}{r},\ Y_1^0(x,y,z) = \sqrt{\frac{3}{4\pi}} \cdot \frac{z}{r},\ Y_1^{-1}(x,y,z) = \sqrt{\frac{3}{8\pi}} \cdot \frac{x-iy}{r} \tag{4.4.46}$$

and

$$Y_2^2 = \sqrt{\frac{15}{32\pi}} \frac{(x+iy)^2}{r^2},\ Y_2^1 = -\sqrt{\frac{15}{8\pi}} \frac{(x+iy)z}{r^2},\ Y_2^0 = \sqrt{\frac{5}{16\pi}} \frac{(3z^2-1)}{r^2},\ etc. \tag{4.4.47}$$

The Y_1^m as a Basis of the $l = 1$ Subspace

The Y_1^m are the $l=1$ eigenstates of L^2 and L_z. But what if we'd chosen to look for the common eigenstates of L^2 and L_x instead? What $l=1$ state has zero angular momentum component in the direction of the *x*-axis? Clearly it will be

$$\sqrt{\frac{3}{4\pi}} \cdot \frac{x}{r} \tag{4.4.48}$$

, in other words the previous Y_1^0 with *z* replaced by *x*, because after all, our labeling of axes was arbitrary.

Now,

$$\sqrt{\frac{3}{4\pi}} \cdot \frac{x}{r}\ is\ just\ (1/\sqrt{2})(-Y_1^1 + Y_1^{-1}). \tag{4.4.49}$$

In fact, *any* $l=1$ state, with a specified component in *any* direction, can be written as

$$\alpha_1|1,1\rangle + \alpha_0|1,0\rangle + \alpha_{-1}|1,-1\rangle = \sum \alpha_m|1,m\rangle. \tag{4.4.50}$$

This can be seen as follows: an $l=1$ state has to be *linear* in $x/r,\ y/r,\ z/r$(any quadratic term would give rise to $e^{2i\phi}$ about an appropriate axis, call that the *z*-axis, so $m=2$ and l must be 2 or greater), and any such state can be written as a linear combination of

$$(x+iy)/\sqrt{2}r,\ (x-iy)/\sqrt{2}r,\ z/r \tag{4.4.51}$$

The bottom line, then, is that the Y_1^m do indeed provide a *complete* basis for the $l=1$ space of eigenstates of L^2.

Representing the Rotation Operator Within the $l = 1$ Subspace

Recall that we originally introduced the angular momentum operator $\vec{J}$ by defining it as the generator of infinitesimal rotations when acting on any wave function, including multicomponent wave functions. We found, using the commutativity properties of ordinary rotations, that the vector components of $\vec{J}$ had to satisfy $[J_x, J_y] = i\hbar J_z$, etc., and from that we deduced the possible sets of eigenvalues of the commuting pair of operators $\vec{J}^2$, J_z were $j(j+1)\hbar^2$ for $\vec{J}^2$, with j an integer of half an odd integer, and for each such j the allowed eigenvalues of J_z were

$$m\hbar,\ m = -j, -j+1, \ldots, +j. \tag{4.4.52}$$

Back to the $l=1$ angular wave functions: we have established that any such function can be written

$$\alpha_1|1,1\rangle + \alpha_0|1,0\rangle + \alpha_{-1}|1,-1\rangle = \sum \alpha_m|1,m\rangle. \tag{4.4.53}$$

and so is a vector in a three-dimensional space spanned by the set $|1,m\rangle,\ m = 1, 0, -1$. In other words, the wave function is a three-component object. The angular momentum operator must therefore be *a matrix operator in this three-dimensional space*, such that, by definition, the effect of an infinitesimal rotation on the multicomponent wave function is:

$$R(\delta\vec{\theta})\Psi_{l=1}(\theta,\phi) = e^{-\frac{i}{\hbar}\delta\vec{\theta}\cdot\hat{\vec{J}}} \begin{pmatrix} \alpha_1 \\ \alpha_0 \\ \alpha_{-1} \end{pmatrix} = \begin{pmatrix} \alpha'_1 \\ \alpha'_0 \\ \alpha'_{-1} \end{pmatrix} \tag{4.4.54}$$

The unitary rotation operator acting in the $l=1$ subspace,

$$U(R(\vec{\theta})) = e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}} \tag{4.4.55}$$

, has to be a 3×3 matrix. The standard notation for its matrix elements is:

$$D^{(1)}_{m'm}(R(\vec{\theta})) = \langle 1,m'|e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}|1,m\rangle \tag{4.4.56}$$

so the rotated ket is

$$\alpha'_{m'} = \sum_{m',m} D^{(1)}_{m'm}\alpha_m,\ \ or\ \alpha' = D\alpha \tag{4.4.57}$$

To evaluate this matrix explicitly, we must expand the exponential and we need the matrix elements of J_z, J_+, J_- between the states $|1,m\rangle$ -- which we already know.

Now, the basis of the three-dimensional space is just the common eigenkets of $\vec{J}^2$, J_z, in this case identical to $\vec{L}^2$, L_z. We know the matrix elements of J_z, J_+, J_- between states $|j,m\rangle$ from the earlier lecture, so it is simple to find the matrix representations of the components of J in this space:

$$J_x^{(1)} = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},\ \ J_y^{(1)} = \frac{i\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},\ \ J_z^{(1)} = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{4.4.58}$$

We have added the superscript (1) because this representation of the infinitesimal rotation operators is specific to $j=1$ (representations for general values of j are as $(2j+1)\times(2j+1)$ matrices, reflecting the dimensionality of the space spanned by the $2j+1$ distinct m values).

Expanding the exponential is not difficult, because by inspection $(J_z^{(1)}/\hbar)^3 = (J_z^{(1)}/\hbar)$, so from spherical symmetry $(\hat{\vec{n}}\cdot\vec{J}^{(1)}/\hbar)^3 = (\hat{\vec{n}}\vec{J}^{(1)}/\hbar)$ for a unit vector in any direction. The result is:

$$D^{(1)}(R(\vec{\theta})) = e^{-\frac{i\theta\hat{\vec{n}}\cdot\vec{J}}{\hbar}} = I + (\cos\theta - 1)\left(\frac{\hat{\vec{n}}\cdot\vec{J}}{\hbar}\right)^2 - i\sin\theta\left(\frac{\hat{\vec{n}}\cdot\vec{J}}{\hbar}\right). \tag{4.4.59}$$

One other point we should note: at the end of the linear algebra lecture, we discussed rotations about the *z*-axis in ordinary (*x*, *y*, *z*) space. Obviously, if we label a point in the (*x*, *y*) plane using the complex number $x+iy$, a rotation by an angle θ about the *z*-axis will move the point in such a way that the new label is $e^{i\theta}(x+iy)$. The angle in this case has the opposite sign to that given by the operator above: the reason is that when we write the eigenstate as

$$-\sqrt{\frac{3}{8\pi}} \cdot \frac{x+iy}{r} \tag{4.4.60}$$

this is a *function* of position in the plane, not a point in the plane, so for the reasons discussed at the beginning of the first Angular Momentum lecture, the sign is opposite.

4.6: The Hydrogen Atom

Factoring Out the Center of Mass Motion

The hydrogen atom consists of two particles, the proton and the electron, interacting via the Coulomb potential $V(\overrightarrow{r_1}-\overrightarrow{r_2})=e^2/r$, where as usual $r=|\overrightarrow{r_1}-\overrightarrow{r_2}|$. Writing the masses of the two particles as $m_1,\ m_2$ Schrödinger’s equation for the atom is:

$$\left(-\frac{\hbar^2}{2m_1}\overrightarrow{\nabla}_1^{\,2}-\frac{\hbar^2}{2m_2}\overrightarrow{\nabla}_2^{\,2}-\frac{e^2}{r}\right)\psi(\overrightarrow{r_1},\overrightarrow{r_2})=E\psi(\overrightarrow{r_1},\overrightarrow{r_2}). \tag{4.6.1}$$

But $\overrightarrow{r_1},\ \ \overrightarrow{r_2}$ are *not* the most natural position variables for describing this system: since the potential depends only on the relative position, a better choice is $\vec{r},\ \ \vec{R}$ defined by:

$$\vec{r}=\overrightarrow{r_1}-\overrightarrow{r_2},\quad \vec{R}=\frac{m_1\overrightarrow{r_1}+m_2\overrightarrow{r_2}}{m_1+m_2} \tag{4.6.2}$$

so $\vec{R}$ is the center of mass of the system. It is convenient at the same time to denote the total mass by $M=m_1+m_2$, and the reduced mass by $m=\frac{m_1m_2}{m_1+m_2}$.

Transforming in straightforward fashion to the variables $\vec{r},\ \ \vec{R}$ Schrödinger’s equation becomes

$$\left(-\frac{\hbar^2}{2M}\overrightarrow{\nabla}_R^{\,2}-\frac{\hbar^2}{2m}\overrightarrow{\nabla}_r^{\,2}-\frac{e^2}{r}\right)\psi(\vec{R},\ \vec{r})=E\psi(\vec{R},\ \vec{r}). \tag{4.6.3}$$

Writing the wave function

$$\psi(\vec{R},\ \vec{r})=\Psi(\vec{R})\psi(\vec{r}) \tag{4.6.4}$$

we can split the equation into two:

$$\begin{gathered}\left(-\tfrac{\hbar^2}{2M}\overrightarrow{\nabla}_R^{\,2}\right)\Psi(\vec{R})=E_R\Psi(\vec{R})\\ \left(-\tfrac{\hbar^2}{2m}\overrightarrow{\nabla}_r^{\,2}+V(\vec{r})\right)\psi(\vec{r})=Er\psi(\vec{r})\end{gathered} \tag{4.6.5}$$

and the total system energy is $E=E_R+E_r$. Note that the motion of the center of mass is (of course) just that of a free particle, having a trivial plane wave solution. From now on, we shall only be concerned with the *relative* motion of the particles. Since the proton is far heavier than the electron, we will almost always ignore the difference between the electron mass and the reduced mass, but it should be noted that the difference is easily detectable spectroscopically: for example, the lines shift if the proton is replaced by a deuteron (heavy hydrogen).

We’re ready to write Schrödinger’s equation for the hydrogen atom, dropping the r suffixes in the second equation above, and writing out $\vec{\nabla}^2$ explicitly in spherical coordinates:

$$\begin{gathered}-\tfrac{\hbar^2}{2m}\left(\tfrac{1}{r^2}\tfrac{\partial}{\partial r}\left(r^2\tfrac{\partial\psi}{\partial r}\right)+\tfrac{1}{r^2\sin\theta}\tfrac{\partial}{\partial\theta}\left(\sin\theta\tfrac{\partial\psi}{\partial\theta}\right)+\tfrac{1}{r^2\sin^2\theta}\tfrac{\partial^2\psi}{\partial\varphi^2}\right)-\tfrac{e^2}{r}\psi\\ =E\psi.\end{gathered} \tag{4.6.6}$$

Factoring Out the Angular Dependence: the Radial Equation - $\boldsymbol{R(r)}$

Since the potential is spherically symmetric, the Hamiltonian H commutes with the angular momentum operators L^2, L_z so we can construct a common set of eigenkets of the three operators H, L^2, L_z. The angular dependence of these eigenkets is therefore that of the Y_l^m’s, so the solutions must be of the form

$$\psi_{Elm}(r,\theta,\phi)=R_{Elm}(r)Y_l^m(\theta,\phi). \tag{4.6.7}$$

Now, notice that in the Schrödinger equation above, the angular part of $\vec{\nabla}^2$ is *exactly* the differential operator $L^2/2mr^2$, so operating on $\psi_{Elm}(r,\theta,\phi)=R_{Elm}(r)Y_l^m(\theta,\phi)$ it will give $\hbar^2l(l+1)/2mr^2$. The spherical harmonic Y_l^m can then be cancelled from the two sides of the equation leaving:

$$-\frac{\hbar^2}{2m}\left(\frac{1}{r^2}\frac{d}{dr}(r^2\frac{d}{dr})-\frac{l(l+1)}{r^2}\right)R_{El}(r)-\frac{e^2}{r}R_{El}(r) \\ =ER_{El}(r) \tag{4.6.8}$$

it now being apparent that $R(r)$ cannot depend on m.

The radial derivatives simplify if one factors out $1/r$ from the function R, writing

$$R_{El}(r)=\frac{u(r)}{r} \tag{4.6.9}$$

and temporarily suppressing the E and l to reduce clutter.

The equation becomes:

$$-\frac{\hbar^2}{2m}\left(\frac{d^2}{dr^2}-\frac{l(l+1)}{r^2}\right)u(r)-\frac{e^2}{r}u(r)=Eu(r). \tag{4.6.10}$$

Rearranging,

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2}+\left(\frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}-\frac{e^2}{r}\right)u(r)=Eu(r). \tag{4.6.11}$$

Note that this is the same as the Schrödinger equation for a particle in one dimension, restricted to $r>0$, in a potential (for $l\neq 0$) going to positive infinity at the origin, then negative and going to zero at large distances, so it always has a minimum for some positive r.

We are interested in *bound states* of the proton-electron system, so E will be a negative quantity. At large separations, the wave equation simplifies to

$$-\frac{\hbar^2}{2m}\frac{d^2u(r)}{dr^2}\cong E\,u(r)(for\ large\ r) \tag{4.6.12}$$

having approximate solutions $e^{\kappa r}$, $e^{-\kappa r}$ where $\kappa=\sqrt{-2mE/\hbar^2}$. The bound states we are looking for, of course, have exponentially *decreasing* wave functions at large distances.

Going to a Dimensionless Variable

To further simplify the equation, we introduce the dimensionless variable

$$\rho=\kappa r,\quad \kappa=\sqrt{-2mE/\hbar^2} \tag{4.6.13}$$

giving

$$\frac{d^2u(\rho)}{d\rho^2}=\left(1-\frac{2\nu}{\rho}+\frac{l(l+1)}{\rho^2}\right)u(\rho) \tag{4.6.14}$$

where (for reasons which will become apparent shortly) we have introduced ν defined by

$$2\nu=e^2\kappa/E. \tag{4.6.15}$$

Notice that in transforming from r to the dimensionless variable ρ *the scaling factor* κ *depends on energy*, so will be different for different energy bound states!

Consider now the behavior of the wave function near the origin. The dominant term for sufficiently small ρ is the centrifugal one, so

$$\frac{d^2u(\rho)}{d\rho^2}\cong\frac{l(l+1)}{\rho^2}u(\rho) \tag{4.6.16}$$

for which the solutions are $u(\rho)\sim\rho^{-l}$, $u(\rho)\sim\rho^{l+1}$. Since the wave function cannot be singular, we choose the second.

We have established that the wave function decays as $e^{-\kappa r}=e^{-\rho}$ at large distances, and goes as ρ^{l+1} close to the origin. Factoring out these two asymptotic behaviors, define $w(\rho)$ by

$$u(\rho)=e^{-\rho}\rho^{l+1}w(\rho). \tag{4.6.17}$$

It is straightforward (if tedious) to establish that $w(\rho)$ satisfies the differential equation:

$$\rho\frac{d^2w(\rho)}{d\rho^2}+2(l+1-\rho)\frac{dw(\rho)}{d\rho}+2(\nu-(l+1))w(\rho)=0. \tag{4.6.18}$$

Putting in a trial series solution $w(\rho)=\sum_{k=0}^{\infty}w_k\rho^k$ gives a recurrence relation between successive coefficients:

$$\frac{w_{k+1}}{w_k}=\frac{2(k+l+1-\nu)}{(k+1)(k+2(l+1))}. \tag{4.6.19}$$

For large values of k, $w_{k+1}/w_k\to 2/k$, so $w_k\sim 2^k/k!$ and therefore $w(\rho)\sim e^2\rho$. This means we have found the diverging radial wavefunction $u(\rho)\sim e^\rho$, which is in fact the correct behavior for general values of the energy.

To find the bound states, we must choose energies such that the series is *not* an infinite one. As long as the series stops somewhere, the exponential decrease will eventually take over, and yield a finite (bound state) wave function. Just as for the simple harmonic oscillator, this can only happen if for some k, $w_{k+1}=0$. Inspecting the ratio w_{k+1}/w_k, evidently the *condition for a bound state is* that

$$\nu=n,\quad an\ integer \tag{4.6.20}$$

in which case the series for $w(\rho)$ terminates at $k=n-l-1$. *From now on*, since we know that for the functions we're interested in ν is an integer, *we replace* ν *by* n.

To find the *energies* of these bound states, recall $2n=2\nu=e^2\kappa/E$ and $\kappa=\sqrt{-2mE/\hbar^2}$, so

$$4n^2=\frac{e^4\kappa_n^2}{E_n^2}=-\frac{e^4}{E_n^2}\frac{2mE_n}{\hbar^2}, \tag{4.6.21}$$

so

$$E_n=-\frac{me^4}{2\hbar^2}\frac{1}{n^2}=-\frac{13.6}{n^2}ev=-\frac{1}{n^2}Rydberg. \tag{4.6.22}$$

(This defines the *Rydberg*, a popular unit of energy in atomic physics.)

Remarkably, this is the very same series of bound state energies found by Bohr from his model! Of course, this had better be the case, since the series of energies Bohr found correctly accounted for the spectral lines emitted by hot hydrogen atoms. Notice, though, that there are some important differences with the Bohr model: the energy here is determined entirely by n, called the *principal quantum number*, but, in contrast to Bohr's model, n is *not* the angular momentum. The *true* ground state of the hydrogen atom, $n=1$, has *zero* angular momentum: since $n=k+l+1$, $n=1$ means both $l=0$ and $k=0$. The ground state wave function is therefore spherically symmetric, and the function $w(\rho)=w_0$ is just a constant. Hence $u(\rho)=\rho e^{-\rho}w_0$ and the actual radial wave function is this divided by r, and of course suitably normalized.

To write the wave function in terms of r, we need to find κ. Putting together $\rho=\kappa_n r$, $\kappa_n=\sqrt{-2mE_n/\hbar^2}$ and $E_n=-\frac{me^4}{2\hbar^2}\frac{1}{n^2}$,

$$\kappa_n=\sqrt{2m\frac{me^4}{2\hbar^2}\frac{1}{n^2}}/\hbar=\frac{me^2}{\hbar^2 n}=\frac{1}{a_0 n}, \tag{4.6.23}$$

where the length

$$a_0=\frac{\hbar^2}{me^2}=0.529\times 10^{-10}m. \tag{4.6.24}$$

is called the *Bohr radius*: it is in fact the radius of the lowest orbit in Bohr's model.

Exercise: check this last statement.

It is worth noting at this point that the energy levels can be written in terms of the Bohr radius a_0:

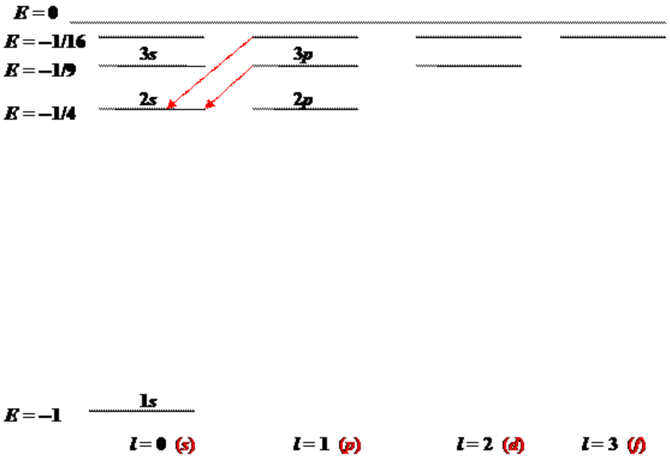
$$E_n = -\frac{e^2}{2a_0}\frac{1}{n^2}. \tag{4.6.25}$$

(This is actually obvious: remember that the energies E_n are identical to those in the Bohr model, in which the radius of the n^{th} orbit is $n^2 a_0$, so the electrostatic potential energy is $-e^2/na_0$, etc.)

Moving on to the excited states: for $n=2$, we have a choice: *either* the radial function $w(\rho)$ can have one term, as before, but now the angular momentum $l=1$ (since $n=k+l+1$); *or* $w(\rho)$ can have *two* terms (so $k=1$), and $l=0$. Both options give the same energy, -0.25 Ry, since n is the same, and the energy only depends on n. In fact, there are *four* states at this energy, since $l=1$ has states with $m=1,\ m=0$ and $m=-1$, and $l=0$ has the one state $m=0$. (For the moment, we are not counting the extra factor of 2 from the two possible *spin* orientations of the electron.)

For $n=3$, there are 9 states altogether: gives one, $l=1$ gives 3 and $l=2$ gives 5 different m values. In fact, for principal quantum number n there are n^2 degenerate states. (n^2 being the sum of the first n odd integers.)

The states can be mapped out, energy vertically, angular momentum horizontally:



The energy $E=-1/n^2$, the levels are labeled nl, n being the principal quantum number and the traditional notation for angular momentum l is given at the bottom of the diagram. The two red vertical arrows are the first two transitions in the spectroscopic Balmer series, four lines of which gave Bohr the clue that led to his model. The corresponding series of transitions to the $1s$ ground state are in the ultraviolet, they are called the Lyman series.

Wave Functions for some Low-*n* States

From now on, we label the wave functions with the quantum numbers, $\psi_{nlm}(r,\theta,\phi)$, so the ground state is the spherically symmetric $\psi_{100}(r)$.

For this state $R(r)=u(r)/r$, where $u(\rho)=e^{-\rho}\rho^{l+1}w(\rho)=e^{-\rho}\rho w_0$, with w_0 a constant, and $\rho=\kappa_1 r=r/a_0$.

So, as a function of r, $\psi_{100}(r)=Ne^{-r/a_0}$ with N an easily evaluated normalization constant:

$$\psi_{100}(r) = \left(\frac{1}{\pi a_0^3}\right)^{1/2} e^{-r/a_0}. \tag{4.6.26}$$

For $n=2,\ l=1$ the function $w(\rho)$ is still a single term, a constant, but now $u(\rho)=e^{-\rho}\rho^{l+1}w(\rho)=e^{-\rho}\rho^2 w_0$, and, for $n=2$, $\rho=\kappa r=r/2a_0$, remembering the energy-dependence of κ.

Therefore $\psi_{210}(r,\theta,\phi)=N\left(\frac{r}{a_0}\right)e^{-r/2a_0}\cos\theta$. Again, evaluating the normalization constant is routine, yielding

$$\psi_{210}(r,\theta,\phi)=\left(\frac{1}{32\pi a_0^3}\right)^{1/2}\left(\frac{r}{a_0}\right)e^{-r/2a_0}\cos\theta. \tag{4.6.27}$$

The wave functions for the other m -values, $\psi_{21\pm1}(r,\theta,\phi)$, have the $\cos\theta$ in ψ_{210} replaced by $\mp(1/\sqrt{2})\sin\theta e^{\pm i\phi}$ respectively (from the earlier discussion of the Y_l^m 's).

The other $n=2$ state has $l=0$, so from $n=k+l+1$, we have $k=1$ and the series for w has two terms, $k=0$ and $k=1$, the ratio being

$$\frac{w_{k+1}}{w_k}=\frac{2(k+l+1-n)}{(k+1)(k+2(l+1))}=-1 \tag{4.6.28}$$

for the relevant values: $k=0,\ l=0,\ n=2$. So $w_1=-w_0$, $w(\rho)=w_0(1-\rho)$. For $n=2,\ \rho=r/2a_0$, the normalized wave function is

$$\psi_{200}(r)=\left(\frac{1}{32\pi a_0^3}\right)^{1/2}\left(2-\frac{r}{a_0}\right)e^{-r/2a_0}. \tag{4.6.29}$$

Note that the zero angular momentum wave functions are nonzero and have nonzero slope at the origin. This means that the full three dimensional wave functions have a slope discontinuity there! But this is fine—the potential is infinite at the origin. (Actually, the proton is not a point charge, so really the kink will be smoothed out over a volume of the size of the proton—a very tiny effect.)

General Solution of the Radial Equation

In practice, the first few radial functions $w(\rho)$ can be constructed fairly easily using the method presented above, but it should be noted that the differential equation for $w(\rho)$

$$\rho\frac{d^2w(\rho)}{d\rho^2}+2(l+1-\rho)\frac{dw(\rho)}{d\rho}+2(n-(l+1))w(\rho)=0 \tag{4.6.30}$$

is in fact Laplace's equation, usually written

$$\left(z\frac{d^2}{dz^2}+(k-1-z)\frac{d}{dz}+p\right)L_p^k(z)=0 \tag{4.6.31}$$

where k,p are integers, and $L_p^k(z)$ is a *Laguerre polynomial* (Messiah, page 482).

The two equations are the same if $z=2\rho$, and the solution to the radial equation is therefore

$$w_{nl}(\rho)=L_{n-l-1}^{2l+1}(2\rho). \tag{4.6.32}$$

Quoting Messiah, the Laguerre polynomials $L_p^0(z)$, and the *associated Laguerre polynomials* $L_p^k(z)$ are given by:

$$\begin{gathered}L_p^0(z)=e^z\tfrac{d^p}{dz^p}e^{-z}z^p\\ L_p^k(z)=(-1)^k\tfrac{d^k}{dz^k}L_{p+k}^0(z).\end{gathered} \tag{4.6.33}$$

(These representations can be found neatly by solving Laplace's equation using – surprise – a Laplace transform. See Merzbacher for details.) The polynomials satisfy the orthonormality relations (with the mathematicians' normalization convention)

$$\int_0^\infty e^{-z}z^kL_p^kL_q^kdz=\frac{[(p+k)!]^3}{p!}\delta_{pq}. \tag{4.6.34}$$

But what do they look like? The function $e^{-z}z^p$ is zero at the origin (apart from the trivial case $p=0$) and zero at infinity, always positive and having nonzero slope except at its maximum value, $z=p$. The p derivatives bring in p separated zeroes, easily checked by sketching the curves generated by successive differentiation. Therefore, $L_p^0(z)$, a polynomial of degree p, has p real positive zeroes, and value at the origin $L_p^0(0)=p!$, since the only nonzero term at $z=0$ is that generated by all p differential operators acting on z^p.

The associated Laguerre polynomial $L_p^k(z)$ is generated by differentiating $L_{p+k}^0(z)$ k times. Now $L_{p+k}^0(z)$ has $p+k$ real positive zeroes, differentiating it gives a polynomial one degree lower, with zeroes which must be one in each interval between the zeroes of

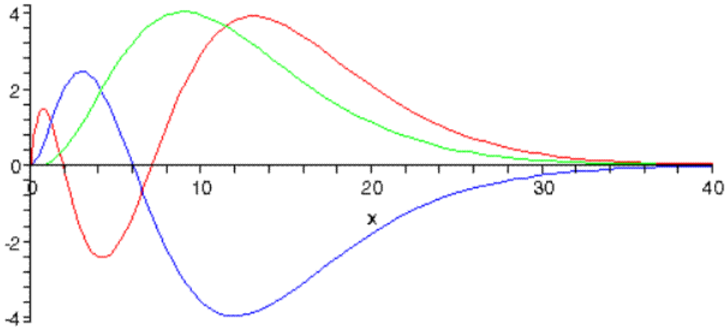
$L^0_{p+k}(z)$. This argument remains valid for successive differentiations, so $L^k_p(z)$ *must have* p *real separate zeroes*.

Putting all this together, and translating back from ρ to r, the radial solutions are:

$$R_{nl}(r) = Ne^{-r/na_0}\left(\frac{r}{na_0}\right)^l L^{2l+1}_{n-l-1}\left(\frac{2r}{na_0}\right) \tag{4.6.35}$$

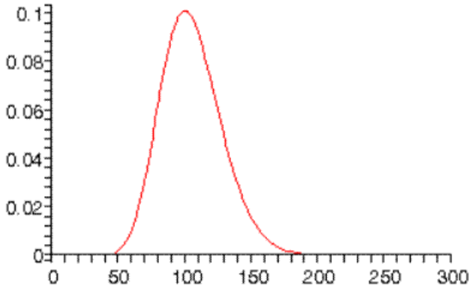
with N the normalization constant. Griffiths (page 141) gives more details, including the normalization constants worked out. We used those to plot the $n=3$ states—plotting here the functions $u(r) = rR(r)$, since the normalization is $4\pi\int_0^\infty |u(r)|^2 dr = 1$, $u(r)$ gives a better idea of at what distance from the proton the electron is most likely to be found.

Here are the three $n=3$ radial wave functions:

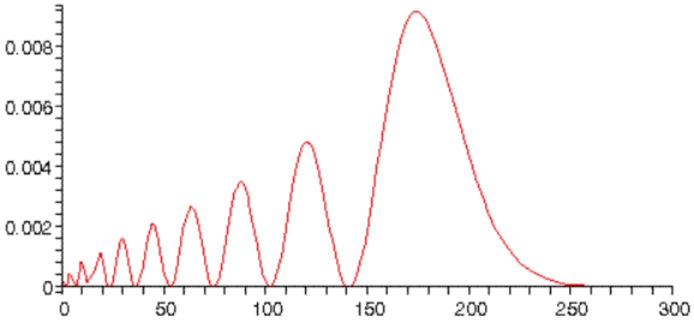


The number of nodes, the radial quantum number, is $3-l-1$. (*Note*: The relative normalizations are correct here, but not the overall normalization.)

For higher n values, the wave functions become reminiscent of classical mechanics. for example, for $n=10$, the highest angular momentum state probability distribution peaks at $r = 100a_0$, the Bohr orbit radius:



whereas for $n=10,\ l=0$, we find:



Notice this peaks just below twice the Bohr radius. This can be understood from classical mechanics: for an inverse square force law, elliptical orbits with the same semimajor axis have the same energy. The $l=n-1$ orbit is a circle, the $l=0$ orbit is a long thin ellipse (one end close to the proton), so it extends almost twice as far from the origin as the circle. Furthermore, the orbiting electron will spend longer at the far distance, since it will be moving very slowly. (*Note*: the normalizations in the above graphs are only approximate.)

4.7: Adding Angular Momenta

Consider a system having two angular momenta, for example an electron in a hydrogen atom having both orbital angular momentum and spin. The ket space for a single angular momentum has an orthonormal basis $|j,m\rangle$ so for two angular momenta an obvious orthonormal basis is the set of direct product kets $|j_1,m_1\rangle \otimes |j_2,m_2\rangle$. What does this mean, exactly? Suppose the first angular momentum $\vec{J}_1$ has magnitude $\vec{J}_1^{\,2} = \hbar^2 j_1(j_1+1)$, and is in the state $\sum_{m_1=-j_1}^{j_1} \alpha_{m_1}|j_1,m_1\rangle$, and similarly the second angular momentum $\vec{J}_2$ is in the state $\sum_{m_2=-j_2}^{j_2} \beta_{m_2}|j_2,m_2\rangle$. Evidently the probability amplitude for finding the first spin in state m_1 and at the same time the second in m_2 is $\alpha_{m_1}\beta_{m_2}$, and we denote that state by $|j_1,m_1\rangle \otimes |j_2,m_2\rangle$. How to handle these direct product spaces will become clear on examining specific examples, as we do below, beginning with two spins one-half.

Now the sum of two angular momenta $\vec{J} = \vec{J}_1 + \vec{J}_2$ is *itself* an angular momentum, operating in a space with a complete basis $|j,m\rangle$. This is easy to prove: the components of $\vec{J}_1$ satisfy $[J_{1i}, J_{1j}] = i\hbar\varepsilon_{ijk}J_{1k}$, and similarly for the components of $\vec{J}_2$. The components of $\vec{J}_1$ commute with the components of $\vec{J}_2$, of course, from which it follows immediately that the vector components of $\vec{J} = \vec{J}_1 + \vec{J}_2$ do indeed obey the angular momentum commutation relations: and recall that the commutation relations were sufficient to determine the allowed sets of eigenvalues. We shall prove later that the eigenstates $|j,m\rangle$ of $\vec{J}^{\,2}, \vec{J}_z$ *are* a complete basis for the product space of the eigenkets of $\vec{J}_1^{\,2},\ \vec{J}_2^{\,2},\ J_{1z},\ J_{2z}$—to establish this, we must first find the possible allowed values of the total angular momentum quantum number j.

Here we have, then, two different orthonormal bases for what is evidently the same vector space. In practical applications, it often turns out that we have to translate from one of these bases to the other. Our present task is to construct the appropriate transformation: we accomplish this by finding the coefficients of any $|j,m\rangle$ in the $|j_1,m_1\rangle \otimes |j_2,m_2\rangle$ basis. These are called the *Clebsch-Gordan* coefficients.

We shall build gradually, beginning with adding two spins one-half, then a spin one-half with an orbital angular momentum, finally two general angular momenta.

Adding Two Spins: the Basis States and Spin Operators

The most elementary example of a system having two angular momenta is the hydrogen atom in its ground state. The orbital angular momentum is zero, the electron has spin angular momentum $\frac{1}{2}\hbar$, and the proton has spin $\frac{1}{2}\hbar$.

The space of possible states of the electron spin has the two basis kets $|\uparrow\rangle_e$ and $|\downarrow\rangle_e$, (also variously written as $|+\rangle, |-\rangle$; $\begin{pmatrix}1\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\end{pmatrix}$; $\chi+ ,\chi-!$) the basis proton spin kets are $|\uparrow\rangle_p$ and $|\downarrow\rangle_p$, so the possible states of the combined system are kets in the direct product space which has a basis of four kets:

$$|\uparrow_e\uparrow_p\rangle,\quad |\uparrow_e\downarrow_p\rangle,\quad |\downarrow_e\uparrow_p\rangle,\quad |\downarrow_e\downarrow_p\rangle \tag{4.7.1}$$

using $|\uparrow_e\uparrow_p\rangle$ as shorthand for $|\uparrow\rangle_e \otimes |\uparrow\rangle_p$.

Note here that we've written the kets in "alphabetical order" with ↑ as the first letter, ↓ as the second. That is to say, we've first written all the kets having ↑ as the first letter, etc.

For the more general case of adding j_1 to j_2, to be considered shortly, we'll order the kets in the same "alphabetical" way, writing first all the kets having $m_1 = j_1$, and so on down to $m_1 = -j_1$:

$$(j_1j_2, j_1(j_2-1), \ldots, j_1(-j_2),\ (j_1-1)j_2, (j_1-1)(j_2-1), \ldots, (j_1-1)(-j_2), \ldots, (-j_1)(-j_2)). \tag{4.7.2}$$

The dimensionality of this space is then $(2j_1+1)\times(2j_2+1)$.

Now the first block of $2j_2+1$ elements all have the same m- component of j_1, that is, $m_1 = j_1$, the next block has $m_1 = j_1 - 1$, and so on. Think about what this means for constructing a *rotation operator* acting on the kets in this space: if it operates only on the angular momentum j_1, it will change the factors m_1 multiplying the blocks, if the operator rotates only j_2, it will operate *within* each block, all the blocks being changed in the same way.

To get a feeling for how this works in practice, we go back to the simplest case, two spins one-half.

The space is four-dimensional, having basis $|\uparrow_e\uparrow_p\rangle,\quad |\uparrow_e\downarrow_p\rangle,\quad |\downarrow_e\uparrow_p\rangle,\quad |\downarrow_e\downarrow_p\rangle$.

Any operator acting on the spins will be represented by a 4×4 matrix, best thought of as a 2×2 matrix made up of 2×2 blocks: an operator acting on the proton spin acts *within* the blocks, one operating on the electron spin affects the overall multiplying factors in front of each block.

Let’s look at a few examples. Recall that the raising operator for a single spin is the 2×2 matrix $S^{+}=\hbar\begin{pmatrix}0&1\\0&0\end{pmatrix}$. So what is the raising operator for the *electron* spin?

$$S_e^{+}\otimes I_p=\hbar\begin{pmatrix}0&I\\0&0\end{pmatrix}=\hbar\begin{pmatrix}0&0&1&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{pmatrix}. \tag{4.7.3}$$

We use bold to denote 2×2 matrices.

The pattern is clear: the big structure (in **bold** above), that of the four 2×2 blocks, reflect the structure of the electron spin operator $S^{+}=\hbar\begin{pmatrix}0&1\\0&0\end{pmatrix}$, within those blocks (of which only one survives) the identity operator $I=\begin{pmatrix}1&0\\0&1\end{pmatrix}$ acts on the proton spin.

The operator that raises the *proton* spin is:

$$I_e\otimes S_p^{+}=\hbar\begin{pmatrix}\sigma^{+}&0\\0&\sigma^{+}\end{pmatrix}=\hbar\begin{pmatrix}0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{pmatrix}. \tag{4.7.4}$$

What about the operator that raises *both* electron and proton spin? In this case, the pattern of blocks, and the pattern within each block, must both be $\begin{pmatrix}0&1\\0&0\end{pmatrix}$, so

$$S_e^{+}\otimes S_p^{+}=\hbar^2\begin{pmatrix}0&\sigma^{+}\\0&0\end{pmatrix}=\hbar^2\begin{pmatrix}0&0&0&1\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}. \tag{4.7.5}$$

There is only one nonzero matrix element because only one member of the base survives this operation.

If two spins interact (via their magnetic moments, for example) in a way that preserves total angular momentum, a possible term in the Hamiltonian would be $S_e^{-}S_p^{+}$, represented by:

$$S_e^{-}S_p^{+}=\hbar\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&1&0&0\\0&0&0&0\end{pmatrix}=\hbar S_e^{-}\otimes S_p^{+}. \tag{4.7.6}$$

Representing the Rotation Operator for Two Spins

Recall from the lecture on spin that the rotation operator on a single spin one-half is

$$D^{(1/2)}(R(\theta\hat{\vec{n}}))=e^{\frac{-i\theta\hat{\vec{n}}\cdot\vec{J}}{\hbar}}=e^{-i(\theta/2)(\hat{\vec{n}}\cdot\vec{\sigma})}=I\cos\frac{\theta}{2}-i(\hat{\vec{n}}\cdot\vec{\sigma})\sin\frac{\theta}{2} \tag{4.7.7}$$

in the 2×2 spin or space. As we established, this matrix operator has the form

$$\begin{pmatrix}a&b\\-b^{*}&a^{*}\end{pmatrix}=\begin{pmatrix}\cos(\theta/2)-in_z\sin(\theta/2)&-(in_x+n_y)\sin(\theta/2)\\(-in_x+n_y)\sin(\theta/2)&\cos(\theta/2)+in_z\sin(\theta/2)\end{pmatrix}. \tag{4.7.8}$$

with $|a|^2+|b|^2=1$.

This set of unitary 2×2 matrices form a *representation of the rotation group* in the sense that the total resulting from two successive rotations is given by the matrix which is the matrix product of those corresponding to the two rotations.

From the discussion in the previous section, it should be clear that in the product space of the two spins, the representation of the rotation operator—both spins of course undergoing the same rotation—is:

$$\begin{pmatrix} a\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} & b\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \\ -b^*\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} & a^*\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a^2 & ab & ab & b^2 \\ -ab^* & aa^* & -bb^* & a^*b \\ -ab^* & -bb^* & aa^* & a^*b \\ b^{*2} & -a^*b^* & -a^*b^* & a^{*2} \end{pmatrix}. \tag{4.7.9}$$

This set of 4×4 matrices, again with $|a|^2+|b|^2=1$, must also form a *representation of the rotation group* over the four-dimensional space. We shall shortly discover that this representation can be simplified, but to achieve that we need to analyze the states in terms of *total* angular momentum.

Representing States of Two Spins in Terms of Total Angular Momentum

We're now ready to look at total spin states for the ground-state (zero *orbital* angular momentum) hydrogen atom.

Consider first the state with both electron and proton spin pointing upwards, $|\uparrow\uparrow\rangle$. The z- component of the total spin is $S_z=S_z^e+S_z^p$, so $S_z=\hbar$. Labeling the total spin state $|s,m\rangle$, we have a state with $m=1$, so $s=1$. (To confirm that this state indeed has $s=1$ we can apply the total-spin raising operator $S_+=S_+^e+S_+^p$. Since both component spins have maximum m value, $S_+|s,1\rangle_{sm}=(S_+^e+S_+^p)|\uparrow\uparrow\rangle=0$, but S_+ only gives zero when acting on the m=s member of a multiplet.)

We find, then, that $|\uparrow\uparrow\rangle=|1,1\rangle_{sm}$ where we've added the suffix sm to make clear that the numbers in the last ket signify $|s,m\rangle$ for the total spin. The total spin $s=1$, being a total angular momentum eigenstate, has a triplet of m values, $m=1,0,-1$, $|1,1\rangle_{sm}$ being the top member. The $m=0$ member is found by applying the lowering operator to $|\uparrow\uparrow\rangle$:

$$S_-|\uparrow\uparrow\rangle=(S_-^e+S_-^p)(|\uparrow\rangle_e\otimes|\uparrow\rangle_p)=\hbar|\downarrow\rangle_e\otimes|\uparrow\rangle_p+\hbar|\uparrow\rangle_e\otimes|\downarrow\rangle_p \tag{4.7.10}$$

which together with

$$S_-|\uparrow\uparrow\rangle=S_-|1,1\rangle_{sm}=\sqrt{2}\hbar|1,0\rangle_{sm}, \tag{4.7.11}$$

gives

$$|1,0\rangle_{sm}=\frac{1}{\sqrt{2}}(|\downarrow\uparrow\rangle+|\uparrow\downarrow\rangle). \tag{4.7.12}$$

Obviously, the third member of the triplet, $|1,-1\rangle_{sm}=|\downarrow\downarrow\rangle$.

But this triplet only accounts for *three* basis states in the $|s,m\rangle$ total angular momentum representation. A fourth state, orthogonal to these three and normalized, is $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle-|\downarrow\uparrow\rangle)$. This has $m=0$, and also has $s=0$, easily checked by noting that the raising operator acting on this state gives zero, so the state has the maximum allowed m for its s value.

To summarize: in the total angular momentum |s,m\rangle representation for two spins one-half, the four basis states are $|1,1\rangle_{sm},\quad |1,0\rangle_{sm},\quad |1,-1\rangle_{sm},\quad |0,0\rangle_{sm}$ This orthonormal basis spans the same space as the other orthonormal set $|\uparrow\uparrow\rangle,\quad |\uparrow\downarrow\rangle,\quad |\downarrow\uparrow\rangle,\quad |\uparrow\uparrow\rangle$. Our construction of the $|s,m\rangle$ states above amounts to finding one set of basis kets in terms of the others.

Note that since both sets of basis kets are orthonormal, mapping a vector from one set to the other is a unitary transformation. But there's more: the coefficients we found expressing one basis ket in the other basis are all *real*. This means that if any ket has real coefficients in one basis, it does in the other. For this special case of all real coefficients, a unitary transformation is termed *orthogonal*.

The orthogonal transformation expressing one base in terms of the other is easy to construct:

$$\begin{pmatrix} |1,1\rangle_{sm} \\ |1,0\rangle_{sm} \\ |0,0\rangle_{sm} \\ |1,-1\rangle_{sm} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} |\uparrow\uparrow\rangle \\ |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \end{pmatrix}. \tag{4.7.13}$$

The matrix is orthogonal and symmetric, so is its own inverse.

Geometrically, $s=1$ means the component spins are parallel, for $s=0$ they are antiparallel. This can be stated more precisely: $2\vec{S}_1\cdot\vec{S}_2=S^2-S_1^2-S_2^2$, so for $s=1$, $\vec{S}_1\cdot\vec{S}_2=\frac{1}{2}\hbar^2\left(2-\frac{3}{4}-\frac{3}{4}\right)=\hbar^2/4$, and for $s=0$ $\vec{S}_1\cdot\vec{S}_2=-\frac{3}{4}\hbar^2$. This makes it easy to construct projection operators into the $s=0$ and $s=1$ subspaces: $P_{s=1}=(\vec{S}_1\cdot\vec{S}_2/\hbar^2)+\frac{3}{4}$.

Physics example: an interesting case of a two-spin system is the hydrogen *molecule*. The electron spins are in the singlet state (otherwise the molecule disassociates) but the two proton spins, which interact through their magnetic moments) can be parallel (total spin one), this is called orthohydrogen, or antiparallel (parahydrogen). The energy difference is sufficiently small that at room temperature the ratio of ortho to para is 3:1, meaning that all spins states are equally probable (effectively infinite temperature), but at lower temperatures the lower energy para form dominates. This is in fact relevant to liquid hydrogen storage technology: the conversion rate from ortho to para is very slow, but when it takes place energy is released. If this happens after storage, additional refrigeration is required. To prevent this, catalysts can be used to hasten the conversion rate during cooling.

Representing the Rotation Operator in the Total Angular Momentum Basis

We've already established that the rotation operator, acting on the two spin system, can be represented by a 4×4 matrix, and that the new (total angular momentum) basis can be reached from the original (two separate spin) basis by the orthogonal transformation given explicitly above. Therefore, pre-and post-multiplying the two-spin rotation operator will in fact give a 4×4 matrix representation of the rotation operator in the new total angular momentum basis.

However, that approach misses the point: first, the singlet state $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle-|\downarrow\uparrow\rangle)$ has *zero* angular momentum, and so is *not changed by rotation*.

Second, the triplet state has angular momentum one, so rotation operators must act on it just as we found earlier for an angular momentum one:

$$D^{(1)}(R(\vec{\theta}))=e^{\frac{-i\theta\hat{\vec{n}}\cdot\vec{J}}{\hbar}}=I+(\cos\theta-1)\left(\frac{\hat{\vec{n}}\cdot\vec{J}}{\hbar}\right)^2-i\sin\theta\left(\frac{\hat{\vec{n}}\cdot\vec{J}}{\hbar}\right). \tag{4.7.14}$$

This means that, as far as rotations are concerned, the space spanned by the four kets $|0,0\rangle_{sm}$, $|1,1\rangle_{sm}$, $|1,0\rangle_{sm}$, $|1,-1\rangle_{sm}$is actually a *sum* of two separate subspaces, the one-dimensional space $|0,0\rangle_{sm}$, and the three-dimensional space having basis $|1,1\rangle_{sm}$, $|1,0\rangle_{sm}$, $|1,-1\rangle_{sm}$ Under rotation, a vector in one of these subspaces stays there: there are no cross terms in the matrix mixing the spaces.

This means that the rotation matrix has the form $\begin{pmatrix}I&0\\0&R_3\end{pmatrix}$ where R_3 is the 3×3 matrix for spin one, I is just the 1×1 trivial matrix in the singlet subspace, in other words 1, and the 0's are 1×3 and 3×1 sets of zeroes.

A state of the spins can of course be a sum of components in the two subspaces, for example

$$|\uparrow\downarrow\rangle=\frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle+|\downarrow\uparrow\rangle)+\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle-|\downarrow\uparrow\rangle)\right). \tag{4.7.15}$$

Reducible and Irreducible Group Representations

We began our discussion of two spins one-half by examining properties of spin operators in the four-dimensional product space of the two two-dimensional spin spaces, and went on to construct a four-dimensional representation of the general rotation operator in that space: a matrix representation of the rotation group. But when the two-spin system is labeled in terms of total angular momentum, we find that in fact this four-dimensional rotation operator is a sum of a three-dimensional rotation, and a trivial identity rotation for an angular momentum zero state. The four-dimensional operator can be "diagonalized": the space split into a three dimensional space and a one-dimensional space that don't mix under rotation, and any state of the system is a sum of kets from the two spaces.

This is often expressed by saying the *product* space of two spins one-half is the *sum* of a spin one space and a spin zero space, and written

$$\frac{1}{2}\otimes\frac{1}{2}=1\oplus0. \tag{4.7.16}$$

Putting in the dimensionalities of the spaces in this equation,

$$2\times 2=3+1. \tag{4.7.17}$$

This simple check on total dimensionality sets the pattern for more complicated product spaces examined below.

The 4×4 representation of the rotation operator is said to be a *reducible* representation: it can be reduced to a sum of smaller dimensional representations. An *irreducible* representation is one in which there are no subspaces invariant under all rotations.

Recall that we constructed the reducible 4×4 representation by taking a direct product of the 2×2 spin one-half representations of the rotation group. The equation $\frac{1}{2}\otimes\frac{1}{2}=1\oplus 0$ we used above to describe the ket spaces is also often used to describe the rotation group representations within those subspaces.

One might wonder why we would bother to build two different bases for the same vector space. The reason is that different problems need different bases. For a system of two spins in an external magnetic field, not interacting with each other, the independent spins basis $|\uparrow\uparrow\rangle$, etc., is natural. On the other hand, for a hydrogen atom in *no* external field, but including an interaction between the spins (which are aligned with the magnetic dipole moments of the particles) the $|j,m\rangle$ basis is the right one: the interaction Hamiltonian is proportional to $\vec{S}^e\cdot\vec{S}^p$, which can be written $\frac{1}{2}(S_x^e+iS_y^e)(S_x^p-iS_y^p)+\frac{1}{2}(S_x^e-iS_y^e)(S_x^p+iS_y^p)+S_z^eS_z^p$, where we recognize the raising and lowering operators for the individual spins. This means that the state $|\uparrow\downarrow\rangle$, for example, cannot be an eigenstate if the Hamiltonian includes $\vec{S}^e\cdot\vec{S}^p$, but for this case the states $|j,m\rangle$ *are* eigenstates because $\vec{S}^e\cdot\vec{S}^p$ commutes with the *total* angular momentum and its components.

But what would be a good basis for a hydrogen atom, including the $\vec{S}^e\cdot\vec{S}^p$ term, *and* in an external magnetic field? That is a nice exercise for the reader.

Adding a Spin to an Orbital Angular Momentum

In this section, we consider a hydrogen atom in a state with nonzero *orbital* angular momentum, $\vec{L}\neq 0$. Such orbital motion is equivalent to an electric current loop and generates a magnetic field. The magnetic dipole moment associated with the electron spin interacts with this field, the appropriate Hamiltonian having a term proportional to $\vec{L}\cdot\vec{S}$, and is termed the spin-orbit interaction. The proton also has a magnetic moment, but that is three orders of magnitude smaller than the electron's, so we'll neglect it for now.

The spin-orbit interaction $\vec{L}\cdot\vec{S}$ is most naturally analyzed in the basis states of *total* angular momentum, $|j,m\rangle$, where $\vec{J}=\vec{L}+\vec{S}$ (see the analogous discussion of the spin-spin interaction above). Write the orbital angular momentum eigenstates $|l,m_l\rangle$ and the spin states $|s,m_s\rangle$ where $|\frac{1}{2},\frac{1}{2}\rangle=|\uparrow\rangle$ and $|\frac{1}{2},-\frac{1}{2}\rangle=|\downarrow\rangle$. The product space $|l,m_l\rangle\otimes|s,m_s\rangle$ is $2(2l+1)$ dimensional: a single ket in this product space would be fully described by $|l,m_l;s,m_s\rangle$, but since both l,s are constant throughout the problem, the only actual *variables* are m_l,m_s so we'll write the ket in the more compact form $|m_l,m_s\rangle_{m_lm_s}$, for example $|2,\frac{1}{2}\rangle_{m_lm_s}$.

The maximum possible angular momentum component in the z- direction is clearly $(l+\frac{1}{2})\hbar$, for the state $|l,\frac{1}{2}\rangle_{m_1m_2}$. In the *total* angular momentum representation, this must be the state $|j,m\rangle=|l+\frac{1}{2},l+\frac{1}{2}\rangle_{jm}$. So the two different bases have a common member:

$$|l+\frac{1}{2},l+\frac{1}{2}\rangle_{jm}=|l,\frac{1}{2}\rangle_{m_lm_s}. \tag{4.7.18}$$

In the total angular momentum $|j,m\rangle$ representation, $|l+\frac{1}{2},l+\frac{1}{2}\rangle_{jm}$ is the top m state of a multiplet having $2(l+\frac{1}{2})+1=2l+2$ members. Just as for the spin-spin case, the next member down of the multiplet is generated by applying the lowering operator:

$$\begin{aligned}J_-|l+\tfrac{1}{2},l+\tfrac{1}{2}\rangle_{jm}&=\sqrt{2l+1}\;\hbar|l+\tfrac{1}{2},l-\tfrac{1}{2}\rangle_{jm}\\&=(L_-+S_-)|l,\tfrac{1}{2}\rangle_{m_lm_s}\\&=\sqrt{2l}\;\hbar|l-1,\tfrac{1}{2}\rangle_{m_lm_s}+\hbar|l,-\tfrac{1}{2}\rangle_{m_lm_s}\end{aligned}\quad. \tag{4.7.19}$$

Therefore

$$|l+\frac{1}{2},l-\frac{1}{2}\rangle_{jm}=\sqrt{\frac{2l}{2l+1}}|l-1,\frac{1}{2}\rangle_{m_lm_s}+\sqrt{\frac{1}{2l+1}}|l,-\frac{1}{2}\rangle_{m_lm_s}. \tag{4.7.20}$$

This state $|l+\frac{1}{2},l-\frac{1}{2}\rangle_{jm}$ lies in the $m=l-\frac{1}{2}$ subspace, which is two-dimensional, having basis vectors $|l-1,\frac{1}{2}\rangle_{m_lm_s}$ and $|l,-\frac{1}{2}\rangle_{m_lm_s}$ in the $|\rangle_{m_lm_s}$ representation. So it must have two basis vectors in the $|\rangle_{jm}$ representation as well. The other $|\rangle_{jm}$ ket must be orthogonal to $|l+\frac{1}{2},l-\frac{1}{2}\rangle_{jm}$ and normalized: it can only be

$$|l-\frac{1}{2},l-\frac{1}{2}\rangle_{jm}=\sqrt{\frac{2l}{2l+1}}|l,-\frac{1}{2}\rangle_{m_lm_s}-\sqrt{\frac{1}{2l+1}}|l-1,\frac{1}{2}\rangle_{m_lm_s}. \tag{4.7.21}$$

We've represented this new ket in $|\rangle_{jm}$ as the top state of a $j=l-\frac{1}{2}$ multiplet. It's easy to check that this is indeed the case: it has $m=l-\frac{1}{2}$, and J_+ acting on it gives zero, so it has to be the top member of its multiplet. The only ambiguity is an overall phase: the Condon-Shortley convention is that the highest m- state of the larger component angular momentum is assigned a positive coefficient.

So $|l-\frac{1}{2},l-\frac{1}{2}\rangle_{jm}$ is the top state of a new multiplet having $2(l-\frac{1}{2})+1=2l$ members. The two multiplets $j=l+\frac{1}{2}$ and $j=l-\frac{1}{2}$ taken together have $2(2l+1)$ members, and therefore span the whole $2(2l+1)$ dimensional space. The rest of the $|\rangle_{jm}$ basis vectors are generated by repeated application of the lowering operator in the two multiplets.

The reason there are only two multiplets in this problem is that there are only two ways the spin one-half can point relative to the orbital angular momentum. Recalling that for the two spins we expressed the product space a sum of a spin 1 space and a spin 0 space, $\frac{1}{2}\otimes\frac{1}{2}=1\oplus 0$, the analogous equation here is

$$\frac{1}{2}\otimes l=(l+\frac{1}{2})\oplus(l-\frac{1}{2}). \tag{4.7.22}$$

For the general case of adding angular momenta $j_1,\ j_2$ with $j_1\geq j_2$, $2j_2+1$ multiplets are generated, corresponding to the number of possible relative orientations of the two angular momenta.

Adding Two Angular Momenta: the General Case

The space of kets describing two angular momenta $j_1,\ j_2$ is the direct product of two spaces each for a single angular momentum, but the direct product nature of the kets is usually not made explicit, $|j_1,m_1\rangle\otimes|j_2,m_2\rangle$ can be written as a single ket $|j_1,m_1;j_2,m_2\rangle$. Just as in the examples above, since $j_1,\ j_2$ are fixed throughout, they don't need to be written into every ket, we'll just write $|m_1,m_2\rangle$, or, when dealing with numerical values, append m_1m_2 as a suffix: $|2,3\rangle_{m_1m_2}$.

The kets $|m_1,m_2\rangle$ form a *complete orthonormal basis* of the $(2j_1+1)(2j_2+1)$ dimensional product space of the two angular momenta: they are the eigenstates of the complete set of commuting variables $\vec{J}^2_1,\quad J_{1z},\quad \vec{J}^2_2,\quad J_{2z}$.

Total Angular Momentum Basis States

There is of course an alternative complete orthogonal basis of the space of the two angular momenta: for total angular momentum $\vec{J}=\vec{J}_1+\vec{J}_2$, a different set of complete commuting variables is: $\vec{J}^2_1,\quad \vec{J}^2_2,\quad \vec{J}^2,\ J_z$. (This is *not* the same set of states as in the previous paragraph: for example, $\vec{J}^2$ does *not* commute with J_{1z}. Check it out!)

This alternative set is a better basis set for two angular momenta interacting with each other—an interaction term like $\vec{J}_1\cdot\vec{J}_2$ can change $m_1,\ m_2$ but not $m=m_1+m_2$, or $\vec{J}^2$.

As always, we're taking $\vec{J}^2_1,\quad \vec{J}^2_2$ to be constants throughout, so the significant variables here are $\vec{J}^2$ and J_z, and we write the states simply as $|j,m\rangle$ or when we have numerical values, $|3,1\rangle_{jm}$, following the notation introduced above. Of course, $\vec{J}^2|j,m\rangle=j(j+1)\hbar^2|j,m\rangle$, and $J_z|j,m\rangle=m\hbar|j,m\rangle$.

Going from One Basis to the Other: the Clebsch-Gordan Coefficients

How do we write a state $|j,m\rangle$ in terms of the states $|m_1,m_2\rangle$? Furthermore, how do we prove the new set of states $|j,m\rangle$ is a complete basis for the space?

We know that the set of states $|m_1,m_2\rangle$ is a complete basis, since the whole space is a product space of the j_1 and j_2 spaces, which are spanned by the sets $|m_1\rangle$, $|m_2\rangle$ respectively. Therefore, the identity operator can be written

$$I=\sum_{m_1=-j_1}^{j_1}\sum_{m_2=-j_2}^{j_2}|m_1,m_2\rangle\langle m_1,m_2|. \tag{4.7.23}$$

It follows that $|j,m\rangle$ can be expressed as a sum over the basis vectors $|m_1,m_2\rangle$:

$$|j,m\rangle=\sum_{m_1=-j_1}^{j_1}\sum_{m_2=-j_2}^{j_2}|m_1,m_2\rangle\langle m_1,m_2|j,m\rangle \tag{4.7.24}$$

The coefficients $\langle m_1,m_2|j,m\ \rangle$ are called the *Clebsch Gordan coefficients*, often written CG coefficients.

One immediate property of the CG coefficients is that $\langle m_1,m_2|j,m\ \rangle=0$ unless $m=m_1+m_2$. This follows from the operator identity $J_z=J_{1z}+J_{2z}$ taken between a bra and a ket from different bases,

$$\langle m_1,m_2|J_z|j,m\rangle=\langle m_1,m_2|J_{1z}+J_{2z}|j,m\rangle \tag{4.7.25}$$

and

$$J_z|j,m\rangle=m\hbar|j,m\rangle,\ \langle m_1,m_2|(J_{1z}+J_{2z})=\langle m_1,m_2|(m_1+m_2)\hbar, \tag{4.7.26}$$

so

$$(m-m_1-m_2)\langle m_1,m_2|j,m\rangle=0. \tag{4.7.27}$$

We already know that the maximum value of m_1 is j_1, and of m_2 is j_2, so the maximum value of m is j_1+j_2 . Therefore, the maximum value of $j=j_1+j_2$, because if it could go any higher, there would be a higher m somewhere in the space, contradicting $m=m_1+m_2$.

For the set $|m_1,m_2\rangle$, there is one ket having this maximal value of *m*: $|j_1,j_2\rangle_{m_1m_2}$. Equally, in the set of states $|j,m\rangle$ there is only one with the maximal *m*: $|j_1+j_2,\ j_1+j_2\rangle_{jm}$. Therefore, these two kets must be identical (setting the arbitrary phase factor equal to one):

$$|j_1,j_2\rangle_{m_1m_2}=|j_1+j_2,\ j_1+j_2\rangle_{jm}. \tag{4.7.28}$$

Now $|j_1+j_2,\ j_1+j_2\rangle_{jm}$ is the top ket in a multiplet having $2(j_1+j_2)+1$ members.

The next-to-top member of the multiplet is generated as before by applying the lowering operator to both representations:

$$J_-|j_1+j_2,\ j_1+j_2\rangle_{jm}=(J_{1-}+J_{2-})|j_1,\ j_2\rangle_{m_1m_2} \tag{4.7.29}$$

giving

$$\sqrt{2(j_1+j_2)}\hbar|j_1+j_2,\ j_1+j_2-1\rangle_{jm}=\sqrt{2j_1}\hbar|j_1-1,\ j_2\rangle_{m_1m_2}+\sqrt{2j_2}\hbar|j_1,\ j_2-1\rangle_{m_1m_2} \tag{4.7.30}$$

so

$$|j_1+j_2,\ j_1+j_2-1\rangle_{jm}=\sqrt{\frac{j_1}{j_1+j_2}}|j_1-1,\ j_2\rangle_{m_1m_2}+\sqrt{\frac{j_2}{j_1+j_2}}|j_1,\ j_2-1\rangle_{m_1m_2} \tag{4.7.31}$$

and by exact analogy with the spin orbit case, the other $|\rangle_{jm}$ basis state in the $m=j_1+j_2-1$ subspace is

$$|j_1+j_2-1,\ j_1+j_2-1\rangle_{jm}=-\sqrt{\frac{j_2}{j_1+j_2}}|j_1-1,\ j_2\rangle_{m_1m_2}+\sqrt{\frac{j_1}{j_1+j_2}}|j_1,\ j_2-1\rangle_{m_1m_2} \tag{4.7.32}$$

with the appropriate sign convention for $j_1>j_2$. This is the top member of a multiplet having $j=j_1+j_2-1$, and so $2(j_1+j_2-1)+1=2(j_1+j_2)-1$ members (checked as usual by applying J_+ and getting zero).

To proceed further, the lowering operator is applied once more, to enter the $m=j_1+j_2-2$ subspace. In the $|\rangle_{m_1m_2}$ representation, this has *three* independent basis vectors (provided $j_2>\frac{1}{2}$): $|j_1-2,\ j_2\rangle_{m_1m_2},\ \ |j_1-1,\ j_2-1\rangle_{m_1m_2},\ \ |j_1,\ j_2-2\rangle_{m_1m_2}$. But only two kets have been lowered in the $|\rangle_{jm}$ representation—the missing third $|\rangle_{jm}$ ket in the $m=j_1+j_2-2$ subspace must be the top member of *another* new multiplet having $j=j_1+j_2-2$, and so $2(j_1+j_2)-3$ members.

Note that the coefficients generated by the lowering operators are all real, so all three $|\rangle_{jm}$ kets in the $m=j_1+j_2-2$ subspace can be written in terms of the $|\rangle_{m_1m_2}$ kets with real coefficients.

This process can be repeated until the $|\rangle_{jm}$ multiplets generated span the space. Recall that the dimensionality of the space, from the $|\rangle_{m_1m_2}$ representation, is $(2j_1+1)(2j_2+1)$. The multiplets in $|\rangle_{jm}$ add to a total dimensionality

$$2(j_1+j_2)+1+2(j_1+j_2)-1+2(j_1+j_2)-3+\ldots \tag{4.7.33}$$

but where do we stop? Common sense suggests that for $j_1>j_2$, the minimum total angular momentum must be $j=j_1-j_2$. Common sense is not necessarily to be trusted, but it is clear that all the members of the multiplets in $|\rangle_{jm}$ generated by using the lowering operator, followed by introducing a new orthogonal multiplet top member each time, as described above, are independent orthonormal kets, and if we stop at $j=j_1-j_2$, the total number generated is

$$\sum_{n=|j_1-j_2|}^{j_1+j_2}(2n+1)=(2j_1+1)(2j_2+1). \tag{4.7.34}$$

(Use $\sum_{n=0}^{m}(2n+1)=(m+1)^2$.) This establishes that including all total angular momenta between $|j_1-j_2|$ and j_1+j_2 does in fact give a complete basis spanning the space, so

$$j_1\otimes j_2=(j_1+j_2)\oplus(j_1+j_2-1)\oplus\cdots\oplus(|j_1-j_2|). \tag{4.7.35}$$

Calculating Clebsch-Gordan Coefficients Using Recursion Relations

The scheme presented above, constructing a succession of multiplets beginning from the highest *m* state and using the Condon-Shortley convention to settle signs, will generate all the CG coefficients. However, another approach proves useful in later work. Recall that by finding matrix elements of $J_z=J_{1z}+J_{2z}$ between a $\langle m_1,m_2|$ bra and a $|j,m\rangle$ ket, we established that the Clebsch-Gordan coefficients are zero unless $m=m_1+m_2$. A parallel evaluation of matrix elements of $J_\pm=J_{1\pm}+J_{2\pm}$ yields a relationship between three CG coefficients:

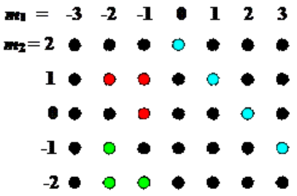
$$\langle m_1,m_2|J_+|j,m\rangle=\langle m_1,m_2|J_{1+}|j,m\rangle+\langle m_1,m_2|J_{2+}|j,m\rangle \tag{4.7.36}$$

yields

$$\begin{gathered}\sqrt{j(j+1)-m(m+1)}\,\langle m_1,m_2|j,m+1\rangle=\\ \sqrt{j_1(j_1+1)-m_1(m_1-1)}\,\langle m_1-1,m_2|j,m\rangle+\sqrt{j_2(j_2+1)-m_2(m_2-1)}\,\langle m_1,m_2-1|j,m\rangle\end{gathered} \tag{4.7.37}$$

where J_{1+} acting to the left *reduces* m_1 by one. (Here, obviously, we must choose $m=m_1+m_2-1$ to have nonzero coefficients.)

To visualize what's going on with all these coefficients, remember m_1 can take $2j_1+1$ values and m_2 can take $2j_2+1$ values, so for given j_1,j_2 every possible state of the two spins can be represented by a dot on a $(2j_1+1)\times(2j_2+1)$ grid: here's $j_1=3,\quad j_2=2$:



Each dot represents an $|m_1,m_2\rangle$ state for $j_1=3, j_2=2$.

The four blue dots have $m=m_1+m_2=2$.

For given total j, the coefficients for the three red dots are connected by a J+ recursion relation.

How do these dots relate to the CG coefficients? the top right-hand dot (3, 2) uniquely represents the $j=5, \quad m=5$ state of total angular momentum. The next dots down, (2, 2) and (3, 1), correspond to two CG coefficients for $j=5$ and two *different* CG coefficients for $j=4$.

If we now pick one value of j less than j_1+j_2, each dot in the grid will correspond to one coefficient.

Note that having fixed j, the grid will be curtailed: let's take $j=3$, so $m=m_1+m_2=3$ at most. Then the grid loses its far corners:

$$|j_1, m_1\rangle \otimes |j_2, m_2\rangle. \tag{4.7.38}$$

Let us examine for this fixed j which CG coefficients are where in this curtailed grid.

There are a total of $2j+1=7$ states for $m=3,\ 2\ ,\dots,-3.$

The top state, $j=3, \quad m=3$, or $|3,3\rangle_{jm}$, is given by three coefficients on the top diagonal line (it's in a three-dimensional subspace, and orthogonal to $j=5$ and $j=4$ multiplet members $|5,3\rangle_{jm}$, $|4,3\rangle_{jm}$ which are also in the $m=3$ subspace). We're not at this point *calculating* these coefficients, we're just trying to find them a home.

Applying the lowering operator to $|3,3\rangle_{jm}$ gives a vector in the four-dimensional $m=2$ subspace, the coefficients would belong to the next diagonal down, which has four elements. (This subspace also includes the top member of the $j=2$ multiplet.) Using the lowering operator one more time we enter the five-dimensional $m=1$ subspace—but that is the maximum number of dimensions in this problem, since angular momenta 3 and 2 cannot be added to give a $j=0$ scalar.

Having now, for this particular j made from j_1+j_2, found where all the CG coefficients for all the $2j+1$ multiplet members are located, we shall see how they can all be systematically calculated using the recursion relations generated by $J_\pm=J_{1\pm}+J_{2\pm}$.

We've mapped the recursion relations on the diagram: given $j,\ j_1,\ j_2$ the three red dots at $(m_1,m_2),\quad (m_1-1,m_2),\quad (m_1,m_2-1)$ (with $m_1=-1,\quad m_2=1$ in this example) locate the three CG coefficients satisfying the linear equation above from

$$\langle m_1,m_2|J_+|j,m\rangle = \langle m_1,m_2|J_{1+}|j,m\rangle + \langle m_1,m_2|J_{2+}|j,m\rangle \tag{4.7.39}$$

so if two of them are known the third is given. Similarly, the parallel equation generated by $J_-=J_{1-}+J_{2-}$ links the three green dots, at $(m_1,m_2),\quad (m_1+1,m_2),\quad (m_1,m_2+1)$.

We begin the computation of the CG coefficients with the blue dot, the point on the leading "arrow" edges. Let us arbitrarily assign a value 1 to this point. If we make it the top member of a "green" triangle, that will link it to the dot below and to a dot to the right which is *off the array*. The dot off the array makes zero contribution, so we have an equation giving the value of the coefficient at the dot below the blue dot as a multiple of the value on the blue dot. We can then continue down to the next dot. We could instead have gone *up* from the blue dot using incomplete red triangles—in fact we can continue around the edge of the whole array. Then, once the values along the edges are fixed, the recursion triangles can be used to move inward and find the rest.

The point of this section is to establish that, apart from an overall multiplicative constant that must be fixed by normalization, *all* the CG coefficients for this value of j can be found *from the recursion relations alone*. The reason this is important is because the *same* algebraic structure, and therefore the same recursion relations, are used to define spherical tensors, so they can *also* be combined using the same CG coefficients. (We still need a sign convention here to present a complete table: so far, the different values of total j have arbitrary relative phases.)

4.8: Tensor Operators

Introduction: Cartesian Vectors and Tensors

Physics is full of vectors: $\vec{x}$, $\vec{L}$, $\vec{S}$ and so on. Classically, a (three-dimensional) vector is defined by its properties under rotation: the three components corresponding to the Cartesian x, y, z axes transform as

$$V_i \rightarrow \sum R_{ij} V_j \tag{4.8.1}$$

with the usual rotation matrix, for example

$$R_z(\theta)=\begin{pmatrix}\cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0\end{pmatrix} \tag{4.8.2}$$

for rotation about the z-axis. (We'll use (x, y, z) and (x_1, y_1, z_1) interchangeably.)

A *tensor* is a generalization of a such a vector to an object with more than one suffix, such as, for example, T_{ij} or T_{ijk} (having 9 and 27 components respectively in three dimensions) with the requirement that these components mix among themselves under rotation by each individual suffix following the vector rule, for example

$$T_{ijk} \rightarrow \sum R_{il} R_{jm} R_{kn} T_{lmn} \tag{4.8.3}$$

where R is the same rotation matrix that transforms a vector. Tensors written in this way are called *Cartesian tensors* (since the suffixes refer to Cartesian axes). The number of suffixes is the *rank* of the Cartesian tensor, a rank n tensor has of course 3^n components.

Tensors are common in physics: they are essential in describing stress, distortion and flow in solids and liquids. The inertial tensor is the basis for analyzing angular motion in classical mechanics. Tensor *forces* are important in the dynamics of the deuteron, and in fact tensors arise for any charge distribution more complicated than a dipole. Going to four dimensions, and generalizing from rotations to Lorentz transformations, Maxwell's equations are most naturally expressed in tensor form, and tensors are central to General Relativity.

To get back to non-relativistic physics, since the defining property of a tensor is its behavior under rotations, spherical polar coordinates are sometimes a more natural basis than Cartesian coordinates. In fact, in that basis tensors (called *spherical tensors*) have rotational properties closely related to those of angular momentum eigenstates, as will become clear in the following sections.

The Rotation Operator in Angular Momentum Eigenket Space

As a preliminary to discussing general tensors in quantum mechanics, we briefly review the rotation operator and quantum vector operators. (A full treatment is given in my 751 lecture.) Recall that the rotation operator turning a ket through an angle $\vec{\theta}$ (the vector direction denotes the axis of rotation, its magnitude the angle turned through) is

$$U(R(\vec{\theta}))=e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}} \tag{4.8.4}$$

Since $\vec{J}$ commutes with the total angular momentum squared $\vec{J}^2=j(j+1)\hbar^2$, we can restrict our attention to a given *total* angular momentum j, having as usual an orthonormal basis set $|j, m\rangle$, or $|m\rangle$ for short, with $2j+1$ components, a general ket $|\alpha\rangle$ in this space is then:

$$|\alpha\rangle=\sum_{m=-j}^{j} \alpha_m |m\rangle. \tag{4.8.5}$$

Rotating this ket,

$$|\alpha\rangle \rightarrow |\alpha'\rangle=e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}|\alpha\rangle \tag{4.8.6}$$

Putting in a complete set of states, and using the standard notation for matrix elements of the rotation operator,

$$|\alpha'\rangle = e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}|\alpha\rangle \tag{4.8.1}$$

$$= \sum_{m',m} \alpha_m |m'\rangle\langle m'|e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}|m\rangle \tag{4.8.2}$$

$$= \sum_{m',m} D^{(j)}_{m'm}(R(\vec{\theta}))\alpha_m|m'\rangle. \tag{4.8.7}$$

$D^{(j)}_{m'm} = \langle m'|e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}|m\rangle$ is standard notation (see the earlier lecture).

So the ket rotation transformation is

$$\alpha'_{m'} = \sum_m D^{(j)}_{m'm}\alpha_m, \quad or\ \alpha' = D\alpha. \tag{4.8.8}$$

with the usual matrix-multiplication rules.

Rotating a Basis Ket

Now suppose we apply the rotation operator to one of the *basis kets* $|j,m\rangle$, what is the result?

$$e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}|j,m\rangle = \sum_{m'}|j,m'\rangle\langle j,m'|e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}|j,m\rangle = \sum_{m'}|j,m'\rangle D^{(j)}_{m'm}(R) \tag{4.8.9}$$

Note the *reversal* of m, m' compared with the operation on the set of component coefficients of the general ket.

(You may be thinking: wait a minute, $|j,m\rangle$ *is* a ket in the space - it can be written $\sum \alpha_{m''}|j,m''\rangle$ with $\alpha_{m''} = \delta_{m''m}$, so we could use the previous rule $\alpha'_{m'} = \sum_m D^{(j)}_{m'm}\alpha_m$ to get

$$\alpha'_{m'} = \sum_m D^{(j)}_{m'm''}\alpha_{m''} = \alpha'_{m'} = \sum_m D^{(j)}_{m'm''}\delta_{m''m} = D^{(j)}_{m'm}. \tag{4.8.3}$$

Reassuringly, this leads to the same result we just found.)

Rotating an Operator, Scalar Operators

Just as in the Schrödinger versus Heisenberg formulations, we can either apply the rotation operator to the kets and leave the operators alone, or we can leave the kets alone, and rotate the operators:

$$A \to e^{\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}Ae^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}} = U^\dagger AU \tag{4.8.10}$$

which will yield the same matrix elements, so the same physics.

A *scalar* operator is an operator which is *invariant* under rotations, for example the Hamiltonian of a particle in a spherically symmetric potential. (There are many less trivial examples of scalar operators, such as the dot product of two vector operators, as in a spin-orbit coupling.)

The transformation of an operator under an infinitesimal rotation is given by:

$$S \to U^\dagger(R)SU(R) \tag{4.8.4}$$

with

$$U(R) = 1 - \frac{i\vec{\varepsilon}\cdot\vec{J}}{\hbar} \tag{4.8.11}$$

from which

$$S \to S + \left[\frac{i\vec{\varepsilon}\cdot\vec{J}}{\hbar}, S\right]. \tag{4.8.12}$$

It follows that a scalar operator S, which does not change at all, must commute with all the components of the angular momentum operator, and hence must have a common set of eigenkets with, say, $\vec{J}^2$ and J_z.

Vector Operators: Definition and Commutation Properties

A quantum mechanical vector operator $\vec{V}$ is *defined* by requiring that the expectation values of its three components in any state *transform like the components of a classical vector* under rotation.

It follows from this that the operator itself must transform vectorially,

$$V_i' = U^\dagger(R) V_i U(R) = \sum R_{ij} V_j \tag{4.8.13}$$

To see what this implies, it is easiest to look at a simple case. For an infinitesimal rotation about the z-axis,

$$R_z(\varepsilon) = \begin{pmatrix} 1 & -\varepsilon & 0 \\ \varepsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.8.14}$$

the vector transforms

$$\begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -\varepsilon & 0 \\ \varepsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} = \begin{pmatrix} V_x - \varepsilon V_y \\ V_y + \varepsilon V_x \\ V_z \end{pmatrix} \tag{4.8.15}$$

The unitary Hilbert space operator U corresponding to this rotation $U(R_z(\varepsilon) = 1 - \frac{i\varepsilon J_z}{\hbar}$, so

$$U^\dagger V_i U \ = (1 + i\varepsilon J_z/\hbar) V_i (1 - i\varepsilon J_z/\hbar) \tag{4.8.5}$$

$$= V_i + \frac{i\varepsilon}{\hbar}[J_z, V_i] \tag{4.8.16}$$

The requirement that the two transformations above, the infinitesimal classical rotation generated by $R_z(\varepsilon)$ and the infinitesimal unitary transformation $U^\dagger(R) V_i U(R)$, are in fact the same thing yields the commutation relations of a vector operator with angular momentum:

$$\begin{aligned} i[J_z, V_x] &= -\hbar V_y \\ i[J_z, V_y] &= +\hbar V_x \end{aligned} \tag{4.8.17}$$

From this result and its cyclic equivalents, the components of *any* vector operator $\vec{V}$ must satisfy:

$$[V_i, J_j] = i\varepsilon_{ijk}\hbar V_k. \tag{4.8.18}$$

Exercise: verify that the components of $\vec{x}$, $\vec{L}$, $\vec{S}$ do in fact satisfy these commutation relations.

(*Note*: Confusingly, there is a slightly different situation in which we need to rotate an operator, and it gives an opposite result. Suppose an operator T acts on a ket $|\alpha\rangle$ to give the ket $|\alpha'\rangle = T|\alpha\rangle$. For kets $|\alpha\rangle$ and $|\alpha'\rangle$ to go to $U|\alpha\rangle$ and $U|\alpha'\rangle$ respectively under a rotation U, T itself must transform as $T \rightarrow UTU^\dagger$ (recall $U^\dagger = U^{-1}$). The point is that this is a Schrödinger rather than a Heisenberg-type transformation: we're rotating the kets, not the operators.)

Warning: Does a vector operator transform like the components of a vector or like the basis kets of the space? You'll see it written both ways, so watch out!

We've already defined it as transforming like the components:

$$V_i' = U^\dagger(R) V_i U(R) = \sum R_{ij} V_j \tag{4.8.13}$$

but if we now take the *opposite* rotation, the unitary matrix $U(R)$ is replaced by its inverse $U^\dagger(R)$ and *vice versa.* Remember also that the ordinary spatial rotation matrix R is orthogonal, so its inverse is its transpose, and the above equation is equivalent to

$$V_i' = U(R) V_i U^\dagger(R) = \sum R_{ji} V_j. \tag{4.8.19}$$

This definition of a vector operator is that its elements transform just as do the basis kets of the space - so it's crucial to look carefully at the equation to figure out which is the rotation matrix, and which is its inverse!

This second form of the equation is the one in common use.

Cartesian Tensor Operators

From the definition given earlier, under rotation the elements of a rank two Cartesian tensor transform as:

$$T_{ij} \to T'_{ij} = \sum\sum R_{ii'} R_{jj'} T_{i'j'}. \tag{4.8.20}$$

where R_{ij} is the rotation matrix for a vector.

It is illuminating to consider a particular example of a second-rank tensor, $T_{ij} = U_i V_j$, where $\vec{U}$ and $\vec{V}$ are ordinary three-dimensional vectors.

The problem with this tensor is that it is *reducible*, using the word in the same sense as in our discussion of group representations is discussing addition of angular momenta. That is to say, combinations of the elements can be arranged in sets such that rotations operate only within these sets. This is made evident by writing:

$$U_i V_j = \frac{\vec{U}\cdot\vec{V}}{3}\delta_{ij} + \frac{(U_i V_j - U_j V_i)}{2} + \left(\frac{U_i V_j + U_j V_i}{2} - \frac{\vec{U}\cdot\vec{V}}{3}\delta_{ij}\right). \tag{4.8.21}$$

The first term, the dot product of the two vectors, is clearly a *scalar* under rotation, the second term, which is an antisymmetric tensor has three independent components which are the *vector* components of the vector product $\vec{U}\times\vec{V}$, and the third term is a *symmetric traceless tensor*, which has five independent components. Altogether, then, there are $1+3+5=9$ components, as required.

Spherical Tensors

Notice the numbers of elements of these irreducible subgroups: $1, 3, 5$ These are exactly the numbers of elements of angular momenta representations for *j* = 0, 1, 2!

This is of course no coincidence: as we shall make more explicit below, a three-dimensional vector is mathematically isomorphic to a quantum spin one, the tensor we have written is therefore a direct product of two spins one, so, exactly as we argues in discussing addition of angular momenta, it will be a reducible representation of the rotation group, and will be a sum of representations corresponding to the possible total angular momenta from adding two spins one, that is, $j = 0,\ 1,\ 2$.

As discussed earlier, the matrix elements of the rotation operator

$$U(R(\vec{\theta})) = e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}} \tag{4.8.22}$$

within a definite j subspace are written

$$D^{j}_{m'm}(R(\vec{\theta})) = \langle j, m'|e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}|j, m\rangle \tag{4.8.23}$$

so under rotation operator a basis state $|j, m\rangle$ transforms as:

$$e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}|j, m\rangle = \sum_{m'} |j, m'\rangle\langle j, m'|e^{-\frac{i\vec{\theta}\cdot\vec{J}}{\hbar}}|j, m\rangle = \sum_{m'} |j, m'\rangle\ D^{(j)}_{m'm}(R). \tag{4.8.24}$$

The essential point is that these irreducible subgroups into which Cartesian tensors decompose under rotation (generalizing from our one example) form a more natural basis set of tensors for problems with rotational symmetries.

Definition: spherical tensor

We define a *spherical tensor* of rank k as a set of $2k+1$ operators T^q_k, $q = k, k-1, \ldots, -k$ such that under rotation they transform among themselves with exactly the same matrix of coefficients as that for the $2j+1$ angular momentum eigenkets $|m\rangle$ for $k = j$, that is,

$$U(R) T^q_k U^\dagger(R) = \sum_{q'} D(k)q'qTq'k. \tag{4.8.25}$$

To see the properties of these spherical tensors, it is useful to evaluate the above equation for infinitesimal rotations, for which

$$D^{(k)}_{q'q}(\vec{\varepsilon}) = \langle k, q'|I - i\vec{\varepsilon}\cdot\vec{J}/\hbar|k, q\rangle = \delta_{q'q} - i\vec{\varepsilon}\cdot\langle k, q'|\vec{J}/\hbar|k, q\rangle. \tag{4.8.26}$$

(The matrix element $\langle k,q'|\vec{J}/\hbar|k,q\rangle$ is just the familiar Clebsch Gordan coefficient in changed notation: the rank k corresponds to the usual j, and q to the “magnetic” quantum number m.)

Specifically, consider an infinitesimal rotation $\vec{\varepsilon}\cdot\vec{J}=\varepsilon J_+$. (Strictly speaking, this is not a real rotation, but the formalism doesn’t care, and the result we derive can be confirmed by rotation about the x and y directions and adding appropriate terms.)

The equation is

$$(1-i\varepsilon J_+/\hbar)T_k^q(1+i\varepsilon J_+/\hbar)=\sum_{q'}(\delta_{q'q}-i\varepsilon\langle k,q'|J_+/\hbar|k,q\rangle)Tq'k \tag{4.8.27}$$

and equating terms linear in ε,

$$\begin{gathered}[J\pm,T_k^q]=\pm\hbar\sqrt{(k\mp q)(k\pm q+1)}\;T_k^{q\pm1}\\ [Jz,T_k^q]=\hbar qT_k^q.\end{gathered} \tag{4.8.28}$$

Sakurai observes that this set of commutation relations could be taken as the *definition* of the spherical tensors.

Notational note: we have followed Shankar here in having the rank k as a subscript, the “magnetic” quantum number q as a superscript, the same convention used for the spherical harmonics (but not for the D matrices!) Sakurai, Baym and others have the rank above, usually in parentheses, and the magnetic number below. Fortunately, all use k for rank and q for magnetic quantum number.

A Spherical Vector

The $j=1$ angular momentum eigenkets are just the familiar spherical harmonics

$$Y_1^0=\sqrt{\frac{3}{4\pi}}\frac{z}{r},\;Y_1^{\pm1}=\mp\sqrt{\frac{3}{4\pi}}\frac{x\pm iy}{\sqrt{2}r}. \tag{4.8.29}$$

The rotation operator will transform (x,y,z) as an ordinary vector in three-space, and this is evidently equivalent to

$$|j=1,m\rangle\rightarrow\sum_{m'}|j=1,m'\rangle\;D_{m'm}^{(j)}(R) \tag{4.8.30}$$

It follows that the spherical representation of a three vector $(V_x,\;V_y,\;V_z)$ has the form:

$$T_1^{\pm1}=\mp\frac{V_x\pm iV_y}{\sqrt{2}}=V_1^{\pm1},\;T_1^0=V_z=V_1^0. \tag{4.8.31}$$

In line with spherical tensor notation, the components $(T_1^1,\;T_1^0,\;T_1^{-1})$ are denoted T_1^q.

Matrix Elements of Tensor Operators between Angular Momentum Eigenkets

By definition, an irreducible tensor operator T_k^q transforms under rotation like an angular momentum eigenket $|k,q\rangle$. Therefore, rotating the ket $T_k^q|j,m\rangle$,

$$U\,T_k^q|j,m\rangle=U\,T_k^qU^{-1}U|j,m\rangle=\sum_{q'}D_{q'q}^{(k)}T_k^{q'}\sum_{m'}D_{m'm}^{(j)}|j,m'\rangle. \tag{4.8.32}$$

The product of the two D matrices appearing is precisely the set of coefficients to rotate *the direct product of eigenkets* $|k,q\rangle\otimes|j,m\rangle$ where $|k,q\rangle$ is the angular momentum eigenket having $j=k,\;\;m=q$.

We have met this direct product of two angular momentum eigenkets before: this is just a system having two angular momenta, such as orbital plus spin angular momenta. So we see that T_k^q acting on $|j,m\rangle$ generates a state having total angular momentum the sum of (k,q) and (j,m).

To link up (more or less) with Shankar’s notation: our direct product state $|k,q\rangle\otimes|j,m\rangle$ is the same as $|k,q;j,m\rangle$ in the notation $|j_1,m_1;j_2,m_2\rangle$ for a product state of two angular momenta (possibly including spins). Such a state can be written as a sum over states of the form $|j_{tot},m_{tot};j_1,j_2\rangle$ where this denotes a state of total angular momentum j_{tot}, z- direction component m_{tot}, made up of two spins having total angular momentum j_1,j_2 respectively.

This is the standard Clebsch-Gordan sum:

$$|j_1,m_1;j_2,m_2\rangle=\sum_{j_{tot}=|j_1-j_2|}^{j_1+j_2}\sum_{m_{tot}=-j_{tot}}^{j_{tot}}|j_{tot},m_{tot};j_1,j_2\rangle\langle j_{tot},m_{tot};j_1,j_2|j_1,m_1;j_2,m_2\rangle. \tag{4.8.33}$$

The summed terms give a unit operator within this $(2j_1+1)(2j_2+1)$ dimensional space, the term $\langle j_{tot},m_{tot};j_1,j_2|j_1,m_1;j_2,m_2\rangle$ is a Clebsch-Gordan coefficient. The only nonzero coefficients have $m_{tot}=m_1+m_2$, and j_{tot} restricted as noted, so for given $m_1,\ m_2$ we just set $m_{tot}=m_1+m_2$, we don't sum over m_{tot}, and the sum over j_{tot} begins at $|m_{tot}|$.

Translating into our $|k,q\rangle\otimes|j,m\rangle$ notation, and cleaning up,

$$|k,q;j,m\rangle=\sum_{j_{tot}=|q+m|}^{k+j}|j_{tot},q+m;k,j\rangle\langle j_{tot},q+m;k,j|k,q;j,m\rangle. \tag{4.8.34}$$

We are now able to evaluate the angular component of the matrix element of a spherical tensor operator between angular momentum eigenkets: we see that it will only be nonzero for $m_{tot}=m_1+m_2$, and j_{tot} at least $|m_{tot}|$.

The Wigner-Eckart Theorem

At this point, we must bear in mind that these tensor operators are not necessarily just functions of angle. For example, the position operator is a spherical vector multiplied by the radial variable r, and kets specifying atomic eigenstates will include radial quantum numbers as well as angular momentum, so the matrix element of a tensor between two states will have the form

$$\langle\alpha_2,j_2,m_2|T^q_k|\alpha 1,j_1,m_1\rangle, \tag{4.8.35}$$

where the j's and m's denote the usual angular momentum eigenstates and the α's are *nonangular* quantum numbers, such as those for radial states.

The basic point of the Wigner-Eckart theorem is that *the angular dependence of these matrix elements can be factored out, and it is given by the Clebsch-Gordan coefficients*.

Having factored it out, the remaining dependence, which is only on the *total* angular momentum in each of the kets, *not* the relative orientation (and of course on the α 's), is traditionally written as a bracket with double lines, that is,

$$\langle\alpha_2,j_2,m_2j+1|T^q_k|\alpha 1,j_1,m_1\rangle=\frac{\langle\alpha_2,j_2||T_k||\alpha 1,j_1\rangle}{\sqrt{2j+1}}\cdot\langle j_2,m_2|k,q;j_1,m_1\rangle. \tag{4.8.36}$$

The denominator is the conventional normalization of the double-bar matrix element. The proof is given in, for example, Sakurai (page 239) and is not that difficult. The basic strategy is to put the defining identities

$$[J\pm,T^q_k]=\pm\hbar\sqrt{(k\mp q)(k\pm q+1)}\ \ T^{q\pm1}k \tag{4.8.37}$$
$$[J_z,T^q_k]=\hbar qT^q_k$$

between $|\alpha,j,m\rangle$ bras and kets, then get rid of the $J_\pm$ and J_z by having them operate on the bra or ket. This generates a series of linear equations for $\langle\alpha_2,j_2,m_2|T^q_k|\alpha 1,j_1,m_1\rangle$ matrix elements with m variables differing by one, and in fact this set of linear equations is *identical* to the set that generates the Clebsch-Gordan coefficients, so we must conclude that these spherical tensor matrix elements, ranging over possible m and j values, are exactly proportional to the Clebsch-Gordan coefficients - and that is the theorem.

A Few Hints for Shankar's problem 15.3.3: that first matrix element comes from adding a spin j to a spin 1, writing the usual maximum m state, applying the lowering operator to both sides to get the total angular momentum $j+1,\ m=j$ state, then finding the same m state orthogonal to that, which corresponds to total angular momentum j (instead of $j+1$).

For the operator J, the Wigner-Eckart matrix element simplifies because J cannot affect α, and also it commutes with J^2, so cannot change the total angular momentum.

So, in the Wigner-Eckart equation, replace T^q_k on the left-hand side by J^0_1, which is just J_z. The result of (1) should follow.

(2) First note that a scalar operator cannot change m. Since c is independent of A we can take $A=J$ to find c.

CHAPTER OVERVIEW

5: Interlude - The Nature of Electrons

5.1: Bosons and Fermions
5.2: Multielectron Atoms

5.1: Bosons and Fermions

Two Electrons in a One Dimensional Well

So far, we have used Schrödinger's equation to see how a single particle, usually an electron, behaves in a variety of potentials. If we are going to think about atoms other than hydrogen, it is necessary to extend the Schrödinger equation so that it describes more than one particle.

As a simple example of a two particle system, let us consider two electrons confined to the same one-dimensional infinite square well,

$$\begin{array}{ll} V(x)=0, & -L/2<x<L/2 \\ V(x)=\infty & \quad otherwise. \end{array} \tag{5.1.1}$$

To make things even simpler, let us assume that the electrons do not interact with each other—we switch off their electrostatic repulsion. Then by analogy with our construction of the Schrödinger equation for a single electron, we can construct the two-electron analogue:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x_1,x_2,t)}{\partial x_1^2}-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x_1,x_2,t)}{\partial x_2^2}=E\Psi(x_1,x_2,t) \tag{5.1.2}$$

inside the well, with the wavefunction going to zero for x_1 or x_2 equal to $L/2$ or $-L/2$.

On looking at Equation 5.1.2, we see it is *the same* as the Schrödinger equation for a *single* electron in a *two* dimensional square well, and so can be solved in the same way, by separation of variables. For example, the wavefunction we plotted for the two dimensional rectangular well is in the square case:

$$\Psi_{(2,3)}(x_1,x_2,t)=A\sin\left(\frac{2\pi x_1}{L}\right)\cos\left(\frac{3\pi x_2}{L}\right)e^{-iEt/\hbar} \tag{5.1.3}$$

has the same energy

$$E=\frac{\hbar^2}{2m}\frac{\pi^2}{L^2}\left(2^2+3^2\right) \tag{5.1.4}$$

as the physically distinct wavefunction:

$$\Psi_{(3,2)}(x_1,x_2,t)=A\sin\left(\frac{2\pi x_2}{L}\right)\cos\left(\frac{3\pi x_1}{L}\right)e^{-iEt/\hbar} \tag{5.1.5}$$

Interpreting the Wavefunction

We have already discussed how the above wavefunctions are to be interpreted if they are regarded as two-dimensional wavefunctions for a single electron: $|\Psi(x_1,x_2)|^2dx_1dx_2$ is the probability of finding the electron in a small area dx_1dx_2 at $(x_1,\ x_2)$.

The natural extension of this interpretation to two electrons is to assume now $|\Psi(x_1,\ x_2)|^2dx_1dx_2$ is the *joint probability* of finding electron 1 in a small line length dx_1 at position x_1, and at the same time finding electron 2 in a small length dx_2 at x_2.

This, however, leads to a real problem. Consider the wavefunction in Equation 5.1.3 again. Let us now take specific points, $x_1=0,\ x_2=L/4$. Then the probability of finding electron 1 in an infinitesimal interval at x_1 and electron 2 similarly at x_2 is zero, because Ψ is zero at $x_1=0$. On the other hand, the probability of finding electron 2 at $x_1=0$ and electron 1 at $x_2=L/4$ is *not* zero.

The problem is this: the electrons are *identical* (we assume their spins point the same way). We can't tell which is which, and nobody else can either. The indistinguishability of elementary particles is not like that of apparently identical macroscopic objects, where one could always place some tiny mark. There is no way to mark an electron. *This means, though, that the best we can do is to talk about the probability of finding one electron at x_1 and another at x_2, we cannot specify which electron we find where.* Therefore, any alleged wavefunction that gives different probabilities for finding electron 1 at x_1, 2 at x_2 and finding 2 at x_1, 1 at x_2 is not physically meaningful.

> *We can formulate the probability of finding one electron at x_1 and another at x_2, but we cannot specify which specific electron we find at either positions.*

That is to say, a wavefunction describing two identical particles must have a symmetric probability distribution

$$|\Psi(x_1, x_2)|^2 = |\Psi(x_2, x_1)|^2 \tag{5.1.6}$$

This is definitely not the case with our function $\Psi_{(2,3)}(x_1, x_2)$, so although it is a solution to the two particle Schrödinger equation (Equation 5.1.2), it is **not a physically meaningful** wavefunction for two particles in a box.

In fact, this is not difficult to fix—recall that the function $\Psi_{(3,2)}(x_1, x_2)$ has the same energy, and in fact just corresponds to the two particles being switched around, that is to say:

$$\Psi_{(3,2)}(x_1, x_2) = \Psi_{(2,3)}(x_2, x_1). \tag{5.1.7}$$

It follows that the symmetric function

$$\Psi^S_{(2,3)}(x_1, x_2) = \frac{1}{\sqrt{2}}\left(\Psi_{(2,3)}(x_1, x_2) + \Psi_{(2,3)}(x_2, x_1)\right) \tag{5.1.8}$$

and the antisymmetric function

$$\Psi^A_{(2,3)}(x_1, x_2) = \frac{1}{\sqrt{2}}\left(\Psi_{(2,3)}(x_1, x_2) - \Psi_{(2,3)}(x_2, x_1)\right) \tag{5.1.9}$$

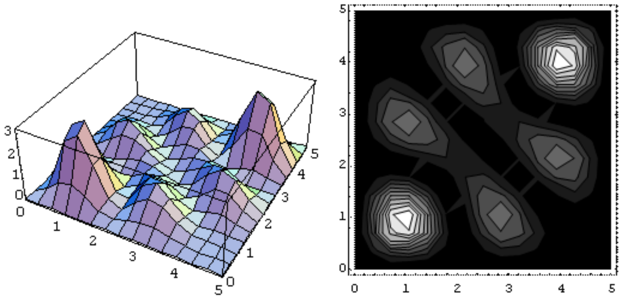
are both solutions to Schrödinger's equation for the energy E, and both satisfy the requirement $|\Psi(x_1, x_2)|^2 = |\Psi(x_2, x_1)|^2$ necessary for identical particles, so these are the appropriate candidate wavefunctions for the two particles in the one-dimensional box.

Bosons, Fermions and the Pauli Exclusion Principle

It turns out that both symmetric and antisymmetric wavefunctions arise in nature in describing identical particles. In fact, all elementary particles are either *fermions*, which have antisymmetric multiparticle wavefunctions, or *bosons*, which have symmetric wavefunctions. Electrons, protons and neutrons are fermions; photons, a-particles and helium atoms are bosons.

It is important to realize that this requirement of symmetry of the probability distribution, arising from the true indistinguishability of the particles, has a large effect on the probability distribution, and, furthermore, the effect is very different for fermions and bosons. The simplest way to see this is just to plot $|\Psi(x_1, x_2)|^2$ for the symmetric and for the antisymmetric wavefunctions, and compare them.

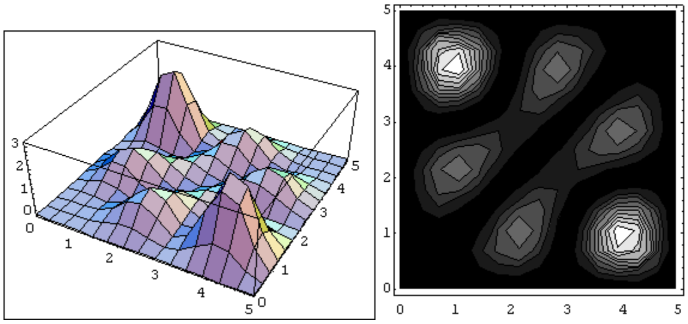
For the *symmetric* case, the probability distribution (unnormalized) looks like:



(The graytone figure is a contour map.)

The main point to note here is that the highest joint probability is that of finding the two particles close together, near (1,1) and (4,4). (These actual points depend on the wavefunction we chose initially, of course).

Now look at the *antisymmetric* case:



The difference is dramatic. For the antisymmetric wavefunction, the particles are most likely to be found *far away* from each other. In fact, there is *zero* probability that they will be found at the same spot, because if $\Psi(x_1, x_2) = -\Psi(x_2, x_1)$, obviously $\Psi(x, x) = 0$. A more general statement is that *two fermions cannot be in the same quantum state*, because if they were, the wavefunction would be of the symmetric form $f(x_1)f(x_2)$, and could not be antisymmetrized. This is the Pauli Exclusion Principle—it is the basis of the periodic table, and consequently of almost everything else.

Pauli Exclusion Principle

Two fermions **cannot** be in the same quantum state (i.e., same set of relevant quantum numbers).

We should perhaps emphasize that these wavefunctions were calculated with the electrostatic repulsion between the electrons switched off. That is *not* what is keeping the electrons apart, although it will increase the separation if it is included. We should also warn against a simple classical picture of the Pauli principle, the thought that two things can't be in the same place, after all, so perhaps it's no surprise two electrons can't be in the same state. Two electrons *can* be in the same identical *space* wavefunction provided that their spins point in opposite ways. Furthermore, two bosons can be in the same state, and although that is perhaps reasonable sounding for photons, it is equally true for heavy atoms. In Bose-Einstein Condensation, a large number of atoms occupy the same quantum state. This happens in liquid ^{4}Helium below about two kelvin, it becomes a **superfluid** and flows without friction. BE condensation has also been achieved with laser cooled collections of large atoms in a trap.

5.2: Multielectron Atoms

The Stern-Gerlach Experiment

We've established that for the hydrogen atom, the angular momentum of the electron's orbital motion has values $\sqrt{l(l+1)}\,\hbar$, where $l=0,1,2,\ldots$, and the component of angular momentum in the z-direction is $m\hbar$, where m takes integer values $-l,-l+1,\ldots,+l$. This means that if we measure the angle between the total angular momentum and the z-axis, there can only be $2l+1$ possible answers, the total angular momentum cannot point in an arbitrary direction relative to the *z*-axis, odd though this conclusion seems. This is sometimes called "space quantization".

> *The total angular momentum of an electron cannot point in an arbitrary direction.*

Is there any way we can actually *see* some effect of this directional quantization? The answer is yes—because the electron moving around its orbit is a tiny loop of electric current, and, therefore, an electromagnet. So, if we switch on a magnetic field in the direction of the *z*-axis, the energy of the atom will depend on the degree of alignment of its magnetic moment with the external applied magnetic field. The magnetic field from a small current loop is like that from a small bar magnet aligned along the axis of the loop.

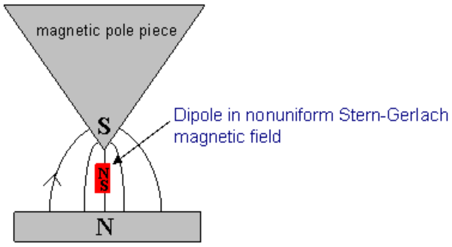
The simplest way to see how the potential energy of the little magnet depends on which way it's pointing relative to the field is to take a little bar with an N pole of strength $+p$ at one end, an S pole of strength $-p$ at the other end. Think of a compass needle, of length d, say. The magnetic moment is defined as pole strength multiplied by distance between the poles, $\mu=pd$, and is considered to be a vector pointing along the axis of the magnet, from S to N. The potential energy of this little magnet in an external field H is $-\mu H$, lowest when the magnet is fully aligned with the field. It is easy to check this: counting as zero potential energy the magnet at right angles with the field, the work needed to point it at an angle θ is $2p(d/2)\cos\theta$.

The magnetic moment of a current I going in a circle around an area A is just IA. The electron has charge e, and speed v, so goes around $v/2\pi r$ times per second. In other words, if you stand at one point in the orbit, the total charge passing you per second is $ev/2\pi r=I$. Hence the magnetic moment, usually denoted $\mu_L=IA$, is $\pi r^2 ev/2\pi r=rev/2$. The angular momentum is $L=mvr$, so

$$\mu_L=\frac{eL}{2m}. \tag{5.2.1}$$

Thus if the electron is in an $l=1$ orbit, the current will generate a magnetic moment $e\hbar/2m$, which is 9.3 × 10^{-24} joules per tesla, or 5.8 × 10^{-5} eV per tesla. Note that this means in a one tesla field an atomic energy level will move ~10^{-4} eV, an easily detectable shift in spectral lines will result.

But there is a more direct way to see how the atoms are oriented, the Stern Gerlach apparatus (1922). In this experiment, a beam of atoms is sent into a *nonuniform* magnetic field. This means the north and south poles of a small bar magnet would feel different strength forces, so there would be a net force on a small magnet, and hence on an atom. Furthermore, the direction of this force would depend on the orientation of the dipole.



Suppose the nonuniform field is pointing upwards, and is stronger at the top. Then a small bar magnet oriented vertically with the north pole on top will be pushed upwards, because the north pole will be experiencing the stronger force. If the south pole is on top, the magnet will be pushed downwards. If the magnet is horizontal, there will be no net force (assuming magnetic field strength varies only negligibly in the horizontal direction).

Imagine, then, a stream of atoms with magnetic moments entering a region of magnetic field as described. Each atom will feel a vertical force depending on the orientation of its magnetic moment. If with no magnetic field present the stream of atoms formed a dot on a screen after passing through the apparatus, on switching on the field one would expect the dot to be stretched into a vertical line, if one assumed equal likelihood of all orientations of the magnetic moment. However, the quantum theory predicts that this is not the case—we have argued that for $l=1$, say, there are only three allowed orientations of the magnet (atom) relative to the field. Therefore, we would predict that three dots (or, more realistically, blobs) would appear on the screen, not a continuous line.

In fact, when the experiment was carried out, there was a very surprising result. Perhaps the most dramatic form of the new result came later, in 1927, when *ground state* $l=0$ *hydrogen atoms* were used (Phipps and Taylor, Phys Rev **29,** 309). Such atoms have *no* orbital angular momentum, and therefore no orbital current, and were not expected to show magnetic effects. Yet on going through the Stern-Gerlach apparatus, the beam of hydrogen atoms *split into two*! This was difficult to interpret, because the least allowed angular momentum, $l=1$, would give three blobs, and $l=0$ would give only one. You might expect a mixture to give one strong blob and two weak ones, but two equal blobs didn't seem possible, theoretically. Stern and Gerlach had themselves seen two blobs with silver atoms in 1922. We mention the hydrogen case first because it was by far the best understood atom (and still is!) so the need for new physics was clearest.

The solution to the problem was suggested by two graduate students, Goudsmit and Uhlenbeck. They suggested that the electron *itself* had a spin. That is to say, the electron both orbited the proton *and* spun on its own axis, just as the earth orbits the sun once a year and also spins on its own axis once a day. If the electron spin is assumed to be $\hbar/2$, and we assume as before that the *z*-component can only change by whole units of $\hbar$, then there are only *two* allowed values of the *z*-component, $\pm\hbar/2$. Of course, this is a hand waving argument—the reason the *z*-component only changed by integers was that the wave function had to fit a whole number of wavelengths on going around the *z*-axis. But our wave function for spin one-half , if it is of the same form as those for angular momentum, must have a term $e^{i\varphi/2}$, and so is multiplied by -1 on rotating through 2π! (In fact, that the *z*-component can only change by whole units of $\hbar$ follows from very general properties of angular momentum.)

Further difficulties arose when people tried to construct models of how a spinning electron would have its own magnetic moment. It's not too difficult to see how this might occur—if the electron is a charged sphere, or has charge on its surface, then its rotation implies that this charge is going around in circles, little current loops, and so will give a magnetic field. The problem was, it was known that the electron was a very small object. It turned out that the equatorial speed of the electron would have to be greater than the speed of light for the magnetic moment to be of the observed strength.

These difficulties in understanding the electron spin and magnetic moment were far from trivial, and in fact were not resolved until around 1930, by Dirac, who gave a fully relativistic treatment of the problem, which, remarkably, predicted the magnetic moment correctly and at the same time treated the electron as a point particle. There is no simple picture presenting this in classical or semiclassical terms, but Dirac's work is the basis of our modern understanding of particle physics. It is unfortunately beyond the scope of this course.

The bottom line, as far as we are concerned, is that assuming the electron has spin one-half and hence two possible spin orientations with respect to a given axis explains the observed Stern-Gerlach results, and also, more importantly, helps us construct the periodic table, as we shall see below.

Building the Periodic Table

The *atomic number* (usually denoted *Z*) of an element denotes its place in the periodic table, so H has *Z* = 1; He, Z = 2; Li = 3, Be = 4, B = 5, C = 6, N = 7, O = 8, F = 9, Ne = 10, and so on. This number is equal to the number of protons in the nucleus, and also equal to the number of electrons orbiting around the nucleus, to preserve electrical neutrality.

To try to understand how the electrons orbit the nucleus, we need to make some simplifying assumptions. We are not going to be able to solve Schrödinger's equation for even two electrons exactly, if we include their repulsion of each other. However, the presence of the other electrons is clearly important—their repulsion to some extent counteracts the attraction the nucleus has for a given electron. For an electron imagined to be in some outer orbit, the electrons in closer to the nucleus orbits lower the effective nuclear charge. Thinking now about the force felt by one electron, a simple approximation is to imagine all the other electrons as changing the electrical attraction the one electron feels from the nucleus to a shielded attraction, so that the further it is away from the nucleus, the weaker an attractive charge it sees.

We then make the naïve assumption that all the electrons see the same potential, this shielded Coulomb potential, so we have *Z* electrons all in the same potential well, but we assume they are independent particles, in the sense that they do not repel each other, except to the extent already taken into account by changing to a shielded potential. So the question is, what are the possible wave functions of *Z* independent electrons in this well? The crucial point is that although they do not interact with each other, they are identical, so the wave function must be antisymmetrized, as we discussed for the two particle case earlier. This means that the electrons must be in different bound states in the well—the Pauli exclusion principle. But what do the bound state wave functions look like in this potential? Since the shielded Coulomb potential is still spherically symmetric, all our arguments about the θ, φ behavior of the ordinary Coulomb potential apply equally to the shielded case, in particular the angular momentum has values $\sqrt{l(l+1)}\,\hbar$, where $l=0,1,2,\ldots$, and the component of angular momentum in the *z*-direction is $m\hbar$, where m takes integer values $-l,\ -l+1,\ \ldots,+l$. Furthermore, each electron has spin $\hbar/2$, and there are two allowed values of the spin *z*-component, $\pm\hbar/2$. The radial wave functions R(r) are clearly somewhat different from those in the pure Coulomb case. The main difference is that states of different angular momentum which were degenerate in the Coulomb case are no longer the same energy in the shielded Coulomb case. If you examine wave functions corresponding to the same energy but different values of *l*, you will see that the higher the *l*-value, the smaller the wave function is near the nucleus. This means the higher l wave functions do not feel the powerful unshielded potential near the nucleus, and so are not as strongly bound as the lower l functions.

Notation

A standard notation is used by atomic physicists to describe these states. The different angular momenta are denoted by letters, *s* for $l=0$, *p* for $l=1$, *d* for $l=2$, *f* for $l=3$, *g* for $l=4$ and then on alphabetically. The Principal Quantum Number n, such that for the hydrogen atom $E=-1/n^2$ in Rydberg units, is given as a number, so the lowest hydrogen atom state is written 1*s*. The two $n=2$ orbital states are 2*s* and 2*p*, then come 3*s*, 3*p* and 3*d* and so on. From the discussion immediately above, 2*s* and 2*p* have the same energy in the hydrogen atom, but for the shielded potential used to approximate for the presence of other electrons in bigger atoms 2*s* would be more tightly bound, and so at a lower energy, than 2*p*.

Filling an Atom with Electrons

Let us now consider taking a bare nucleus, charge *Z*, and adding *Z* electrons to it one by one. From the Pauli Exclusion Principle, each electron must be in a different state. But remember that having a different spin counts as different (you could tell them apart) so we can put two electrons, with opposite spins, into each orbital state. Thus He has two electrons in the 1*s* state. Li must have two electrons in 1*s*, and one electron in 2*s*. This suggests a picture of one electron outside of a “closed shell” of two 1s electrons. The next occurrence of a similar picture is Na, having *Z* = 11, which is chemically very similar to Li. This means that 10 electrons fill closed shells. We can understand this because 2 go into 1*s*, 2 go into 2*s* and 6 fill 2*p*. But notice by saying it takes 6 electrons to fill 3*p*, we are saying there are three distinct $l=1$ orbitals. In other words, *the chemical properties of the elements support and confirm the hypothesis of “space quantization*”—that there are only three distinct $l=1$ angular wave functions, those given by $m=1,0$ and -1.

Atoms interact chemically by sharing or partially transferring electrons. It’s easier to transfer an electron that is loosely bound, and easier to accept one if there’s a “hole” in a shell. Not surprisingly, atoms with filled shells only, like He and Ne, are chemically unreactive. The valency, roughly speaking, is the number of electrons available for transfer (so Li and Na have valency 1) or available sites for reception of electrons—fluorine has an outer shell with one vacancy, so a valency of 1. To some extent, valency can vary depending on the strength of attraction of other atoms in the chemical environment.

Filling a Box with Electrons

When many Li atoms are put together to form a solid, it is found that the loosely attached outer electrons leave their original atoms and wander freely throughout the metal. Their wave functions are well represented by standing plane waves in a box (let’s take a cube of metal, of side L). Each such plane wave state in the box can be represented by three numbers $n_x,\ n_y$ and n_z, representing the number of nodes of the standing wave in the *x*, *y* and *z* directions respectively. Extending slightly our analysis of an electron in a two-dimensional box, the energy of such a state will be

$$E=\frac{\hbar^2}{2m}\frac{\pi^2}{L^2}\left(n_x^2+n_y^2+n_z^2\right). \tag{5.2.2}$$

Thus if we imagine pouring electrons into an empty lattice of Li atoms each with one electron missing (not a physically realistic procedure!) two electrons (opposite spins) will go into each state, first (0, 0, 0) then (1, 0, 0) or equally (0, 1, 0) etc., and from the form of the energy we can see that in (n_x,n_y,n_z) space, the electrons will fill up all the positive integer points within a sphere up to some maximum energy determined by how many electrons we put in. Notice that since the n’s are all positive integers, the filled space is only the one-eighth of the sphere's volume corresponding to $n_x > 0$, $n_y > 0$ and $n_z > 0$ for the sphere centered at the origin.

Physicists sometimes formulate this filling of electron states slightly differently, by imposing periodic boundary conditions on a piece of metal, like replacing a finite line by a ring. This is not easy to do in three dimensions, but is convenient to talk about. The advantage is that instead of standing waves, all the electrons have definite momenta. The allowed momenta form a grid in "momentum space" a lot like the allowed integers in the standing waves above. In fact it turns out that there are the same number of allowed momenta up to a certain energy as there are allowed standing wave states. The difference is that in momentum space, if momentum k is allowed, so is $-k$, and in the ground state momentum states are filled up to a spherical surface, called the "Fermi Surface"—an energy equipotential at the "Fermi energy". Typical Fermi energies are of the order of electron volts. By spreading out through the metal in this way the electrons attain an overall lower energy state than if each stayed with its own atom. This is why the solid is stable. On applying heat to the electrons, even 1000K is only 0.1eV, so only those near the surface of the filled sphere are free to move, because of the exclusion principle, the others are locked in. This means the heat capacity of the electrons is much less than $(3/2)kT$ per particle, as would be predicted classically. This was another long standing classical puzzle solved by the advent of quantum mechanics.

CHAPTER OVERVIEW

6: Charged Particle in Magnetic Field

6.1: Charged Particle in a Magnetic Field

Thumbnail: Trajectory of a particle with a positive or negative charge q under the influence of a magnetic field B, which is directed perpendicularly out of the screen. (CC SA-BY 3.0; Jaro.p).

6.1: Charged Particle in a Magnetic Field

Classically, the force on a charged particle in electric and magnetic fields is given by the Lorentz force law:

$$\vec{F}=q\left(\vec{E}+\frac{\vec{v}\times\vec{B}}{c}\right) \tag{6.1.1}$$

This velocity-dependent force is quite different from the conservative forces from potentials that we have dealt with so far, and the recipe for going from classical to quantum mechanics—replacing momenta with the appropriate derivative operators—has to be carried out with more care. We begin by demonstrating how the Lorentz force law arises classically in the Lagrangian and Hamiltonian formulations.

Laws of Classical Mechanics

Recall first that the Principle of Least Action leads to the Euler-Lagrange equations for the Lagrangian L:

$$\frac{d}{dt}\left(\frac{\partial L(q_i,\dot{q}_i)}{\partial\dot{q}_i}\right)-\frac{\partial L(q_i,\dot{q}_i)}{\partial q_i}=0 \tag{6.1.2}$$

with q_i and $\dot{q}_i$ being coordinates and velocities. The canonical momentum p_i is defined by the equation

$$p_i=\frac{\partial L}{\partial\dot{q}_i} \tag{6.1.3}$$

and the Hamiltonian is defined by performing a Legendre transformation of the Lagrangian:

$$H(q_i,p_i)=\sum_i\left(p_i\dot{q}_i-L(q_i,\dot{q}_i)\right) \tag{6.1.4}$$

It is straightforward to check that the equations of motion can be written:

$$\dot{q}_i=\frac{\partial H}{\partial p_i},\ \dot{p}_i=-\frac{\partial H}{\partial q_i} \tag{6.1.5}$$

These are known as *Hamilton's Equations*. Note that if the Hamiltonian is independent of a particular coordinate q_i, the corresponding momentum p_i remains constant. (Such a coordinate is termed *cyclic*, because the most common example is an angular coordinate in a spherically symmetric Hamiltonian, where angular momentum remains constant.)

For the conservative forces we have been considering so far,

$$L=T-V \tag{6.1.6}$$

and

$$H=T+V \tag{6.1.7}$$

with T the kinetic energy, V the potential energy.

Poisson Brackets

Any dynamical variable f in the system is some function of the q_i's and p_i's and (assuming it does not depend explicitly on time) its development is given by:

$$\frac{d}{dt}f(q_i,p_i)=\frac{\partial f}{\partial q_i}\dot{q}_i+\frac{\partial f}{\partial p_i}\dot{p}_i=\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i}-\frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i}=\{f,H\}. \tag{6.1.8}$$

The curly brackets are called *Poisson Brackets*, and are defined for any dynamical variables as:

$$\{A,B\}=\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i}-\frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i}. \tag{6.1.9}$$

We have shown from Hamilton's equations that for any variable $\dot{f}=\{f,H\}$.

It is easy to check that for the coordinates and canonical momenta,

$$q_i, q_j = 0 = p_i, p_j,\;\; q_i, p_j = \delta_{ij}. \tag{6.1.10}$$

This was the classical mathematical structure that led Dirac to link up classical and quantum mechanics: he realized that the Poisson brackets were the classical version of the commutators, so a classical canonical momentum must correspond to the quantum differential operator in the corresponding coordinate.

> *Poisson brackets are the classical version of the commutators*

Particle in a Magnetic Field

The Lorentz force is velocity dependent, so cannot be just the gradient of some potential. Nevertheless, the classical particle path is still given by the Principle of Least Action. The electric and magnetic fields can be written in terms of a scalar and a vector potential:

$$\vec{B} = \vec{\nabla} \times \vec{A} \tag{6.1.11}$$

$$\vec{E} = -\vec{\nabla}\varphi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}. \tag{6.1.12}$$

The right Lagrangian turns out to be:

$$L = \frac{1}{2}m\vec{v}^2 - q\varphi + \frac{q}{c}\vec{v}\cdot\vec{A}. \tag{6.1.13}$$

Relativity Effects

If you're familiar with Relativity, the interaction term here looks less arbitrary: the relativistic version would have the relativistically invariant $(q/c)\int A^\mu dx_\mu$ added to the action integral, where the four-potential $A_\mu = (\vec{A}, \varphi)$ and $dx_\mu = (dx_1, dx_2, dx_3, cdt)$. This is the simplest possible invariant interaction between the electromagnetic field and the particle's four-velocity. Then in the nonrelativistic limit, $(q/c)\int A^\mu dx_\mu$ just becomes $\int q(\vec{v}\cdot\vec{A}/c - \varphi)\, dt$.

The derivation of the Lorentz force from the Hamilton equations is straightforward.

Note that for zero vector potential, the Lagrangian has the usual $T - V$ form.

For this one-particle problem, the general coordinates q_i are just the Cartesian co-ordinates $x_i = (x_1, x_2, x_3)$, the position of the particle, and the $\dot{q}_i$ are the three components $\dot{x}_i = v_i$ of the particle's velocity.

The important *new* point is that the canonical momentum

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial \dot{x}_i} = mv_i + \frac{q}{c}A_i \tag{6.1.14}$$

is no longer mass × velocity—there is an extra term!

The Hamiltonian is

$$\begin{aligned} H(q_i, p_i) &= \sum p_i \dot{q}_i - L(q_i, \dot{q}_i) \\ &= \sum (mv_i + \tfrac{q}{c}A_i)v_i - \tfrac{1}{2}m\vec{v}^2 + q\varphi - \tfrac{q}{c}\vec{v}\cdot\vec{A} \\ &= \tfrac{1}{2}m\vec{v}^2 + q\varphi \end{aligned} \tag{6.1.15}$$

Reassuringly, the Hamiltonian just has the familiar form of kinetic energy plus potential energy. However, to get Hamilton's equations of motion, the Hamiltonian has to be expressed solely in terms of the coordinates and canonical momenta. That is,

$$H = \frac{(\vec{p} - q\vec{A}(\vec{x}, t)/c)^2}{2m} + q\varphi\ (\vec{x}, t) \tag{6.1.16}$$

where we have noted explicitly that the potentials mean those at the position $\vec{x}$ of the particle at time t.

Let us now consider Hamilton's equations

$$\dot{x}_i = \frac{\partial H}{\partial p_i},\;\; \dot{p}_i = -\frac{\partial H}{\partial x_i} \tag{6.1.17}$$

It is easy to see how the first equation comes out, bearing in mind that

$$p_i = mv_i + \frac{q}{c}A_i = m\dot{x}_i + \frac{q}{c}A_i. \tag{6.1.18}$$

The second equation yields the Lorentz force law, but is a little more tricky. The first point to bear in mind is that dp/dt is *not* the acceleration, the A term also varies in time, and in a quite complicated way, since it is *the field at a point moving with the particle*. That is,

$$\dot{p}_i = m\ddot{x}_i + \frac{q}{c}\dot{A}_i = m\ddot{x}_i + \frac{q}{c}\left(\frac{\partial A_i}{\partial t} + v_j\nabla_j A_i\right). \tag{6.1.19}$$

The right-hand side of the second Hamilton equation $\dot{p}_i = -\frac{\partial H}{\partial x_i}$ is

$$\begin{aligned} -\tfrac{\partial H}{\partial x_i} &= \tfrac{(\vec{p}-q\vec{A}(\vec{x},t)/c)}{m}\cdot\tfrac{q}{c}\cdot\tfrac{\partial\vec{A}}{\partial x_i} - q\tfrac{\partial\varphi(\vec{x},t)}{\partial x_i} \\ &= \tfrac{q}{c}v_j\nabla_i Aj - q\nabla_i\varphi. \end{aligned} \tag{6.1.20}$$

Putting the two sides together, the Hamilton equation reads:

$$m\ddot{x}_i = -\frac{q}{c}\left(\frac{\partial A_i}{\partial t} + v_j\nabla_j A_i\right) + \frac{q}{c}v_j\nabla_i A_j - q\nabla_i\varphi\ . \tag{6.1.21}$$

Using $\vec{v}\times(\vec{\nabla}\times\vec{A}) = \vec{\nabla}(\vec{v}\cdot\vec{A}) - (\vec{v}\cdot\vec{\nabla})\vec{A}$, $\vec{B} = \vec{\nabla}\times\vec{A}$, and the expressions for the electric and magnetic fields in terms of the potentials, the Lorentz force law emerges:

$$m\ddot{\vec{x}} = q\left(\vec{E} + \frac{\vec{v}\times\vec{B}}{c}\right) \tag{6.1.22}$$

Quantum Mechanics of a Particle in a Magnetic Field

We make the standard substitution:

$$\vec{p} = -i\hbar\vec{\nabla},\ \ so\ that\ [x_i, p_j] = i\hbar\delta_{ij}\ as\ usual:\ \ but\ now\ p_i \neq mv_i. \tag{6.1.23}$$

This leads to the novel situation that the velocities in different directions *do not commute*. From

$$mv_i = -i\hbar\nabla_i - qA_i/c \tag{6.1.24}$$

it is easy to check that

$$[v_x, v_y] = \frac{iq\hbar}{m^2c}B \tag{6.1.25}$$

To actually solve Schrödinger's equation for an electron confined to a plane in a uniform perpendicular magnetic field, it is convenient to use the Landau gauge,

$$\vec{A}(x,y,z) = (-By, 0, 0) \tag{6.1.26}$$

giving a constant field B in the *z* direction. The equation is

$$H\psi(x,y) = \left[\frac{1}{2m}(p_x + qBy/c)^2 + \frac{p_y^2}{2m}\right]\psi(x,y) = E\psi(x,y). \tag{6.1.27}$$

Note that *x does not appear in this Hamiltonian*, so it is a cyclic coordinate, and p_x is conserved. In other words, this H commutes with p_x, so H and p_x have a common set of eigenstates. We know the eigenstates of p_x are just the plane waves $e^{ip_x x/\hbar}$, so the common eigenstates must have the form:

$$\psi(x,y) = e^{ip_x x/\hbar}\chi(y). \tag{6.1.28}$$

Operating on this wavefunction with the Hamiltonian, the operator p_x appearing in H simply gives its eigenvalue. That is, the p_x in H just becomes a number! Therefore, writing $p_y = -i\hbar d/dy$, the *y*-component $\chi(y)$ of the wavefunction satisfies:

$$-\frac{\hbar^2}{2m}\frac{d^2}{dy^2}\chi(y)+\frac{1}{2}m\left(\frac{qB}{mc}\right)^2(y-y_0)^2\chi(y)=E\chi(y) \tag{6.1.29}$$

where

$$y_0=-cp_x/qB. \tag{6.1.30}$$

We now see that the conserved canonical momentum p_x in the *x*-direction is actually the coordinate of the center of a simple harmonic oscillator potential in the *y*-direction! This simple harmonic oscillator has frequency $\omega=|q|B/mc$, so the allowed values of energy for a particle in a plane in a perpendicular magnetic field are:

$$E=(n+\frac{1}{2})\hbar\omega=(n+\frac{1}{2})\hbar|q|B/mc. \tag{6.1.31}$$

The frequency is of course the cyclotron frequency—that of the classical electron in a circular orbit in the field (given by $mv^2/r=qvB/c,\ \omega=v/r=qB/mc$).

Let us confine our attention to states corresponding to the lowest oscillator state, $E=\frac{1}{2}\hbar\omega$. How many such states are there? Consider a square of conductor, area $A=L_x\times L_y$, and, for simplicity, take periodic boundary conditions. The center of the oscillator wave function y_0 must lie between 0 and L_y. But remember that $y_0=-cp_x/qB$, and with periodic boundary conditions $e^{ip_xL_x/\hbar}=1$, so $p_x=2n\pi\hbar/L_x=nh/L_x$. This means that y_0 takes a series of evenly-spaced discrete values, separated by

$$\Delta y_0=ch/qBL_x. \tag{6.1.32}$$

So the total number of states $N=L_y/\Delta y_0$,

$$N=\frac{L_xL_y}{\left(\frac{hc}{qB}\right)}=A\cdot\frac{B}{\Phi_0}, \tag{6.1.33}$$

where Φ_0 is called the “flux quantum”. So the total number of states in the lowest energy level $E=\frac{1}{2}\hbar\omega$ (usually referred to as the lowest Landau level, or *LLL*) is exactly equal to the total number of flux quanta making up the field B penetrating the area A.

It is instructive to find y_0 from a purely classical analysis.

Writing $m\dot{\vec{v}}=\frac{q}{c}\vec{v}\times\vec{B}$ in components,

$$\begin{aligned} m\ddot{x}&=\tfrac{qB}{c}\dot{y},\\ m\ddot{y}&=-\tfrac{qB}{c}\dot{x}. \end{aligned} \tag{6.1.34}$$

These equations integrate trivially to give:

$$\begin{aligned} m\dot{x}&=\tfrac{qB}{c}(y-y_0),\\ m\dot{y}&=-\tfrac{qB}{c}(x-x_0) \end{aligned}. \tag{6.1.35}$$

Here (x_0,y_0) are the coordinates of the center of the classical circular motion (the velocity vector $\dot{\vec{r}}=(\dot{x},\dot{y})$ is always perpendicular to $(\vec{r}-\vec{r}_0)$), and $\vec{r}_0$ is given by

$$\begin{aligned} y_0&=y-cmv_x/qB=-cp_x/qB\\ x_0&=x+cmv_y/qB=x+cp_y/qB. \end{aligned} \tag{6.1.36}$$

(Recall that we are using the gauge $\vec{A}(x,y,z)=(-By,0,0)$, and $p_x=\frac{\partial L}{\partial\dot{x}}=mv_x+\frac{q}{c}A_x$, etc.)

Just as y_0 is a conserved quantity, so is x_0: it commutes with the Hamiltonian since

$$[x+cp_y/qB,p_x+qBy/c]=0. \tag{6.1.37}$$

However, x_0 and y_0 *do not commute with each other*:

$$[x_0,y_0]=-i\hbar c/qB. \tag{6.1.38}$$

This is why, when we chose a gauge in which y_0 was sharply defined, x_0 was spread over the sample. If we attempt to localize the point (x_0,y_0) as well as possible, it is fuzzed out over an area essentially that occupied by one flux quantum. The natural length scale of the problem is therefore the magnetic length defined by

$$l=\sqrt{\frac{\hbar c}{qB}}. \tag{6.1.39}$$

References: the classical mechanics at the beginning is similar to Shankar's presentation, the quantum mechanics is closer to that in Landau.

CHAPTER OVERVIEW

7: The Density Matrix

7.1: The Density Matrix

Thumbnail: measured density matrix of a thermal state. (CC SA-BY 3.0 unported; measured density matrix of a thermal state).

7.1: The Density Matrix

Pure States and Mixed States

Let's being with two imagined Stern-Gerlach experiments. In that experiment, a stream of (non-ionized) silver atoms from an oven is directed through an inhomogeneous vertical magnetic field, and the stream splits into two. The silver atoms have nonzero magnetic moments, and a magnetic moment in an inhomogeneous magnetic field experiences a nonzero force, causing the atom to veer from its straight line path, the magnitude of the deflection being proportional to the component of the atom's magnetic moment in the vertical (field) direction. The observation of the beam splitting into two, and no more, means that the vertical component of the magnetic moment, and therefore the associated angular momentum, can only have *two* different values. From the basic analysis of rotation operators and the properties of angular momentum that follow, this observation forces us to the conclusion that the total angular momentum of a silver atom is $\frac{1}{2}\hbar$. Ordinary orbital angular momenta cannot have half-integer values; this experiment was one of the first indications that the electron has a spin degree of freedom, an angular momentum that cannot be interpreted as orbital angular momentum of constituent parts. The silver atom has 47 electrons, 46 of them have total spin and orbital momenta that separately cancel, the 47th has no orbital angular momentum, and its spin is the entire angular momentum of the atom.

Here we shall use the Stern-Gerlach stream as an example of a large collection of quantum systems (the atoms) to clarify just how to describe such a collection, often called an *ensemble*. To avoid unnecessary complications, we only consider the *spin* degrees of freedom. We begin by examining two different streams:

Suppose experimentalist A prepares a stream of silver atoms such that each atom is in the spin state ψ_A:

$$|\psi_A\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle). \tag{7.1.1}$$

Meanwhile, experimentalist B prepares a stream of silver atoms which is a *mixture*: half the atoms are in state $|\uparrow\rangle$ and half are in the state $|\downarrow\rangle$: call this mix B.

Example 7.1.1

Can we distinguish the A stream from the B stream?

Solution

Evidently, not by measuring the spin in the z-direction! Both will give up 50% of the time, down 50%.

But: we *can* distinguish them by measuring the spin in the x-direction: the ψ_A quantum state is in fact just that of a spin in the x-direction, so it will give "up" in the x-direction every time—from now on we call it $|\uparrow_x\rangle$, whereas the state $|\uparrow\rangle$ ("up" in the z-direction) will yield "up" in the x-direction only 50% of the time, as will $|\downarrow\rangle$.

The state $\psi_A = |\uparrow_x\rangle$ is called a ***pure*** state, it's the kind of quantum state we've been studying this whole course.

The stream B, in contrast, is in a ***mixed*** state: the kind that actually occurs to a greater or lesser extent in a real life stream of atoms, different pure quantum states occurring with different probabilities, but with no phase coherence between them. In other words, these relative probabilities in B of different quantum states do *not* derive from probability amplitudes, as they do in finding the probability of spin up in stream A: the probabilities of the different quantum states in the mixed state B are exactly like classical probabilities.

That being said, though, to find the probability of measuring spin up in some such mixed state, one *first* uses the classical-type probability for each component state, *then* for each quantum state in the mix, one finds the probability of spin up *in that state* by the standard quantum technique.

Therefore, for a mixed state in which the system is in state $|\psi_i\rangle$ with probability w_i, $\sum w_i = 1$, the expectation value of an operator $\hat{A}$ is

$$\langle\hat{A}\rangle = \sum w_i \langle\psi_i|\hat{A}|\psi_i\rangle \tag{7.1.2}$$

and we should emphasize that these $|\psi_i\rangle$ do *not* need to be orthogonal (but they are of course normalized): for example one could be $|\uparrow_x\rangle$, another $|\uparrow_z\rangle$. (We put the usually omitted z in for emphasis.) The reason we put a hat on $\hat{A}$ here is to emphasize that this is an operator, but the w_i are just numbers.

The Density Matrix

The equation for the expectation value $\langle\hat{A}\rangle$ can be written:

$$\langle\hat{A}\rangle = Trace(\hat{\rho}\hat{A}) \tag{7.1.3}$$

where

$$\hat{\rho} = \sum w_i |\psi_i\rangle\langle\psi_i| \ . \tag{7.1.4}$$

To see exactly how this comes about, recall that for an operator $\hat{B}$ in a finite-dimensional vector space with an orthonormal basis set $|j\rangle$, $Tr\hat{B} = \sum_{j=1}^{n}\langle j|\hat{B}|j\rangle = B_{jj}$, where the repeated suffix implies summation of the diagonal matrix elements of the operator.

Therefore,

$$Tr(\hat{\rho}\hat{A}) = \sum_{j=1}^{n}\sum_{i=1}^{n} w_i\langle j|\psi_i\rangle\langle\psi_i|\hat{A}|j\rangle \tag{7.1.5}$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{n} w_i\langle\psi_i|\hat{A}|j\rangle\langle j|\psi_i\rangle \tag{7.1.6}$$

$$= \sum_{i=1}^{n} w_i\langle\psi_i|\hat{A}|\psi_i\rangle \tag{7.1.7}$$

since $\sum|j\rangle\langle j| = I$, the identity.

This $\hat{\rho}$ is called the *density matrix*: its matrix form is made explicit by considering states $|\psi_i\rangle$ in a finite N-dimensional vector space (such as spins or angular momenta)

$$|\psi_i\rangle = \sum_j (V_i)_j|j\rangle \tag{7.1.8}$$

where the $|j\rangle$ are an orthonormal basis set, and $(V_i)_j$ is the j^{th} component of a normalized vector V_i. It is convenient to express $\hat{\rho}$ in terms of kets and bras belonging to this orthonormal basis,

$$\hat{\rho} = \sum w_i|\psi_i\rangle\langle\psi_i| = \sum_{i,j,k} w_i(V_i)_j(V^{\dagger}i)_k|j\rangle\langle k| = \sum_{j,k}\rho_{jk}|j\rangle\langle k| \tag{7.1.9}$$

and evidently

$$\langle\hat{A}\rangle = Trace(\hat{\rho}\hat{A}) = \sum_{n,j,k}\langle n|\rho_{jk}|j\rangle\langle k|\hat{A}|n\rangle = \sum_{j,k}\rho_{jk}\langle k|\hat{A}|j\rangle = \sum_{j,k}\rho_{jk}A_{kj}. \tag{7.1.10}$$

(Since ρ_{jk} is just a *number*, $\langle n|\rho_{jk}|j\rangle = \rho_{jk}\langle n|j\rangle = \rho_{jk}\delta_{nj}$.)

$Trace(\hat{\rho}\hat{A})$ is *basis-independent*, the trace of a matrix being unchanged by a unitary transformation, since it follows from $Tr(ABC) = Tr(BCA)$ that

$$TrU^{\dagger}AU = TrAUU^{\dagger} = TrA \ for \ UU^{\dagger} = 1. \tag{7.1.11}$$

Note that since the vectors V_i are normalized, $\sum_j(V_i)_j(V_i^{\dagger})_j = 1$, with the i not summed over, and $\sum w_i = 1$, it follows that

$$Tr\hat{\rho} = 1 \tag{7.1.12}$$

(also evident by putting $A = 1$ in the equation for $\langle A\rangle$).

For a system in a ***pure*** quantum state $|\psi\rangle$, $\hat{\rho} = |\psi\rangle\langle\psi|$, just the projection operator into that state, and

$$\hat{\rho}^2 = \hat{\rho}, \tag{7.1.13}$$

as for all projection operators.

It’s worth spelling out how this differs from the mixed state by looking at the form of the density matrix.

For the pure state $|\psi\rangle$, if a basis is chosen so that $|\psi\rangle$ is a member of the basis (this can always be done), $\hat{\rho}$ is a matrix with every element zero except the one diagonal element corresponding to $|\psi\rangle\langle\psi|$, which will be unity. Obviously, $\hat{\rho}^2 = \hat{\rho}$. This is less obvious in a general basis, where $\hat{\rho}$ will not necessarily be diagonal. But the statement $\hat{\rho}^2 = \hat{\rho}$ remains true under a transformation to a new basis.

For a mixed state, let’s say for example a mixture of orthogonal states $|\psi_1\rangle, \ |\psi_2\rangle$, if we choose a basis including both states, the density matrix will be diagonal with just two entries $w_1, \ w_2$. Both these numbers must be less than unity, so $\hat{\rho}^2 \neq \hat{\rho}$. A mix of nonorthogonal states is left as an exercise for the reader.

Example 7.1.1: Pure State (case A)

First, our case A above (pure state): all spins in state $|\uparrow_x\rangle = (1/\sqrt{2})(|\uparrow\rangle + |\downarrow\rangle)$.

In the standard $|\uparrow\rangle, |\downarrow\rangle$ basis,

$$\hat{\rho} = |\uparrow_x\rangle\langle\uparrow_x| = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \tag{7.1.14}$$

and

$$\begin{aligned} \langle s_x \rangle &= Tr(\hat{\rho} s_x) = \frac{\hbar}{2} Tr \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{\hbar}{2} \\ \langle s_z \rangle &= Tr(\hat{\rho} s_z) = \frac{\hbar}{2} Tr \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 0. \end{aligned} \tag{7.1.15}$$

Notice that $\hat{\rho}^2 = \hat{\rho}$.

Example 7.1.1: 50-50 mixed up and down (case B)

50% in the state $|\uparrow\rangle$, 50% $|\downarrow\rangle$.

The density matrix is

$$\begin{aligned} \hat{\rho} &= \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{2}|\downarrow\rangle\langle\downarrow| \\ &= \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{7.1.16}$$

This is proportional to the unit matrix, so

$$Tr\hat{\rho} s_x = \frac{1}{2}\frac{\hbar}{2} Tr\sigma_x = 0, \tag{7.1.17}$$

and similarly for s_y and s_z, since the Pauli σ -matrices are all traceless. Note also that $\hat{\rho}^2 = \frac{1}{2}\hat{\rho} \neq \hat{\rho}$, as is true for all mixed states.

Example 7.1.3: Finally, a 50-50 mixed state relative to the x-axis (case C)

That is, 50% of the spins in the state $|\uparrow_x\rangle = (1/\sqrt{2})(|\uparrow\rangle + |\downarrow\rangle)$, “up” along the x- axis, and 50% in $|\downarrow_x\rangle = (1/\sqrt{2})(|\uparrow\rangle - |\downarrow\rangle)$, “down” in the x-direction.

It is easy to check that

$$\hat{\rho} = \frac{1}{2}|\uparrow_x\rangle\langle\uparrow_x| + \frac{1}{2}|\downarrow_x\rangle\langle\downarrow_x| = \frac{1}{2}\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{7.1.18}$$

This is exactly the same density matrix we found for 50% in the state $|\uparrow\rangle$, 50% $|\downarrow\rangle$!

The reason is that both formulations describe a state about which we know nothing—we are in a state of *total ignorance*, the spins are completely random, all directions are equally likely. The density matrix describing such a state cannot depend on the direction we choose for our axes.

Another two-state quantum system that can be analyzed in the same way is the polarization state of a beam of light, the basis states being polarization in the x-direction and polarization in the y-direction, for a beam traveling parallel to the z- axis. Ordinary unpolarized light corresponds to the random mixed state, with the same density matrix as in the last example above.

Time Evolution of the Density Matrix

In the mixed state, the quantum states evolve independently according to Schrödinger's equation, so

$$i\hbar\frac{d\hat{\rho}}{dt}=\sum w_i H|\psi_i\rangle\langle\psi_i|-\sum w_i|\psi_i\rangle\langle\psi_i|H=[H,\hat{\rho}]. \tag{7.1.19}$$

Note that this has the opposite sign from the evolution of a Heisenberg operator, not surprising since the density operator is made up of Schrödinger bras and kets.

The equation is the quantum analogue of *Liouville's theorem* in statistical mechanics. Liouville's theorem describes the evolution in time of an ensemble of identical classical systems, such as many boxes each filled with the same amount of the same gas at the same temperature, but the positions and momenta of the individual atoms are randomly different in each. Each box can be classically described by a single point in a huge dimensional space, a space having six dimensions for each atom (position and momentum, we ignore possible internal degrees of freedom). The whole ensemble, then, is a gas of these points in this huge space, and the rate of change of local density of this gas, from Hamilton's equations, is $\partial\rho/\partial t=-\{\rho,H\}$, the bracket now being a Poisson bracket (see my *Classical Mechanics* notes). Anyway, this is the classical precursor of, and the reason for the name of, the density matrix.

Thermal Equilibrium

A system in thermal equilibrium is represented in statistical mechanics by a *canonical ensemble*. If the eigenstate $|i\rangle$ of the Hamiltonian has energy E_i, the relative probability of the system being in that state is $e^{-E_i/kT}=e^{-\beta E_i}$ in the standard notation. Therefore the density matrix is:

$$\hat{\rho}=\frac{1}{Z}\sum_i e^{-\beta E_i}|i\rangle\langle i|=\frac{e^{-\beta H}}{Z}, \tag{7.1.20}$$

where

$$Z=\sum_i e^{-\beta E_i}=Tre^{-\beta H}. \tag{7.1.21}$$

Notice that in this formulation, apart from the normalization constant Z, the density operator is analogous to the propagator $U(t)=e^{-iHt/\hbar}$ for an imaginary time $t=-i\hbar\beta$. Incidentally, for interacting quantum fields, the propagator can be constructed as a set of Feynman diagrams corresponding to all possible sequences of particle scatterings by interaction. To find the thermodynamic properties of a field theory at finite temperature, essentially the same set of diagrams is used to find the free energy: the diagrams now describe the system propagating for a finite imaginary time, the same mathematical tools can be used.

At zero temperature ($\beta=\infty$) the probability coefficients $w_i=e^{-\beta E_i}/Z$ are all zero except for the ground state: the system is in a pure state, and the density matrix has every element zero except for a single element on the diagonal. At infinite temperature, all the w_i are equal: the density matrix is just $1/N$ times the unit matrix, where N is the total number of states available to the system. In fact, the *entropy* of the system can be expressed in terms of the density matrix: $S=-kTr(\hat{\rho}\ln\hat{\rho})$. This is not as bad as it looks: both operators are diagonal in the energy subspace.

CHAPTER OVERVIEW

8: Approximate Methods

So far, we have concentrated on problems that were analytically solvable, such as the simple harmonic oscillator, the hydrogen atom, and square well type potentials. In fact, we shall soon be confronted with situations where an exact analytic solution is unknown: more general potentials, or atoms with more than one electron. To make progress in these cases, we need approximation methods. The best known method is perturbation theory, which has proved highly successful over a wide range of problems (but by no means all).

8.1: Variational Methods
8.2: The WKB Approximation
8.3: Note on the WKB Connection Formula

Thumbnail: Two (or more) wave functions are mixed by linear combination. The coefficients c_1, c_2 determine the weight each of them is given. The optimum coefficients are found by searching for minima in the potential landscape spanned by c_1 and c_2. (CC BY-SA 3.0; Rudolf Winter at Aberystwyth University).

8.1: Variational Methods

So far, we have concentrated on problems that were analytically solvable, such as the simple harmonic oscillator, the hydrogen atom, and square well type potentials. In fact, we shall soon be confronted with situations where an exact analytic solution is unknown: more general potentials, or atoms with more than one electron. To make progress in these cases, we need approximation methods. The best known method is perturbation theory, which has proved highly successful over a wide range of problems (but by no means all). We shall soon be discussing perturbation methods at length. First, though, we shall review two other approximation methods: in this lecture, the variational method, then in the next lecture the semiclassical WKB method. The variational method works best for the ground state, and in some circumstances (to be described below) for some other low lying states; the WKB method is good for higher states.

Variational Method for Finding the Ground State Energy

The idea is to guess the ground state wave function, but the guess must have an adjustable parameter, which can then be varied (hence the name) to minimize the expectation value of the energy, and thereby find the best approximation to the true ground state wave function. This crude sounding approach can in fact give a surprisingly good approximation to the ground state energy (but usually not so good for the wave function, as will become clear).

We'll begin with a single particle in a potential, $H=p^2/2m+V(\vec{r})$. If the particle is restricted to one dimension, and we're looking for the ground state in any fairly localized potential well, we can start with the family of normalized Gaussians, $|\psi,\alpha\rangle=\left(\frac{\alpha}{\pi}\right)^{1/4}e^{-\alpha x^2/2}$: just find $\langle\psi,\alpha|H|\psi,\alpha\rangle$, differentiate the result with respect to α, setting this to zero (and checking that you have in fact found a minimum.) Not surprisingly, this gives the exact ground state for the simple harmonic oscillator potential, and for nothing else. What is perhaps surprising is that the result is only off by 30% or so for the attractive delta-function potential, even though the wave function looks a lot different (solved in detail in Griffiths, page 258). Obviously, the Gaussian family cannot be used if there is an infinite wall anywhere: one must find a family of wave functions vanishing at the wall.

To gain some insight into what we're doing, suppose the Hamiltonian $H=p^2/2m+V(\vec{r})$ has the set of (unknown to us) eigenstates

$$H|n\rangle=E_n|n\rangle. \tag{8.1.1}$$

Since the Hamiltonian is Hermitian, these states span the space of possible wave functions, including our variational family, so:

$$|\psi,\alpha\rangle=\sum a_n(\alpha)|n\rangle. \tag{8.1.2}$$

From this,

$$\frac{\langle\psi,\alpha|H|\psi,\alpha\rangle}{\langle\psi,\alpha|\psi,\alpha\rangle}=\sum|a_n|^2E_n\geq E_0 \tag{8.1.3}$$

for *any* $|\psi,\alpha\rangle$. (We don't need the denominator if we've chosen a family of normalized wave functions, as we did with the Gaussians above.) Evidently, minimizing $\frac{\langle\psi,\alpha|H|\psi,\alpha\rangle}{\langle\psi,\alpha|\psi,\alpha\rangle}$ as a function of α gives us an upper bound on the ground state energy, hopefully not too far from the true value.

We can see immediately that this will probably be better for finding for the ground state energy than for mapping the ground state wave function: suppose the optimum state in our family is actually $|\alpha_{min}\rangle=N(|0\rangle+0.2|1\rangle)$, with the normalization constant $N\cong 0.98$, a 20% admixture of the first excited state. Then the wave function is off by of order 20%, but the energy estimate will be too high by $0.04(E_1-E_0)$ usually a much smaller error.

To get some idea of how well this works, Messiah applies the method to the ground state of the hydrogen atom. We know it's going to be spherically symmetric, so it amounts to a one-dimensional problem: just the radial wave function. Using standard notation,

$$a_0=\hbar^2/me^2,\quad E_0=me^4/2\hbar^2,\quad \rho=r/a_0 \tag{8.1.4}$$

and for a trial wave function *u*

$$E(u)=-E_0\frac{\int u\left(\frac{d^2}{d\rho^2}+\frac{2}{\rho}\right)ud\rho}{\int u^2d\rho} \tag{8.1.5}$$

(we’re going to take *u* real).

Messiah tries three families:

$$\begin{aligned} u_1 &= \rho e^{-\alpha\rho} \\ u_2 &= \frac{\rho}{\alpha^2+\rho^2} \\ u_3 &= \rho^2 e^{-\alpha\rho} \end{aligned} \tag{8.1.6}$$

and finds $\alpha_{min} = 1,\ \pi/4,\ 3/2$ respectively. The first family, u_1, includes the exact result, and the minimization procedure finds it.

For the three families, then the energy of the best state is off by 0, 25%, 21% respectively.

The wave function error is defined as how far the square of the overlap with the true ground state wave function falls short of unity. For the three families, $\varepsilon = 1 - |\langle\psi_0|\psi_{var}\rangle|^2$:0, 0.21, 0.05. Notice here that our hand waving argument that the energies would be found much more accurately than the wave functions comes unstuck. The third family has far better wave function overlap than the second, but only a slightly better energy estimate. Why? A key point is that the potential is singular at the origin, there is a big contribution to potential energy from a rather small region, and the third family wave function is the least accurate of the three there. The second family functions are very inaccurate at large distances: the expectation value $\langle r\rangle = 1.5a_0,\ \infty,\ 1.66a_0$ for the three families. But at large distances, both kinetic and potential energies are small, so the result can still look reasonable. These examples reinforce the point that the variational method should be used cautiously.

Variational Method for Higher States

In some cases, the approach can be used easily for higher states: specifically, in problems having some symmetry. For example, if the one dimensional attractive potential is symmetric about the origin, and has more than one bound state, the ground state will be even, the first excited state odd. Therefore, we can estimate the energy of the first excited state by minimizing a family of odd functions, such as

$$\psi(x,\alpha) = (\sqrt{\pi}/2\alpha^{3/2})xe^{-\alpha x^2/2}. \tag{8.1.7}$$

Ground State Energy of the Helium Atom by the Variational Method

We know the ground state energy of the hydrogen atom is -1 Ryd, or -13.6 ev. The He$^+$ ion has $Z=2$, so will have ground state energy, proportional to Z^2, equal to -4 Ryd. Therefore for the He atom, if we neglect the electron-electron interaction, the ground state energy will be -8 Ryd, -109 ev., the two electrons having opposite spins will both be in the lowest spatial state. Actually, experimentally, the He atom ground state energy is only -79 ev, because the repulsion between the electrons loosens things.

To get a better value for the ground state energy still using tractable wave functions, we change the wave functions from the ionic wave function $(Z^3/\pi a_0^3)^{1/2}e^{-Z_r/a_0}$ with $Z=2$ to $(Z'^3/\pi a_0^3)^{1/2}e^{-Z'_r/a_0}$ with Z' now a variable parameter. In other words, we are trying to allow for the electron-electron repulsion, which must push the wave functions out a bit, by keeping exactly the same shaped wave function but lessening the effective nuclear charge as reflected in the spread of the wave function from Z to Z', and we’ll determine Z' by varying it to find the minimum total energy, including the term from electron-electron repulsion.

To find the potential energy from the nuclear-electron interactions, we of course use the actual nuclear charge $Z=2$, but the Z' wave function, so the nuclear *P.E.* for the two electrons is:

$$\begin{aligned} P.E. &= -2Ze^2\int_0^\infty \frac{1}{r}4\pi r^2 dr(Z'^3/\pi a_0^3)e^{-2Z'_r/a_0} \\ &= -4ZZ'(e^2/2a_0) \\ &= -8Z'\ Ryd\ \ (Z=2). \end{aligned} \tag{8.1.8}$$

This could have been figured out from the formula for the one-electron ion, where the potential energy for the one electron is $-2Z^2$ Ryd, one factor of Z being from the nuclear charge, the other from the consequent shrinking of the orbit.

The kinetic energy is even easier: it depends entirely on the shape of the wave function, not on the actual nuclear charge, so for our trial wave function it has to be Z'^2 Ryds per electron.

The tricky part is the *P.E.* for the electron-electron interaction. This is positive.

Each electron has a wave function $(Z'^3/\pi a_0^3)^{1/2}e^{-Z'_r/a_0}$, a spherical charge probability distribution.

Denoting charge probability density by $\rho(r)$, we need

$$\begin{gathered} I=\int\int d\,\vec{r}_1 d\,\vec{r}_2 \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{|\vec{r}_1-\vec{r}_2|} \\ =16\pi^2 \int_0^\infty r_1^2 dr_1 \int_0^\infty r_2^2 dr_2 \frac{\rho(r_1)\rho(r_2)}{r_>}, \cdot \\ r_> = max(r_1, r_2). \end{gathered} \tag{8.1.9}$$

Collecting terms, the total energy (for $Z=2$) is:

$$E=-2(4Z'-Z'^2-\frac{5}{8}Z')Ryd \tag{8.1.10}$$

and this is minimized by taking $Z'=2-\frac{5}{16}$, giving an energy of -77.5 ev, off the true value by about 1 ev, so indeed the presence of the other electron is taken care of as far as total energy is concerned by shielding the nuclear charge by an amount (5/16)e.

8.2: The WKB Approximation

The WKB (Wentzel, Kramers, Brillouin) approximation is, in sense to be made clear below, a quasi-classical method for solving the one-dimensional (and effectively one-dimensional, such as radial) time-independent Schrödinger equation. The nontrivial step in the method is the connection formulas (see below), that problem was first solved by Lord Rayleigh (Proc. Roy. Soc. A, 86, 1912, 207) and as Jeffries notes (*Mathematical Physics*, p 526) “it has been rediscovered by several later writers” presumable referring to W, K and B. By the way, the English call it the Jeffries approximation, or, if feeling sufficiently ecumenical, the WKBJ approximation. (In this lecture, we only consider bound states: the most famous application of WKB, α- decay, was covered in detail in the undergraduate quantum mechanics course based on Griffiths’ book.)

We’ll follow the development in Landau and Lifshitz, who consider this all sufficiently obvious that they don’t mention any of these people. In fact, they call it

The Semiclassical Approximation to Leading Order

Consider a particle moving along in a slowly varying one-dimensional potential. By “slowly varying” we mean here that in any small region the wave function is well approximated by a plane wave, and that the wavelength only changes over distances long compared with a wavelength. We’re also assuming for the moment that the particle has positive kinetic energy in the region. Under these conditions, it’s easy to see the general form of the solution to the time independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}+V(x)\psi(x)=E\psi(x) \tag{8.2.1}$$

Very approximately, $\psi(x)$ will look like $A(x)e^{\pm ip(x)x/\hbar}$ where $p(x)$ is the “local momentum” we define classically by

$$p(x)^2/2m+V(x)=E, \tag{8.2.2}$$

and $A(x)$ is slowly varying compared with the phase factor.

> *Clearly this is a semiclassical limit: $\hbar$ has to be sufficiently small that there are many oscillations in the typical distance over which the potential varies.*

To handle this a little more precisely, we emphasize the rapid *phase* variation in this semiclassical limit by writing the wave function

$$\psi(x)=e^{(i/\hbar)\sigma(x)} \tag{8.2.3}$$

and writing Schrödinger ’s equation for $\sigma(x)$.

So from

$$i\hbar\psi'(x)=-\sigma'(x)e^{(i/\hbar)\sigma(x)}, \tag{8.2.4}$$

and

$$-\hbar^2\psi''(x)=-i\hbar\sigma''(x)e^{(i/\hbar)\sigma(x)}+(\sigma'(x))^2e^{(i/\hbar)\sigma(x)}, \tag{8.2.5}$$

Schrödinger’s equation written for the phase function is:

$$-i\hbar\sigma''(x)+(\sigma'(x))^2=(p(x))^2. \tag{8.2.6}$$

And, since we’re assuming the system is close to classical, it makes sense to expand σ as a series in $\hbar$ (following Landau and Lifshitz):

$$\sigma=\sigma_0+(\hbar/i)\sigma_1+(\hbar/i)^2\sigma_2+\dots \tag{8.2.7}$$

The zeroth order approximation is

$$(\sigma'_0)^2=p^2 \tag{8.2.8}$$

and fixing the sign of p by

$$p(x)=+\sqrt{2m(E-V(x))} \tag{8.2.9}$$

we conclude that

$$\sigma_0=\pm\int p(x)dx. \tag{8.2.10}$$

(As we discussed in the lecture on path integrals, in the classical limit one path dominates, and the phase of the wave function is $(i/\hbar)$ times the classical action S along that path. In the present case, $S=-Et\pm\int pdx$, we've already factored out the Et since we're dealing here with the time-independent wave function.)

Region of Validity of the Approximation

From the Schrödinger equation $-i\hbar\sigma''(x)+(\sigma'(x))^2=(p(x))^2$, it is evident that this approximate solution is *only valid if we can ignore that first term*. That is to say, we must have

$$|\hbar\sigma''(x)/(\sigma'(x))^2|\ll 1, \tag{8.2.11}$$

or

$$\left|\frac{d(\hbar/\sigma')}{dx}\right|\ll 1. \tag{8.2.12}$$

But in leading approximation $\sigma'=p$, and $p=2\pi\hbar/\lambda$, so the condition is

$$\frac{1}{2\pi}\left|\frac{d\lambda}{dx}\right| \ll 1. \tag{8.2.13}$$

This just means the change in wavelength over a distance of one wavelength must be small. Obviously, this cannot always be the case: if the particle is confined by an attractive potential, at the edge of the classically allowed region, that is, where $E=V(x)$, p is zero and the wavelength is infinite. The approximation is only good well away from that point, to which we shall return shortly.

Next to Leading Order Correction

The second term in the $\hbar$ expansion of the phase,

$$\sigma=\sigma_0+(\hbar/i)\sigma_1+\dots \tag{8.2.1}$$

satisfies

$$-i\hbar\sigma''_0+2\sigma'_0(\hbar/i)\sigma'_1=0 \tag{8.2.14}$$

so

$$\sigma'_1=-\sigma''_0/2\sigma'_0=-p'/2p, \tag{8.2.15}$$

and

$$\sigma_1=-\frac{1}{2}\ln p. \tag{8.2.16}$$

So the wave function to this order is:

$$\psi(x)=\frac{C_1}{\sqrt{p(x)}}e^{(i/\hbar)\int pdx}+\frac{C_2}{\sqrt{p(x)}}e^{-(i/\hbar)\int pdx}. \tag{8.2.17}$$

(Recall we fixed the sign of p to be positive.)

To interpret the $\sqrt{p(x)}$ factor, consider the first term, a wave moving to the right. Since p is real, the exponential has modulus unity, and the local amplitude squared is proportional to $1/p$, that is, $1/v$, where v is the velocity of the particle. This is simple to understand physically: the probability of finding the particle in any given small interval is proportional to the time it spends there, hence inversely proportional to its speed.

We turn now to the wave function in the classically forbidden region,

$$p(x)^2/2m = E - V(x) < 0. \tag{8.2.18}$$

Here p is of course pure imaginary, but the same formal phase solution of the Schrödinger equation works, again provided that the particle is well away from the points where $E = V(x)$.

The wave function is:

$$\psi(x) = \frac{C'_1}{\sqrt{|p(x)|}} e^{-(1/\hbar)\int |p| dx} + \frac{C'_2}{\sqrt{|p(x)|}} e^{(1/\hbar)\int |p| dx}. \tag{8.2.19}$$

Connection Formulas, Boundary Conditions and Quantization Rules

Let us assume we're dealing with a one-dimensional potential, and the classically allowed region is $b \le x \le a$. (I'm just following Landau's notation here.) Clearly, in the forbidden region to the right, $x > a$, only the first term in the above equation for $\psi(x)$ appears, and for $x < b$ only the second term. Furthermore, in the "inside" (classically allowed) region, $b \le x \le a$, the wave function has the oscillating form discussed earlier.

But how do we connect the three regions together? We make an assumption: we take it that the potential varies sufficiently smoothly that it's a good approximation to take it to be linear in the vicinity of the classical turning points. That is to say, we assume that a linear potential is a sufficiently good approximation out to the point where the short wavelength (or decay length for tunneling regions) description is adequate.

Therefore, near $x = a$, we take the potential to be

$$E - V(x) \cong F_0(x-a) \tag{8.2.20}$$

(so F_0 would be the force) and then approximate the wave function by the known exact solution for a linear everywhere potential: the Airy function.

It is known that for the Airy function, the solution having the form

$$\psi(x) = \frac{C}{2\sqrt{|p(x)|}} e^{-(1/\hbar)\int_a^x |p| dx} \tag{8.2.21}$$

to the right becomes

$$\begin{aligned} \psi(x) &= \tfrac{C}{|p(x)|}\cos\left((1/h)\textstyle\int_a^x p dx + \tfrac{1}{4}\pi\right) \\ &= \tfrac{C}{|p(x)|}\sin\left((1/h)\textstyle\int_x^a p dx + \tfrac{1}{4}\pi\right) \end{aligned} \tag{8.2.22}$$

(The derivation of this "connection" is given in my notes here.)

At b, the same argument gives

$$\psi(x) = \frac{C}{|p(x)|}\sin\left((1/h)\int_b^x p dx + \frac{1}{4}\pi\right). \tag{8.2.23}$$

For these two expressions to be consistent, we must have

$$\frac{1}{\hbar}\int_b^a p dx + \frac{1}{2}\pi = (n+1)\pi, \ or \ \oint p dx = 2\pi\hbar\left(n + \frac{1}{2}\right). \tag{8.2.24}$$

where the latter integral is over a complete cycle of the classical motion.

Here \(n\) is the number of zeros of the wave function: this is the *quantization condition*.

Relating Classical Circuit time to Quantized Energy Levels

The time for a complete classical circuit is

$$T = 2\int_b^a dx/v = 2m\int_b^a dx/p \tag{8.2.2}$$

is the area of the classical path in phase space, so we see each state has an element of phase space $2\pi\hbar$. From this, we can figure out the approximate energy splitting between levels in the quasi-classical limit: the change in the integral with energy ΔE

corresponding to one level must be $2\pi\hbar$. That is,

$$\Delta E\oint(\partial p/\partial E)dx=2\pi\hbar. \tag{8.2.25}$$

Now $(\partial E/\partial p)=v$, so

$$\oint(\partial p/\partial E)dx=\oint dx/v=T. \tag{8.2.3}$$

Therefore, $\Delta E=2\pi\hbar/T=\hbar\omega$.

This is just saying that if the particle emits one photon and drops to the next level, the frequency of the photon emitted is just the orbital frequency of the particle, a very natural conclusion in the quasi-classical limit.

The Radial Case

In the above analysis for a particle confined to one dimension, the connection formulas can be understood with a simple picture: the wave function “spills over” into the forbidden regime, and its twisting there counts as an extra $\frac{1}{4}\pi$ of phase change, so in the lowest state the total phase change in the allowed region need only be $\frac{1}{2}\pi$. In the radial case, assuming the potential is well behaved at the origin, the wave function goes to zero there. A bound state will still spill over beyond the classical turning point at r_0, say, but clearly there must be a total phase change of $\frac{3}{4}\pi$ in the allowed region for the lowest state, since there can be no spill over to negative r.

The general formula will be

$$\frac{1}{\hbar}\int_0^{r_0}p(r)dr=(n+34)\pi,\ \ n=0,1,2,\ldots, \tag{8.2.26}$$

the series terminating if and when the turning point reaches infinity.

Warning: actually, some potentials, including the Coulomb potential and the centrifugal barrier for $l\neq 0$, are in fact singular at $r=0$. These cases require special treatment.

8.3: Note on the WKB Connection Formula

Semiclassical Analysis of a Particle Trapped in a Well in One Dimension

The WKB semiclassical approximate solution to Schrödinger’s equation,

$$\psi(x)=\psi(x_0)\sqrt{\frac{p(x_0)}{p(x)}}\exp\left(\pm\frac{i}{\hbar}\int_{x_0}^{x}p(x')dx'\right) \tag{8.3.1}$$

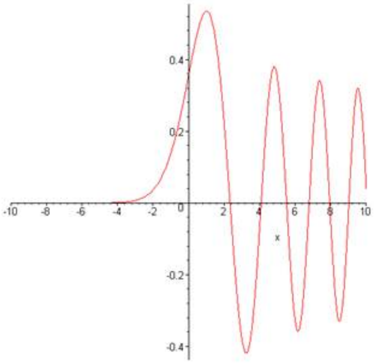
is reliable in regions where the wavelength (for oscillating solutions) or the decay length (for exponential solutions) changes only slightly over a distance of one wavelength or decay length respectively. For a particle trapped in a (one-dimensional) potential well, classically the particle would bounce back and forth between the two turning points where its kinetic energy vanishes. In the quantum case, these are precisely the points where the wavelength becomes infinite, so the WKB solution fails.

Taking the Potential Near the Turning Points to be Linear…

However, for a reasonably smooth potential it may be an adequate approximation to treat a turning point region as one where the potential is increasing linearly with distance over a sufficient range that beyond this point the WKB approximation can be used in both directions. The solution of Schrödinger’s equation for a linearly increasing or decreasing potential is well known, it is the Airy function, the solution of the differential equation

$$\frac{d^2y}{dx^2}+xy=0 \tag{8.3.2}$$

plotted here at the left-hand turning point :



The strategy is to evaluate this function for large x, both positive and negative, so that we can join together the two WKB solutions, valid in the far regions, in a quantitative fashion.

Following Mathews and Walker (page 116) the differential equation is most simply solved by taking its Fourier transform. If

$$g(\omega)=\int_{-\infty}^{\infty}y(x)e^{-i\omega x}dx, \tag{8.3.3}$$

then

$$-\omega^2g(\omega)+i\frac{dg}{d\omega}=0,\ \ so\ g(\omega)=Ae^{-i(\omega^3/3)}. \tag{8.3.4}$$

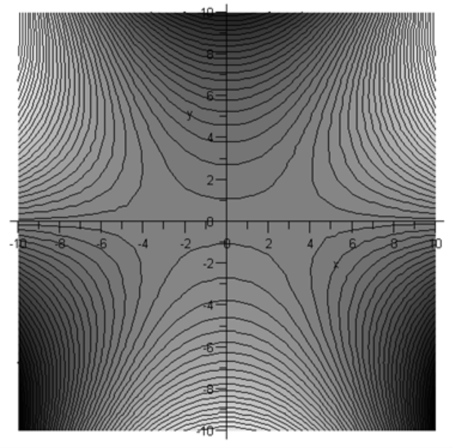
Therefore

$$y(x)=A\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\exp\left[i\left(\omega x-\frac{\omega^3}{3}\right)\right]. \tag{8.3.5}$$

Keeping a Low Profile …

This is an exact result, but a nontrivial integral! Fortunately, we are only interested in its value for *large* values of $|x|$, and this is precisely where saddlepoint methods become accurate: the contour is distorted to be as low-lying as possible, then the only significant contributions to the integral (thanks to the exponential variation of the amplitude) are from those places where we must go over a saddlepoint to get from one valley to the next. To see where the saddlepoints are, and how they relate to an integral along the real axis, we plot below contour maps of the absolute value of the term in the exponential.

Large positive x: In the map below, we take $x=10$ (rather small), so the saddlepoints are at $\pm\sqrt{10}$. If the path of integration is moved down from the real axis into the (dark) valleys (so the integrand becomes exponentially smaller) the contour from $-\infty$ will come up from the bottom left to the left saddlepoint, over the saddle into the top center valley, then back over the second saddlepoint into the bottom right valley and on to $+\infty$.



Lighter color means higher ground.

Writing the integral as

$$\int_{-\infty}^{\infty} e^{f(\omega)} d\omega \tag{8.3.6}$$

(dropping irrelevant overall constants) then

$$f(\omega)=i\left(\omega x-\frac{\omega^3}{3}\right),\ f'(\omega)=i(x-\omega^2)\ \mathit{and}\ f\text{ ” }(\omega)=-2i\omega. \tag{8.3.7}$$

(Mathews and Walker take the “large” parameter x out of f, we’ve left it in—this doesn’t affect the final result.) Near the positive saddlepoint

$$f(\omega)=f(\sqrt{x})+\frac{1}{2}f''(\sqrt{x})(\omega-\sqrt{x})^2=i\frac{2}{3}x^{3/2}-i\sqrt{x}(\omega-\sqrt{x})^2 \tag{8.3.8}$$

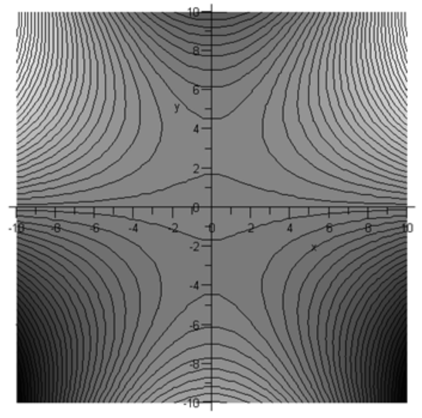
dropping higher-order terms. Since $f''$is pure imaginary at the saddlepoint, the appropriate path for a real exponent in the Gaussian integral is at $\pi/4$to the x- axis. So in the path integral

$$dz=e^{\pm i\pi/4}ds \tag{8.3.9}$$

where ds is a real parameter measuring incremental path length, and the sign in the exponent is positive for the saddlepoint on the left. The contributions from the two saddlepoints give the asymptotic (large positive x) solutions as:

$$y(x)\sim\frac{2\sqrt{\pi}}{x^{1/4}}\cos\left(\frac{2}{3}x^{3/2}-\frac{\pi}{4}\right). \tag{8.3.10}$$

Large negative x: In this case, the saddlepoint geography is quite different, although *the distant geography is the same*, being dominated by the ω^3 term.



Just as before, the real axis integration path can be moved down from the real axis into the valleys at bottom left and bottom right. (The extra contributions from linking up the new path with the real axis at infinity are zero.) It is clear from the map above that to get from the valley on the left to the one on the right means just going over the saddlepoint on the negative imaginary axis. Note from the darker shading that the other saddlepoint is at higher elevation.

The integration through the saddlepoint is parallel to the real axis, and gives

$$y(x) \sim \frac{\sqrt{\pi}}{(-x)^{1/4}} \exp\left[-\frac{2}{3}(-x)^{3/2}\right]. \tag{8.3.11}$$

This, then, is the decaying wavefunction solution of the Airy equation that we are looking for, and it is clear that it goes smoothly from this exponential to the cosine form

$$y(x) \sim \frac{2\sqrt{\pi}}{x^{1/4}} \cos\left(\frac{2}{3}x^{3/2} - \frac{\pi}{4}\right). \tag{8.3.10}$$

as x is taken along the real axis from large negative to large positive values.

Landau's Quick Method

Incidentally, Landau gives a quick way to see how these formulas connect. These are asymptotic solutions to the original Airy equation, valid as long as x is well away from the origin. But we can go from one to the other, avoiding the origin, if we regard x as a complex variable and move out into the complex plane—Landau takes a semicircle in the upper half plane, begins with the exponentially decaying solution now written

$$y(\rho e^{i\phi}) \sim \frac{\sqrt{\pi}}{(-\rho e^{i\phi})^{1/4}} \exp\left[-\frac{2}{3}\rho^{3/2}\left(\cos\frac{3}{2}\phi + i\sin\frac{3}{2}\phi\right)\right] \tag{8.3.12}$$

the phase ϕ varying from 0 to π. The exponential factor at first increases in modulus, then becomes pure imaginary, exactly equal to one of the two terms in the cosine form for positive x.

If one takes the cosine formula and continues it into the upper half complex plane, one term grows exponentially, the other goes to zero. Suppose one begins at some large value of x and moves in a circle around $x = 0$ back to the now negative real axis. This will give

$$y(x) \sim \frac{\sqrt{\pi}}{(x)^{1/4}} \exp i\left(\frac{2}{3}(x)^{3/2} - \sqrt{\pi}4\right). \tag{8.3.13}$$

This is actually the same as the expression we already derived: the $\pi/4$ cancels against the minus sign inside the fourth root, the i against the minus sign raised to the $3/2$ power.

CHAPTER OVERVIEW

9: Perturbation Theory

The Schrödinger equation for realistic systems quickly becomes unwieldy, and analytical solutions are only available for very simple systems - the ones we have described as fundamental systems in this module. Numerical approaches can cope with more complex problems, but are still (and will remain for a good while) limited by the available computer power. Approximations are necessary to cope with real systems. Perturbation theory is one such approximation that is best used for small changes to a known system, whereby the Hamiltonian is modified.

9.1: Time-Independent Perturbation Theory
9.2: The Peierls Transition - an Unexpected Insulator
9.3: Van Der Waals Forces between Atoms
9.4: The Interaction Representation
9.5: Time-Dependent Perturbation Theory
9.6: The Photoelectric Effect in Hydrogen
9.7: Quantizing Radiation

Thumbnail: The unperturbed Hamiltonian (blue curve) of a known system is modified by adding a perturbation (red curve) with a variable control parameter λ, which governs the extent to which the system is perturbed. (CC BY-SA 3.0; Rudolf Winter at Aberystwyth University).

9.1: Time-Independent Perturbation Theory

If an atom (not necessarily in its ground state) is placed in an external electric field, the energy levels shift, and the wavefunctions are distorted. This is called the *Stark effect*. The new energy levels and wavefunctions could in principle be found by writing down a complete Hamiltonian, including the external field, and finding the eigenkets. This actually can be done in one case: the hydrogen atom, but even there, if the external field is small compared with the electric field inside the atom (which is billions of volts per meter) it is easier to compute the changes in the energy levels and wavefunctions with a scheme of successive corrections to the zero-field values. This method, termed *perturbation theory*, is the single most important method of solving problems in quantum mechanics, and is widely used in atomic physics, condensed matter and particle physics.

> It should be noted that there *are* problems which cannot be solved using perturbation theory, even when the perturbation is very weak, although such problems are the exception rather than the rule. One such case is the one-dimensional problem of free particles perturbed by a localized potential of strength λ. As we found earlier in the course, switching on an arbitrarily weak attractive potential causes the $k=0$ free particle wavefunction to drop below the continuum of plane wave energies and become a localized bound state with binding energy of order λ^2. However, changing the sign of λ to give a repulsive potential there is no bound state, the lowest energy plane wave state stays at energy zero. Therefore the energy shift on switching on the perturbation cannot be represented as a power series in λ, the strength of the perturbation. This particular difficulty does not in general occur in three dimensions, where arbitrarily weak potentials do not give bound states—except for certain many-body problems (like the Cooper pair problem) where the exclusion principle reduces the effective dimensionality of the available states.

The Perturbation Series

We begin with a Hamiltonian H^0 having known eigenkets and eigenenergies:

$$H^0|n^0\rangle = E_n^0|n^0\rangle. \tag{9.1.1}$$

The task is to find how these eigenkets and eigenenergies change if a small term H^1 (an external field, for example) is added to the Hamiltonian, so:

$$(H^0+H^1)|n\rangle = E_n|n\rangle. \tag{9.1.2}$$

That is to say, on switching on H^1,

$$|n^0\rangle \to |n\rangle, \quad E_n^0 \to E_n. \tag{9.1.3}$$

The basic assumption in perturbation theory is that H^1 is sufficiently small that the leading corrections are the same order of magnitude as H^1 itself, and the true energies can be better and better approximated by a successive series of corrections, each of order H^1/H^0 compared with the previous one.

The strategy, then, is to expand the true wavefunction and corresponding eigenenergy as series in H^1/H^0. These series are then fed into $(H^0+H^1)|n\rangle = E_n|n\rangle$, and terms of the same order of magnitude in H^1/H^0 on the two sides are set equal. The equations thus generated are solved one by one to give progressively more accurate results.

To make it easier to identify terms of the same order in H^1/H^0 on the two sides of the equation, it is convenient to introduce a dimensionless parameter λ which always goes with H^1, and then expand $|n\rangle$, E_n as power series in λ, $|n\rangle = |n^0\rangle + \lambda|n^1\rangle + \lambda^2|n^2\rangle + \dots$, etc. The ket $|n^m\rangle$ multiplied by λ^m is therefore of order $(H^1/H^0)^m$.

> This λ is purely a bookkeeping device: we will set it equal to 1 when we are through! It's just there to keep track of the orders of magnitudes of the various terms.

Putting the series expansions for $|n\rangle, \quad E_n$ in

$$(H^0+\lambda H^1)|n\rangle = E_n|n\rangle \tag{9.1.4}$$

we have

$$\begin{gathered}(H^0+\lambda H^1)(|n^0\rangle + \lambda|n^1\rangle + \lambda^2|n^2\rangle + \dots) \\ = (E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \dots)(|n^0\rangle + \lambda|n^1\rangle + \lambda^2|n^2\rangle + \dots).\end{gathered} \tag{9.1.5}$$

We're now ready to match the two sides term by term in powers of λ.

Zeroth Order of Energy

The zeroth-order term ($\lambda = 0$), of course, just gives back

$$H^0|n^0\rangle = E_n^0|n^0\rangle \tag{9.1.6}$$

Zeroth Order of Eigenstates

$$|n^0\rangle = |n^0\rangle \tag{9.1.7}$$

First-Order of Energy

Matching the terms linear in λ on both sides:

$$H^0|n^1\rangle + H^1|n^0\rangle = E_n^0|n^1\rangle + E_n^1|n^0\rangle. \tag{9.1.8}$$

This equation is the key to finding the first-order change in energy E_n^1. Taking the inner product of both sides with $\langle n^0|$:

$$\langle n^0|H^0|n^1\rangle + \langle n^0|H^1|n^0\rangle = \langle n^0|E_n^0|n^1\rangle + \langle n^0|E_n^1|n^0\rangle, \tag{9.1.9}$$

then using $\langle n^0|H^0 = \langle n^0|E_n^0$, and $\langle n^0|n^0\rangle = 1$, we find

$$E_n^1 = \langle n^0|H^1|n^0\rangle. \tag{9.1.10}$$

Taking now $\lambda = 1$, *we have established that the first-order change in the energy of a state resulting from adding a perturbing term* H^1 *to the Hamiltonian is just the expectation value of* H^1 *in that state.*

For example, we can estimate the ground state energy of the helium atom by treating the electrostatic repulsion between the electrons as a perturbation. The zeroth-order ground state has the two (opposite spin) electrons in the ground state hydrogen-atom wavefunction (scaled for the doubling of nuclear charge). The first-order energy correction E_0^1 is then given by computing the expectation value $\langle e^2/r_{12}\rangle$ for this ground state wavefunction.

First-Order of Eigenstates

The general expression for the first-order change in the *wavefunction* is found by taking the inner product of the first-order equation with the bra $\langle m^0|, \quad m \neq n$,

$$\langle m^0|H^0|n^1\rangle + \langle m^0|H^1|n^0\rangle = \langle m^0|E_n^0|n^1\rangle + \langle m^0|E_n^1|n^0\rangle. \tag{9.1.11}$$

The last term is zero, since $\langle m^0|n^0\rangle = 0$, and in the first term $\langle m^0|H^0 = \langle m^0|E_m^0$, so

$$\langle m^0|n^1\rangle = \frac{\langle m^0|H^1|n^0\rangle}{E_n^0 - E_m^0} \tag{9.1.12}$$

and therefore the wavefunction correct to first order is:

$$\begin{aligned} |n\rangle &= |n^0\rangle + |n^1\rangle && (9.1.13) \\ &= |n^0\rangle + \sum_{m\neq n} \frac{|m^0\rangle\langle m^0|H^1|n^0\rangle}{E_n^0 - E_m^0}. && (9.1.14) \end{aligned}$$

Second-Order of Energy

To find the *second*-order correction to the energy, it is necessary to match the second-order terms in

$$\begin{gathered} (H^0 + \lambda H^1)(|n^0\rangle + \lambda|n^1\rangle + \lambda^2|n^2\rangle + \ldots) \\ = (E_n^0 + \lambda E_n^1 + \lambda^2 E_n^2 + \ldots)(|n^0\rangle + \lambda|n^1\rangle + \lambda^2|n^2\rangle + \ldots) \end{gathered} \tag{9.1.15}$$

giving:

$$H^0|n^2\rangle + H^1|n^1\rangle = E_n^0|n^2\rangle + E_n^1|n^1\rangle + E_n^2|n^0\rangle. \tag{9.1.16}$$

Taking the inner product with $\langle n^0|$ yields:

$$\langle n^0|H^0|n^2\rangle + \langle n^0|H^1|n^1\rangle = E_n^0\langle n^0|n^2\rangle + E_n^1\langle n^0|n^1\rangle + E_n^2\langle n^0|n^0\rangle. \tag{9.1.17}$$

The leading terms on the two sides cancel as before. What about the term $E_n^1\langle n^0|n^1\rangle$? Since $|n\rangle = |n^0\rangle + |n^1\rangle$, and both $|n\rangle$, $|n^0\rangle$ are normalized, $\langle n^0|n^1\rangle + \langle n^1|n^0\rangle = 0$ in leading order—that is to say, $\langle n^0|n^1\rangle$ is pure imaginary. That just means that if to this order $|n\rangle$ has a component parallel to $|n^0\rangle$, that component has a small pure imaginary amplitude, and $|n\rangle$ can be written (to this order) as $|n\rangle = e^{i\alpha}|n^0\rangle + kets \perp |n^0\rangle$, with α small. But the phase factor can be eliminated by redefining the phase of $|n\rangle$, and with that redefinition $|n^1\rangle$ has *no* component in the $|n^0\rangle$ direction, we can therefore drop the term $E_n^1\langle n^0|n^1\rangle$.

So the second-order correction to the energy is:

$$E_n^2 = \langle n^0|H^1|n^1\rangle \tag{9.1.18}$$

$$= \langle n^0|H^1 \sum_{m\neq n} \frac{|m^0\rangle\langle m^0|H^1|n^0\rangle}{E_n^0 - E_m^0} \tag{9.1.19}$$

$$= \sum_{m\neq n} \frac{|\langle m^0|H^1|n^0\rangle|^2}{E_n^0 - E_m^0}. \tag{9.1.20}$$

Selection Rules

Perturbation theory involves evaluating matrix elements of operators. Very often, many of the matrix elements in a sum are zero—obvious tests are parity and the Wigner-Eckart theorem. These are examples of *selection rules*: tests to find if a matrix element may be nonzero.

The Quadratic Stark Effect

When a hydrogen atom in its ground state is placed in an electric field, the electron cloud and the proton are pulled different ways, an electric dipole forms, and the overall energy is lowered. The perturbing Hamiltonian from the electric field is

$$H^1 = e\,\mathscr{E}\,z = e\mathscr{E}\,r\cos\theta \tag{9.1.21}$$

where $\mathscr{E}$ is the electric field strength, the field is in the z- direction, the electron charge e is negative.

We shall denote the unperturbed eigenenergies of the hydrogen atom by $E_n = E_{nlm} = -1/n^2$, so in particular we *denote the ground state energy by* E_1**.**

The first-order correction to the ground state energy $E_1^1 = \langle 100|e\mathscr{E}z|100\rangle$, where

$$|100\rangle \equiv \psi_{100}(r) = \left(\frac{1}{\pi a_0^3}\right)^{1/2} e^{-r/a_0}. \tag{9.1.22}$$

This first-order term is zero since there are equal contributions from positive and negative z.

The second-order term is

$$E_1^2 = \sum_{n\neq 1;l,m} \frac{|\langle nlm|e\mathscr{E}z|100\rangle|^2}{E_1 - E_n} \tag{9.1.23}$$

where we are now using $|nlm\rangle$ to denote the unperturbed hydrogen atom wavefunctions, and here the $E_n = -1/n^2$ (in Rydbergs) are the unperturbed energies.

Most of the terms in this infinite series are zero—the selection rules help get rid of them as follows: since $e\mathscr{E}z$ is the $m=0$ component of a spherical vector and $|100\rangle$ is a zero angular momentum state, it follows from the Wigner-Eckart theorem that $\langle nlm|$ can only be $\langle n10|$. This reduces the second-order sum over states to:

$$E_1^2 = \sum_{n\neq 1} \frac{|\langle n^10|e\mathscr{E}z|100\rangle|^2}{E_1 - E_n}. \tag{9.1.24}$$

This is still not easy to evaluate, but an *upper bound* can be found by observing that $|E_1 - E_n| \geq |E_1 - E_2|$, so

$$\begin{aligned} |E_0^2| &< \tfrac{1}{E_2-E_1}\textstyle\sum_{n\neq1}|\langle n10|e\mathscr{E}z|100\rangle|^2 \\ &= \tfrac{1}{E_2-E_1}\textstyle\sum_{n\neq1;l,m}\langle 100|e\mathscr{E}z|nlm\rangle\langle nlm|e\mathscr{E}z|100\rangle \end{aligned} \tag{9.1.25}$$

where we have temporarily *restored* the full sum over n,l,m that is, we've put back all the zero terms. The reason for this seeming backward step is that, having taken the energy-difference denominator outside the sum, we can even include $|100\rangle$ in the $|nlm\rangle$ sum (it's another zero term) and in fact we can even include the plane-wave (ionized) states as well as the bound states, since the plane waves all have energy greater than zero. At this point, the $\sum_{n,l,m}$ sum becomes a sum over all states, and therefore just becomes the unit operator,

$$\sum_{n,l,m}|nlm\rangle\langle nlm| = I, \tag{9.1.26}$$

so

$$|E_1^2| < \frac{1}{E_2-E_1}\langle 100|(e\mathscr{E}z)^2|100\rangle. \tag{9.1.27}$$

For the ground state hydrogen wavefunction, $\langle 100|z^2|100\rangle = a_0^2,\ E_1 = -e^2/2a_0,\quad E_2 = E_1/4$, so

$$|E_1^2| < \frac{1}{(\frac{3}{4}e^2/2a_0)}(e\mathscr{E})^2a_0^2 = \frac{8}{3}\mathscr{E}^2a_0^3. \tag{9.1.28}$$

Furthermore, since all the terms in the series for E_1^2 are negative, *the first term sets a lower bound on* $|E_1^2|$:

$$|E_1^2| > \frac{|\langle 210|e\mathscr{E}z|100\rangle|^2}{E_1-E_2}. \tag{9.1.29}$$

This can be evaluated in straightforward fashion to find $|E_1^2| > 0.55\times\frac{8}{3}\mathscr{E}^2a_0^3$.

So, even though we have not actually evaluated the second-order correction to the energy explicitly, we have it bracketed between two values, the lower one being more than half the upper one. Other ingenious methods have been developed (see Shankar or Sakurai) to find that the true answer is $|E_1^2| = \frac{9}{4}\mathscr{E}^2a_0^3$, but in fact the whole problem can be solved exactly using parabolic coordinates.

Degenerate Perturbation Theory: Distorted 2-D Harmonic Oscillator

The above analysis works fine as long as the successive terms in the perturbation theory form a convergent series. A necessary condition is that the matrix elements of the perturbing Hamiltonian must be smaller than the corresponding energy level differences of the original Hamiltonian. If H^0 has different states with the same energy, in other words degenerate energy levels, and the perturbation has nonzero matrix elements *between these degenerate levels*, then obviously the theory breaks down. To see just how it breaks down, and how to fix it, we consider the two-dimensional simple harmonic oscillator:

$$H^0 = \frac{p_x^2+p_y^2}{2m}+\frac{1}{2}m\omega^2(x^2+y^2). \tag{9.1.30}$$

Recall that for the *one*-dimensional simple harmonic oscillator the ground state wavefunction is

$$|0\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}e^{-m\omega x^2/2\hbar} = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}e^{-\xi^2/2}\ \ with\ \ \xi=\sqrt{\frac{m\omega}{\hbar}}x,\ \ and\ \ |1\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\sqrt{2}\xi e^{-\xi^2/2}. \tag{9.1.31}$$

The two-dimensional oscillator is simply a product of two one-dimensional oscillators, so, writing $\eta = \sqrt{\frac{m\omega}{\hbar}}y$, the ground state is $|0\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/2}e^{-(\xi^2+\eta^2)/2}$, and the two (degenerate) next states, energy $\hbar\omega$ above the ground state, are

$$|1,0\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/2}\sqrt{2}\xi e^{-(\xi^2+\eta^2)/2},\ \ |0,1\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/2}\sqrt{2}\eta e^{-(\xi^2+\eta^2)/2}. \tag{9.1.32}$$

Suppose now we add a small perturbation

$$H^1 = \alpha m\omega^2 xy, \tag{9.1.33}$$

with α a small parameter.

Notice that

$$\langle 0|H^1|0\rangle = \langle 1,0|H^1|1,0\rangle = \langle 0,1|H^1|0,1\rangle = 0 \tag{9.1.34}$$

so according to naïve perturbation theory, there is *no* first-order correction to the energies of these states.

However, on going to second-order in the energy correction, the theory breaks down. The matrix element $\langle 1,0|H^1|0,1\rangle$is nonzero, but the two states $|0,1\rangle,\quad |1,0\rangle$have the same energy! This gives an infinite term in the series for E_n^2.

Yet we know that a small term of this type will not wreck a two-dimensional simple harmonic oscillator, so what is wrong with our approach? It is helpful to plot the original harmonic oscillator potential $\frac{1}{2}m\omega^2(x^2+y^2)$ together with the perturbing potential $\alpha m\omega^2 xy$. The first of course has circular symmetry, the second has axes in the directions $x=\pm y$, climbing most steeply from the origin along $x=y$, falling most rapidly in the directions $x=-y$. If we combine the two potentials into a single quadratic form, the original circles of constant potential become ellipses, with their axes aligned along $x=\pm y$.

The problem arises even in the classical two-dimensional oscillator: picture a ball rolling backwards and forwards in a smooth saucer, a circular bowl. Now imagine the saucer is made slightly elliptical. The ball will still roll backwards and forwards through the center if it is released along one of the axes of the ellipse, although with different periods, as the axes differ in steepness. However, if it is released at a point *off* the axes, it will describe a complex path resolvable into components in the two axis directions having different periods.

For the quantum oscillator as for the classical one, as soon as the perturbation is introduced, the eigenkets are in the direction of the new elliptic axes. This is a large change from the original x and y axes, and definitely *not* proportional to the small parameter α. But the original unperturbed problem had circular symmetry, and there was no particular reason to choose the x and y axes as we did. If we had instead chosen as our original axes the lines $x=\pm y$, the kets would *not* have undergone large changes on switching on the perturbation.

The resolution of the problem is now clear: *before* switching on the perturbation, *choose a set of basis kets in a degenerate subspace such that the perturbation is diagonal in that subspace*.

In fact, for the simple harmonic oscillator example above, the problem can be solved exactly:

$$\begin{gathered}\frac{1}{2}m\omega^2(x^2+y^2)+\alpha m\omega^2 xy \\ =\frac{1}{2}m\omega^2\left[(1+\alpha)(\frac{x+y}{\sqrt{2}})^2+(1-\alpha)(\frac{x-y}{\sqrt{2}})^2\right]\end{gathered} \tag{9.1.35}$$

and it is clear that, despite the results of naïve first-order theory, there is indeed *a first order shift* in the energy levels,

$$\hbar\omega \rightarrow \hbar\omega\sqrt{1\pm\alpha} \approx \hbar\omega(1\pm\alpha/2). \tag{9.1.36}$$

The Linear Stark Effect

The hydrogen atom, like the two-dimensional harmonic oscillator discussed above, has a nondegenerate ground state but degeneracy in its lowest excited states. Specifically, there are four $n=2$ states, all having energy -1/4 Ryd :

$$\begin{gathered}\psi_{200}(r)=\left(\frac{1}{32\pi a_0^3}\right)^{1/2}\left(2-\frac{r}{a_0}\right)e^{-r/2a_0}, \\ \psi_{210}(r,\theta,\phi)=\left(\frac{1}{32\pi a_0^3}\right)^{1/2}\left(\frac{r}{a_0}\right)e^{-r/2a_0}cos\theta, \\ \psi_{21\pm1}(r,\theta,\phi)=\left(\frac{1}{32\pi a_0^3}\right)^{1/2}\left(\frac{r}{a_0}\right)e^{-r/2a_0}sin\theta e^{\pm i\phi}.\end{gathered} \tag{9.1.37}$$

Perturbing this system with an electric field in the z- direction, $H^1=e\mathscr{E}z=e\mathscr{E}rcos\theta$, note first that naïve perturbation theory predicts *no* first-order shift in any of these energy levels. However, to second order, there is a nonzero matrix element between two degenerate levels $\langle 200|H^1|210\rangle$ All the other matrix elements between these basis kets in the four-dimensional degenerate subspace are zero, so the only diagonalization necessary is within the*two*-dimensional degenerate subspace spanned by $|200\rangle$, $|210\rangle$, where

$$H^1=\begin{pmatrix}0 & \Delta \\ \Delta & 0\end{pmatrix} \tag{9.1.38}$$

with

$$\Delta = \langle 200|H^1|210\rangle \tag{9.1.39}$$

$$= e\mathscr{E}\left(\frac{1}{32\pi a_0^3}\right)\int_0^\infty \left(2-\frac{r}{a_0}\right)\left(\frac{rcos\theta}{a_0}\right)^2 e^{-r/a_0} r^2 dr\sin\theta d\theta d\phi \tag{9.1.40}$$

$$= -3e\mathscr{E}a_0. \tag{9.1.41}$$

Diagonalizing H^1 within this subspace, then, the new basis states are $(|200\rangle \pm |210\rangle)/\sqrt{2}$ with energy shifts $\pm\Delta$, linear in the perturbing electric field.

The states $|21\pm1\rangle$ are not changed by the presence of the field to this approximation, so the complete energy map of the $n=2$ states in the electric field has two states at the original energy of -1/4Ryd, one state moved up from that energy by Δ, and one down by Δ.

Notice that the new eigenstates $(|200\rangle \pm |210\rangle)/\sqrt{2}$ are *not* eigenstates of the parity operator—a sketch of their wavefunctions reveals that in fact they have nonvanishing electric dipole moment $\vec{\mu}$, indeed this is the reason for the energy shift, $\pm\Delta = \mp 3e\mathscr{E}a_0 = \mp\vec{\mu}\cdot\vec{\mathscr{E}}$.

9.2: The Peierls Transition - an Unexpected Insulator

Introduction: Looking for Superconductors and Finding Insulators

In 1964, Little suggested (Phys. Rev 134, A1416) that it might be possible to synthesize a room temperature superconductor using organic materials in which the electrons traveled along certain kinds of chains, effectively confined to one dimension. The first satisfactory theory of "ordinary" superconductivity, that of Bardeen, Cooper, and Schrieffer (BCS) had appeared a few years earlier, in 1957. The key point was that electrons became bound together in opposite spin pairs, and at sufficiently low temperatures these bound pairs, being boson like, formed a coherent condensate—all the pairs had the same total momentum, so all traveled together, a supercurrent. The locking of the electrons into this condensate effectively eliminated the usual single-electron scattering by impurities that degrades ordinary currents in conductors.

But what could bind the electrostatically repelling electrons? The answer turned out to be lattice distortions, as first suggested by Fröhlich in 1950. An electron traveling through the crystal attracts the positive ions, the consequent excess of local positive charge attracts another electron. The strength of this binding, and hence the temperature at which the superconducting transition takes place, depends on the rapidity of the lattice response. This was confirmed by the isotope effect: lattice response time obviously depends on the inertia of the lattice, the BCS theory predicted that for a superconducting element with different isotopic varieties, the ratio of the superconducting transition temperatures for pure isotopes T_2/T_1 was equal to $\sqrt{M_1/M_2}$, M_1, M_2 being the ion masses, the lighter isotope having the higher transition temperature. This was indeed the case.

Little's idea was that the build up of positive charge by a passing electron could be speeded up dramatically if instead of having to move ions, it need only rearrange other electrons. Unfortunately, there were no obvious three-dimensional candidate materials. However, if the conduction electrons moved along a one-dimensional chain, polarizable side chains might be attached, and rearrangement of the electronic charge distribution in these side chains would respond very rapidly to a passing conduction electron, building up a local positive charge. If this worked, order of magnitude arguments suggested possible enhancement of the transition temperature by a factor $\sqrt{M/m}$ over ordinary superconductors, m being the electron mass.

In the 1970's, various organic materials were synthesized and tested, beginning with one called TTF-TCNQ, in which a set of polymer-like long molecules donated electrons to another set, leaving one-dimensional conductors with partially filled bands (see later), seemingly good candidates for superconductivity. Unfortunately, on cooling, these materials surprisingly became *insulators* rather than superconductors! This was the first example of a *Peierls transition*, a widespread phenomenon in quasi one-dimensional systems.

The basic mechanism of the Peierls transition can be understood with a simple model. It is a nice example of applied second-order perturbation theory, including the degenerate case. We examine the model and the result below. It should be added that in some newer materials the Peierls transition is (unexpectedly) suppressed under high pressure, and superconductivity has in fact been observed in organic salts, but so far only at transition temperatures around one Kelvin: Little's dream is not yet realized.

Second-Order Perturbation Theory: a Periodic Potential in One Dimension

To understand how a one-dimensional conductor might turn into an insulator at low temperatures, we must first become familiar with the simplest model of a one-dimensional conductor:

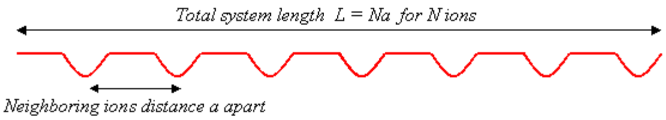
$$H = H^0 + V = \frac{p^2}{2m} + V(x) \tag{9.2.1}$$

with H^0 a gas of noninteracting electrons on a line, and V periodic, that is, $V(x+a) = V(x)$,

the potential from a line of ions spaced a apart. We'll take the system to have N ions in a total length L, so

$$L = Na \tag{9.2.2}$$

and to keep the math simple, we'll require periodic boundary conditions.



Ionic potential seen by electrons in one-dimensional system.

The physics here is that *without* the potential, the electron eigenstates are *plane waves*. The effect of the lattice potential is to partially reflect the waves, like a diffraction grating, generating components at different wavelengths. This effect becomes particularly important when the electron wavelength matches twice the ion spacing. For that case, the reflected and original waves have the same strength, the electron is at a standstill. We'll explore just how this happens later.

The eigenstates of H^0 are then

$$|k\rangle^{(0)} = \frac{1}{\sqrt{L}} e^{ikx}, \quad \textit{with} \quad e^{ikL} = 1, \quad \textit{so} \quad k = \frac{2\pi n}{L}, \tag{9.2.3}$$

n being an integer. The unperturbed energy eigenvalues,

$$H^0 |k\rangle^{(0)} = E^0 |k\rangle^{(0)}, \ \textit{are just} \ E^0 = \frac{\hbar^2 k^2}{2m}. \tag{9.2.4}$$

This is to be understood as

$$H^0 |k_n\rangle^{(0)} = E^0_n |k_n\rangle^{(0)} \tag{9.2.5}$$

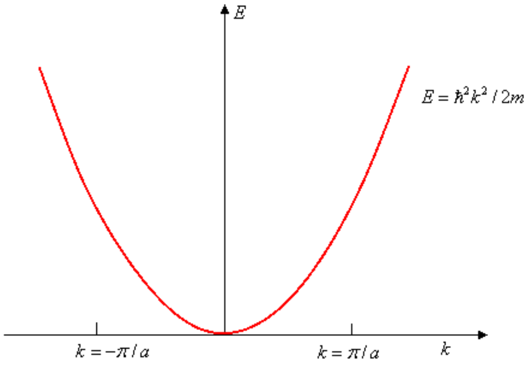
$$E^0_n = \frac{\hbar^2 k^2_n}{2m} \tag{9.2.6}$$

and

$$k_n = 2\pi n L. \tag{9.2.7}$$

We are following standard practice here. We shall also write $\sum_k f(k)$ meaning $\sum_n f(k_n)$.)

It's worth plotting the (E, k) curve:



Energy momentum (E, k) curve for a free electron in one dimension.

Suppose we have ions with two electrons each to contribute to this one-dimensional (supposed) conductor. Assuming they move into these plane wave states, in the system ground state they will fill up the lowest energy states up to a maximum k- value denoted by $\pm k_F$ (F stands for Fermi, this is the *Fermi* momentum.) Where is it?

We know there will be a total of $2N$ electrons. We also know that the allowed values of k, from the boundary conditions, are $k_n = 2\pi n/L$, with n an integer. In other words, the allowed k 's are uniformly spaced $2\pi/L$ apart, meaning they have a *density* of $L/2\pi$ in k- space, so the total number between $\pm k_F$ is Lk_F/π. The $2N$ electrons will have N of each spin, each k- state can take two electrons (one of each spin), so $Lk_F/\pi = N = L/a$, and

$$k_F = \pi/a. \tag{9.2.8}$$

To do perturbation theory, we must find the matrix elements of $V(x)$ between eigenstates of H^0:

$$^{(0)}\langle k'|V|k\rangle^{(0)} = 1L\int ei(k-k')xV(x)dx. \tag{9.2.9}$$

This is just the Fourier component $V_{k-k'}$ of $V(x)$.

If $V(x)$ is periodic with period a,

$$V_k \neq 0 \;\; only\, if \;\; k = nK, \;\; n\; an\; integer, \;\; K = 2\pi/a. \tag{9.2.10}$$

In other words, if a function is periodic with spatial period a, the only nonzero Fourier components are those having the same spatial period a.

Therefore

$$V(x) = \sum_n V_{nK}e^{inKx}, \;\; and \;\; V_{-nK} = V^*_{nK} \;\; since\; V(x)\; is\; real; \;\; K = 2\pi/a. \tag{9.2.11}$$

The $n=0$ component of $V(x)$ is of no interest—it is just a constant potential, and so can be taken to be zero. Note that this eliminates the trivial first order correction $E^1_k = ^{(0)}\langle k|V|k\rangle^{(0)}$ to the energy eigenvalues.

We shall consider *only* the components $n=+1$ and $n=-1$ of $V(x)$, it turns out that the other components can be treated in similar fashion. For $n=+1$, $n=-1$, the potential only has nonzero matrix elements between the plane wave state k and $k+K$, $k-K$ respectively.

So, the second order correction to the energy is:

$$\begin{aligned} E^2_k &= \textstyle\sum_{k'\neq k} \frac{|^{(0)}\langle k|V|k'\rangle^{(0)}|^2}{E^0_k - E^0_{k'}} \\ &= \frac{|^{(0)}\langle k|V|k+K\rangle^{(0)}|^2}{E^0_k - E^0_{k+K}} + \frac{|^{(0)}\langle k|V|k-K\rangle^{(0)}|^2}{E^0_k - E^0_{k-K}} \\ &= \frac{|V_K|^2}{E^0_k - E^0_{k+K}} + \frac{|V_{-K}|^2}{E^0_k - E^0_{k-K}}. \end{aligned} \tag{9.2.12}$$

This result is reasonable provided the terms are small, that is, the energy differences appearing in the denominators are large compared to the relevant Fourier component V_K. However, this cannot always be true! Notice that the state $k = \pi/a$ has exactly the same unperturbed energy E^0 as the state $k-K = -\pi/a$: in this case, nondegenerate perturbation theory is clearly wrong. In fact, even for states close to $k = \pi/a$, the energy denominator $E^0_k - E^0_{k-K}$ is small compared with the numerator $|V_{-K}|^2$, so the series is not converging.

Quasi-Degenerate Perturbation Theory near the Critical Wavelength

The good news is that, despite the many states near $k = \pi/a$ and $k = -\pi/a$ that are close together in energy, for any *one* state k near π/a the potential only has a nonzero matrix element to *one* other state close in energy, the state $k-K$, that is, $k - 2\pi/a$.

The strategy now is to do what might be called *quasidegenerate* perturbation theory: to diagonalize the full Hamiltonian in the subspace spanned by these two states $|k\rangle^{(0)}, \quad |k-K\rangle^{(0)}$. Other states with nonzero matrix elements to these states are relatively much further away in energy, and can be treated using ordinary perturbation theory.

The matrix elements of the full Hamiltonian in the subspace spanned by these two states are:

$$\begin{vmatrix} E^0_k & V^*_K \\ V^*_K & E^0_{k-K} \end{vmatrix}. \tag{9.2.13}$$

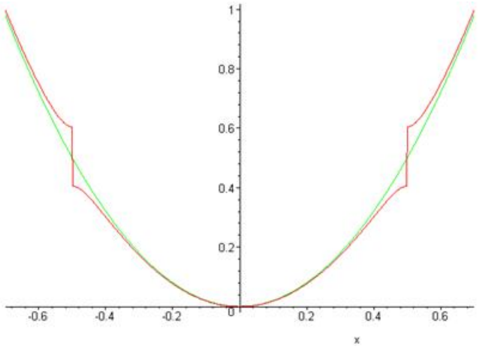
Diagonalizing *within this subspace* gives energy eigenvalues:

$$E_\pm = \frac{1}{2}(E^0_k + E^0_{k-K}) \pm \sqrt{\left(\frac{E^0_k - E^0_{k-K}}{2}\right)^2 + |V_K|^2}. \tag{9.2.14}$$

Notice that, provided $|E_k^0 - E_{k-K}^0| \gg |V_K|$, to leading order this gives back $E_\pm = E_k^0, \quad E_{k-K}^0$, the order depending on k. However, as k approaches π/a, $|E_k^0 - E_{k-K}^0|$ becomes of order $|V_K|$, and the energies deviate from the unperturbed values. If k is approaching π/a from below, $E_k^0 < E_{k-K}^0$, and the lower energy is pushed *downwards* by the perturbation: $E_k = E_- < E_k^0$. This is a common occurrence with almost degenerate states, perturbations cause the energy levels to "repel" each other.

For $k = \pi/a, \; E_{k-K}^0 = E_{k-2\pi/a}^0 = E_k^0$. At this value of k, the unperturbed states are exactly degenerate, and the perturbation lifts the degeneracy to give $E_\pm = E_{\pi/a}^0 \pm |V_K|$.

In the graph below, the green (continuous) curve is the unperturbed energy as a function of k, the red curve (with the step) the calculated energy including the leading correction from the periodic potential.



Energy Gaps and Bands

The energy jump, or gap, of $2|V_K|$ at $|k| = \pi/a$ means that there are no plane wave type eigenstates with energies in that range—attempting to integrate Schrödinger's equation in the periodic potential for such an energy gives exponentially growing and decaying solutions. Such energy gaps in fact are present in real crystalline solids, the allowed energies are said to be in "bands". The lowest band for our model is from $k = -\pi/a$ to π/a. Since the allowed values of k are given by $k = 2\pi n/L$, the spacing between adjacent k's is $2\pi/L$ and the total number of k's in the lowest band is $L/a = N$, the same as the number of atoms. Since each electron has two spin states, this implies that a one-dimensional crystal of *divalent* atoms will *just fill* the lowest band with electrons. Therefore, any outside field can only excite an electron to a different state if an energy of at least $2|V_K|$ is supplied—for a small electric field, the filled band of electrons will remain in the ground state, there will be no current. This material is an insulator.

On the other hand, if *monovalent* atoms are used, it is clear that the lowest band is *only half full*, adjacent empty electron states are available. The electrons are free to accelerate if an external field is applied. Barring the unexpected, this one-dimensional crystal would be a metal.

Let us now examine how the periodic potential alters the eigenstates. Ignoring the small corrections from plane waves outside the $|k\rangle^{(0)}, \quad |k-K\rangle^{(0)}$ subspace, the eigenstates to this order have the form

$$|k\rangle = a_k |k\rangle^{(0)} + a_{k-K} |k-K\rangle^{(0)} \tag{9.2.15}$$

where

$$\frac{a_{k-K}}{a_k} = \frac{E_- - E_k^0}{V_K^*} \tag{9.2.16}$$

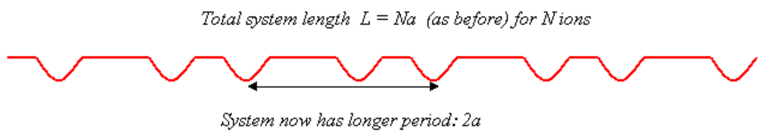
from the diagonalization of the 2×2 matrix representing the Hamiltonian in the subspace.

As k increases from 0 towards π/a, the plane wave initially proportional to e^{ikx} has a gradually increasing admixture of $e^{i(k-2\pi/a)x}$, until at $k = \pi/a$ the two have equal weight—meaning that the eigenfunction is now a standing wave. In fact, there are *two* standing wave solutions at $k = \pi/a$, corresponding to the energies below and above the gap. Taking the atoms to have an attractive potential, the lower energy wave has a probability distribution peaking at the atomic positions. The diffractive scattering that gives a left-moving component to a right moving wave is known as Bragg scattering. It also manifests itself in the *group velocity* of the electronic excitations, $v_{group} = d\omega/dk = (1/\hbar) dE/dk$. An electron injected into a one-dimensional metal would not be a plane wave state, but a wavepacket traveling at the group velocity. It is evident that for an injected electron with mean value of k close to π/a, the electron will move very slowly into the metal. This is to be expected—the eigenstates become standing waves as $k \rightarrow \pi/a$.

For three-dimensional crystals, the situation is far more complicated, but many of the same ideas are relevant. Electron waves are now diffracted by whole planes of atoms, and the three-dimensional momentum space is divided into Brillouin zones, with planes having an energy gap across them.

The Peierls Transition: how Cooling a Conductor Can Give an Insulator

As mentioned in the Introduction, substances very close to monovalent one-dimensional crystals have been synthesized, and it has been found—surprisingly—that at low temperatures many of them undergo a transition from metallic to *insulating* behavior. *What happens is that the atoms in the lattice rearrange slightly, moving from an equally-spaced crystal to one in which the spacing alternates, that is, the atoms form pairs*. This is called *dimerization*, and costs some elastic energy, since for identical atoms the lowest state must be one of equal spacing for any reasonable potential. However, the *electrons* are able to move to a lower energy state by this maneuver.

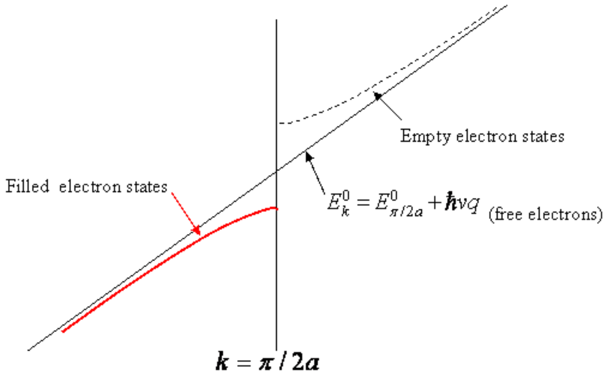


Ionic potential seen by electrons in a dimerized one-dimensional system.

Just how this happens can be understood using the perturbation theory analysis above. For equally spaced atoms, the electrons *half-fill* the band, that is, they fill it up (two electrons, one of each spin, per state) to $|k| = \pi/2a$.

The crucial point is that if the atoms move together slightly into pairs, *the crystal has a new period $2a$ instead of a*. This means that the *potential* now has a nonzero component at $K = -\pi/a$, with a nonzero matrix element between the states $k = \pi/2a$ and $k = -\pi/2a$, and so on. From this point, we can rerun the analysis above, except that now the gaps open up at $|k| = \pi/2a$ instead of at $|k| = \pi/a$.

The important point is that if the electrons fill all the states to $|k| = \pi/2a$, and none beyond (as would be the case for monovalent atoms) then the opening of a gap at $|k| = \pi/2a$ means that *all the electrons are in states whose energy is lowered*. To find the *total* energy benefit we need to integrate over k.



Change in electronic energy levels near $k = \pi/2a$ from dimerization: in this region, the free electron curve is approximated with a straight line: $q = k - \pi/2a$.

Calculating the Electronic Energy Gained by Doubling the Lattice Period

It is evident from the above that most of the contribution comes from fairly close to $k = \pi/2a$ (and of course symmetrically $k = -\pi/2a$). Since we want to find the total lowering in energy, let us study first the *bare* energy as a function of k, that is, the energy with no potential present. Of course, there isn't much to say: $E^0_k = \hbar^2 k^2/2m$. However, the physics of these one-dimensional systems concerns only excitations near the "Fermi surface", the boundary between filled (low energy) states at zero temperature and empty states. This “Fermi surface” is in fact just two points in one dimension: $k=\pm\pi/2a$. In the neighborhood of these two Fermi points, it is an excellent approximation to replace the gently curving $E^0_k=\hbar^2k^2/2m$ by *straight line approximations*—the slope being $dE/dk=\hbar^2k/m=\hbar p/m=\hbar v$.

Linearizing in the neighborhood of $k=\pi/2a$, then, we take

$$E^0_k=E^0_{\pi/2a}+\hbar v(k-\pi/2a)=E^0_{\pi/2a}+\hbar vq,\tag{9.2.17}$$

where

$$q=k-\pi/2a,\tag{9.2.18}$$

just k measured from the Fermi point $\pi/2a$.

The variable q is *negative* for the relevant states, since they are on the lower energy side.

The density of states in k-space is a constant $2\times L/2\pi=L/\pi$, remembering the two spin states per k- value.

Recall

$$E_\pm=\frac{1}{2}(E^0_k+E^0_{k-K})\pm\sqrt{\left(\frac{E^0_k-E^0_{k-K}}{2}\right)^2+|V_K|^2}\tag{9.2.19}$$

but now

$$K=\pi/a\tag{9.2.20}$$

and the lowering of energy of the electrons (counting it as a positive quantity) is:

$$2\int_0^{\pi/2a}(E^0_k-E_-)Ldk/\pi=2\int\left(\frac{1}{2}(E^0_k-E^0_{k-K})+\sqrt{\left(\frac{E^0_k-E^0_{k-K}}{2}\right)^2+|V_K|^2}\right)Ldk/\pi\tag{9.2.21}$$

where the extra factor of 2 counts the symmetrical contribution from the left-hand gap. (In examining the above expression, recall that for the $k>0$ states we are interested in, $k<\pi/2a$, $E^0_k-E^0_{k-K}$ is *negative*. The integrand on the right-hand side is still positive, very small for small k, reaching a maximum of $|V_K|$ at $k=\pi/2a$.)

Putting in our linearized energy approximation,

$$E^0_k=E^0_{\pi/2a}+\hbar v(k-\pi/2a)=E^0_{\pi/2a}+\hbar vq,\tag{9.2.22}$$

and remembering that now $K=\pi/a$,

$$E^0_{k-K}=E^0_{-\pi/2a}-\hbar v(k+\pi/2a)=E^0_{-\pi/2a}-\hbar vq.\tag{9.2.23}$$

Since $E^0_{\pi/2a}=E^0_{-\pi/2a}$,

$$(E^0_k-E^0_{k-K})=2v\hbar q.\tag{9.2.24}$$

Substituting these linearized values in the integral for the total energy lowering:

$$2\int_0^{\pi/2a}(E^0_k-E_-)Ldk/\pi=2\int_{-D}^{0}(v\hbar q+\sqrt{(v\hbar q)^2+|V_K|^2})Ldq/\pi\tag{9.2.25}$$

where in terms of the variable q we have set the lower limit of integration at $-D$: we can safely be vague about this lower limit, as the integral turns out to be logarithmic.

Since the integral is over negative numbers, and we have taken the positive square root, it is zero for zero V_K, as it must be.

The integral can be done exactly, but it is more illuminating to divide the range of integration into $|v\hbar q|\le|V_K|$ and $|v\hbar q|>|V_K|$, then estimate the contributions from these two ranges separately.

First, consider $|v\hbar q|\le|V_K|$. Here the integrand is of order $|V_K|$, and the region Δq of integration corresponding to $|v\hbar q|\le|V_K|$ is of order $|V_K|/\hbar v$, so the integral over this range is of order $(L/\hbar v)|V_K|^2$.

Second, in the region $|v\hbar q|>|V_K|$, we can write

$$2\int\left(v\hbar q+\sqrt{(v\hbar q)^2+|V_K|^2}\right)\ Ldq/\pi=2\int\left(v\hbar q+|v\hbar q|\sqrt{1+\frac{|V_K|^2}{(v\hbar q)^2}}\right)\ Ldq/\pi \tag{9.2.26}$$

and expand the square root term. The leading terms cancel since q is negative, and the main contribution comes from the next term. This gives:

$$\Delta E\approx 2|V_K|^2\int_{-D}^{-|V_K|}\frac{1}{2\hbar v}\frac{Ldq}{\pi|q|}=\frac{L|V_K|^2}{\hbar v}\ln\frac{|V_K|}{D}. \tag{9.2.27}$$

The important thing here is the logarithm. For sufficiently small $|V_K|$, this large (negative) term will dominate any term which is just proportional to V_K^2. But the elastic energy cost of the lattice "dimerizing"—the atoms forming pairs, so that the distance between atoms alternates on going along the chain—must be proportional to V_K^2. This leads to the conclusion that some, probably small, dimerization is always going to happen—*a one-dimensional equally spaced chain with one electron per ion is unstable*.

This dimerization is known as a *Peierls transition*. Peierls discovered it in the 1930's when writing a section on one-dimensional models in an introductory solid-state textbook. He put it in the book, but didn't publish it otherwise. As mentioned in the Introduction, it became very relevant later when some theories suggested that quasi-one-dimensional conductors, materials made up of loosely connected chains, each chain having one electron per atom for a half-filled lowest band, might be high-temperature superconductors. It was found instead that many such materials actually became *insulators* on cooling: the reason was that at high temperatures, the electrons filled states above and below the point $\pi/2a$ fairly equally, so dimerization did not lower the overall energy much. On lowering the temperature, a point was reached where the Peierls transition gave a lower energy state, and the material became an insulator.

9.3: Van Der Waals Forces between Atoms

Introduction

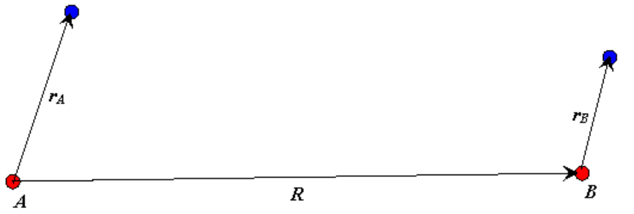
The perfect gas equation of state $PV = NkT$ is manifestly incapable of describing actual gases at low temperatures, since they undergo a discontinuous change of volume and become liquids. In the 1870's, the Dutch physicist Van der Waals came up with an improvement: a gas law that recognized the molecules interacted with each other. He put in two parameters to mimic this interaction. The first, an attractive intermolecular force at long distances, helps draw the gas together and therefore reduces the necessary outside pressure to contain the gas in a given volume—the gas is a little thinner near the walls. The attractive long range force can be represented by a negative potential $-aN/V$ on going away from the walls—the molecules near the walls are attracted inwards, those in the bulk are attracted equally in all directions, so effectively the long range attraction is equivalent to a potential well extending throughout the volume, ending close to the walls. Consequently, the gas density N/V near the walls is decreased by a factor $e^{-E/kT} = e^{-aN/VkT} \cong 1 - aN/VkT$. Therefore, the pressure measured at the containing wall is from slightly diluted gas, so $P = (N/V)kT$ becomes $P = (N/V)(1 - aN/VkT)kT$, or $(P + a(N/V)^2)V = NkT$. The second parameter van der Waals added was to take account of the finite molecular volume. A real gas cannot be compressed indefinitely—it becomes a liquid, for all practical purposes incompressible. He represented this by replacing the volume V with $V - Nb$, Nb is referred to as the "excluded volume", roughly speaking the volume of the molecules. Putting in these two terms gives his famous equation

$$\left[P + a\left(\frac{N}{V}\right)^2\right](V - Nb) = NkT. \tag{9.3.1}$$

This rather crude approximation does in fact give sets of isotherms representing the basic physics of a phase transition quite well. (For further details, and an enlightening discussion, see for example Appendix D of *Thermal Physics*, by R. Baierlein.)

Ground State Hydrogen Atoms

Our interest here is in understanding the van der Waals long-range attractive force between electrically neutral atoms and molecules in quantum mechanical terms. We begin with the simplest possible example, two hydrogen atoms, both in the ground state:



We label the atoms A and B, the vectors from the protons to the electron position are denoted by $\overrightarrow{r_A}$ and $\overrightarrow{r_B}$ respectively, and $\vec{R}$ is the vector from proton A to proton B.

Then the Hamiltonian $H = H^0 + V$, where

$$H^0 = -\frac{\hbar^2}{2m}(\nabla_A^2 + \nabla_B^2) - \frac{e^2}{r_A} - \frac{e^2}{r_B} \tag{9.3.2}$$

and the electrostatic interaction between the two atoms

$$V = \frac{e^2}{R} + \frac{e^2}{|\vec{R} + \overrightarrow{r_B} - \overrightarrow{r_A}|} - \frac{e^2}{|\vec{R} + \overrightarrow{r_B}|} - \frac{e^2}{|\vec{R} - \overrightarrow{r_A}|} \tag{9.3.3}$$

The ground state of H^0 is just the product of the ground states of the atoms A, B, that is,

$$|0\rangle = |100\rangle_A \otimes |100\rangle_B. \tag{9.3.4}$$

Assuming now that the distance between the two atoms is much greater than their size, we can expand the interaction V in the small parameters $r_A/R, \quad r_B/R$. As one might suspect from the diagram above, the leading order terms in the electrostatic energy are just those of a dipole-dipole interaction:

$$V=-e^2(\overrightarrow{r_A}\cdot\vec{\nabla})(\overrightarrow{r_B}\cdot\vec{\nabla})\frac{1}{R}=e^2\left[\frac{\overrightarrow{r_A}\cdot\overrightarrow{r_B}}{R^3}-\frac{3(\overrightarrow{r_A}\cdot\vec{R})(\overrightarrow{r_B}\cdot\vec{R})}{R^5}\right] \tag{9.3.5}$$

Taking now the z- axis in the direction $\vec{R}$, this interaction energy is

$$V=\frac{e^2}{R^3}(x_Ax_B+y_Ay_B-2z_AZ_B) \tag{9.3.6}$$

Now the first-order correction to the ground state energy of the two-atom system from this interaction is $E_n^1=\langle n^0|H^1|n^0\rangle$, where here $H^1=V$ and $|n^0\rangle=|100\rangle_A\otimes|100\rangle_B$. Beginning with the first term x_Ax_B in V

$$({}_A\langle100|\otimes_B\langle100|)(x_Ax_B)(|100\rangle_A\otimes|100\rangle_B)=({}_A\langle100|x_A|100\rangle_A)({}_B\langle100|x_B|100\rangle_B) \tag{9.3.7}$$

is clearly zero since the ground states are spherically symmetric. Similarly, the other terms in V are zero to first order.

Recall that the second-order energy correction is $E_n^2=\sum_{m\neq n}\frac{|\langle m0|H^1|n^0\rangle|^2}{E_n^0-E_m^0}$.

That is,

$$E^{(2)}=\sum_{\substack{n,l,m\\n',l',m'}}\frac{|({}_A\langle nlm|\otimes_B\langle n'l'm'|)V(|100\rangle_A\otimes|100\rangle_B)|^2}{2E_1-E_n-E_{n'}}. \tag{9.3.8}$$

A typical term here is

$$({}_A\langle nlm|\otimes_B\langle n'l'm'|)(x_Ax_B)(|100\rangle_A\otimes|100\rangle_B)=({}_A\langle nlm|x_A|100\rangle_A)({}_B\langle n'l'm'|x_B|100\rangle_B), \tag{9.3.9}$$

so the single-atom matrix elements are exactly those we discussed for the Stark effect (as we would expect—this is an electrostatic interaction!). As before, only $l=1,\quad l'=1$ contribute. To make a rough estimate of the size of $E^{(2)}$, we can use the same trick used for the quadratic Stark effect: replace the denominators by the constant $2E_1$ (the other terms are a lot smaller for the bound states, and continuum states have small overlap terms in the numerator). The sum over intermediate states n,l,m,n',l',m'can then be taken to be completely unrestricted, including even the ground state, giving

$$\sum_{\substack{n,l,m\\n',l',m'}}(|nlm\rangle_A\otimes|n'l'm'\rangle_B)({}_A\langle nlm|\otimes_B\langle n'l'm'|)=I, \tag{9.3.10}$$

the identity operator. In this approximation, then, just as for the Stark effect,

$$E^{(2)}\simeq\frac{e^4}{R^6}\frac{1}{2E_1}({}_A\langle100|\otimes_B\langle100|)(x_Ax_B+y_Ay_B-2z_AZ_B)^2(|100\rangle_A\otimes|100\rangle_B) \tag{9.3.11}$$

where $E_1=-1$ Ryd., so this is a *lowering* of energy.

In multiplying out $(x_Ax_B+y_Ay_B-2z_AZ_B)^2$, the cross terms will have expectation values of zero. The ground state wave function is symmetrical, so all we need is $\langle100|x^2|100\rangle=a_0^2$, where a_0 is the Bohr radius.

This gives

$$E^{(2)}\simeq\frac{e^4}{R^6}\frac{1}{2E_1}6a_0^4\simeq-6\frac{e^2}{R}\left(\frac{a_0}{R}\right)^5 \tag{9.3.12}$$

using $E_1=-e^2/2a_0$. Bear in mind that this is an approximation, but a pretty good one—a more accurate calculation replaces the 6 by 6.5.

Forces between a 1*s* Hydrogen Atom and a 2*p* Hydrogen Atom

With one atom in the $|100\rangle$ and the other in $|210\rangle$, say, a typical leading order term would be

$$({}_A\langle100|\otimes_B\langle210|)(x_Ax_B)(|100\rangle_A\otimes|100\rangle_B)=({}_A\langle100|x_A|100\rangle_A)({}_B\langle210|x_B|100\rangle_B), \tag{9.3.13}$$

and this is certainly zero, as are all the other leading terms. Baym (*Lectures on Quantum Mechanics*) concluded from this that there is *no* leading order energy correction between two hydrogen atoms if one of them is in the ground state. This is incorrect: the first excited state of the two-atom system (without interaction) is degenerate, so, exactly as for the 2-D simple harmonic oscillator treated in the previous lecture, we must diagonalize the perturbation in the subspace of these degenerate first excited states. (For this section, we follow fairly closely the excellent treatment in *Quantum Mechanics*, by C. Cohen-Tannoudji et al.)

The space of the degenerate first excited states of the two noninteracting atoms is spanned by the product-space kets:

$$\begin{array}{llll} (|100\rangle_A \otimes |200\rangle_B), & (|200\rangle_A \otimes |100\rangle_B), & (|100\rangle_A \otimes |211\rangle_B), & (|211\rangle_A \otimes |100\rangle_B), \\ (|100\rangle_A \otimes |210\rangle_B), & (|210\rangle_A \otimes |100\rangle_B), & (|100\rangle_A \otimes |21-1\rangle_B), & (|21-1\rangle_A \otimes |100\rangle_B). \end{array} \tag{9.3.14}$$

The task, then, is to diagonalize $V = \frac{e^2}{R^3}(x_A x_B + y_A y_B - 2z_A Z_B)$ in this eight-dimensional subspace.

We begin by representing V as an 8×8 matrix using these states as the basis. First, note that all the diagonal elements of the matrix are zero—in all of them, we're finding the average of x,y or z for one of the atoms in the ground state. Second, writing $V = \frac{e^2}{R^3}(\vec{r_A} \cdot \vec{r_B} - 3z_A Z_B)$, it is evident that V is unchanged if the system is rotated around the z- axis (the line joining the two protons). This means that the commutator $[V, L_z] = 0$, where L_z is the total angular momentum component in the z- direction, so V will only have nonzero matrix elements between states having the same total L_z. Third, from parity (or Wigner-Eckart) all matrix elements in the subspace spanned by $(|100\rangle_A \otimes |200\rangle_B), \ (|200\rangle_A \otimes |100\rangle_B)$ must be zero.

This reduces the nonzero part of the 8×8 matrix to a direct product of three 2×2 matrices, corresponding to the three values of $L_z = m$. For example, the $m = 0$ subspace is spanned by $(|100\rangle_A \otimes |210\rangle_B), \ (|210\rangle_A \otimes |100\rangle_B)$. The diagonal elements of the 2×2 matrix are zero, the off-diagonal elements are equal to $-2\frac{e^2}{R^3}({}_A\langle 100|z_A|210\rangle_A)({}_B\langle 210|Z_B|100\rangle_B)$, where we have kept the unnecessary labels A, B to make clear where this term comes from. (The x_A and y_A terms will not contribute for $m = 0$.)

This is now a straightforward integral over hydrogen wave functions. The three 2×2 matrices have the form

$$\begin{pmatrix} 0 & k_m/R^3 \\ k_m/R^3 & 0 \end{pmatrix} \tag{9.3.15}$$

(following the notation of Cohen-Tannoudji) where $k_m \sim e^2 a_0^2$, and the energy eigenvalues are $\pm k_m/R^3$, with corresponding eigenkets $(1/\sqrt{2})[(|100\rangle_A \otimes |210\rangle_B) \pm (|210\rangle_A \otimes |100\rangle_B)]$.

So for two hydrogen atoms, one in the ground state and one in the first excited state, the van der Waal interaction energy goes as $1/R^3$, much more important than the $1/R^6$ energy for two hydrogen atoms in the ground state. Notice also that the $1/R^3$ can be *positive or negative*, depending on whether the atoms are in an even or an odd state—so the atoms sometimes repel each other.

Finally, if two atoms are initially in a state $(|100\rangle_A \otimes |210\rangle_B)$, note that this is *not* an eigenstate of the Hamiltonian when the interaction is included. Writing the state as a sum of the even and odd states, which have slightly different phase frequencies from the energy difference, we find the excitation moves back and forth between the two atoms with a period $hR^3/2k_{m=0}$.

9.4: The Interaction Representation

Recall that in the first part of this course sequence, we discussed the Schrödinger and Heisenberg representations of quantum mechanics here. In the Schrödinger representation, the operators are time-independent (except for explicitly time-dependent potentials) the kets representing the quantum states develop in time. In the Heisenberg representation, the kets stay the same, the time dependence is in the operators. These differing representations describe the same physics—matrix elements of operators between kets must be the same in both. The most natural to use depends on the problem at hand. In the classical limit, for example, the Heisenberg operators have the time dependence of the corresponding classical operators.

In fact, for perturbation theory problems with a time-dependent potential, an intermediate representation, the *interaction representation*, is very convenient. Using a subscript S to denote the Schrödinger representation,

$$i\hbar\frac{d}{dt}|\psi_S(t)\rangle = H_S|\psi_S(t)\rangle = (H_S^0 + V_S(t))|\psi_S(t)\rangle, \tag{9.4.1}$$

we define the *interaction representation* by the unitary transformation

$$|\psi_I(t)\rangle \;= e^{iH_S^0 t/\hbar}|\psi_S(t)\rangle \tag{9.4.2}$$

so the interaction representation kets and the Schrödinger representation kets coincide at $t=0$, and if the interaction were zero, the interaction representation kets would be constant in time, like those in the Heisenberg representation.

For nonzero $V(t)$, then, the time development of the interaction representation kets is entirely due to $V(t)$, and is easily found by differentiating both sides of the equation:

$$\begin{aligned} i\hbar\tfrac{d}{dt}|\psi_I(t)\rangle &= -H^0|\psi_I(t)\rangle + e^{iH_S^0 t/\hbar} i\hbar\tfrac{d}{dt}|\psi_S(t)\rangle \\ &= -H^0|\psi_I(t)\rangle + e^{iH_S^0 t/\hbar}(H_S^0 + V_S(t))e^{-iH_S^0 t/\hbar}|\psi_I(t)\rangle \\ &= e^{iH_S^0 t/\hbar}V_S(t)e^{-iH_S^0 t/\hbar}|\psi_I(t)\rangle \\ &= V_I(t)|\psi_I(t)\rangle, \end{aligned} \tag{9.4.3}$$

where we have introduced the interaction representation operator $V_I(t)$, defined by

$$V_I(t) = e^{iH_S^0 t/\hbar}V_S(t)e^{-iH_S^0 t/\hbar}. \tag{9.4.4}$$

Operators in this representation must have this time dependence relative to the Schrödinger operators to ensure that matrix elements, the only quantities of physical significance, are the same in the two representations. That is to say, we must have

$$\langle f_I^0|O_I|i_I^0\rangle = \langle f_S^0|O_S|i_S^0\rangle, \tag{9.4.5}$$

the two representations must predict the same probability amplitude for any transition. Integrating both sides of the differential equation,

$$|\psi_I(t)\rangle \;= |\psi_I(0)\rangle - \frac{i}{\hbar}\int_0^t dt'V_I(t')|\psi_I(t')\rangle\ . \tag{9.4.6}$$

This is not a solution—we've just gone from a differential equation to an integral equation. This is only worth doing if V_I is small, in which case the integral equation can be solved iteratively.

The zeroth approximation is then

$$|\psi_I(t)\rangle \;= |\psi_I(0)\rangle. \tag{9.4.7}$$

Putting this value into the small term on the right hand side of the integral equation gives the first order solution,

$$|\psi_I(t)\rangle \;= |\psi_I(t_0)\rangle - \frac{i}{\hbar}\int_0^t dt'V_I(t')|\psi_I(0)\rangle\ . \tag{9.4.8}$$

The second order solution is now given by putting the first order solution into the integral on the right:

$$|\psi_I(t)\rangle \;= |\psi_I(0)\rangle - \frac{i}{\hbar}\int_0^t dt'V_I(t')\left(|\psi_I(0)\rangle - \frac{i}{\hbar}\int_0^{t'} dt''V_I(t'')|\psi_I(0)\rangle\right)\ . \tag{9.4.9}$$

The T symbol means that on expanding out the exponential, the operators at different times are arranged in order of time, the latest on the left, without worrying about commutators. If we just blindly expand the exponential, we will get, for example, a third-order term

$$T\frac{1}{3!}\left(-\frac{i}{\hbar}\int_0^t dt'V_I(t')\right)\left(-\frac{i}{\hbar}\int_0^t dt''V_I(t'')\right)\left(-\frac{i}{\hbar}\int_0^t dt'''V_I(t''')\right). \tag{9.4.10}$$

The T operator tells us to rearrange the $V_I(t)$'s in chronological order. Since there are three of them, they clearly appear in all possible orders before T operates, that is to say, there are 3! different ordered terms that T makes the same. This just nicely cancels the 3! in the exponential expansion, to give us the expression we found by iteration.

This time-ordered exponential is therefore the *interaction representation propagator*:

$$|\psi_I(t)\rangle = U_I(t,0)|\psi_I(0)\rangle,\;\; U_I(t,0) = T\exp\left(-i\hbar\int_0^t dt'V_I(t')\right). \tag{9.4.11}$$

9.5: Time-Dependent Perturbation Theory

Introduction: General Formalism

We look at a Hamiltonian with $V(t)$ some time-dependent perturbation

$$H = H^0 + V(t) \tag{9.5.1}$$

so now the wavefunction will have perturbation-induced time dependence. Our starting point is the set of eigenstates $|n\rangle$ of the unperturbed Hamiltonian $H^0|n\rangle = E_n|n\rangle$, notice we are not labeling with a zero, no E_n^0, because with a time-dependent Hamiltonian, energy will not be conserved, so it is pointless to look for energy corrections. What happens instead, provided the perturbation is not too large, is that the system makes transitions between the eigenstates $|n\rangle$ of H^0.

Of course, even for $V = 0$, the wavefunctions have the usual time dependence,

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_nt/\hbar}|n\rangle \tag{9.5.2}$$

with the c_n 's constant. What happens on introducing $V(t)$ is that the c_n's *themselves* acquire time dependence,

$$|\psi(t)\rangle = \sum_n c_n(t) e^{-iE_nt/\hbar}|n\rangle \tag{9.5.3}$$

and this time dependence is determined by Schrödinger's equation with the Hamiltonian in Equation 9.5.1

$$i\hbar\frac{\partial}{\partial t}\sum_n c_n(t)e^{-iE_nt/\hbar}|n\rangle = (H^0 + V(t))\sum_n c_n(t)e^{-iE_nt/\hbar}|n\rangle \tag{9.5.4}$$

so

$$i\hbar\sum_n \dot{c_n}(t)e^{-iE_nt/\hbar}|n\rangle = V(t)\sum_n c_n(t)e^{-iE_nt/\hbar}|n\rangle \tag{9.5.5}$$

Taking the inner product with the bra $\langle m|e^{iE_mt/\hbar}$, and introducing $\omega_{mn} = \dfrac{E_m - E_n}{\hbar}$,

$$i\hbar\dot{c}_m = \sum_n \langle m|V(t)|n\rangle c_n e^{i\omega_{mn}t} = \sum_n V_{mn}e^{i\omega_{mn}t}c_n \tag{9.5.6}$$

This is a matrix differential equation for the c_n's :

$$i\hbar\begin{pmatrix}\dot{c}_1\\ \dot{c}_2\\ \dot{c}_3\\ .\\ .\end{pmatrix} = \begin{pmatrix} V_{11} & V_{12}e^{i\omega_{12}t} & . & . & .\\ V_{21}e^{i\omega_{12}t} & V_{22} & . & . & .\\ . & . & V_{33} & . & .\\ . & . & . & . & .\\ . & . & . & . & . \end{pmatrix}\begin{pmatrix}c_1\\ c_2\\ c_3\\ .\\ .\end{pmatrix} \tag{9.5.7}$$

and solving this set of coupled equations will give us the $c_n(t)$'s, and hence the probability of finding the system in any particular state at any later time.

If the system is in initial state $|i\rangle$ at $t = 0$, the probability amplitude for it being in state $|f\rangle$ at time t is *to leading order* in the perturbation

$$c_f(t) = \delta_{fi} - \frac{i}{\hbar}\int_0^t V_{fi}(t')e^{i\omega_{fi}t'}dt'. \tag{9.5.8}$$

The probability that the system is in fact in state $|f\rangle$ at time t is therefore

$$|c_f(t)|^2 = \frac{1}{\hbar^2}\left|\int_0^t V_{fi}(t')e^{i\omega_{fi}t'}dt'\right|^2. \tag{9.5.9}$$

Obviously, this is only going to be a good approximation if it predicts that the probability of transition is small—otherwise we need to go to higher order, using the Interaction Representation (or an exact solution like that in the next section).

Example 9.5.1: Kicking an Oscillator

Suppose a simple harmonic oscillator is in its ground state $|0\rangle$ at $t=-\infty$. It is perturbed by a small time-dependent potential $V(t)=-eExe^{-t^2/\tau^2}$. What is the probability of finding it in the first excited state $|1\rangle$ at $t=+\infty$?

Solution

Here

$$V_{fi}(t')=-eE\langle 1|x|0\rangle e^{-t'^2/\tau^2} \tag{9.5.10}$$

and

$$x=\sqrt{\hbar/2m\omega}(a+a^\dagger) \tag{9.5.11}$$

from which the probability can be evaluated. It is

$$(e^2E^2/\hbar^2)(\hbar/2m\omega)\pi\tau^2e^{-\omega^2\tau^2/2}. \tag{9.5.12}$$

It’s worth thinking through the physical interpretations for very long and for very short times, and explaining the significance of the time for which the probability is a maximum.

The Two-State System: an Exact Solution

For the particular case of a two-state system perturbed by a periodic external field, the matrix equation above can be solved exactly. Of course, real physical systems have more than two states, but in fact for some important cases two of the states may be strongly coupled to each other, but only weakly coupled to other states, and the analysis then becomes relevant. A famous example, the ammonia maser, is discussed at the end of the section.

For a two-state system, then, the most general wavefunction is

$$|\psi(t)\rangle=c_1(t)e^{-iE_1t/\hbar}|1\rangle+c_2(t)e^{-iE2t/\hbar}|2\rangle \tag{9.5.13}$$

and the differential equation for the $c_n(t)$’s is:

$$i\hbar\begin{pmatrix}\dot{c}_1\\ \dot{c}_2\end{pmatrix}=\begin{pmatrix}0 & Ve^{i\omega t}e^{i\omega_{12}t}\\ Ve^{-i\omega t}e^{-i\omega_{12}t} & 0\end{pmatrix}\begin{pmatrix}c_1\\ c_2\end{pmatrix}. \tag{9.5.14}$$

Writing $\omega+\omega_{12}=\alpha$ for convenience, the coupled equations are:

$$\begin{aligned}i\hbar\dot{c}_1&=Ve^{i\alpha t}c_2\\ i\hbar\dot{c}_2&=Ve^{-i\alpha t}c_1.\end{aligned} \tag{9.5.15}$$

These two first-order equations can be transformed into a single second-order equation by differentiating the second one, then substituting $\dot{c}_1$ from the first one and c_1 from the second one to give

$$\ddot{c}_2=-i\alpha\,\dot{c}_2-\frac{V^2}{\hbar^2}c_2. \tag{9.5.16}$$

This is a standard second-order differential equation, solved by putting in a trial solution $c_2(t)=c_2(0)e^{i\Omega t}$. This satisfies the equation if

$$\Omega=-\frac{\alpha}{2}\pm\sqrt{\frac{\alpha^2}{4}+\frac{V^2}{\hbar^2}}, \tag{9.5.17}$$

so, reverting to the original $\omega+\omega_{12}=\alpha$, the general solution is:

$$c_2(t)=e^{-i\frac{(\omega-\omega_{21})}{2}t}\left(Ae^{i\sqrt{\left(\frac{\omega-\omega_{21}}{2}\right)^2+\frac{V^2}{\hbar^2}}\,t}+Be^{-i\sqrt{\left(\frac{\omega-\omega_{21}}{2}\right)^2+\frac{V^2}{\hbar^2}}\,t}\right). \tag{9.5.18}$$

Taking the initial state to be $c_1(0)=1,\quad c_2(0)=0$ gives $A=-B$.

To fix the overall constant, note that at $t=0$,

$$\dot{c}_2(0) \;=\; \frac{V}{i\hbar}c_1(0) \;=\frac{V}{i\hbar}. \tag{9.5.19}$$

Therefore

$$|c_2(t)|^2=\frac{\frac{V^2}{\hbar^2}}{\left(\frac{\omega-\omega_{21}}{2}\right)^2+\frac{V^2}{\hbar^2}}\sin^2\left(\sqrt{\left(\frac{\omega-\omega_{21}}{2}\right)^2+\frac{V^2}{\hbar^2}}\;t\right). \tag{9.5.20}$$

Note in particular the result if $\omega=\omega_{12}$:

$$|c_2(t)|^2=\sin^2\left(\frac{Vt}{\hbar}\right). \tag{9.5.21}$$

Assuming $E_2>E_1$, and the two-state system to be initially in the ground state $|1\rangle$, this means that after a time $h/4V$ the system will *certainly* be in state $|2\rangle$, and will oscillate back and forth between the two states with period $h/2V$.

That is to say, a precisely timed period spent in an oscillating field can drive a collection of molecules all in the ground state to be all in an excited state. The ammonia *maser* works by sending a stream of ammonia molecules, traveling at known velocity, down a tube having an oscillating field for a definite length, so the molecules emerging at the other end are all (or almost all, depending on the precision of ingoing velocity, etc.) in the first excited state. Application of a small amount of electromagnetic radiation of the same frequency to the outgoing molecules will cause some to decay, generating intense radiation and therefore a much shorter period for all to decay, emitting coherent radiation.

A “Sudden” Perturbation

A sudden perturbation is defined here as a sudden switch from one time-independent Hamiltonian H_0 to another one H'_0, the time of switching being much shorter than any natural period of the system. In this case, perturbation theory is irrelevant: if the system is initially in an eigenstate $|n\rangle$ of H_0, one simply has to write it as a sum over the eigenstates of H'_0, $|n\rangle=\sum_{n'}|n'\rangle\langle n'|n\rangle$. The nontrivial part of the problem is in establishing that the change *is* sudden enough, by estimating the actual time taken for the Hamiltonian to change, and the periods of motion associated with the state $|n\rangle$ and with its transitions to neighboring states.

Harmonic Perturbations: Fermi’s Golden Rule

Let us consider a system in an initial state $|i\rangle$ perturbed by a periodic potential $V(t)=Ve^{-i\omega t}$ switched on at $t=0$. For example, this could be an atom perturbed by an external oscillating electric field, such as an incident light wave.

What is the probability that at a later time t the system be in state $|f\rangle$?

Recall the matrix differential equation for the c_n’s (Equation 9.5.7)

$$i\hbar\begin{pmatrix}\dot{c}_1\\ \dot{c}_2\\ \dot{c}_3\\ .\\ .\end{pmatrix}=\begin{pmatrix}V_{11} & V_{12}e^{i\omega_{12}t} & . & . & .\\ V_{21}e^{i\omega_{12}t} & V_{22} & . & . & .\\ . & . & V_{33} & . & .\\ . & . & . & . & .\\ . & . & . & . & .\end{pmatrix}\begin{pmatrix}c_1\\ c_2\\ c_3\\ .\\ .\end{pmatrix}$$

Since the system is definitely in state $|i\rangle$ at $t=0$, the ket vector on the right is initially $c_i=1,\quad c_{j\neq i}=0$.

The first-order approximation to keep the vector $c_i=1,\quad c_{j\neq i}=0$ on the right, that is, to solve the equations

$$i\hbar\dot{c}_f(t)=V_{fi}e^{i\omega_{fi}t}. \tag{9.5.22}$$

Integrating this equation, the probability amplitude for an atom in initial state $|i\rangle$ to be in state $|f\rangle$ after time t is, to first order:

$$c_f(t) = -\frac{i}{\hbar}\int_0^t \langle f|\ V|i\rangle e^{i(\omega_{fi}-\omega)t'}dt' \tag{9.5.23}$$

$$= -\frac{i}{\hbar}\langle f|V|i\rangle \frac{e^{i(\omega_{fi}-\omega)t}-1}{i(\omega_{fi}-\omega)}. \tag{9.5.24}$$

The probability of transition is therefore

$$P_{i\to f}(t) = |c_f|^2 \tag{9.5.25}$$

$$= \frac{1}{\hbar^2}|\langle f|V|i\rangle|^2\left(\frac{\sin((\omega_{fi}-\omega)t/2)}{(\omega_{fi}-\omega)/2}\right)^2 \tag{9.5.26}$$

and we are interested in the large t limit.

Writing $\alpha = (\omega_{fi}-\omega)/2$, our function has the form $\frac{\sin^2\alpha t}{\alpha^2}$. This function has a peak at $\alpha = 0$, with maximum value t^2, and width of order $1/t$, so a total weight of order t. The function has more peaks at $\alpha t = (n+1/2)\pi$. These are bounded by the denominator at $1/\alpha^2$. For large t their contribution comes from a range of order $1/t$ also, and as $t\to\infty$ the function tends to a δ function at the origin, but multiplied by t.

This divergence is telling us that there is a finite probability *rate* for the transition, so the likelihood of transition is proportional to time elapsed. Therefore, we should divide by t to get the transition rate.

To get the quantitative result, we need to evaluate the weight of the δ function term. We use the standard result

$$\int_{-\infty}^{\infty}\left(\frac{\sin\xi}{\xi}\right)^2 d\xi = \pi \tag{9.5.27}$$

to find

$$\int_{-\infty}^{\infty}\left(\frac{\sin\alpha t}{\alpha}\right)^2 d\alpha = \pi t \tag{9.5.28}$$

and therefore

$$\lim_{t\to\infty}\frac{1}{t}\left(\frac{\sin\alpha t}{\alpha}\right)^2 = \pi\delta(\alpha). \tag{9.5.29}$$

Now, the transition rate is the probability of transition divided by t in the large t limit, that is,

$$R_{i\to f}(t) = \lim_{t\to\infty}\frac{P_{i\to f}(t)}{t} \tag{9.5.30}$$

$$= \lim_{t\to\infty}\frac{1}{t}\frac{1}{\hbar^2}|\langle f|V|i\rangle|^2\left[\frac{\sin((\omega_{fi}-\omega)t/2)}{(\omega_{fi}-\omega)/2}\right] \tag{9.5.31}$$

$$= \frac{1}{\hbar^2}|\langle f|V|i\rangle|^2\pi\delta(\frac{1}{2}(\omega_{fi}-\omega)) \tag{9.5.32}$$

$$= \frac{2\pi}{\hbar^2}|\langle f|V|i\rangle|^2\delta(\omega_{fi}-\omega) \tag{9.5.33}$$

This last line is Fermi’s Golden Rule: we shall be using it a lot. You might worry that in the long time limit we have taken the probability of transition is in fact diverging, so how can we use first order perturbation theory? The point is that for a transition with $\omega_{fi}\neq\omega$, “long time” means $(\omega_{fi}-\omega)t\gg 1$, this can still be a very short time compared with the mean transition time, which depends on the matrix element. In fact, Fermi’s Rule agrees extremely well with experiment when applied to atomic systems.

Another Derivation of the Golden Rule

Actually, when light falls on an atom, the full periodic potential is not suddenly switched on, on an atomic time scale, but builds up over many cycles (of the atom and of the light). Baym re-derives the Golden Rule assuming the limit of a very slow switch on,

$$V(t)=e^{\varepsilon t}Ve^{-i\omega t} \tag{9.5.34}$$

with ε very small, so V switched on very gradually in the past, and we are looking at times much smaller than $1/\varepsilon$. We can then take the initial time to be $-\infty$, that is,

$$c_f(t)=-\frac{i}{\hbar}\int_{-\infty}^{t}\langle f|\ V|i\rangle e^{i(\omega_{fi}-\omega-i\varepsilon)t'}dt'=-\frac{1}{\hbar}\frac{e^{i(\omega_{fi}-\omega-i\varepsilon)t}}{\omega_{fi}-\omega-i\varepsilon}\langle f|V|i\rangle \tag{9.5.35}$$

so

$$|c_f(t)|^2=\frac{1}{\hbar^2}\frac{e^{2\varepsilon t}}{(\omega_{fi}-\omega)^2+\varepsilon^2}|\langle f|V|i\rangle|^2 \tag{9.5.36}$$

and the time rate of change

$$\frac{d}{dt}|c_f(t)|^2=\frac{1}{\hbar^2}\frac{2\varepsilon e^{2\varepsilon t}}{(\omega_{fi}-\omega)^2+\varepsilon^2}|\langle f|V|i\rangle|^2. \tag{9.5.37}$$

In the limit $\varepsilon\to 0$, the function

$$\frac{2\varepsilon}{(\omega_{fi}-\omega)^2+\varepsilon^2}\to 2\pi\delta(\omega_{fi}-\omega) \tag{9.5.38}$$

giving the Golden Rule again (Equation 9.5.33).

Harmonic Perturbations: Second-Order Transitions

Sometimes the first order matrix element $\langle f|V|i\rangle$ is identically zero (parity, Wigner-Eckart, etc.) but other matrix elements are nonzero—and the transition can be accomplished by an indirect route. In the notes on the interaction representation, we derived the probability amplitude for the second-order process,

$$c_n^{(2)}(t)=\left(\frac{1}{i\hbar}\right)^2\sum_n\int_0^t\int_0^{t'}dt'dt''e^{-i\omega_f(t-t')}\langle f|V_S(t')|n\rangle e^{-i\omega_n(t'-t'')}\langle n|V_S(t'')|i\rangle e^{-i\omega_i t''}, \tag{9.5.39}$$

Taking the gradually switched-on harmonic perturbation

$$V_S(t)=e^{\varepsilon t}Ve^{-i\omega t} \tag{9.5.40}$$

and the initial time $-\infty$, as above,

$$c_n^{(2)}(t)=\left(\frac{1}{i\hbar}\right)^2\sum_n\langle f|V|n\rangle\langle n|V|i\rangle e^{-i\omega_f t}\int_{-\infty}^{t}dt'\int_{-\infty}^{t'}dt''e^{i(\omega_f-\omega_n-\omega-i\varepsilon)t'}e^{i(\omega_n-\omega_i-\omega-i\varepsilon)t''}. \tag{9.5.41}$$

Exactly as in the first-order Golden Rule, we can find the transition rate:

$$\frac{d}{dt}|c_n^{(2)}(t)|^2=\frac{2\pi}{\hbar^4}\left|\sum_n\frac{\langle f|V|n\rangle\langle n|V|i\rangle}{\omega_n-\omega_i-\omega-i\varepsilon}\right|^2\delta(\omega_f-\omega_i-2\omega). \tag{9.5.42}$$

(The $\hbar^4$ in the denominator goes to $\hbar$ on replacing the frequencies ω with energies E, both in the denominator and the delta function, remember that if $E=\hbar\omega,\ \delta(\omega)=\hbar\delta(E)$.)

This is a transition in which the system gains energy $2\hbar\omega$ from the beam, in other words *two* photons are absorbed, the first taking the system to the intermediate energy ω_n, which is short-lived and therefore not well defined in energy—there is no energy conservation requirement into this state, only between initial and final states.

Of course, if an atom in an arbitrary state is exposed to monochromatic light, other second order processes in which two photons are emitted, or one is absorbed and one emitted (in either order) are also possible.

9.6: The Photoelectric Effect in Hydrogen

In the photoelectric effect, incoming light causes an atom to eject an electron. We consider the simplest possible scenario: that the atom is hydrogen in its ground state. The interesting question is: for an ingoing light wave of definite frequency and amplitude, what is the probability of ionization of a hydrogen atom in a given time? In other words, assuming we can use time-dependent perturbation theory, what is the ionization rate?

Formally, we know what to do. We must find the interaction Hamiltonian H^1, then use **Fermi's Golden Rule** for the transition rate with a periodic perturbation:

$$R_{i\to f}=\frac{2\pi}{\hbar}|\langle f|H^1|i\rangle|^2\delta(E_f-E_i-\hbar\omega) \tag{9.6.1}$$

But it's not that easy! For one thing, the outgoing electron will be in some kind of plane wave state, so whatever convention we adopt for normalizing such states appears in the rate. But also the δ function is tricky for excitation into the continuum: just how many of these plane wave states satisfy $E_f=E_i+\hbar\omega$? We shall discover that with a consistent formalism, these two difficulties cancel each other.

The Interaction Hamiltonian

Taking the incoming wave to be an electromagnetic field having vector potential

$$\vec{A}(\vec{r},t)=\vec{A}_0\cos(\vec{k}\cdot\vec{r}-\omega t) \tag{9.6.2}$$

The interaction Hamiltonian is given by replacing the electron kinetic energy term $\vec{p}^2/2m$ with $(\vec{p}-q\vec{A}/c)^2/2m$. The relevant new term is

$$-(1/2m)(q/c(\vec{p}\cdot\vec{A}+\vec{A}\cdot\vec{p}))=(e/mc)\vec{A}\cdot\vec{p} \tag{9.6.3}$$

since $q=-e$ and $\vec{\nabla}\cdot\vec{A}=0$ in our gauge.

Therefore

$$H^1=\left(\frac{e}{mc}\right)\cos(\vec{k}\cdot\vec{r}-\omega t)\vec{A}_0\cdot\vec{p}=\left(\frac{e}{2mc}\right)\left(e^{i(\vec{k}\cdot\vec{r}-\omega t)}+e^{-i(\vec{k}\cdot\vec{r}-\omega t)}\right)\vec{A}_0\cdot\vec{p}. \tag{9.6.4}$$

The two different terms in this expression, having time dependences $e^{-i\omega t}$ and $e^{i\omega t}$ will give δ functions $\delta(E_f-E_i-\hbar\omega)$ and $\delta(E_f-E_i+\hbar\omega)$ respectively in the transition rate. The $e^{-i\omega t}$ term therefore corresponds to absorption of a photon, since we are looking at a process in which the electron gains energy, $E_f>E_i$. The $e^{i\omega t}$ term is for the process where an atom in an excited state emits a photon into the beam and *drops* in energy.

So the relevant interaction Hamiltonian is

$$H^1(t)=H^1e^{-i\omega t}\quad where\quad H^1=\left(\frac{e}{2mc}\right)e^{i\vec{k}\cdot\vec{r}}\vec{A}_0\cdot\vec{p}. \tag{9.6.5}$$

Plane Waves: Density of States

We make the assumption that the final state is a plane wave state $|\vec{k}_f\rangle\propto e^{i\vec{k}_f\cdot\vec{r}}$.

The most straightforward way of handling the plane wave states is to confine the whole system to an extremely large cubical box of side L, and impose *periodic* boundary conditions (so that plane traveling wave states are allowed).

The big box has volume $V=L^3$, so the appropriately normalized plane wave states are

$$|\vec{k}\rangle=\frac{1}{L^{3/2}}e^{i\vec{k}\cdot\vec{r}}=\frac{1}{\sqrt{V}}e^{i\vec{k}\cdot\vec{r}} \tag{9.6.6}$$

As will become apparent, we need to count how thickly these states are distributed, both in momentum space (or k- space) and in energy. We'll begin by reviewing the one-dimensional problem—the three-dimensional case is a simple generalization.

Recall that for particles in *one* dimension confined to a line of length L with periodic boundary conditions, the allowed values of wave number k were given by $e^{ikL}=1$, so $k=2n\pi/L$ with n an integer. Thus considering only intervals $\Delta k\gg 2\pi/L$, the “density of states” in k is $L/2\pi$: an interval of length Δk contains $(L/2\pi)\Delta k=\rho(k)\Delta k$ states, where here $\rho(k)=L/2\pi$. The density of states in *energy*, $\rho(E)$, follows from differentiating $E=\hbar^2k^2/2m$. Writing $\Delta E=(\hbar^2k/m)\Delta k$, gives the incremental change ΔE in E for a given incremental change Δk in k, so the two intervals ΔE and Δk must contain the same number of states, that is, $\rho(E)\Delta E=\rho(k)\Delta k$. It then follows from $\rho(k)=L/2\pi$ that the one-dimensional density of states in energy

$$\rho_{1D}(E)=(L/2\pi)(m/\hbar^2k)=(L/2\pi\hbar)\sqrt{m/2E}. \tag{9.6.7}$$

Note this *one-dimensional* density of states goes to infinity as E goes to zero.

In *three* dimensions, with a cube of side L and periodic boundary conditions, the density of states in k- space is $(L/2\pi)^3$. The allowed states can be visualized as the points of a cubic lattice, $(k_x,k_y,k_z)=2\pi L(n_x,n_y,n_z)$, the n ’s being integers, so each allowed state has associated with it the volume of a small cube $(2\pi/L)^3$.

To find the three-dimensional density of states in energy, using $E=\hbar^2\vec{k}2/2m$, again $\Delta E=(\hbar^2k/m)\Delta k$ but now to find the number of states in a small energy range we must multiply by $4\pi k^2$, since the states in the energy range lie between two close concentric spheres in k-space. This gives

$$\rho(E)=(L/2\pi)^3 4\pi k^2(m/\hbar^2k)=(L/2\pi)^3 4\pi k(m/\hbar^2)=(V/2\pi^2)(m/\hbar^3)\sqrt{2mE}. \tag{9.6.8}$$

Notice that in contrast to the one-dimensional case, the three-dimensional density of states goes to *zero* at zero energy. (*Exercise*: What happens in two dimensions?)

(Of course, if we are detecting the ejected electron with apparatus restricted to a solid angle $d\Omega$, the 4π is replaced by $d\Omega$.)

The orthogonality condition between the plane wave states is

$$\langle\vec{k}|\vec{k}'\rangle=\delta_{\vec{k},\vec{k}'} \tag{9.6.9}$$

the ordinary Kronecker delta function—not Dirac’s—since the k ’s are an enumerated set,

$$(k_x,k_y,k_z)=\frac{2\pi}{L}(n_x,n_y,n_z) \tag{9.6.10}$$

the n’s being integers.

Finding the Matrix Element

The ground state wave function for hydrogen is

$$|100\rangle=\sqrt{\frac{1}{\pi a_0^3}}e^{-r/a_0}. \tag{9.6.11}$$

The matrix element entering Fermi’s Golden Rule (Equation 9.6.1) is therefore:

$$\langle\vec{k}_f|\left(\frac{e}{2mc}\right)e^{i\vec{k}\cdot\vec{r}}\vec{A}_0\cdot\vec{p}|100\rangle=\int d^3r(1/L)^{3/2}e^{-i\vec{k}_f\cdot\vec{r}}\left(\frac{e}{2mc}\right)e^{i\vec{k}\cdot\vec{r}}\vec{A}_0\cdot(-i\hbar\vec{\nabla})\sqrt{\frac{1}{\pi a_0^3}}e^{-r/a_0} \tag{9.6.12}$$

Actually the $e^{i\vec{k}\cdot\vec{r}}$ term is not very important—the wavelength of incoming photons for the usual photoelectric effect is far greater than the size of the hydrogen atom in its ground state (which our integral is limited to) so $e^{i\vec{k}\cdot\vec{r}}\cong 1$, and we can drop that term.

One point we’ve overlooked is that the electromagnetic wave has a *magnetic* field just as strong as the electric field, so what about the interaction of this magnetic field with the electron’s magnetic moment? This turns out to be much weaker than the $\left(\frac{e}{2mc}\right)\vec{A}_0\cdot\vec{p}$ term: the magnetic interaction

$$\vec{\mu}_B\cdot\vec{B}=\left(\frac{e}{2mc}\right)\vec{S}\cdot\vec{B}, \tag{9.6.13}$$

and the ratio of this magnetic contribution to the electric one is

$$\frac{\left(\frac{e}{2mc}\right)\vec{S}\cdot\vec{B}}{\left(\frac{e}{2mc}\right)\vec{A}_0\cdot\vec{p}} \simeq \frac{\hbar\vec{\sigma}\cdot\vec{\nabla}\times\vec{A}}{\vec{A}\cdot\vec{p}} \simeq \frac{\hbar k}{p} \tag{9.6.14}$$

with $p\sim\hbar/a_0$, so this ratio is of order a_0/λ, λ being the wavelength of the incoming light, around 100 nm to ionize hydrogen. So, we can safely ignore the magnetic interaction.

This interaction Hamiltonian $H^1=\left(\frac{e}{2mc}\right)\vec{A}_0\cdot\vec{p}e^{-i\omega t}$ is called the *dipole approximation*, because it can also be written in terms of the atom's dipole moment $e\vec{r}$. To see how this comes about, taking $|i\rangle$, $|f\rangle$ to be eigenstates of $H=\vec{p}^2/2m+V(\vec{r})$, and using $[\vec{r},\vec{p}]=i\hbar$, $[\vec{r},H]=(i\hbar/m)\vec{p}$, we find the matrix elements of $\vec{p},\vec{r}$ for this ionization transition are simply related

$$\begin{aligned}\langle f|\vec{p}|i\rangle &= (m/i\hbar)\langle f|\vec{r}H-H\vec{r}|i\rangle\\ &=(m/i\hbar)(E_i-E_f)\langle f|\vec{r}|i\rangle\\ &=im\omega\langle f|\vec{r}|i\rangle.\end{aligned} \tag{9.6.15}$$

Therefore

$$\langle f|H^1(t)|i\rangle=\left(\frac{e}{2mc}\right)\vec{A}_0e^{-i\omega t}\cdot\langle f|\vec{p}|i\rangle=\left(\frac{e}{2mc}\right)im\omega\vec{A}_0e^{-i\omega t}\cdot\langle f|\vec{r}|i\rangle, \tag{9.6.16}$$

and $\vec{E}=-(1/c)\partial\vec{A}/\partial t=(i\omega/2c)\vec{A}_0e^{-i\omega t}$, from which $\langle f|H^1(t)|i\rangle=\langle f|-\vec{\mu}\cdot\vec{E}(t)|i\rangle$, with $\vec{\mu}=-e\vec{r}$, the atom's electric dipole moment.

Nevertheless, for the particular interaction we are considering here, the $\vec{p}=-i\hbar\vec{\nabla}$ representation proves more convenient. (We'll use the $\vec{\mu}=-e\vec{r}$ representation in later work.)

We must evaluate:

$$(1/L)^{3/2}\left(\frac{e}{2mc}\right)\sqrt{\frac{1}{\pi a_0^3}}\int d^3re^{-i\vec{k}_f\cdot\vec{r}}\vec{A}_0\cdot(-i\hbar\vec{\nabla})e^{-r/a_0}. \tag{9.6.17}$$

Integration by parts gives the gradient operator acting on the plane wave state,

$$\int d^3re^{-i\vec{k}_f\cdot\vec{r}}\vec{A}_0\cdot(-i\hbar\vec{\nabla})e^{-r/a_0}=-(\vec{A}_0\cdot\vec{p}_f)\int d^3re^{-i\vec{k}_f\cdot\vec{r}}e^{-r/a_0}. \tag{9.6.18}$$

The integral is now a Fourier transform of the hydrogen ground state wave function, and is straightforward: choose the z- axis in the direction of $\vec{k}_f$, the φ -integration gives 2π, the θ -integration has $\sin\theta d\theta=-d(cos\theta)$, etc. The result is $(8\pi/a_0)/(a_0^{-2}+k_f^2)^2$.

Finally, we can put this into Fermi's Golden Rule:

$$\begin{gathered}R_{i\to f}=\\ \frac{2\pi}{\hbar}|(1/L)^{3/2}\left(\frac{e}{2mc}\right)\sqrt{\frac{1}{\pi a_0^3}}(\vec{A}_0\cdot\vec{p}_f)\left(\frac{8\pi/a_0}{(a_0^{-2}+k_f^2)^2}\right)|^2\delta(E_f-E_i-\hbar\omega).\end{gathered} \tag{9.6.19}$$

To detect the ejected electron, we will have a detector sensitive to some small solid angle, $d\Omega$, not to some precise value of $\vec{p}_f$. There will also be some tiny uncertainty in $|\vec{p}_f|$, equivalent to an energy uncertainty, because for one thing the ejection takes place after a finite time. This means the δ -function actually has finite width, and by taking our normalizing box big enough, there will be many states within this width—so, effectively, the δ -function is measuring the density of possible outgoing states (see discussion at end). Recall the density of states in energy for outgoing solid angle $d\Omega$ is

$$\rho(E,d\Omega)=(L/2\pi)^3k^2(m/\hbar^2k)d\Omega=(L/2\pi)^3k(m/\hbar^2)d\Omega, \tag{9.6.20}$$

giving

$$R_{i\to f} = \frac{2\pi}{\hbar}\left|(1/L)^{3/2}\left(\frac{e}{2mc}\right)\sqrt{\frac{1}{\pi a_0^3}}(\vec{A}_0\cdot\vec{p}_f)\left(\frac{8\pi/a_0}{(a_0^{-2}+k_f^2)^2}\right)\right|^2(L/2\pi)^3 k_f(m/\hbar^2)d\Omega. \tag{9.6.21}$$

Notice first that the L^3 terms cancel, reassuringly, our result cannot depend on the size of the box chosen for the plane wave states. Writing $p_f=\hbar k_f$, and of course $p_f^2/2m=E_i+\hbar\omega$, we find

$$R_{i\to f}=\frac{4mp_f}{\pi a_0^5\hbar^4}\left(\frac{e}{mc}\right)^2(\vec{A}_0\cdot\vec{p}_f)^2\left(\frac{1}{a_0^{-2}+(p_f/\hbar)^2}\right)^4 d\Omega. \tag{9.6.22}$$

Note that the rate is angle-dependent, since $(\vec{A}_0\cdot\vec{p}_f)^2=A_0^2p_f^2\cos^2\theta$: ejection is most likely parallel to the electric field. The *total* ionization rate is given by integrating the rate over all angles, and on the unit sphere $\overline{\cos^2\theta}=\bar{z}^2=1/3$, so in the above, $(\vec{A}_0\cdot\vec{p}_f)^2d\Omega\to 4\pi A_0^2p_f^2/3$.

The Photoelectric Cross Section

Imagine now sending this radiation into a gas of hydrogen atoms, many of them, but not enough to shade each other from the radiation significantly. Energy will be absorbed from the beam as atoms ionize. What is the rate at which the beam loses energy? A convenient way of visualizing this rate of loss of energy is to replace each atom by a tiny perfectly absorbent disc oriented with its normal parallel to the beam, the size of these discs such that the beam loses energy at the same rate as it would by ionization. The area of the disc equivalent to one atom is called the *photoelectric cross section*.

The energy density in the beam of radiation is

$$\frac{1}{8\pi}(|\vec{E}|^2+|\vec{B}|^2)=\frac{1}{8\pi}\left(2\frac{\omega^2}{c^2}\vec{A}_0^2\cos^2(\vec{k}\cdot\vec{r}-\omega t)\right) \tag{9.6.23}$$

Denoting the photoelectric cross section by σ,

$$\text{energy absorbed per second}=\sigma\times c\times\text{energy density}, \tag{9.6.24}$$

and averaging $\cos^2$, this gives the energy absorption rate per atom to be $A_0^2\omega 2\sigma/8\pi c$

However, if the rate of ionization of one atom is $R_{i\to f}$, and that ionization takes energy $\hbar\omega$ from the beam, the rate of energy absorption is just $\hbar\omega R_{i\to f}$, so the ionization cross section is given by

$$A_0^2\omega 2\sigma/8\pi c=\hbar\omega R_{i\to f} \tag{9.6.25}$$

This gives

$$\sigma=\frac{8\pi c}{A_0^2\omega^2}\hbar\omega R_{i\to f} \tag{9.6.26}$$

$$=\frac{8\pi c}{A_0^2\omega^2}\hbar\omega\frac{4mp_f}{\pi a_0^5\hbar^4}\left(\frac{e}{mc}\right)^2\frac{4\pi A_0^2p_f^2}{3}\left(\frac{1}{a_0^{-2}+(p_f/\hbar)^2}\right)^4 \tag{9.6.27}$$

$$=\frac{128}{\omega}\frac{e^2}{a_0^5\hbar^3}\frac{\pi p_f^3}{3mc}\left(\frac{1}{a_0^{-2}+(pf/\hbar)^2}\right)^4. \tag{9.6.28}$$

Appendix: the Golden Rule Delta Function and the Density of States

For our big box model, the states are infinite in number, but can be counted by going outwards from the origin in k- space, and adopting some convention for ordering those of equal energy. We can label the states with $\vec{n}=(n_x,n_y,n_z)$, a vector with integer components placing the state in k- space, and denote its energy $E_{\vec{n}}$. The contribution of this state to the density of states is a Dirac δ-function $\delta(E-E_{\vec{n}})$, that is to say, this state contributes 1 to the density of states at the point $E_{\vec{n}}$ on the energy axis. Therefore the density of states in energy is

$$\rho(E)=\sum_{\vec{n}} \delta(E-E_{\vec{n}}) \tag{9.6.29}$$

This is well approximated by the smooth function we derived above.

Now consider the integral over final plane wave states needed in the evaluation of the Golden Rule formula. That δ -function has finite width (from the time-energy uncertainty principle), so by taking our big box big enough we can have many plane wave states within the width of the Golden Rule δ -function: to picture this, let’s represent it by a function equal to Δ over an interval $1/\Delta$, zero otherwise. Then integrating this Golden Rule δ -function with $\rho(E)$ will give a contribution Δ from each state inside the interval of width $1/\Delta$. If the states were uniformly distributed in energy, this would give the total number of states in an interval of unit energy—and that is the definition of the density of states.

9.7: Quantizing Radiation

Introduction

In analyzing the photoelectric effect in hydrogen, we derived the rate of ionization of a hydrogen atom in a monochromatic electromagnetic wave of given strength, and the result we derived is in good agreement with experiment. Recall that the interaction Hamiltonian was

$$\begin{aligned} H^1 &= \left(\frac{e}{mc}\right) cos(\vec{k}\cdot\vec{r}-\omega t)\vec{A}_0\cdot\vec{p} \\ &= \left(\frac{e}{2mc}\right)(e^{i(\vec{k}\cdot\vec{r}-\omega t)}+e^{-i(\vec{k}\cdot\vec{r}-\omega t)})\vec{A}_0\cdot\vec{p}. \end{aligned} \tag{9.7.1}$$

and we dropped the $e^{i\omega t}$ term because it would correspond to the atom giving energy to the field, and our atom was already in its ground state. However, if we go through the same calculation for an atom *not* initially in the ground state, then indeed an electromagnetic wave of appropriate frequency will cause a transition rate to a lower energy state, and $e^{i\omega t}$ is the relevant term.

But this is not the whole story. An atom in an excited state will eventually emit a photon and go to a lower energy state, even if there is *zero* external field. Our analysis so far does not predict this -- obviously, the interaction written above is only nonzero if $\vec{A}$ is nonzero! So what are we missing?

Essentially, the answer is that the electromagnetic field itself is quantized. Of course, we know that, it's made up of photons. Recall Planck's successful analysis of radiation in a box: he considered all possible normal modes for the radiation, and asserted that a mode of energy ω could only gain or lose energy in amounts $\hbar\omega$. This led to the correct formula for blackbody radiation, then Einstein proved that the same assumption, with the same $\hbar$, accounted for the photoelectric effect. We now understand that these modes of oscillation of radiation are just simple harmonic oscillators, with energy $(n+\frac{1}{2})\hbar\omega$, and, just as a mass on a spring oscillator has fluctuations in the ground state, $\langle x\rangle=0$ but $\langle x^2\rangle\neq 0$, for these electromagnetic modes $\langle\vec{A}\rangle=0$ but $\langle\vec{A}^2\rangle\neq 0$.

> *The electromagnetic field itself is quantized.*

These fluctuations in $\vec{A}$ mean the interaction Hamiltonian is momentarily nonzero, and therefore can cause a transition.

Therefore, to find the spontaneous transition rate (as it's called) for an atom in a zero (classically speaking) electromagnetic field, we need to express the electromagnetic field in terms of normal modes (we'll take a big box), then quantize these modes as quantum simple harmonic oscillators, introducing raising and lowering operators for each oscillator (these will be photon creation and annihilation operators) then construct the appropriate quantum operator expression for $\vec{A}$ to put in the electron-radiation interaction Hamiltonian.

The bras and kets will now be quantum states of the electron *and* the radiation field, in contrast to our analysis of the classical field above, where the radiation field didn't change. (Of course, it did, really, in that it lost one photon, but in the classical limit there are infinitely many photons in each mode, so that wouldn't register.)

We use the Coulomb gauge $\vec{\nabla}\cdot\vec{A}=0$ satisfies

$$\nabla^2\vec{A}-\frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2}=0. \tag{9.7.2}$$

Taking for convenience periodic boundary conditions in the big box, we can write $\vec{A}$ (classically) as a Fourier series at $t = 0$:

$$\vec{A}(\vec{r},t=0)=\frac{1}{\sqrt{V}}\sum_{\vec{k}}\sum_{\alpha=1,2}(c_{\vec{k},\alpha}(0)\vec{\varepsilon}_\alpha e^{i\vec{k}\cdot\vec{r}}+c^*_{\vec{k},\alpha}(0)\vec{\varepsilon}^*_\alpha e^{-i\vec{k}\cdot\vec{r}}) \tag{9.7.3}$$

The time-dependence is given by putting in the whole plane wave: $e^{i\vec{k}\cdot\vec{r}}\to e^{i(\vec{k}\cdot\vec{r}-\omega t)}$, which time dependence can be taken into the coefficient, $c_{\vec{k},\alpha}(t)=c_{\vec{k},\alpha}(0)e^{-i\omega t}$, so

$$\vec{A}(\vec{r},t)=\frac{1}{\sqrt{V}}\sum_{\vec{k}}\sum_{\alpha=1,2}(c_{\vec{k},\alpha}(t)\vec{\varepsilon}_\alpha e^{i\vec{k}\cdot\vec{r}}+c^*_{\vec{k},\alpha}(t)\vec{\varepsilon}^*_\alpha e^{-i\vec{k}\cdot\vec{r}}) \tag{9.7.4}$$

The vector $\vec{\varepsilon}_\alpha$ is the polarization of the plane wave. It's in the same direction as the electric field. Actually it varies with $\vec{k}$, because from $\vec{\nabla}\cdot\vec{A}=0$, it's perpendicular to $\vec{k}$. That is, for a given $\vec{k}$ there are two independent polarizations. For $\vec{k}$ along the z -axis, they could be along the x -and y -axes, these would be called linear polarization, and is the standard approach. But we could also take the vectors $(1/\sqrt{2})(1,\ \pm i,\ 0)$. These correspond to *circular polarization*: equal x - and y -components but with the *y*-component 90 degrees ahead in phase. You may recognize the vectors $(1/\sqrt{2})(1,\ \pm i,\ 0)$ as the eigenvectors for the rotation operator around the z -axis -- the circularly polarized beam carries angular momentum,$\pm\hbar$ per photon, pointed along the direction of motion.

The energy density $\frac{1}{8\pi}(\overline{|\vec{E}|^2+|\vec{B}|^2})$ can be expressed as a sum over the individual $(\vec{k},\vec{\varepsilon})$ modes.

Writing the electric and magnetic fields in terms of the vector potential,

$$\vec{E}=-(1/c)\partial\vec{A}/\partial t,\quad \vec{B}=\vec{\nabla}\times\vec{A}. \tag{9.7.5}$$

where

$$\vec{A}(\vec{r},t)=\frac{1}{\sqrt{V}}\sum_{\vec{k}}\sum_{\alpha=1,2}(c_{\vec{k},\alpha}(t)\vec{\varepsilon}_\alpha e^{i\vec{k}\cdot\vec{r}}+c^*_{\vec{k},\alpha}(t)\vec{\varepsilon}^*_\alpha e^{-i\vec{k}\cdot\vec{r}}) \tag{9.7.6}$$

and thereby expressing the total energy

$$\frac{V}{8\pi}(\overline{|\vec{E}|^2+|\vec{B}|^2})=\frac{V}{4\pi}\left(\frac{\omega}{c}\right)^2\overline{|\vec{A}|^2} \tag{9.7.7}$$

in terms of the $(\vec{k},\vec{\varepsilon})$ amplitudes $c^*_{\vec{k},\alpha}(t),\ \ c_{\vec{k},\alpha}(t)$, then integrating the energy density over the whole large box the cross terms disappear from the orthogonality of the different modes and the total energy in the box -- the Hamiltonian -- is:

$$H=\frac{1}{2\pi}\sum_{\vec{k}}\sum_{\alpha}\left(\frac{\omega}{c}\right)^2 c^*_{\vec{k},\alpha}c_{\vec{k},\alpha}. \tag{9.7.8}$$

Note that although the Hamiltonian is (of course) time independent, the *coefficients* $c_{\vec{k},\alpha}$ here *are* time dependent, $c_{\vec{k},\alpha}(t)=c_{\vec{k},\alpha}(0)e^{-i\omega t}$.

But this is formally *identical* to a set of simple harmonic oscillators! Recall that for the classical oscillator, $p^2+(m\omega x)^2=2mE$, the vector $z=m\omega x+ip$ has time dependence $z(t)=z_0e^{-i\omega t}$, and the oscillator energy is proportional to z^*z $(x,p$ are the usual conjugate variables). Clearly, $c_{\vec{k},\alpha}(t)$ here corresponds to $z(t)$: same time dependence, same Hamiltonian. Therefore the real and imaginary parts of $c_{\vec{k},\alpha}(t)$ must *also* be conjugate variables, which can therefore be quantized exactly as for the simple harmonic oscillator.

From

$$\vec{A}(\vec{r},t)=\frac{1}{\sqrt{V}}\sum_{\vec{k}}\sum_{\alpha=1,2}(c_{\vec{k},\alpha}(t)\vec{\varepsilon}_\alpha e^{i\vec{k}\cdot\vec{r}}+c^*_{\vec{k},\alpha}(t)\vec{\varepsilon}^*_\alpha e^{-i\vec{k}\cdot\vec{r}}) \tag{9.7.9}$$

we see that the real part of $c_{\vec{k},\alpha}(t)$ basically gives the contribution of the $\vec{k},\alpha$ and, recalling the time dependence $c_{\vec{k},\alpha}(t)=c_{\vec{k},\alpha}(0)e^{-i\omega t}$, the imaginary part is proportional to the contribution to $\partial\vec{A}(\vec{r},t)/\partial t$, that is, to $\vec{E}(\vec{r},t)$. Essentially, then, the real part of $c_{\vec{k},\alpha}(t)$, proportional to the $\vec{k},\alpha$ Fourier component of the vector potential $\vec{A}$, is what corresponds to displacement x in a 1-D simple harmonic oscillator, and the imaginary part of $c_{\vec{k},\alpha}(t)$, the $\vec{k},\alpha$ Fourier component of $\vec{E}$, corresponds to the momentum in the simple harmonic oscillator.

To carry out the quantization, we must express the classical Hamiltonian

$$H=\frac{1}{2\pi}\sum_{\vec{k}}\sum_{\alpha}\left(\frac{\omega}{c}\right)^2 c^*_{\vec{k},\alpha}c_{\vec{k},\alpha} \tag{9.7.10}$$

in the form

$$H=\sum_{\vec{k}}\sum_{\alpha}\frac{1}{2}\left(P^2_{\vec{k},\alpha}+\omega^2 Q^2_{\vec{k},\alpha}\right) \tag{9.7.11}$$

with $P_{\vec{k},\alpha}$, $Q_{\vec{k},\alpha}$ being the imaginary and real parts of the oscillator amplitude $c_{\vec{k},\alpha}(t)$ (scaled appropriately) exactly parallel to the standard treatment of the simple harmonic oscillator:

$$Q_{\vec{k},\alpha}=\frac{1}{c\sqrt{4\pi}}\left(c_{\vec{k},\alpha}+c^*_{\vec{k},\alpha}\right),\quad P_{\vec{k},\alpha}=-i\omega c4\pi\surd\left(c_{\vec{k},\alpha}-c^*_{\vec{k},\alpha}\right). \tag{9.7.12}$$

From the time-dependence $c_{\vec{k},\alpha}(t)=c_{\vec{k},\alpha}(0)e^{-i\omega t}$, these (classical) variables P,Q are *canonical*:

$$\frac{\partial H}{\partial Q_{\vec{k},\alpha}}=-\dot{P}_{\vec{k},\alpha},\quad \frac{\partial H}{\partial P_{\vec{k},\alpha}}=\dot{Q}_{\vec{k},\alpha}. \tag{9.7.13}$$

The Hamiltonian can now be quantized by the standard procedure. The pairs of canonical variables P,Q (one pair to each mode $\vec{k},\alpha$) become operators, the Poisson brackets become commutators, the scale determined by Planck's constant:

$$[Q_{\vec{k},\alpha},P_{\vec{k}',\alpha'}]=i\hbar\delta_{\vec{k},\vec{k}'}\delta_{\alpha,\alpha'}. \tag{9.7.14}$$

The next step is to express the electron radiation interaction $(e/mc)\vec{A}\cdot\vec{p}$ in terms of these field operators. Since the electromagnetic field is quantized, the interaction with the electron must be that the electron emits or absorbs quanta (photons). This is most directly represented by writing the interaction in terms of creation and annihilation (raising and lowering) operators:

$$\begin{aligned} a_{\vec{k},\alpha}&=\frac{1}{\sqrt{2\hbar\omega}}\left(\omega Q_{\vec{k},\alpha}+iP_{\vec{k},\alpha}\right)\\ a^{\dagger}_{\vec{k},\alpha}&=\frac{1}{\sqrt{2\hbar\omega}}\left(\omega Q_{\vec{k},\alpha}-iP_{\vec{k},\alpha}\right)\end{aligned} \tag{9.7.15}$$

These satisfy $[a,a^{\dagger}]=1$.

(Notice that the annihilation operator $a_{\vec{k},\alpha}$ is nothing but the operator representation of the classical complex amplitude $c_{\vec{k},\alpha}$, with an extra factor to make it dimensionless, $c_{\vec{k},\alpha}\to c\sqrt{\frac{2\pi\hbar}{\omega}}\,a_{\vec{k},\alpha}$. We discussed this same equivalence in the lecture on coherent states, which were eigenstates of the annihilation operator.)

Following the standard simple harmonic oscillator development, the operator $\hat{n}_{\vec{k},\alpha}=a^{\dagger}_{\vec{k},\alpha}a_{\vec{k},\alpha}$ has eigenstates with integer eigenvalues, $\hat{n}|n\rangle=n|n\rangle$, the contribution to the Hamiltonian from the mode $\vec{k},\alpha$ is just $H_{\vec{k},\alpha}=(\hat{n}_{\vec{k},\alpha}+\frac{1}{2})\hbar\omega$, and $a^{\dagger}|n\rangle=\sqrt{n+1}\,|n+1\rangle$, $a|n\rangle=\sqrt{n}|n-1\rangle$.

The bottom line is: the classical plane wave expansion of $\vec{A}$, with wave amplitudes $c_{\vec{k},\alpha}(t)$

$$\vec{A}(\vec{r},t)=\frac{1}{\sqrt{V}}\sum_{\vec{k}}\sum_{\alpha=1,2}\left(c_{\vec{k},\alpha}(t)\vec{\varepsilon}_{\alpha}e^{i\vec{k}\cdot\vec{r}}+c^*_{\vec{k},\alpha}(t)\vec{\varepsilon}^{\,*}_{\alpha}e^{-i\vec{k}\cdot\vec{r}}\right) \tag{9.7.16}$$

is replaced on quantization by a parallel *operator* expansion, the wave amplitude $c_{\vec{k},\alpha}(t)$ becoming the (scaled) annihilation operator:

$$\vec{A}(\vec{r},t)=\frac{1}{\sqrt{V}}\sum_{\vec{k}}\sum_{\alpha=1,2}c\sqrt{\frac{2\pi\hbar}{\omega}}\left(a_{\vec{k},\alpha}(t)\vec{\varepsilon}_{\alpha}e^{i\vec{k}\cdot\vec{r}}+a^{\dagger}_{\vec{k},\alpha}(t)\vec{\varepsilon}^{\,*}_{\alpha}e^{-i\vec{k}\cdot\vec{r}}\right). \tag{9.7.17}$$

Revisiting the Photoelectric Effect, now with a Quantized Field

Recall now that for the photoelectric effect in hydrogen, following Shankar we wrote the ingoing electromagnetic field $\vec{A}(\vec{r},t)=\vec{A}_0\cos(\vec{k}\cdot\vec{r}-\omega t)$. The only relevant component was that going as $e^{i(\vec{k}\cdot\vec{r}-\omega t)}$. In this section, following standard usage (including Shankar) we take an ingoing field $\vec{A}_0e^{i(\vec{k}\cdot\vec{r}-\omega t)}$ -- an irritating change by a factor of 2, but apparently unavoidable if we want to follow Shankar’s nonquantized photoelectric effect, then go on to the quantized case. Anyway, recall the matrix element to calculate the rate was (with ingoing wave now $\vec{A}=\vec{A}_0\, e^{i(\vec{k}\cdot\vec{r}-\omega t)}$)

$$\langle\vec{k}_f|\left(\frac{e}{mc}\right)\vec{A}_0\, e^{i(\vec{k}\cdot\vec{r}-\omega t)}\cdot\vec{p}|100\rangle \tag{9.7.18}$$

On quantizing the field, from the end of the previous section

$$\vec{A}_0 e^{i(\vec{k}\cdot\vec{r}-\omega t)}=\frac{c_{\vec{k},\alpha}(0)\vec{\varepsilon}}{\sqrt{V}}e^{i(\vec{k}\cdot\vec{r}-\omega t)}\to c\sqrt{\frac{2\pi\hbar}{\omega}}a_{\vec{k},\alpha}\frac{\vec{\varepsilon}}{\sqrt{V}}e^{i(\vec{k}\cdot\vec{r}-\omega t)} \tag{9.7.19}$$

(the c at the beginning here being the speed of light).

Now that the electromagnetic field amplitude $\vec{A}_0$ is expressed as an annihilation operator, appropriate (photon number) bras and kets must be supplied for it to operate on. The relevant photon mode is $\vec{k},\alpha$, so labeling the corresponding photon number states $|n_{\vec{k},\alpha}\rangle=|n\rangle_{\vec{k},\alpha}$ the matrix element that must appear in the Golden Rule is

$$\begin{gathered}(\langle\vec{k}_f|\otimes\langle n-1|_{\vec{k},\alpha})\left(\frac{e}{mc}\right)e^{i\vec{k}\cdot\vec{r}}\vec{A}_0\cdot\vec{p}(|100\rangle\otimes|n\rangle_{\vec{k},\alpha})\\ =\langle\vec{k}_f;n-1|\left(\frac{e}{mc}\right)e^{i\vec{k}\cdot\vec{r}}c\sqrt{\frac{2\pi\hbar}{\omega}}a_{\vec{k},\alpha}\frac{\vec{\varepsilon}\cdot\vec{p}}{\sqrt{V}}|100;n\rangle.\end{gathered} \tag{9.7.20}$$

(We’ve removed the $e^{i\omega t}$, that just contributes to the δ -function in the Golden Rule.)

Since $a_{\vec{k},\alpha}|n\rangle_{\vec{k},\alpha}=\sqrt{n_{\vec{k},\alpha}}|n-1\rangle_{\vec{k},\alpha}$, it is clear that quantizing the incoming electromagnetic wave amounts to replacing the classical vector potential for this wave

$$\vec{A}_0\to c\vec{\varepsilon}\sqrt{\frac{2\pi\hbar n_{\vec{k},\alpha}}{\omega V}} \tag{9.7.21}$$

At the photon occupation level $n_{\vec{k},\alpha}$ the (macroscopic) energy in this single mode $\frac{1}{2\pi}\left(\frac{\omega}{c}\right)^2 c^*_{\vec{k},\alpha}c_{\vec{k},\alpha}$ becomes

$$\frac{1}{2\pi}\left(\frac{\omega}{c}\right)^2 c^2\frac{2\pi\hbar}{\omega}a^\dagger_{\vec{k},\alpha}a_{\vec{k},\alpha}=n_{\vec{k},\alpha}\hbar\omega. \tag{9.7.22}$$

(Recall the Hamiltonian for the classical electromagnetic field is $H=\frac{1}{2\pi}\sum_{\vec{k}}\sum_\alpha\left(\frac{\omega}{c}\right)^2 c^*_{\vec{k},\alpha}c_{\vec{k},\alpha}$ in terms of the $c_{\vec{k},\alpha}$ ’s.)

From $a|n\rangle=\sqrt{n}|n-1\rangle$, the Golden Rule matrix element

$$\langle\vec{k}_f;n-1|\left(\frac{e}{mc}\right)e^{i\vec{k}\cdot\vec{r}}c\sqrt{\frac{2\pi\hbar}{\omega}}a_{\vec{k},\alpha}\frac{\vec{\varepsilon}\cdot\vec{p}}{\sqrt{V}}|100;n\rangle \tag{9.7.23}$$

is proportional to $\sqrt{n_{\vec{k},\alpha}}$, so the Golden Rule rate, which includes the square of the matrix element, will be exactly proportional to $n_{\vec{k},\alpha}$. But from $\vec{A}_0\to c\vec{\varepsilon}\sqrt{\frac{2\pi\hbar n_{\vec{k},\alpha}}{\omega V}}$, this is proportional to $|\vec{A}_0|^2$, and in fact *the quantum rate of absorption of radiation is exactly equal to the classical rate* over the whole range of field strengths.

Spontaneous Emission

However, this exact correspondence with the classical result does *not* hold for photon emission! In that case, the atom *adds* a photon to a mode which already contains *n* photons, say, and the relevant matrix element is $a^\dagger|n\rangle=\sqrt{n+1}|n+1\rangle$, so the equivalent classical vector $\vec{A}_0$ is $c\sqrt{\frac{(n_{\vec{k},\alpha}+1)2\pi\hbar}{\omega V}}\vec{\varepsilon}_\alpha$. This is nonzero even if $n_{\vec{k},\alpha}$ is zero— hence *spontaneous* emission.

For spontaneous emission, then, the relevant matrix element is

$$\langle 100;1|\left(\frac{e}{mc}\right)e^{-i\vec{k}\cdot\vec{r}}c\sqrt{\frac{2\pi\hbar}{\omega}}a^\dagger_{\vec{k},\alpha}\frac{\vec{\varepsilon}\cdot\vec{p}}{\sqrt{V}}|21m;0\rangle. \tag{9.7.24}$$

The density of outgoing states for the emitted photon, taking box normalization with periodic boundary conditions as usual, is

$$\frac{V}{(2\pi)^3}k^2 dk d\Omega=\frac{V}{(2\pi)^3}\frac{\omega^2 d\omega d\Omega}{c^3}=\frac{V}{(2\pi)^3}\frac{\omega^2 dE d\Omega}{\hbar c^3} \tag{9.7.25}$$

so the density of states in energy contribution to the Golden Rule delta function is $\frac{V}{(2\pi)^3}\frac{\omega^2 d\Omega}{\hbar c^3}$, and the photon emission rate with polarization $\vec{\varepsilon}$ into a solid angle $d\Omega$ will be:

$$\frac{2\pi}{\hbar}\left|\langle 100;1|\left(\frac{e}{mc}\right)e^{-i\vec{k}\cdot\vec{r}}c\sqrt{\frac{2\pi\hbar}{\omega}}\,a^{\dagger}_{\vec{k},\alpha}\frac{\vec{\varepsilon}\cdot\vec{p}}{\sqrt{V}}|21m;0\rangle\right|^2\frac{V}{(2\pi)^3}\frac{\omega^2 d\Omega}{\hbar c^3}. \tag{9.7.26}$$

One slight difference in evaluating the matrix element from our treatment of the photoelectric effect is in the representation of the dipole interaction. Recall that there we gave the equivalent forms

$$\langle f|H^1|i\rangle=\left(\frac{e}{mc}\right)\vec{A}_0\cdot\langle f|\vec{p}|i\rangle e^{-i\omega t}=\left(\frac{e}{mc}\right)im\omega\vec{A}_0\cdot\langle f|\vec{r}|i\rangle e^{-i\omega t} \tag{9.7.27}$$

and used the $\vec{p}$ representation because the outgoing photoelectron was taken to be in a plane wave state, an eigenstates of $\vec{p}$. But for spontaneous emission, the electron goes from one bound state to another, so the $\vec{r}$ form gives a more immediate picture of the interacting dipole with the external field, and in fact the integration between the states is generally a little more direct.

So in the matrix element we make the substitution $\vec{\varepsilon}\cdot\vec{p}\to im\omega\vec{\varepsilon}\cdot\vec{r}$, and must then evaluate the atomic matrix element $\langle 100|\vec{\varepsilon}\cdot\vec{r}|21m\rangle$. The natural way to do this is to express the vectors in terms of spherical harmonics, that is, to write them as spherical vectors,

$$r_1^{\pm 1}=\mp(x\pm iy)/\sqrt{2}=r\sqrt{4\pi/3}Y_1^{\pm 1},\quad r_1^0=z=r\sqrt{4\pi/3}Y_1^0 \tag{9.7.28}$$

and similarly for $\vec{\varepsilon}$. The integrals are then straightforward, but tedious.

An amusing point made by Sakurai is that the total transition probability for spontaneous emission is

$$\frac{1}{137}\frac{4}{3}\frac{\omega^3}{c^2}|\langle 100|\vec{x}|21m\rangle|^2 \tag{9.7.29}$$

and this same expression was obtained using the Correspondence Principle by Heisenberg, before quantum field theory was invented.

The calculated lifetime of the $n=2$ state is 1.6×10^{-9} seconds.

CHAPTER OVERVIEW

10: Scattering Theory

Almost everything we know about nuclei and elementary particles has been discovered in scattering experiments, from Rutherford’s surprise at finding that atoms have their mass and positive charge concentrated in almost point-like nuclei, to the more recent discoveries, on a far smaller length scale, that protons and neutrons are themselves made up of apparently point-like quarks.

10.1: Scattering Theory
10.2: More Scattering Theory - Partial Waves
10.3: Scattering Amplitudes, Bound States, Resonances
10.4: Identical Particles- Symmetry and Scattering

Thumbnail: Collimated homogeneous beam of monoenergetic particles, long wavepacket which is approximately a planewave, but strictly does not extend to infinity in all directions, is incident on a target and subsequently scattered into the detector subtending a solid angle. The detector is assumed to be far away from the scattering center. (Department of Physics Wiki @ Florida State University).

10.1: Scattering Theory

Almost everything we know about nuclei and elementary particles has been discovered in scattering experiments, from Rutherford's surprise at finding that atoms have their mass and positive charge concentrated in almost point-like nuclei, to the more recent discoveries, on a far smaller length scale, that protons and neutrons are themselves made up of apparently point-like quarks.

The simplest model of a scattering experiment is given by solving Schrödinger's equation for a plane wave impinging on a localized potential. A potential $V(r)$ might represent what a fast electron encounters on striking an atom, or an alpha particle a nucleus. Obviously, representing any such system by a potential *is* only a beginning, but in certain energy ranges it is quite reasonable, and we have to start somewhere!

The basic scenario is to shoot in a stream of particles, all at the same energy, and detect how many are deflected into a battery of detectors which measure angles of deflection. We assume all the ingoing particles are represented by wavepackets of the same shape and size, so we should solve Schrödinger's time-dependent equation for such a wave packet and find the probability amplitudes for outgoing waves in different directions at some later time after scattering has taken place. But we adopt a simpler approach: we assume the wavepacket has a well-defined energy (and hence momentum), so it is *many* wavelengths long. This means that during the scattering process it looks a lot like a plane wave, and for a period of time the scattering is time independent. We assume, then, that the problem is well approximated by solving the time-*independent* Schrödinger equation with an ingoing plane wave. This is much easier!

All we can detect are outgoing waves far outside the region of scattering. For an ingoing plane wave $e^{i\vec{k}\cdot\vec{r}}$, the wavefunction *far away from the scattering region* must have the form

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} + f(\theta,\varphi)\frac{e^{ikr}}{r} \tag{10.1.1}$$

where θ, φ are measured with respect to the ingoing direction.

Note that the *scattering amplitude* $f(\theta,\varphi)$ has the dimensions of length.

We don't worry about overall normalization, because what is relevant is the *fraction* of the incoming beam scattered in a particular direction, or, to be more precise, into a small solid angle $d\Omega$ in the direction θ, φ. The ingoing particle current (with the above normalization) is $\hbar k/m = v$ through unit area perpendicular to the ingoing beam, the outgoing current into the small angle $d\Omega$ is $(\hbar k/m)|f(\theta,\varphi)|^2 d\Omega$. It is evident that this outgoing current corresponds to the original ingoing current flowing through a perpendicular area of size $d\sigma(\theta,\varphi) = |f(\theta,\varphi)|^2 d\Omega$, and

$$\frac{d\sigma}{d\Omega} = |f(\theta,\varphi)|^2 \tag{10.1.2}$$

is called the ***differential cross section*** for scattering in the direction θ, φ.

The Time-Independent Description

We shall review the time-independent formulation of scattering theory, first as it is presented in Baym, in terms of the standard Schrödinger equation wavefunctions, then do the same thing a la Sakurai, in the more formal, but of course equivalent, language of bras and kets. The Schrödinger wavefunction approach is an easier introduction, but the formal language is more convenient for analyzing the structure of higher order terms.

Actually, Baym's treatment isn't *quite* time-independent, in that he uses an ingoing wavepacket, but it is one of great length, well approximated by a plane wave. Sakurai goes straight to the plane wave, and we do too. This case is very reminiscent of one-dimensional scattering, in which a plane wave from the left generates outgoing waves in both directions, and the amplitudes can be calculated from the Schrödinger equation for a single energy eigenstate. The only difference is that in 3D there will be outgoing waves in all directions.

Following Baym, Schrödinger's equation is:

$$\left(\frac{\hbar^2}{2m}\nabla^2 + E_k\right)\psi_{\vec{k}}(\vec{r}) = V(\vec{r})\psi_{\vec{k}}(\vec{r}), \quad where \;\; E_k = \frac{\hbar^2 k^2}{2m}. \tag{10.1.3}$$

This $\psi_{\vec{k}}$ we take to have an incoming plane wave component $e^{i\vec{k}\cdot\vec{r}}$. Overall normalization is irrelevant, since the differential cross-section depends only on the *ratio* of the scattered wave amplitude to that of the ingoing wave.

The standard approach to an equation like the one above is to transform it into an integral equation using Green's functions. If $V(\vec{r})$ is small (just how small it has to be will become clear later) the integral equation can then be solved by iteration.

The Green's function $G(\vec{r}, \vec{k})$ is essentially the inverse of the differential operator,

$$\left(\frac{\hbar^2}{2m}\nabla^2 + E_k\right) G(\vec{r}, \vec{k}) = \delta(\vec{r}). \tag{10.1.4}$$

This is not a mathematically unique definition: clearly, we can add to $G(\vec{r}, \vec{k})$ any solution of the homogeneous equation

$$\left(\frac{\hbar^2}{2m}\nabla^2 + E_k\right) \varphi\,(\vec{r}, \vec{k}) = 0, \tag{10.1.5}$$

for example, the incoming plane wave.

If we write the integral equation

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} + \int d^3r' G(\vec{r}-\vec{r}\,')V(\vec{r}\,')\ \psi_{\vec{k}}(\vec{r}\,') \tag{10.1.6}$$

this $\psi_{\vec{k}}(\vec{r})$ is certainly a solution to the original Schrödinger equation, as is easily checked by applying the operator

$$\left(\frac{\hbar^2}{2m}\nabla^2 + E_k\right) \tag{10.1.7}$$

to both sides of the equation.

The integral equation can be formally solved by iteration, and for “small” V the solution will converge. But this won't really do—remember, we haven't a unique $G(\vec{r}, \vec{k})$! We have to fix $G(\vec{r}, \vec{k})$ by connecting better with the scattering problem we're trying to solve.

We know our solution has a *single* ingoing plane wave, and outgoing waves in all other directions, generated by the interaction of the plane wave with the potential. But the Schrödinger equation could equally describe *ingoing* waves in the other directions. In defining the Green's function and writing the integral equation, we have nowhere specified the distant form of the wavefunction, that is, we have not required that the Green's function on the right hand side of the integral equation only generate *outgoing* waves. To see how to do this, we must write the Green's function itself as a sum over waves, in other words a Fourier transform, and see how to eliminate the unphysical (for the present problem) incoming waves in that sum.

The explicit form of the Green's function is

$$G(r,k) = \int \frac{d^3k'}{(2\pi)^3} \frac{e^{i\vec{k}'\cdot\vec{r}}}{E_k - \frac{\hbar^2 k'^2}{2m}} = -\frac{m}{2\pi^2 ir\hbar^2}\int_{-\infty}^{\infty}\frac{k'dk'e^{ik'r}}{k'^2-k^2}. \tag{10.1.8}$$

Note that $G(r,k)$ only depends on $\vec{k}$ through E_k, and only on $\vec{r}$ through r, since the integration over $\vec{k}'$ is over all directions. It is easy to verify that this Green's function satisfies the differential equation, by applying the differential operator to the first integral above: the result is to cancel the denominator in the integral, leaving just $\int \frac{d^3k'}{(2\pi)^3} e^{i\vec{k}'\cdot\vec{r}}$, which is the δ- function in $\vec{r}$.

To get the second form of $G(r,k)$ in the equation above, we first do the angular integration $d(\cos\theta)$ to get $(e^{ik'r} - e^{-ik'r})/ik'r$, then rearrange the integral over the $-e^{-ik'r}$ term by switching the sign of k', so it becomes an integral from $-\infty$ to 0 instead of 0 to ∞. Then we add the two terms (the $e^{ik'r}$ and the $-e^{-ik'r}$) together to give an integral from $-\infty$ to ∞. This integral from $-\infty$ to ∞ is then done by contour integration—at least, after we've figured out what to do about the singularities at $k' = \pm k$.

For the integral to be defined, the contour must be distorted slightly so it bypasses these poles.

It is at this point we feed in our physical knowledge of the situation: that in the scattering process, the second term in

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} + \int d^3r' G(\vec{r}-\vec{r}\,')V(\vec{r}\,')\ \psi_{\vec{k}}(\vec{r}\,'), \tag{10.1.9}$$

that is, the Green's function term, has to be a sum over *outgoing waves only*. And, we can guarantee this by distorting the contour of integration in the right direction, as follows.

The contour integral has to be evaluated by closing the contour. Since r is positive $e^{ik'r}$ goes to zero in the *upper* half k' plane, but diverges in the lower half, so we must close the contour in the upper half plane to ensure no contribution from the semicircle at infinity. Therefore, to get the desired outgoing waves, e^{ikr} but not e^{-ikr}, our contour closed in the upper half plane must encircle the pole at $k'=+k$ but *not* the one at $k'=-k$. (e^{ikr} *does* represent outgoing waves: the suppressed time dependence is $e^{-iEt/\hbar}=e^{-i\omega t}$, giving $e^{i(kr-\omega t)}$.) In other words, the relative configuration of the real-axis part of the contour and the two poles has to be:

x (pole)

x (pole at $k'=-k-i\varepsilon$) (pole at $k'=+k+i\varepsilon$)

Instead of moving the contour slightly off the real axis to avoid the poles, we've moved the *poles* slightly instead. These movements are infinitesimal, so which gets moved makes no difference to the value of the integral. It is more convenient to move the poles, as shown, because this move can be efficiently included in the integral just by adding an infinitesimal imaginary part to the denominator:

$$G_+(r,k)=\int\frac{d^3k'}{(2\pi)^3}\frac{e^{i\vec{k}'\cdot\vec{r}}}{E_k-\frac{\hbar^2k'^2}{2m}+i\varepsilon}=-\frac{m}{2\pi^2ir\hbar^2}\int_{-\infty}^{\infty}\frac{k'dk'e^{ik'r}}{k'^2-k2-i\varepsilon}. \tag{10.1.10}$$

Notice that we have written G_+ instead of G, because G can denote any solution of

$$\left(\frac{\hbar^2}{2m}\nabla^2+E_k\right)G(\vec{r},\vec{k})=\delta(\vec{r}) \tag{10.1.11}$$

and we are specifying the particular solution having only *outgoing* waves. *In contrast to* G, G_+ *is well-defined and unique*. (There is another perfectly valid solution having only ingoing waves, but it is irrelevant to the scattering problem. The difference between the ingoing and outgoing solutions satisfies the homogeneous equation having zero on the right-hand side.)

Once we move the poles slightly as described above, the pole at $k'=+k+i\varepsilon$ is in fact the *only* singularity of the integrand lying inside the contour of integration (closed in the upper half plane), so the value of the integral is just the contribution from this pole, that is,

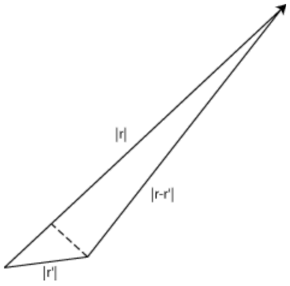
$$G_+(r,k)=-\frac{m}{2\pi^2ir\hbar^2}(2\pi i)\frac{ke^{ikr}}{2k}=-\frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r} \tag{10.1.12}$$

.

Therefore the $i\varepsilon$ prescription (as it's sometimes called) in G_+ does indeed give us what we want: a solution having *only outgoing waves*, and the integral equation becomes:

$$\psi_{\vec{k}}(\vec{r})=e^{i\vec{k}\cdot\vec{r}}-\frac{m}{2\pi\hbar^2}\int d^3r'\frac{e^{ik|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}V(\vec{r}')\,\psi_{\vec{k}}(\vec{r}'). \tag{10.1.13}$$

This can be written more simply if we assume the potential to be localized, so that we can take $|\vec{r}|\gg|\vec{r}'|$. In this case, it is a good approximation to take $|\vec{r}-\vec{r}'|=r$ in the denominator. However, this approximation *cannot* be made in the exponential, because to leading order (see diagram) and although the second term is much smaller than the first, it is a *phase*, which may be of order unity. Such a factor must of course be included so that the contributions to the integral from different regions of the potential are added with the correct relative phases.



$$k|\vec{r}-\vec{r}'| = kr - k\hat{\vec{r}}\cdot\vec{r}' = kr - \vec{k}_f\cdot\vec{r}' \tag{10.1.14}$$

Therefore, assuming the detector distance r is much larger than the range of the potential, we can write

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} - \frac{m}{2\pi\hbar^2}\frac{e^{i\,kr}}{r}\int d^3r'\,e^{-i\,\vec{k}_f\cdot\vec{r}'}V(\vec{r}')\;\psi_{\vec{k}}(\vec{r}'). \tag{10.1.15}$$

The Born Approximation

From the above equation, the first order approximation to the scattering is given by replacing ψ in the integral on the right with the zeroth-order term $e^{i\vec{k}\cdot\vec{r}}$,

$$\psi_{\vec{k}(Born)}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} - \frac{m}{2\pi\hbar^2}\frac{e^{i\,kr}}{r}\int d^3r'\,e^{i\vec{k}\cdot\vec{r}'-\vec{k}_f\cdot\vec{r}'}V(\vec{r}')\;. \tag{10.1.16}$$

This is the *Born approximation*. In terms of the scattering amplitude $f(\theta,\varphi)$, which we defined in terms of the asymptotic wave function:

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} + f(\theta,\varphi)\frac{e^{ikr}}{r} \tag{10.1.17}$$

the Born approximation is:

$$f_{Born}(\theta,\varphi) = -\frac{m}{2\pi\hbar^2}\int d^3r'\,e^{i\,\vec{k}\cdot\vec{r}'-\vec{k}_f\cdot\vec{r}'}V(\vec{r}') = -\frac{m}{2\pi\hbar^2}\int d^3r'\,e^{-i\,\vec{q}\cdot\vec{r}'}V(\vec{r}') \tag{10.1.18}$$

where $\hbar\vec{q}$ is the *momentum transfer*, $\hbar\vec{q} = \hbar(\vec{k}_f - \vec{k})$. (Since the incoming and outgoing momenta have equal magnitude, it is easy to check that $q = 2k\sin\theta/2$.)

The essential physics here is that a particle scattered with momentum change $\hbar\vec{q}$ is scattered by the $\vec{q}$ -Fourier component of the potential—one can imagine the potential as built up of Fourier components each of which acts like a diffraction grating. Higher order corrections to the Born approximation correspond to successive scatterings off these gratings—these higher orders are generated by iteration of

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} - \frac{m}{2\pi\hbar^2}\int d^3r'\,\frac{e^{ik|\,\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}V(\vec{r}')\;\psi_{\vec{k}}(\vec{r}'). \tag{10.1.19}$$

It is important to establish when the Born approximation is a good one: sometimes it isn't. Actually, we are just doing perturbation theory in disguise, so we need the perturbation to be small, that is to say, replacing $\psi_{\vec{k}}(\vec{r}')$ by $e^{i\vec{k}\cdot\vec{r}}$ in the integral on the right in the equation above should only make a small difference to the value of $\psi_{\vec{k}}(\vec{r})$ given by doing the integral. This is of course a rather tricky exercise in self-consistency.

Let us attempt to estimate what difference the replacement of $\psi_{\vec{k}}(\vec{r}')$ by $e^{i\vec{k}\cdot\vec{r}'}$ in the integrand *does* make for the common case of a spherically symmetric potential $V(r)$ parameterized by depth V_0 and range r_0. The integral is effectively only over a region of size r_0 around the origin.

First consider *low energy scattering*, $kr_0 < 1$ say, so for estimation purposes we can replace the exponential term by 1 in the region of integration. We also assume that where ψ appears in the integral on the right-hand side of the equation $|\psi_{\vec{k}}(\vec{r}')|$ is also pretty close to 1 (remember the integral is only over a volume within r_0 or so of the origin) and so we just replace it by 1. In other words, we're assuming that the ingoing plane wave, the $e^{i\vec{k}\cdot\vec{r}'}$, is not dramatically distorted inside that volume where the potential is significant.

Now, we've assumed the wave function near the origin is close to 1, so putting that value in the integrand on the right had better give a value for $\psi_{\vec{k}}(\vec{r})$ on the left hand side of the equation which is pretty close to 1. The approximations give:

$$\psi_k(0) \approx 1 - \frac{m}{2\pi\hbar^2}\int d^3r' \frac{V(r')}{r'}, \tag{10.1.20}$$

so the Born approximation will be reasonable at low energies ($kr_0 < 1$) *if* the second term on the right hand side is a lot less than unity.

When is this true for a real potential? Taking $V(r)$ to have depth V_0 and range r_0, the Born approximation is good if:

$$\frac{m}{2\pi\hbar^2}\int_0^{r_0} 4\pi r^2 \frac{V_0}{r} dr \ll 1, \quad \textit{or} \quad V_0 \ll \frac{\hbar^2}{mr_0^2}. \tag{10.1.21}$$

Notice that the right hand side of this inequality is of order the kinetic energy of a particle confined to a volume equal to the range of the potential, so *the Born approximation is valid at low energies provided the potential is well below the strength necessary for a bound state*.

In fact, the Born approximation works better at higher energies, because the oscillating phase term in $-\frac{m}{2\pi\hbar^2}\int d^3r' e^{-i\vec{q}\cdot\vec{r}'} V(\vec{r}')$ cuts down the value of the integral by a factor of order of magnitude $1/(kr_0)$. This means the condition becomes $V_0 \ll kr_0 \frac{\hbar^2}{mr_0^2}$, always satisfied at high enough energies.

The Lippmann-Schwinger Equation

It proves illuminating, especially in understanding scattering beyond the Born approximation, to recast the Green's function derivation of the scattering amplitude in the more formal language of bras, kets and operators. The Green's function was introduced in the previous section as the (non-unique) inverse of the operator

$$E_k - H_0 = \left(\frac{\hbar^2}{2m}\nabla^2 + E_k\right). \tag{10.1.22}$$

(Parenthetical remark: in numerical computation, the wavefunction might be specified at points on a lattice in space, and a differential operator like this would be represented as a difference operator, that is, as a large but finite matrix operating on a large vector whose elements were the wavefunction values at points on the lattice. The Green's function would then be the inverse matrix with appropriate boundary conditions specified to ensure uniqueness.)

Purely formally (and following Sakurai), writing $H = H_0 + V$, with H_0 the kinetic energy operator $\vec{p}^2/2m$, the ingoing plane wave state is a solution of

$$H_0|\vec{k}\rangle = E_k|\vec{k}\rangle. \tag{10.1.23}$$

We want to solve

$$(H_0 + V)|\psi\rangle = E_k|\psi\rangle. \tag{10.1.24}$$

The transformation from a differential equation to an integral equation in this language is:

$$|\psi\rangle = |\vec{k}\rangle + \frac{1}{E_k - H_0}V|\psi\rangle. \tag{10.1.25}$$

This gives the undisturbed incoming wave for $V = 0$, and by operating on both sides of the equation with $E - H_0$, we find $|\psi\rangle$ does indeed satisfy the full Schrödinger equation. But of course this transformation from a differential to an integral equation has the same flaw as the earlier treatment: H_0 has a continuum of eigenvalues in the infinite volume limit, so the operator equation becomes ill-defined for those eigenstates with energy arbitrarily close to the incoming energy, and those are precisely the states of physical relevance.

To make explicit that this is indeed the problem we've already solved, let us translate it into the earlier language. First take the inner product with the bra $\langle\vec{r}|$:

$$\langle\vec{r}|\psi\rangle=\langle\vec{r}|\vec{k}\rangle+\langle\vec{r}|\frac{1}{E_k-H_0}V|\psi\rangle. \tag{10.1.26}$$

Next, insert a representation of unity as a sum over eigenstates of momentum (and therefore of H_0) into the last term:

$$\begin{aligned}\langle\vec{r}|\psi\rangle&=\langle\vec{r}|\vec{k}\rangle+\int\tfrac{dk'}{(2\pi)^3}\langle\vec{r}|\vec{k}'\rangle\langle\vec{k}'|\tfrac{1}{E_k-H_0}V|\psi\rangle\\&=\langle\vec{r}|\vec{k}\rangle+\int\tfrac{dk'}{(2\pi)^3}\langle\vec{r}|\vec{k}'\rangle\tfrac{1}{E_k-E_k'}\langle\vec{k}'|V|\psi\rangle.\end{aligned} \tag{10.1.27}$$

Finally, insert *another* representation of unity as a sum over eigenstates of *position* in the last term:

$$\langle\vec{r}|\psi\rangle=\langle\vec{r}|\vec{k}\rangle+\int d^3r'\int\frac{dk'}{(2\pi)^3}\langle\vec{r}|\vec{k}'\rangle\frac{1}{E_k-E_k'}\langle\vec{k}'|\vec{r}'\rangle\langle\vec{r}'|V|\psi\rangle. \tag{10.1.28}$$

Comparing this expression with the integral equation in the earlier discussion, it is evident that they are indeed equivalent, and therefore the correct i\varepsilon prescription to give the scattered wave function,

$$|\psi\rangle=|\vec{k}\rangle+\frac{1}{E_k-H_0+i\varepsilon}V|\psi\rangle=|\vec{k}\rangle+G_+V|\psi\rangle \tag{10.1.29}$$

where

$$G_+=\frac{1}{E_k-H_0+i\varepsilon}=\int\frac{d^3k'}{(2\pi)^3}\frac{|\vec{k}'\rangle\langle\vec{k}'|}{E_k-E_k'+i\varepsilon} \tag{10.1.30}$$

in which form it is evident that $\langle\vec{r}|G_+|\vec{r}'\rangle$ is the same as $G_+(\vec{r}-\vec{r}')$ in the previous work.

This equation for the scattered wave $|\psi\rangle$ is called the *Lippmann-Schwinger equation*.

Note : Sakurai defines his Green’s function as

$$G_+(Sakurai)=\frac{\hbar^2}{2m}\frac{1}{E-H_0+i\varepsilon}. \tag{10.1.31}$$

Now that we have a well-defined Green’s function operator G_+, the Lippmann-Schwinger equation can be solved formally:

$$|\psi\rangle=|\vec{k}\rangle+G_+V|\psi\rangle,\quad so\quad|\psi\rangle=\frac{1}{1-G_+V}|\vec{k}\rangle, \tag{10.1.32}$$

with a series solution

$$|\psi\rangle=|\vec{k}\rangle+G_+V|\vec{k}\rangle+G_+VG_+V|\vec{k}\rangle+G_+VG_+VG_+V|\vec{k}\rangle+\dots \tag{10.1.33}$$

just a formal version of the solution we found earlier.

The Transition Matrix

Operating on both sides of the above the equation with V,

$$V|\psi\rangle=V|\vec{k}\rangle+VG_+V|\vec{k}\rangle+VG_+VG_+V|\vec{k}\rangle+\dots=T|\vec{k}\rangle \tag{10.1.34}$$

defining the “transition matrix” T by

$$T=V+VG_+V+VG_+VG_+V+\dots=V+V\frac{1}{E-H_0+i\varepsilon}V+\dots \tag{10.1.35}$$

In terms of this transition matrix operator, the scattered wave can be written

$$|\psi\rangle=|\vec{k}\rangle+G_+T|\vec{k}\rangle. \tag{10.1.36}$$

Comparing this with

$$|\psi\rangle=|\vec{k}\rangle+G_+V|\psi\rangle, \tag{10.1.37}$$

and recalling that the Born approximation is given by

$$|\psi\rangle_{Born} = |\vec{k}\rangle + G_+ \, V|\vec{k}\rangle, \tag{10.1.38}$$

we see that T is a kind of generalized potential, including all the higher order terms, so that just as the Born approximation gave the scattering amplitude in terms of V,

$$f^{Born}(\theta,\varphi) = -\frac{m}{2\pi\hbar^2}\int d^3r' e^{i\vec{k}\cdot\vec{r}' - \vec{k}_f\cdot\vec{r}'} V(\vec{r}') \tag{10.1.39}$$

the *exact* result including *all* higher order terms must have the same structure with T replacing V. Of course, unlike $V(\vec{r})$,T is not a diagonal matrix in r- space: it depends on two space variables, and its Fourier transform is a therefore function of two momenta, that is, the incoming $\vec{k}$ and the scattered $\vec{k}'$. Thus we find:

$$f(\theta,\varphi) = -\frac{m}{2\pi\hbar^2}\int d^3r \int d^3r' e^{i\vec{k}\cdot\vec{r} - \vec{k}'\cdot\vec{r}'} T(\vec{r}',\vec{r}) = -\frac{m}{2\pi\hbar^2}\langle\vec{k}'|T|\vec{k}\rangle. \tag{10.1.40}$$

We have replaced the $\vec{k}_f$ in the Born expression with $\vec{k}'$. Sakurai has an extra $(2\pi)^3$ in the term on the right, because he uses $\langle\vec{k}|\vec{k}'\rangle = \delta(\vec{k} - \overrightarrow{k'})$, $\langle\vec{r}|\vec{k}\rangle = \frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^{3/2}}$, we use $\langle\vec{k}|\vec{k}'\rangle = (2\pi)^3\delta(\vec{k} - \overrightarrow{k'})$, $\langle\vec{r}|\vec{k}\rangle = e^{i\vec{k}\cdot\vec{r}}$.

The Optical Theorem

The Optical Theorem relates the imaginary part of the forward scattering amplitude to the total cross-section,

$$Im\, f(\theta = 0) = \frac{k\sigma_{tot}}{4\pi}. \tag{10.1.41}$$

The physical content of this initially mysterious theorem will become a lot clearer after we discuss partial waves and some geometric effects. It *does* tell us that f cannot be real in all directions, and that in particular f has a positive imaginary part in the forward direction. We've included the proof here for the record, but you can skip it for now. But note that this proof is more general than the simple one given (later) in the section on partial waves, in that we do not here assume the potential to have spherical symmetry.

From the expression for $f(\theta,\varphi)$ above, we see that we must find the imaginary part of $\langle\vec{k}|T|\vec{k}\rangle$.

Recall that

$$V|\psi\rangle = T|\vec{k}\rangle, \tag{10.1.42}$$

so we need to find

$$Im\,\langle\vec{k}|V|\psi\rangle = Im\left[\left(\langle\psi| - \langle\psi|V\frac{1}{E - H_0 - i\varepsilon}\right)V|\psi\rangle\right]. \tag{10.1.43}$$

Since V is hermitian, the only imaginary part of the above matrix element comes from the $i\varepsilon$, recalling that

$$\frac{1}{E - H_0 - i\varepsilon} = \frac{P}{E - H_0} + i\pi\delta(E - H_0). \tag{10.1.44}$$

Therefore,

$$Im\,\langle\vec{k}|V|\psi\rangle = -\pi\langle\psi|V\delta(E - H_0)V|\psi\rangle. \tag{10.1.45}$$

Again using

`V|\psi\rangle = T|\vec{k}\rangle \label{10.1.42}`

we can rewrite the equation

$$Im\,\langle\vec{k}|T|\vec{k}\rangle = Im\,\langle\vec{k}|V|\psi\rangle = -\pi\langle\psi|V\delta(E - H_0)V|\psi\rangle = -\pi\langle\vec{k}|T^\dagger\delta(E - H_0)T|\vec{k}\rangle. \tag{10.1.46}$$

Inserting a complete set of plane wave states in the final matrix element above gives

$$\begin{aligned} Im\,\langle\vec{k}|T|\vec{k}\rangle &= -\pi\langle\vec{k}|T^{\dagger}\delta(E-H_0)T|\vec{k}\rangle \\ &= -\pi\int\frac{d^3k'}{(2\pi)^3}\langle\vec{k}|T^{\dagger}|\vec{k}'\rangle\langle\vec{k}'|T|\vec{k}\rangle\delta(E-\frac{\hbar^2k'^2}{2m}) \\ &= -\pi\int\frac{d\Omega'}{(2\pi)^3}\frac{mk}{\hbar^2}|\langle\vec{k}'|T|\vec{k}\rangle|^2. \end{aligned} \tag{10.1.47}$$

(This is the same formula as Sakurai's in **7.3**: our extra $(2\pi)^3$ in the denominator is only apparent, because our plane wave states differ from his by a factor $(2\pi)^{3/2}$.)

Time-Dependent Formulation of Scattering Theory

In the time-independent formulation presented above, we solved the Lippmann-Schwinger equation to find

$$|\psi\rangle = \ |\vec{k}\rangle + G_+\,V|\vec{k}\rangle + G_+\,VG_+\,V|\vec{k}\rangle + G_+\,VG_+\,VG_+\,V|\vec{k}\rangle + \dots \tag{10.1.48}$$

where

$$G_+(E) = \frac{1}{E-H_0+i\varepsilon} = \int\frac{d^3k'}{(2\pi)^3}\frac{|\vec{k'}\rangle\langle\vec{k'}|}{E-E_{k'}+i\varepsilon} \tag{10.1.49}$$

and $E=E_k$.

(*Reminder on our wave function normalization convention*: we always have a denominator 2π for an integral dk. This means the identity operator as a sum over plane wave projection operators is $I=\int\frac{d^3k}{(2\pi)^3}|\vec{k}\rangle\langle\vec{k}|$. The normalization is $\langle\vec{k}|\vec{k}'\rangle=(2\pi)^3\delta(\vec{k}-\vec{k'})$, and $\langle\vec{r}|\vec{k}\rangle=e^{i\vec{k}\cdot\vec{r}}$. Sakurai uses $\langle\vec{k}|\vec{k}'\rangle=\delta(\vec{k}-\vec{k'})$, $\langle\vec{r}|\vec{k}\rangle=\frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^{3/2}}$ and $I=\int d^3k|\vec{k}\rangle\langle\vec{k}|$ as does Shankar in Chapter 1, for one dimension, page 67, but later, Chapter 21 page 585, Shankar has *switched* to our notation—so watch out! Our convention is also used by Baym and by Peskin.)

In fact, this function G_+ is the Fourier transform of the propagator we discussed last semester. To see how this comes about, take the matrix element between two position eigenstates and Fourier transform from energy to time:

$$\begin{aligned} G_+(\vec{r},\vec{r'},t) &= \frac{1}{2\pi}\int e^{-iEt/\hbar}dE\langle\vec{r}|G_+|\vec{r'}\rangle \\ &= \frac{1}{2\pi}\int e^{-iEt/\hbar}dE\int\frac{d^3k'}{(2\pi)^3}\frac{\langle\vec{r}|k'\rangle\langle k'|\vec{r}'\rangle}{E-E_{k'}+i\varepsilon} \\ &= \frac{1}{2\pi}\int e^{-iEt/\hbar}dE\int\frac{d^3k'}{(2\pi)^3}\frac{e^{i\vec{k'}\cdot(\vec{r}-\vec{r'})}}{E-E_{k'}+i\varepsilon}. \end{aligned} \tag{10.1.50}$$

The integral over E is along the real axis, and the contour is closed in the half plane where the integrand goes to zero for in the imaginary direction, that is, in the lower half plane for $t>0$ and the upper half plane for $t<0$. But with the $i\varepsilon$ term shown, all the singularities of the integrand are in the *lower* half plane. *Hence G_+ is identically zero for $t<0$.*

For $t>0$, G_+ is just the free particle propagator between the two points (apart from the phase factor $-i$):

$$G_+(\vec{r},\vec{r'},t) = -i\int\frac{d^3k'}{(2\pi)^3}e^{i\vec{k'}\cdot(\vec{r}-\vec{r'})-iE_{k'}t/\hbar}. \tag{10.1.51}$$

To summarize: terms in the series solution of the Lippmann-Schwinger equation can be interpreted as successive scatterings off the Fourier components of a potential, with plane wave propagation in between, with the sign of the $i\varepsilon$ term ensuring that there are only outgoing waves from each scattering. In the Fourier transformed version above, the sum is over scattering at all possible points where the potential is nonzero, with G_+ propagation in between, the $i\varepsilon$ ensuring that the scattering path only moves forward in time.

Last semester, we defined the free-particle propagator as the operator $U(t)=e^{-iH_0t/\hbar}$. The propagator describes development of the free-particle wave function in time, so naturally $U(t)=0$ for $t<0$. Then Fourier transforming the propagator from t to E, and inserting an infinitesimal exponentially decaying factor to define the integral at infinity, we find

$$U(E) = \int_0^{\infty} e^{iEt/\hbar}e^{-iH_0t/\hbar}e^{-\varepsilon t/\hbar}dt = \frac{i\hbar}{E-H_0+i\varepsilon}. \tag{10.1.52}$$

Note that the propagators U and G_+ differ by a factor of $i\hbar$, specifically

$$G_+(t) = \frac{-i}{\hbar}\theta(t)e^{-iH_0t/\hbar}. \tag{10.1.53}$$

We follow Sakurai's (section 7.11) notation, this is the correctly normalized Green's function for the time-dependent free-particle Schrödinger equation: it is the solution of

$$\left(i\hbar\frac{\partial}{\partial t} - H_0\right)G_+(t) = \delta(t) \tag{10.1.54}$$

which propagates forwards in time. The reason the propagators U and G_+ differ by a factor of $i\hbar$ is that the Lippmann-Schwinger equation can be generated as a time sequence using the higher-order perturbation theory interaction representation described earlier in the course, essentially expanding $e^{-i(H_0+V)t/\hbar}$ as a time-ordered series expansion in V, and each factor V has an accompanying $1/(i\hbar)$, these factors are taken care of by using G_+ instead of U.

Exercise: in that earlier lecture, we gave the second-order term as:

$$c_n^{(2)}(t) = \left(\frac{1}{i\hbar}\right)^2 \sum_n \int_0^t \int_0^{t'} dt'dt''e^{-i\omega_f(t-t')}\langle f|V_S(t')|n\rangle e^{-i\omega_n(t'-t'')}\langle n|V_S(t'')|i\rangle e^{-i\omega it''} \tag{10.1.55}$$

Assume that the potential V is constant in time. Fourier transform this expression from t to E, $E = \hbar\omega$, and establish that it has the structure $G_+(E)VG_+(E)VG_+(E)$.

The Born Cross-Section from Time-Dependent Theory

We established in the lecture on Time-Dependent Perturbation Theory that to leading order in the perturbation, the transition rate from an initial state i to a final state f is given by Fermi's Golden Rule:

$$R_{i\to f} = \frac{2\pi}{\hbar}|\langle f|V|i\rangle|^2\delta(E_f - E_i). \tag{10.1.56}$$

We can use this result to find—in leading order—the rate of scattering from an incoming plane wave into any outgoing plane wave state having the same energy, and hence by adding the rate over the plane wave directions pointing within a given small solid angle $d\Omega$, rederive the Born approximation.

Conceptually, though, this is a bit tricky. From the above solution of the Schrödinger equation, we know the outgoing wave is a spherical one, so in a particular direction the amplitude decreases. But that doesn't happen with a plane wave! The clearest way to handle this is to put the system in a big box, a cube of side L, with periodic boundary conditions. This makes it easier to count states and normalize the plane waves properly—of course, in the limit of a large box, the plane waves form a complete set, so any spherical wave *can* be expressed as a sum over these plane waves.

In this section, then, we use box-normalized plane waves:

$$|\vec{k}\rangle = \frac{1}{L^{3/2}}e^{i\vec{k}\cdot\vec{r}}, \quad \langle\overrightarrow{k'}|\vec{k}\rangle = \delta_{\vec{k}\overrightarrow{k'}}\ . \tag{10.1.57}$$

So

$$\langle f|V|i\rangle = \frac{1}{L^3}\int d^3re^{-i\vec{k}_f\cdot\vec{r}}V(\vec{r})e^{i\vec{k}_i\cdot\vec{r}} = \frac{1}{L^3}\int d^3re^{-i\vec{q}\cdot\vec{r}}V(\vec{r}) \tag{10.1.58}$$

where the momentum transfer to the particle $\hbar\vec{q} = \hbar(\vec{k}_f - \vec{k}_i)$.

It is important to note that we are taking the incoming wave to be just *one* of the normalized plane wave states satisfying the box periodic boundary conditions, so now the incoming current, being from just one of these plane waves, is

$$j_{in} = |\psi|^2v = \frac{1}{L^3}\frac{p}{m}. \tag{10.1.59}$$

The Golden Rule becomes

$$R_{i\to(f\ in\ d\Omega)} = \frac{2\pi}{\hbar}|\langle f|V|i\rangle|^2\delta(E_f - E_i)d\Omega \tag{10.1.60}$$

f denoting a plane wave going outwards within the solid angle $d\Omega$.

Now, the δ- function simply counts the number of states available at the correct (initial) energy, within the specified final solid angle of direction. The density of states in momentum space (for volume L^3 of real space) is one state in each momentum-space volume $(2\pi\hbar)^3/L^3$, so using $dE/dp = p/m$, the density of states in *energy* for outgoing solid angle $d\Omega$ is $L^3 mpd\Omega/(2\pi\hbar)^3$.

Putting this all together

$$R_{i\to(f\ in\ d\Omega)} = \frac{2\pi}{\hbar}\left|\frac{1}{L^3}\int d^3xe^{-i\vec{q}\cdot\vec{x}}V(\vec{x})\right|^2 L^3 mpd\Omega/(2\pi\hbar)^3. \tag{10.1.61}$$

The transition rate, the rate of scattering into $d\Omega$, is just the incident current multiplied by the infinitesimal scattering cross-section $d\sigma(\theta,\varphi)$ (that was our definition of $d\sigma$),

$$j_{in}\left(\frac{d\sigma(\theta,\varphi)}{d\Omega}\right)d\Omega = R_{i\to(f\ in\ d\Omega)} \tag{10.1.62}$$

because our definition of $R_{i\to(f\ in\ d\Omega)}$ included the appropriately normalized ingoing wave.

So finally

$$\frac{d\sigma(\theta,\varphi)}{d\Omega} = \frac{R_{i\to(f\ in\ d\Omega)}}{j_{in}d\Omega} = \frac{m}{p}\frac{2\pi}{\hbar}\left|\int d^3re^{-i\vec{q}\cdot\vec{r}}V(\vec{r})\right|^2 mp/(2\pi\hbar)^3 = \left|\frac{m}{2\pi\hbar^2}\int d^3re^{-i\vec{q}\cdot\vec{r}}V(\vec{r})\right|^2. \tag{10.1.63}$$

Footnote: the continuum version.

R_{i\to(f\; in\; d\Omega)}=\frac{2\pi}{\hbar} |\langle f|V|i\rangle |^2\delta(E_f-E_i)d\Omega. \label{10.1.60}

In the continuum version, $\langle\vec{r}|\vec{k}\rangle = e^{i\vec{k}\cdot\vec{r}}$, so the matrix element term is just $\left|\int d^3re^{-i\vec{q}\cdot\vec{r}}V(\vec{r})\right|^2$. The energy δ- function is only meaningful inside an integral, in this case over the small volume of outgoing scattering states in the solid angle $d\Omega$ and energy equal to the ingoing energy. But this integral over k' - space *must* include the $1/(2\pi)^3$ factor, according to our rule, giving an outgoing phase space term

$$\int\frac{d^3k'}{(2\pi)^3}\delta(E_{k'}-E_k) = d\Omega\int\frac{k'^2dk'}{(2\pi)^3}\delta\left(\frac{\hbar^2k'^2}{2m}-\frac{\hbar^2k^2}{2m}\right) = d\Omega\frac{k^2}{(2\pi)^3}\frac{m}{\hbar^2k} = d\Omega\frac{mp}{(2\pi\hbar)^3}. \tag{10.1.64}$$

This establishes that our continuum normalization conventions give the same result as that obtained from box normalization.

Electrons Scattering from Atoms

This same approach, using the Golden Rule to derive the leading order scattering rate, is useful is analyzing the scattering of fast electrons by atoms. The problem with slow electrons is that the wave function needs to be antisymmetric with respect to all electrons present. We assume fast electrons have little overlap with the atomic electron wave functions in momentum space, so we don't have to worry about symmetry.

With this approximation, following Sakurai (page 431) the scattering amplitude matrix element is

$$\int d^3re^{i\vec{q}\cdot\vec{r}}\langle n|\left(-\frac{Ze^2}{r}+\sum_i\frac{e^2}{|\vec{r}-\vec{r}_i|}\right)|0\rangle \tag{10.1.65}$$

where the potential term $V(\vec{x})$ is that from the nucleus, plus the repulsion from the other electrons at positions $\vec{x}_i$. Taking the final atomic state n allows for the possibility of inelastic scattering.

Since the distance r of the scattered electron from the nucleus has nothing to do with the atomic state, $n=0$ for the nuclear contribution, which is then just Coulomb scattering, and

$$\int d^3r\frac{e^{i\vec{q}\cdot\vec{r}}}{r} = \frac{4\pi}{q^2}. \tag{10.1.66}$$

(To do this integral, put in a convergence factor $e^{-\varepsilon r}$ then let $\varepsilon\to 0$.)

The term involving the atomic electrons is another matter: for the i^{th} electron, integrating over the coordinate of the scattered electron gives a factor $(4\pi/q^2)e^{i\vec{q}\cdot\vec{r}_i}$, but the hard part is finding the value of the matrix element of this operator between the atomic states. Notice that this is just the Fourier transform of the electrostatic potential from the i^{th} electron's charge density, $\nabla^2 V_i(\vec{r}i) = 4\pi\rho_i(\vec{r}_i)$ transforms to $V_i(\vec{q}) = (4\pi/q^2)\rho_i(\vec{q})$ and $\rho_i(\vec{r}) = e\delta(\vec{r} - \vec{r}_i)$ Fourier transforms to $e(e^{i\vec{q}\cdot\vec{r}_i})$.

The Form Factor

For elastic scattering, then, the contribution of the atomic electrons is simply interpreted: their charge density gives rise to a potential by the usual electrostatic equation, and the (fast) electron is scattered by this potential. For inelastic scattering, the Fourier transform of the electron density is evaluated between different atomic states. In both cases, the matrix element is called the *form factor* $F_n(\vec{q})$ for the scattering, actually $ZF_n(\vec{q}) = \langle n|\sum_i e^{i\vec{q}\cdot\vec{r}_i}|0\rangle$. The normalizing factor Z is introduced so that for elastic scattering, $F_n(\vec{q}) \to 1$ as $q \to 0$.

So the form factor is a map of the charge density in q- space. By measuring the scattering rate at different angles, and Fourier analyzing, it is possible to delineate the charge distribution in ordinary space. The same technique works for nuclei, and in fact for particles—the neutron, for example, although electrically neutral, has a nontrivial electrical charge distribution within its volume, revealed by scattering very fast electrons.

More general form factors describe distribution of spin, and also time dependence of distributions of charge or spin in excited systems. These can all be measured with suitably designed scattering experiments.

10.2: More Scattering Theory - Partial Waves

Plane Waves and Partial Waves

We are considering the solution to Schrödinger’s equation for scattering of an incoming plane wave in the z -direction by a potential localized in a region near the origin, so that the total wave function beyond the range of the potential has the form

$$\psi(r,\theta,\varphi)=e^{ikr\cos\theta}+f(\theta,\varphi)\frac{e^{ikr}}{r}. \tag{10.2.1}$$

The overall normalization is of no concern, we are only interested in the *fraction* of the ingoing wave that is scattered. Clearly the outgoing current generated by scattering into a solid angle $d\Omega$ at angle $\theta,\ \varphi$ is $|f(\theta,\varphi)|^2 d\Omega$ multiplied by a velocity factor that also appears in the incoming wave.

Many potentials in nature are spherically symmetric, or nearly so, and from a theorist’s point of view it would be nice if the experimentalists could exploit this symmetry by arranging to send in spherical waves corresponding to different angular momenta rather than breaking the symmetry by choosing a particular direction. Unfortunately, this is difficult to arrange, and we must be satisfied with the remaining azimuthal symmetry of rotations about the ingoing beam direction.

In fact, though, a full analysis of the outgoing scattered waves from an ingoing plane wave yields the same information as would spherical wave scattering. This is because a plane wave can actually be written as a *sum over spherical waves*:

$$e^{i\vec{k}\cdot\vec{r}}=e^{ikr\cos\theta}=\sum_l i^l(2l+1)j_l(kr)P_l(\cos\theta) \tag{10.2.2}$$

Visualizing this plane wave flowing past the origin, it is clear that in spherical terms the plane wave contains both incoming *and* outgoing spherical waves. As we shall discuss in more detail in the next few pages, the real function $j_l(kr)$ is a standing wave, made up of incoming and outgoing waves of equal amplitude.

We are, obviously, interested only in the outgoing spherical waves *that originate by scattering from the potential*, so we must be careful not to confuse the pre-existing outgoing wave components of the plane wave with the *new* outgoing waves generated by the potential.

The radial functions $j_l(kr)$ appearing in the above expansion of a plane wave in its spherical components are the *spherical Bessel functions*, discussed below. The azimuthal rotational symmetry of plane wave + spherical potential around the direction of the ingoing wave ensures that the angular dependence of the wave function is just $P_l(\cos\theta)$, not $Y_{lm}(\theta,\varphi)$. The coefficient $i^l(2l+1)$ is derived in Landau and Lifshitz, §34, by comparing the coefficient of $(kr\cos\theta)^n$ on the two sides of the equation: as we shall see below, $(kr)^n$ does not appear in $j_l(kr)$ for l greater than n, and $(\cos\theta)^n$ does not appear in $P_l(\cos\theta)$ for l less than n, so the *combination* $(kr\cos\theta)^n$ only occurs once—in the n^{th} term, and the coefficients on both sides of the equation can be matched. (To get the coefficient right, we must of course specify the normalizations for the Bessel function—see below—and the Legendre polynomial.)

Mathematical Interval: The Spherical Bessel and Neumann Functions

The plane wave $e^{i\vec{k}\cdot\vec{r}}$ is a trivial solution of Schrödinger’s equation with zero potential, and therefore, since the $P_l(\cos\theta)$ form a linearly independent set, each term $j_l(kr)P_l(\cos\theta)$ in the plane wave series must be itself a solution to the zero-potential Schrödinger’s equation. It follows that $j_l(kr)$ satisfies the zero-potential radial Schrödinger equation:

$$\frac{d^2}{dr^2}R_l(r)+\frac{2}{r}\frac{d}{dr}R_l(r)+\left(k^2-\frac{l(l+1)}{r^2}\right)R_l(r)=0. \tag{10.2.3}$$

The standard substitution $R_l(r)=u_l(r)/r$ yields

$$\frac{d^2u_l(r)}{dr^2}+\left(k^2-\frac{l(l+1)}{r^2}\right)u(r)=0 \tag{10.2.4}$$

For the simple case $l=0$ the two solutions are $u_0(r)=\sin kr,\ \ \cos kr.$ The corresponding radial functions $R_0(r)$ are (apart from overall constants) the zeroth-order *Bessel* and *Neumann* functions respectively.

The standard normalization for the zeroth-order Bessel function is

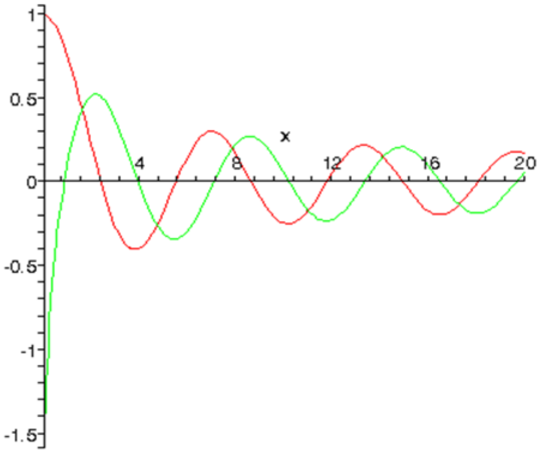
$$j_0(kr)=\frac{\sin kr}{kr}, \tag{10.2.5}$$

and the zeroth-order Neumann function

$$n_0(kr)=-\frac{\cos kr}{kr}. \tag{10.2.6}$$

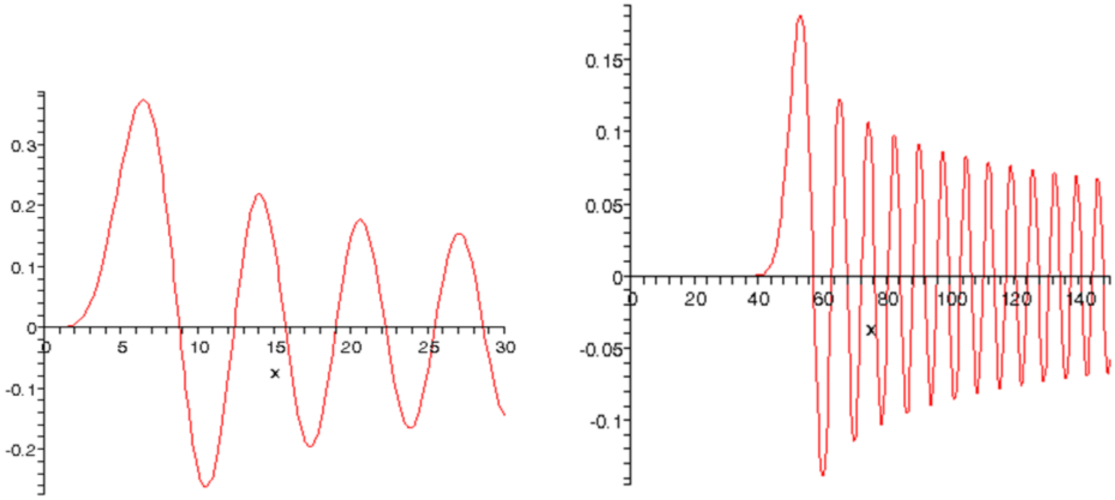
Note that the Bessel function is the one well-behaved at the origin: it could be generated by integrating out from the origin with initial boundary conditions of value one, slope zero.

Here is a plot of $j_0(kr)$ and $n_0(kr)$ from $kr=0.1\ to\ 20$:



For nonzero l, near the origin $R_l(r)\sim r^l$ or $\sim r^{-(l+1)}$. The well-behaved $\sim r^l$ solution is the Bessel function, the singular function the Neumann function. The standard normalizations of these functions are given below.

Here are $j_5(kr)$ and $j_{50}(kr)$:



Detailed Derivation of Bessel and Neumann Functions

This subsection is just here for completeness. We use the dimensionless variable $\rho=kr$.

To find the higher l solutions, we follow a clever trick given in Landau and Lifshitz (§33).

Factor out the ρ^l behavior near the origin by writing

$$R_l = (\rho)^l \chi_l(\rho). \tag{10.2.7}$$

The function $\chi_l(\rho)$ satisfies

$$\frac{d^2}{d\rho^2}\chi_l(\rho) + \frac{2(l+1)}{\rho}\frac{d}{d\rho}\chi_l(\rho) + \chi_l(\rho) = 0. \tag{10.2.8}$$

The trick is to differentiate this equation with respect to ρ:

$$\frac{d^3}{d\rho^3}\chi_l(\rho) + \frac{2(l+1)}{\rho}\frac{d^2}{d\rho^2}\chi_l(\rho) + \left(1 - \frac{2(l+1)}{\rho^2}\right)\frac{d}{d\rho}\chi_l(\rho) = 0. \tag{10.2.9}$$

Writing purely formally $\frac{d}{d\rho}\chi_l(\rho) = \rho\chi_{l+1}(\rho)$, the equation becomes

$$\frac{d^2}{d\rho^2}\chi_{l+1}(\rho) + \frac{2(l+2)}{\rho}\frac{d}{d\rho}\chi_{l+1}(\rho) + \chi_{l+1}(\rho) = 0. \tag{10.2.10}$$

But this *is* the equation that $\chi_{l+1}(\rho)$ must obey! So we have a *recursion formula* for generating *all* the $j_l(\rho)$ from the zeroth one: $\chi_{l+1}(\rho) = \frac{1}{\rho}\frac{d}{d\rho}\chi_l(\rho)$, and $j_l(\rho) = (\rho)^l\chi_l(\rho)$, up to a normalization constant fixed by convention.

In fact, the standard normalization is

$$j_l(\rho) = (-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l \left(\frac{\sin\rho}{\rho}\right). \tag{10.2.11}$$

Now

$$(\sin\rho)/\rho = \sum_0^\infty (-1)^n \rho^{2n}/(2n+1)! \tag{10.2.12}$$

This is a sum of only *even* powers of ρ. It is easily checked that operating on this series with $\left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l$ can never generate any negative powers of ρ. It follows that $j_l(\rho)$, written as a power series in ρ, has leading term proportional to ρ^l. The coefficient of this leading term can be found by applying the differential operator to the series for $(\sin\rho)/\rho$,

$$j_l(\rho) \sim \frac{\rho^l}{(2l+1)!!} \quad as\ \rho \to 0. \tag{10.2.13}$$

This r^l behavior near the origin is the usual well-behaved solution to Schrödinger's equation in the region where the centrifugal term dominates.

Note that the small ρ behavior is *not* immediately evident from the usual presentation of the $j_l(\rho)$'s, written as a mix of powers and trigonometric functions, for example

$$j_1(\rho) = \frac{\sin\rho}{\rho^2} - \frac{\cos\rho}{\rho}, \quad j_2(\rho) = \left(\frac{3}{\rho^3} - \frac{1}{\rho}\right)\sin\rho - \frac{3\cos\rho}{\rho^2}, \ etc. \tag{10.2.14}$$

Turning now to the behavior of the $j_l(\rho)$'s for *large* ρ, from

$$j_l(\rho) = (-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l \left(\frac{\sin\rho}{\rho}\right) \tag{10.2.15}$$

it is evident that the dominant term in the large ρ regime (the one of order $1/\rho$) is generated by differentiating *only* the trigonometric function at each step. Each such differentiation can be seen to be equivalent to multiplying by (-1) and subtracting $\pi/2$ from the argument, so

$$j_l(\rho) \to \frac{1}{\rho}\sin\left(\rho - \frac{l\pi}{2}\right) \ as\ \rho \to \infty. \tag{10.2.16}$$

These $j_l(\rho)$, then, are the physical partial-wave solutions to the Schrödinger equation with zero potential. When a potential is turned on, the wave function near the origin is still $\sim \rho^l$ (assuming, as we always do, that the potential is negligible compared with the $l(l+1)/\rho^2$ term sufficiently close to the origin). The wave function beyond the range of the potential can be found numerically in principle by integrating out from the origin, and in fact will be like $j_l(\rho)$ above *except* that there will be an extra phase factor, called the "phase shift" and denoted by δ) in the sine. The significance of this is that in the far region, the wave function is a linear combination of the Bessel function *and* the Neumann function (the solution to the zero-potential Schrödinger equation singular at the origin). It is therefore necessary to review the Neumann functions as well.

As stated above, the $l=0$ Neumann function is

$$n_0(\rho) = -\frac{\cos\rho}{\rho}, \tag{10.2.17}$$

the minus sign being the standard convention.

An argument parallel to the one above for the Bessel functions establishes that the higher-order Neumann functions are given by:

$$n_l(\rho) = (-\rho)^l \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^l \left(-\frac{\cos\rho}{\rho}\right). \tag{10.2.18}$$

Near the origin

$$n_l(\rho) \sim \frac{(2l-1)!!}{\rho^{l+1}}\ as\ \rho \to 0 \tag{10.2.19}$$

and for large ρ

$$n_l(\rho) \to -\frac{1}{\rho}\cos\left(\rho - \frac{l\pi}{2}\right)\ as\ \rho \to \infty, \tag{10.2.20}$$

so a function of the form $\frac{1}{\rho}\sin\left(\rho - \frac{l\pi}{2} + \delta\right)$ asymptotically can be written as a linear combination of Bessel and Neumann functions in that region.

Finally, the spherical *Hankel* functions are just the combinations of Bessel and Neumann functions that look like outgoing or incoming plane waves in the asymptotic region:

$$h_l(\rho) = j_l(\rho) + i n_l(\rho), \quad h_l^*(\rho) = j_l(\rho) - i n_l(\rho), \tag{10.2.21}$$

so for large ρ,

$$h_l(\rho) \to \frac{e^{i(\rho - l\pi/2)}}{i\rho}, \quad h_l^*(\rho) \to -\frac{e^{-i(\rho - l\pi/2)}}{i\rho}. \tag{10.2.22}$$

The Partial Wave Scattering Matrix

Let us imagine for a moment that we could just send in a (time-independent) spherical wave, with \theta variation given by $P_l(\cos\theta)$. For this l^{th} partial wave (dropping overall normalization constants as usual) the radial function far from the origin for zero potential is

$$j_l(kr) \to \frac{1}{kr}\sin\left(kr - \frac{l\pi}{2}\right) = \frac{i}{2k}\left(\frac{e^{-i(kr - l\pi/2)}}{r} - \frac{e^{+i(kr - l\pi/2)}}{r}\right). \tag{10.2.23}$$

If now the (spherically symmetric) potential is turned on, the only possible change to this standing wave solution in the faraway region (where the potential is zero) is a phase shift δ:

$$\sin\left(kr - \frac{l\pi}{2}\right) \to \sin\left(kr - \frac{l\pi}{2} + \delta_l(k)\right). \tag{10.2.24}$$

This is what we would find on integrating the Schrödinger equation out from nonsingular behavior at the origin.

But in practice, the ingoing wave is given, and *its phase cannot be affected by switching on the potential*. Yet we must still have the solution to the same Schrödinger equation, so to match with the result above we multiply the whole partial wave function by the phase factor $e^{i\delta_l(k)}$. The result is to put *twice* the phase change onto the outgoing wave, so that when the potential is switched on the change in the asymptotic wave function must be

$$\frac{i}{2k}\left(\frac{e^{-i(kr-l\pi/2)}}{r}-\frac{e^{+i(kr-l\pi/2)}}{r}\right)\rightarrow\frac{i}{2k}\left(\frac{e^{-i(kr-l\pi/2)}}{r}-\frac{S_l(k)e^{+i(kr-l\pi/2)}}{r}\right). \tag{10.2.25}$$

This equation introduces the *scattering matrix*

$$S_l(k)=e^{2i\delta_l(k)}, \tag{10.2.26}$$

which must lie on the unit circle $|S|=1$ to conserve probability—the outgoing current must equal the ingoing current. If there is no scattering, that is, zero phase shift, the scattering matrix is unity.

It should be noted that when the radial Schrödinger's equation is solved for a nonzero potential by integrating out from the origin, with $\psi=0$ and $\psi'=1$ initially, the real function thus generated differs from the wave function given above by an *overall* phase factor $e^{i\delta_l(k)}$.

Scattering of a Plane Wave

We're now ready to take the ingoing plane wave, break it into its partial wave components corresponding to different angular momenta, have the partial waves individually phase shifted by l- dependent phases, and add it all back together to get the original plane wave plus the scattered wave.

We are only interested here in the wave function far away from the potential. In this region, the original plane wave is

$$\begin{gathered}e^{i\vec{k}\cdot\vec{r}}=e^{ikr\cos\theta}=\sum_l i^l(2l+1)j_l(kr)P_l(\cos\theta)\\ =\sum_l i^l(2l+1)\tfrac{i}{2k}\left(\tfrac{e^{-i(kr-l\pi/2)}}{r}-\tfrac{e^{+i(kr-l\pi/2)}}{r}\right)P_l(\cos\theta).\end{gathered} \tag{10.2.27}$$

Switching on the potential phase shifts factor the outgoing wave:

$$\frac{e^{+i(kr-l\pi/2)}}{r}\rightarrow\frac{S_l(k)e^{+i(kr-l\pi/2)}}{r} \tag{10.2.28}$$

The actual *scattering* by the potential is the *difference* between these two terms. The complete wave function in the far region (including the incoming plane wave) is therefore:

$$\psi(r,\theta,\varphi)=e^{ikr\cos\theta}+\left(\sum_l(2l+1)\frac{(S_l(k)-1)}{2ik}P_l(\cos\theta)\right)\frac{e^{ikr}}{r}. \tag{10.2.29}$$

The i^l factor cancelled the $e^{-il\pi/2}$. The -1 in $(S_l(k)-1)$ is there because zero scattering means $S=1$. Alternatively, it could be regarded as subtracting off the outgoing waves already present in the plane wave, as discussed above. There is no φ- dependence since with the potential being spherically-symmetric the whole problem is azimuthally-symmetric about the direction of the incoming wave.

It is perhaps worth mentioning that for scattering in just one partial wave, the outgoing current is equal to the ingoing current, whether there is a phase shift or not. So, if switching on the potential does not affect the total current scattered in any partial wave, how can it cause any scattering? The point is that for an ingoing *plane wave* with zero potential, the ingoing and outgoing components have the right relative phase to add to a component of a plane wave—a tautology, perhaps. But if an extra phase is introduced into the outgoing wave *only*, the ingoing + outgoing will *no longer give a plane wave*—there will be an extra outgoing part proportional to $(S_l(k)-1)$.

Recall that the scattering amplitude $f(\theta,\varphi)$ was defined in terms of the solution to Schrödinger's equation having an ingoing plane wave by

$$\psi(r,\theta,\varphi)=e^{ikr\cos\theta}+f(\theta,\varphi)\frac{e^{ikr}}{r}. \tag{10.2.30}$$

We're now ready to express the scattering amplitude in terms of the partial wave phase shifts (for a spherically symmetric potential, of course):

$$f(\theta,\varphi)=f(\theta)=\sum_l(2l+1)\frac{(S_l(k)-1)}{2ik}P_l(\cos\theta)=\sum_l(2l+1)f_l(k)P_l(\cos\theta) \tag{10.2.31}$$

where

$$f_l(k)=\frac{1}{k}e^{i\delta_l(k)}\sin\delta_l(k) \tag{10.2.32}$$

is called the *partial wave scattering amplitude*, or just the *partial wave amplitude*.

So the total scattering amplitude is the sum of these partial wave amplitudes:

$$f(\theta)=\frac{1}{k}\sum_l(2l+1)e^{i\delta_l(k)}\sin\delta_l(k)P_l(\cos\theta). \tag{10.2.33}$$

The total scattering cross-section

$$\begin{gathered}\sigma=\int|f(\theta)|^2d\Omega\\ =2\pi\int_0^\pi|f(\theta)|^2\sin\theta d\theta\\ =2\pi\int_0^\pi|\tfrac{1}{k}\sum_l(2l+1)e^{i\delta_l(k)}\sin\delta_l(k)P_l(\cos\theta)|^2\sin\theta d\theta\end{gathered} \tag{10.2.34}$$

gives

$$\sigma=4\pi\sum_{l=0}^{\infty}(2l+1)|f_l(k)|^2=\frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l. \tag{10.2.35}$$

So the total cross-section is the sum of the cross-sections for each l value. This does *not* mean, though, that the differential cross-section for scattering *into a given solid angle* is a sum over separate l values—the different components interfere. It is only when *all* angles are integrated over that the orthogonality of the Legendre polynomials guarantees that the cross-terms vanish.

Notice that the maximum possible scattering cross-section for particles in angular momentum state l is $(4\pi/k^2)(2l+1)$, which is *four times* the classical cross section for that partial wave impinging on, say, a hard sphere! (Imagine semiclassically particles in an annular area: angular momentum $L=rp$, say, but $L=\hbar l$ and $p=\hbar k$ so $l=rk$. Therefore the annular area corresponding to angular momentum “between” l and $l+1$ has inner and outer radii l/k and $(l+1)/k$ and therefore area $\pi(2l+1)/k^2$.) The quantum result is essentially a diffractive effect, we’ll discuss it more later.

It’s easy to prove the *optical theorem* for a spherically-symmetric potential: just take the imaginary part of each side of the equation

$$f(\theta)=\frac{1}{k}\sum_l(2l+1)e^{i\delta_l(k)}\sin\delta_l(k)P_l(\cos\theta) \tag{10.2.36}$$

at $\theta=0$, using $P_l(1)=1$,

$$Im\ f(\theta=0)=\frac{1}{k}\sum_l(2l+1)\sin^2\delta_l(k) \tag{10.2.37}$$

from which the optical theorem $Im\ f(0)=k\sigma/4\pi$ follows immediately.

It’s also worth noting what the unitarity of the l^{th} partial wave scattering matrix $S_l^\dagger S_l=1$ implies for the partial wave amplitude $f_l(k)=\frac{1}{k}e^{i\delta_l(k)}\sin\delta_l(k)$. Since $S_l(k)=e^{2i\delta_l(k)}$, it follows that

$$S_l(k)=1+2ikf_l(k). \tag{10.2.38}$$

From this, $S_l^\dagger S_l=1$ gives:

$$Im\ f_l(k)=k|f_l(k)|^2. \tag{10.2.39}$$

This can be put more simply:

$$Im\ \frac{1}{f_l(k)}=-k. \tag{10.2.40}$$

In fact,

$$f_l(k)=\frac{1}{k(\cot\delta_l(k)-i)}. \tag{10.2.41}$$

Phase Shifts and Potentials: Some Examples

We assume in this section that the potential can be taken to be zero beyond some boundary radius b. This is an adequate approximation for all potentials found in practice *except* the Coulomb potential, which will be discussed separately later.

Asymptotically, then,

$$\begin{aligned}\psi_l(r)&=\frac{i}{2k}\left(\frac{e^{-i(kr-l\pi/2)}}{r}-\frac{e^{2i\delta_l(k)}e^{+i(kr-l\pi/2)}}{r}\right)\\&=\frac{e^{i\delta_l(k)}}{kr}\sin(kr+\delta_l(k)-l\pi/2)\\&=\frac{e^{i\delta_l(k)}}{kr}(\sin(kr-l\pi/2)\cos\delta_l(k)+\cos(kr-l\pi/2)\sin\delta_l(k)).\end{aligned} \tag{10.2.42}$$

This expression is only exact in the limit $r\to\infty$, but since the potential can be taken zero beyond $r=b$, the wave function must have the form

$$\psi_l(r)=e^{i\delta_l(k)}(\cos\delta_l(k)j_l(kr)-\sin\delta_l(k)n_l(kr)) \tag{10.2.43}$$

for $r>b$.

(The - sign comes from the standard convention for Bessel and Neumann functions—see earlier.)

The Hard Sphere

The simplest example of a scattering potential:

$$\begin{gathered}V(r)=\infty\ for\ r<R,\\V(r)=0\ for\ r\geq R.\end{gathered} \tag{10.2.44}$$

The wave function must equal zero at $r=R$, so from the above form of $\psi_l(r)$,

$$\tan\delta_l(k)=\frac{j_l(kR)}{n_l(kR)}. \tag{10.2.45}$$

For $l=0$,

$$\tan\delta_0(k)=-\frac{(\sin kR)/kR}{(\cos kR)/kR}=-\tan kR, \tag{10.2.46}$$

so $\delta_0(k)=-kR$. This amounts to the wave function being effectively moved over to begin at R instead of at the origin:

$$\frac{\sin kr}{kr}\to\frac{\sin k(r+\delta)}{kr}=\frac{\sin k(r-R)}{kr} \tag{10.2.47}$$

for $r>R$, of course $\psi=0$ for $r<R$.

For higher angular momentum states at low energies $(kR\ll 1)$,

$$\tan\delta_l(k)=\frac{j_l(kR)}{n_l(kR)}\approx\frac{-(kR)^l/(2l+1)!!}{(2l-1)!!/(kR)^{l+1}}=-\frac{(kR)^{2l+1}}{(2l+1)((2l-1)!!)^2}. \tag{10.2.48}$$

Therefore at low enough energy, only $l=0$ scattering is important—as is obvious, since an incoming particle with momentum $p=\hbar k$ and angular momentum $l\hbar$ is most likely at a distance l/k from the center of the potential at closest approach, so if this is much greater than R, the phase shift will be small.

The Born Approximation for Partial Waves

From the definition of $f(\theta,\varphi)$

$$\psi_{\vec{k}}(\vec{r})=e^{i\vec{k}\cdot\vec{r}}+f(\theta,\varphi)\frac{e^{ikr}}{r} \tag{10.2.49}$$

and

$$\psi_{\vec{k}}(\vec{r})=e^{i\vec{k}\cdot\vec{r}}-\frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r}\int d^3r' e^{-i\vec{k}_f\cdot\vec{r}'}V(\vec{r}')\,\psi_{\vec{k}}(\vec{r}') \tag{10.2.50}$$

recall the Born approximation amounts to replacing the wave function $\psi_{\vec{k}}(\vec{r}')$ in the integral on the right by the incoming plane wave, therefore ignoring rescattering.

To translate this into a partial wave approximation, we first take the incoming $\vec{k}$ to be in the z- direction, so *in the integrand* we replace $\psi_{\vec{k}}(\vec{r}')$ by

$$e^{ikr'\cos\theta}=\sum_l i^l(2l+1)j_l(kr')P_l(\cos\theta'). \tag{10.2.51}$$

Labeling the angle between $\vec{k}_f$ and $\vec{r}'$ by γ,

$$e^{-i\vec{k}_f\cdot\vec{r}'}=\sum_l(-i)^l(2l+1)j_l(kr')P_l(\cos\gamma). \tag{10.2.52}$$

Now $\vec{k}_f$ is in the direction (θ,φ) and $\vec{r}'$ in the direction (θ',φ'), and γ is the angle between them. For this situation, there is an *addition theorem for spherical harmonics*:

$$P_l(\cos\gamma)=\frac{4\pi}{2l+1}\sum_{m=-l}^{l}Y^*_{lm}(\theta',\varphi')Y_{lm}(\theta,\varphi). \tag{10.2.53}$$

On inserting this expression and integrating over θ',φ', the nonzero m terms give zero, in fact the only nonzero term is that with the same l as the term in the $\psi_{\vec{k}}(\vec{r}')$ expansion, giving

$$f(\theta)=\frac{-2m}{\hbar^2}\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta)\int_0^\infty r^2drV(r)(j_l(kr))^2 \tag{10.2.54}$$

and remembering

$$f(\theta)=\frac{1}{k}\sum_l(2l+1)e^{i\delta_l(k)}\sin\delta_l(k)P_l(\cos\theta) \tag{10.2.55}$$

it follows that for small phase shifts (the only place it's valid) the partial-wave Born approximation reads

$$\delta_l(k)\approx-\frac{2mk}{\hbar^2}\int_0^\infty r^2drV(r)(j_l(kr))^2. \tag{10.2.56}$$

Low Energy Scattering: the Scattering Length

From

$$f_l(k)=\frac{1}{k(\cot\delta_l(k)-i)}, \tag{10.2.57}$$

the $l=0$ cross section is

$$\sigma_{l=0}=\frac{4\pi}{k^2|\cot\delta_0(k)-i|^2}. \tag{10.2.58}$$

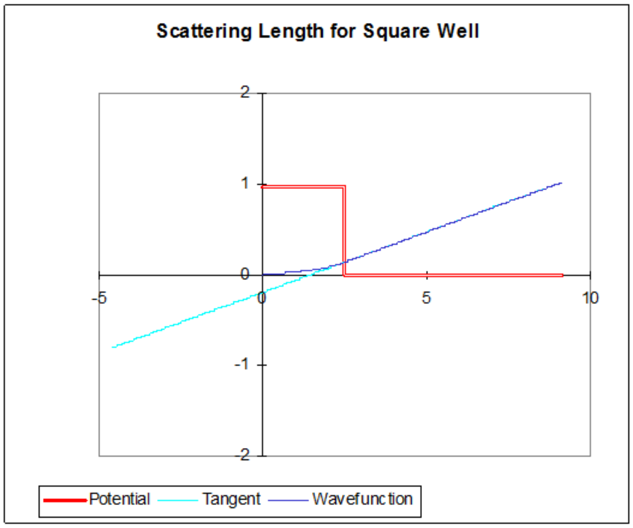
At energy $E\to 0$, the radial Schrödinger equation for $u=r\psi$ away from the potential becomes $d^2u/dr^2=0$, with a straight line solution $u(r)=C(r-a)$. This must be the $k\to 0$ limit of $u(r)=C'\sin(kr+\delta_0(k))$, which can only be correct if δ_0 is itself linear in k for sufficiently small k, and then it must be $\delta_0(k)=-ka$, a being the point at which the extrapolated external wavefunction intersects the axis (maybe at negative r!) So, as k goes to zero, the cot term dominates in the denominator and

$$\sigma_{l=0}(k\to 0)=4\pi a^2. \tag{10.2.59}$$

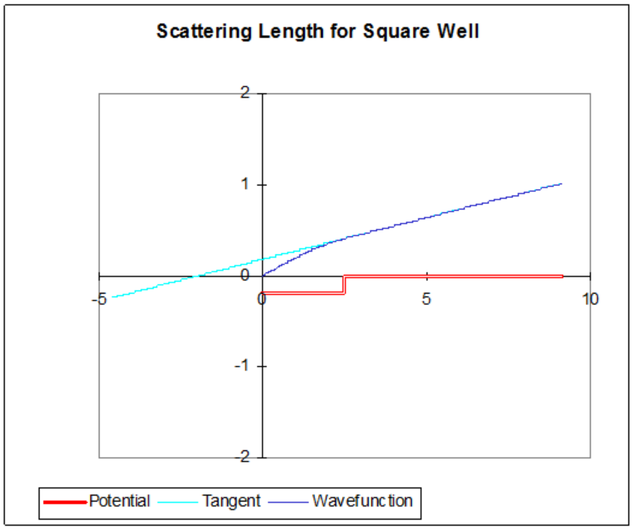
The quantity a is called the *scattering length*.

Integrating the zero-energy radial Schrödinger equation out from $u(r)=0$ at the origin for a weak (spherical) square well potential, it is easy to check that a is positive for a repulsive potential, negative for an attractive potential.

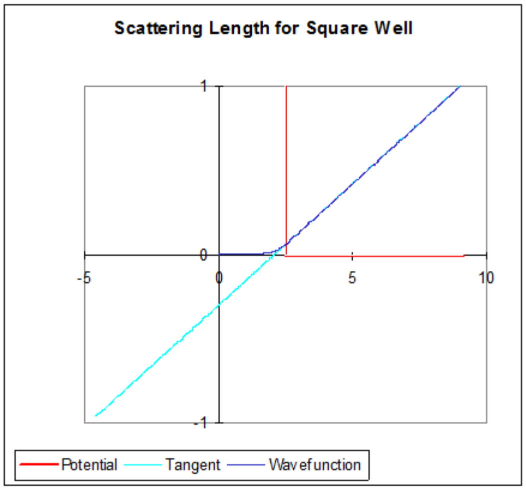
Repulsive potential, zero-energy wave function (so it's a *straight line* outside of the well!):



Attractive potential:

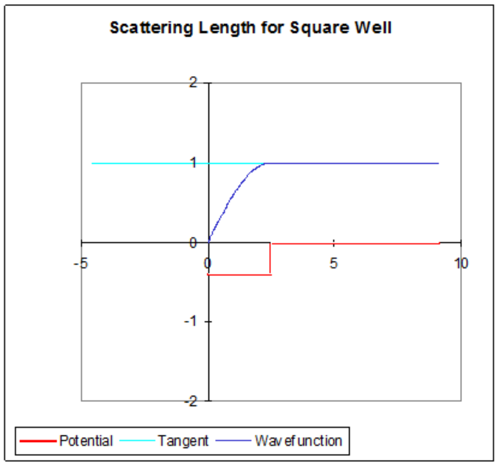


On increasing the strength of the repulsive potential, still solving for the zero-energy wave function, a tends to the potential wall—here's the zero-energy wavefunction for a barrier of height 6:

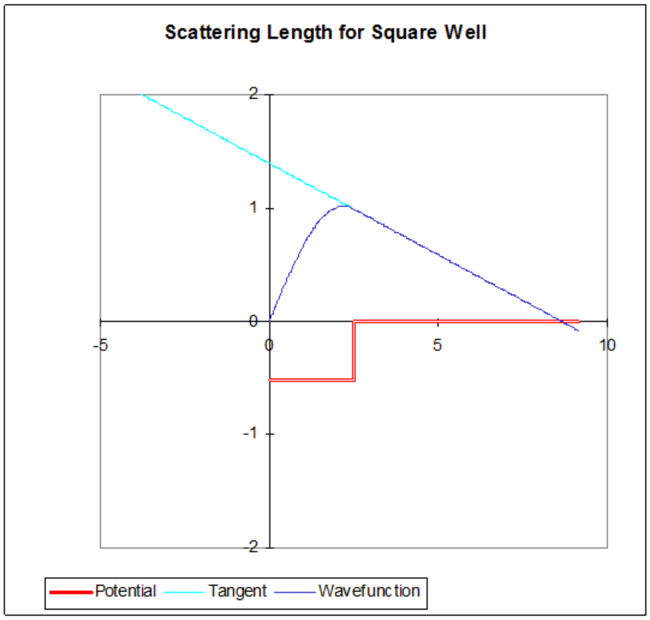


For an infinitely high barrier, the wave function is pushed out of the barrier completely, and the hard sphere result is recovered: scattering length a, cross-section $4\pi a^2$.

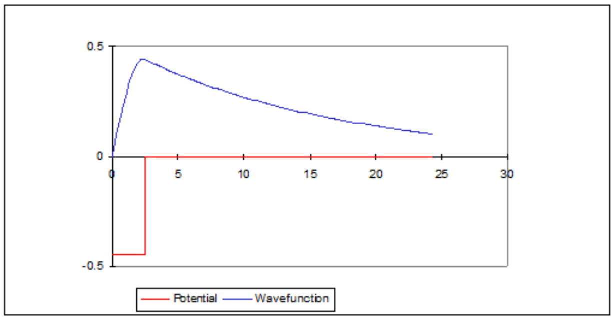
On increasing the strength of the *attractive* well, if there is a phase change greater that $\pi/2$ within the well, a will become positive. In fact, right at $\pi/2$, a is infinite!



And a little more depth to the well gives a *positive* scattering length:



In fact, a well deep enough to have a positive scattering length will also have a bound state. This becomes evident when one considers that the depth at which the scattering length becomes infinite can be thought of as formally having a zero energy bound state, in that although the wave function outside is not normalizable, it is equivalent to an exponentially decaying function with infinite decay length. If one now deepens the well a little, the zero-energy wave function inside the well curves a little more rapidly, so the slope of the wave function at the edge of the well becomes negative, as in the picture above. With this slightly deeper well, we can now *lower* the energy slightly to negative values. This will have little effect on the wave function inside the well, but make possible a fit at the well edge to an exponential decay outside—a genuine bound state, with wave function $\sim e^{-\kappa r}$ outside the well.



If the binding energy of the state is really low, the zero-energy scattering wave function inside the well is almost identical to that of this very low energy bound state, and in particular the logarithmic derivative at the wall will be very close, so $\kappa \cong 1/a$, taking a to be much larger than the radius of the well.

This connects the large scattering length to the energy of the weakly bound state,

$$B.E. = \hbar^2 k^2/2m = \hbar^2/2ma^2. \tag{10.2.60}$$

Wigner was the first to use this to estimate the binding energy of the deuteron from the observed cross section for low energy neutron-proton scattering.

10.3: Scattering Amplitudes, Bound States, Resonances

Low Energy Approximations for the S Matrix

In this section, we examine the properties of the partial-wave scattering matrix

$$S_l(k)=1+2ikf_l(k) \tag{10.3.1}$$

for *complex* values of the momentum variable k. Of course, general complex values of k do not correspond to physical scattering, but it turns out that the scattering of physical waves can often be most simply understood in terms of dominant singularities in the complex k plane.

We begin with the complex k connection between (positive energy) scattering and (negative energy) bound states. The asymptotic form of the $l=0$ partial wavefunction in a scattering experiment is (from the previous lecture)

$$\frac{i}{2k}\left(\frac{e^{-ikr}}{r}-\frac{S_0(k)e^{+ikr}}{r}\right). \tag{10.3.2}$$

An $l=0$ bound state has asymptotic wavefunction

$$\frac{Ce^{-\kappa r}}{r} \tag{10.3.3}$$

where C is a normalization constant.

Notice that this resembles an “outgoing wave” with imaginary momentum $k=i\kappa$. If we analytically continue the *scattering* wavefunction from real k into the complex k- plane, we get *both* exponentially increasing and decreasing wavefunctions, making no physical sense. But there is an exception to this general observation: if the scattering matrix $S_0(k)$ becomes *infinite* at some complex value of k, the exponentially decreasing term will be infinitely larger than the exponentially increasing term. In other words, we’ll only have a decreasing wavefunction—a bound state. We know that the energy of a bound state has to be real and negative, and is also equal to $\hbar^2k^2/2m$, so this can only happen for k *pure imaginary*, $k=i\kappa$.

Now, the existence of a low energy bound state means that the S- matrix has a pole (on the imaginary axis) close to the origin, so this will strongly affect low energy (near the origin, but real k) scattering. Let’s see how that works using the low-energy approximation discussed previously. Recall that the $l=0$ partial wave amplitude

$$f_0(k)=\frac{1}{k(\cot\delta_0(k)-i)}, \tag{10.3.4}$$

and at low energy $\delta_0(k)=-ka$, so

$$f_0(k)=\frac{1}{k((-1/ka)-i)}=-\frac{1}{ik+1/a}, \tag{10.3.5}$$

and

$$S_0(k)=1+2ikf_0(k)=-\frac{k+(i/a)}{k-(i/a)}. \tag{10.3.6}$$

Note that $S\to 1$ as $k\to 0$, as it should, since $\delta_0(k)=-ka\to 0$ and $S_0(k)=e^{2i\delta_0(k)}$. Note also that this approximation correctly gives $|S_0(k)|=1$.

This $S_0(k)$ has a pole in the complex plane at $k=i/a$, and if this corresponds to a bound state having $\kappa=1/a$, then the binding energy $\hbar^2\kappa^2/2m=\hbar^2/2ma^2$. In fact, though, we run into a problem here: we get the same form of $S_0(k)$ at low energies even for a *repulsive* potential, which certainly doesn’t have a bound state! The pole in $S_0(k)$ only means that we can have an asymptotic wavefunction of the right form, but it does *not* guarantee that this asymptotic form will go smoothly to nonsingular behavior at the origin. For a repulsive potential, it’s easy to see that the zero (or negative) energy wavefunction on integrating out from the origin slopes more and more steeply *upwards*, so could never, with increasing r, go over to asymptotic decay.

Effective Range

The low energy approximation above can be written

$$k\cot\delta_0(k)\simeq -1/a. \tag{10.3.7}$$

We shall now derive a better approximation,

$$k\cot\delta_0(k)\simeq -(1/a)+\frac{1}{2}r_0k^2, \tag{10.3.8}$$

where r_0, called the "effective range", gives some measure of the extent of the potential (in contrast to a, which, as we have seen can be arbitrarily large, even for a short-range potential).

A useful mathematical tool needed at this point is the *Wronskian*. For two functions $f(x), g(x)$ the Wronskian is defined as

$$W(f,g)=fg'-f'g, \tag{10.3.9}$$

the prime denoting differentiation as usual. From this, $W'(f,g)=fg''-f''g$, and if $f(x),\ g(x)$ satisfy the same second-order differential equation (like the Schrödinger equation with the same energy) then $W'=0$, so $W(f,g)$ is *constant*, independent of x.

For the radial Schrödinger equation, asymptotically

$$u(k,r)\to C\sin(kr+\delta_0)\equiv v(k,r) \tag{10.3.10}$$

where we now show k explicitly. This asymptotic function $v(k,r)$ satisfies the Schrödinger equation for *zero potential*, but does *not* have the correct physical boundary behavior at $r=0$.

Since in the low-energy limit $\delta_0(k)=-ka$, the $k=0$ asymptotic wavefunction

$$v(0,r)=1-r/a \tag{10.3.11}$$

(taking $C=1/sin\delta_0$).

From the Schrödinger equation

$$-u''(k,r)+(2mV(r)/\hbar^2)u(k,r)=k^2u(k,r) \tag{10.3.12}$$

it is easy to check that the Wronskian of $u(k,r)$ with the corresponding *zero energy* function $u(0,r)$ satisfies:

$$\frac{d}{dr}W[u(k,r),u(0,r)]=k^2u(k,r)u(0,r). \tag{10.3.13}$$

(The term involving the potential has canceled out: dW/dr is nonzero here because these two functions don't satisfy the same differential equation, the energy terms are different.)

The corresponding functions $v(k,r),\ v(0,r)$ satisfy the same Wronskian equation:

$$\frac{d}{dr}W[v(k,r),v(0,r)]=k^2v(k,r)v(0,r). \tag{10.3.14}$$

We can find a formula for the effective range r_0 by integrating the difference between these two equations from $r=0$ to infinity: the two solutions u,v differ only within the range of the potential, and appropriately normalizing them, then taking the difference, gives a measure of this range.

So

$$\begin{aligned}\{W[v(k,r),v(0,r)]-W[u(k,r),u(0,r)]\}_{r=0}^{r=\infty}\\ =k^2\int_0^\infty[v(k,r)v(0,r)-u(k,r)u(0,r)]\ dr\end{aligned} \tag{10.3.15}$$

For r large, $u(k,r)\to v(k,r)$, so there is zero contribution from the upper end. For $r\to 0$, the properly-behaved u functions go to zero, the v functions are

$$v(k,r)=C\sin(kr+\delta_0)=\sin(kr+\delta_0)/sin\delta_0 \tag{10.3.16}$$

from which, with

$$\delta_0(k)=-ka,\quad v(0,r)=1-r/a. \tag{10.3.17}$$

it follows immediately that

$$W[v(k,r),v(0,r)]_{r=0}=-\frac{1}{a}-k\cot\delta_0 \tag{10.3.18}$$

and therefore

$$k\cot\delta_0=-\frac{1}{a}+k^2\int_0^\infty [v(k,r)v(0,r)-u(k,r)u(0,r)]\,dr \tag{10.3.19}$$

with a low-energy limit

$$k\cot\delta_0=-\frac{1}{a}+\frac{1}{2}k^2r_0 \tag{10.3.20}$$

where

$$r_0=2\int_0^\infty [v^2(0,r)-u^2(0,r)]\,dr. \tag{10.3.21}$$

Now, by definition $u(0,r),\ v(0,r)$ coincide outside the range of the potential, but moving from that region towards the origin, they part company when the potential kicks in, with $u\to 0, v\to 1$ as $r\to 0$. Therefore the integral above is a rough measure of the actual *range* of the potential— about half of it (hence the factor of 2 in defining r_0). Note again the contrast with a, which can be infinite for a short range potential.

Coulomb Scattering and the Hydrogen Atom Bound States

One particular set of bound states in a potential we've spent a good deal of time on are the states of the hydrogen atom, and it is interesting to see how they relate to scattering. Recall that the asymptotic form of the bound state wavefunction is:

$$R_{nl}(r)\sim r^n\frac{e^{-r/na_0}}{r}. \tag{10.3.22}$$

But this doesn't have the bound-state form we found above from the analytic continuation argument, there's an extra r^n! What's going on? The problem is that in all our previous work, we assumed that if we looked far enough away from the center of the potential, the radial Schrödinger equation could be taken to be that for zero potential, to any desired accuracy. The Coulomb potential, though, does not decay fast enough with distance for this to be true. For instance, it has bound states having arbitrarily large radii.

Writing

$$\kappa=\frac{1}{na_0}=\frac{me^2}{n\hbar^2} \tag{10.3.23}$$

we have

$$R_{nl}(r)\sim\frac{1}{r}e^{-\kappa r+(me^2/\hbar^2\kappa)\ln r}. \tag{10.3.24}$$

Note that the extra term in the exponent keeps on growing, without limit! We are never free of the potential.

But how does this analysis of hydrogen atom wavefunctions relate to positive-energy scattering states? We can just analytically continue this result back to real k to find out. Replacing $-\kappa$ by ik gives:

$$R_{nl}(r)\sim\frac{1}{r}e^{i(kr+(me^2/\hbar^2k)\ln r)}. \tag{10.3.25}$$

So we have scattering states that are not of the standard form either—the phase shift is infinite, and not well-defined. But we found this result be analytically continuing from the hydrogen atom bound states. Let's check it: let us look at the radial Schrödinger equation for positive energies at large r. Writing $R(r)=u(r)/r$ as usual, let us also put $u(r)=e^{ikr}v(r)$ for large r, and we can also ignore the centrifugal barrier term, so

$$-\frac{\hbar^2}{2m}u''-\frac{e^2}{r}u=Eu=\frac{\hbar^2k^2}{2m}u. \tag{10.3.26}$$

The equation for $v(r)$ is

$$-\frac{\hbar^2}{2m}(2ikv'+v'')-\frac{e^2}{r}v=0 \tag{10.3.27}$$

and since $v(r)$ is slowly varying, the second derivative can be ignored, so

$$v'\cong\frac{ime^2}{\hbar^2kr}v \tag{10.3.28}$$

This leads immediately to the same form we found by analytic continuation.

Resonances and Associated Zeros

Recall in the first semester we discussed α- decay: an α particle in a heavy nucleus can be thought of as trapped in a potential well generated by the attractive nuclear forces. A spherical square well is a workable approximation, except that this square well is at the top of a hill—outside the nucleus, the repulsive electrostatic potential is $(Z-2)2e^2/r$, sloping down from the well edge to zero as $r\to\infty$. This means that for a radioactive nucleus, although the energy level would be at negative energy for the square well on level ground, actually it's above the bottom of the electrostatic hill, and the $r\to\infty$ wavefunction will not be decaying but oscillating. This asymptotic wave is of course very tiny, since typically the chances of detecting the α well outside the nucleus, that is, of decay, is one in millions of years.

Now consider the reverse process: imagine we bombard a decayed nucleus with α particles. If we sent in one at exactly the right energy (very difficult—this is a thought experiment!) the wavefunction would be exactly the same as that for α decay. The wavefunction inside the nucleus would be huge compared with that outside, we'd never see our α again. Less dramatically, if we sent in one close to that energy, the wavefunction would still be very large inside the nucleus, meaning that the particle would spend a long time inside before coming out again. (Recall for a particle in a roller coaster potential in one dimension, the wavefunction is large where the particle spends a lot of time—that's where you're most likely to find it.) This is a *resonance*: at just the right energy, the amplitude of the wavefunction within the potential becomes very large, analogous to the amplitude of a classical driven oscillator as the driving frequency is adjusted to the natural oscillator frequency.

Can we understand this in terms of poles in the S- matrix? Considered as a function of energy, the S- matrix has poles at negative energies corresponding to bound states. But this is a positive energy—and $|S(k)|=1$ for positive energies: that's the regime of real physical scattering. What it can have is a pole *near* a positive energy, in the complex plane. To keep $|S(k)|=1$, it would then have to have a zero at the mirror image point, that is, be locally of the form

$$S_l(E)=e^{2i\delta_l(E)}=\frac{E-E_0-i\Gamma/2}{E-E_0+i\Gamma/2}. \tag{10.3.29}$$

From this,

$$\tan\delta_l(E)=-\Gamma/2(E-E_0), \tag{10.3.30}$$

so

$$\delta_l(E_0)=\pi/2, \tag{10.3.31}$$

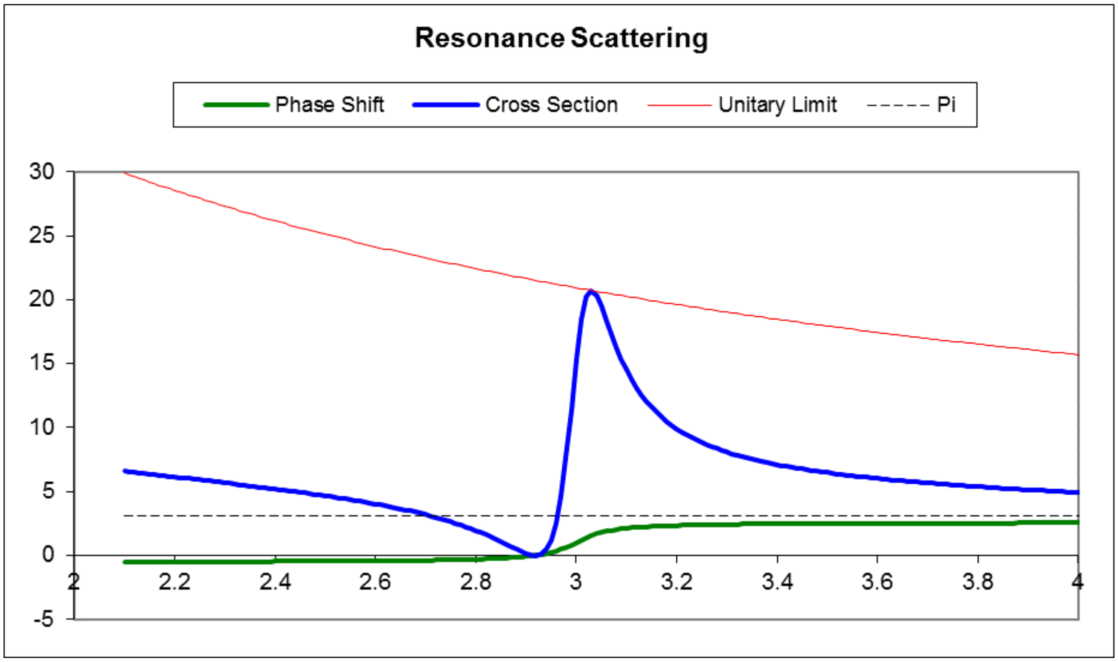
and the scattering cross section reaches its maximum possible value, recall

$$\sigma=4\pi\sum_{l=0}^{\infty}(2l+1)|f_l(k)|^2=\frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l, \tag{10.3.32}$$

so

$$\sigma_l^{max}=\frac{4\pi}{k^2}(2l+1). \tag{10.3.33}$$

For a narrow resonance (small Γ) the phase shift $\delta_l(E)$ goes rapidly from 0 to π as the energy is increased through E_0. Most of the variation occurs within an energy range Γ of E_0, Γ is called the *width* of the resonance. If the resonance is superimposed on a slowly varying background phase shift δ, then it causes an increase from δ to $\delta+\pi$. This will pass through 0 or π, depending on the initial sign of δ, so the maximum scattering at phase shift $\pi/2$ will have associated with it an energy at which there is zero scattering. For substantial background δ, the zero could be close to the peak, as illustrated below:


Resonance Scattering
Phase Shift
Cross Section
Unitary Limit
Pi
30
25
20
15
10
5
0
-5
2
2.2
2.4
2.6
2.8
3
3.2
3.4
3.6
3.8
4

10.4: Identical Particles- Symmetry and Scattering

Introduction

For two identical particles confined to a one-dimensional box, we established earlier that the normalized two-particle wavefunction $\psi(x_1, x_2)$, which gives the probability of finding simultaneously one particle in an infinitesimal length dx_1 at x_1 and another in dx_2 at x_2 as $|\psi(x_1, x_2)|^2 dx_1 dx_2$, only makes sense if $|\psi(x_1, x_2)|^2 = |\psi(x_2, x_1)|^2$, since we don't know which of the two indistinguishable particles we are finding where. It follows from this that there are two possible wave function symmetries: $\psi(x_1, x_2) = \psi(x_2, x_1)$ or $\psi(x_1, x_2) = -\psi(x_2, x_1)$. It turns out that if two identical particles have a symmetric wave function in some state, particles of that type always have symmetric wave functions, and are called *bosons*. (If in some other state they had an antisymmetric wave function, then a linear superposition of those states would be neither symmetric nor antisymmetric, and so could not satisfy $|\psi(x_1, x_2)|^2 = |\psi(x_2, x_1)|^2$.) Similarly, particles having antisymmetric wave functions are called *fermions*. (Actually, we could in principle have $\psi(x_1, x_2) = e^{i\alpha}\psi(x_2, x_1)$, with α a constant phase, but then we wouldn't get back to the original wave function on exchanging the particles twice. Some two-dimensional theories used to describe the quantum Hall effect do in fact have excitations of this kind, called *anyons*, but all ordinary particles are bosons or fermions.)

To construct wave functions for three or more fermions, we assume first that the fermions do not interact with each other, and are confined by a spin-independent potential, such as the Coulomb field of a nucleus. The Hamiltonian will then be symmetric in the fermion variables,

$$H = \vec{p}_1^2/2m + \vec{p}_2^2/2m + \vec{p}_3^2/2m + \dots + V(\vec{r}_1) + V(\vec{r}_2) + V(\vec{r}_3) + \dots \tag{10.4.1}$$

and the solutions of the Schrödinger equation are products of eigenfunctions of the single-particle Hamiltonian $H = \vec{p}^2/2m + V(\vec{r})$. However, these products, for example $\psi_a(1)\psi_b(2)\psi_c(3)$, do not have the required antisymmetry property. Here $a, b, c, \dots$ label the single-particle eigenstates, and $1, 2, 3, \dots$ denote both space and spin coordinates of single particles, so 1 stands for $(\vec{r}_1, s_1)$. The necessary antisymmetrization for the particles 1, 2 is achieved by subtracting the same product wave function with the particles 1 and 2 interchanged, so $\psi_a(1)\psi_b(2)\psi_c(3)$ is replaced by $\psi_a(1)\psi_b(2)\psi_c(3) - \psi_a(2)\psi_b(1)\psi_c(3)$, ignoring overall normalization for now.

But of course the wave function needs to be antisymmetrized with respect to *all* possible particle exchanges, so for 3 particles we must add together all 3! permutations of 1, 2, 3 in the state a, b, c with a factor -1 for each particle exchange necessary to get to a particular ordering from the original ordering of 1 in a, 2 in b, and 3 in c. In fact, such a sum over permutations is precisely the *definition* of the determinant, so, with the appropriate normalization factor:

$$\psi_{abc}(1,2,3) = \frac{1}{\sqrt{3!}}\begin{vmatrix} \psi_a(1) & \psi_b(1) & \psi_c(1) \\ \psi_a(2) & \psi_b(2) & \psi_c(2) \\ \psi_a(3) & \psi_b(3) & \psi_c(3) \end{vmatrix} \tag{10.4.2}$$

where a, b, c label three (different) quantum states and 1, 2, 3 label the three fermions. The determinantal form makes clear the antisymmetry of the wave function with respect to exchanging any two of the particles, since exchanging two rows of a determinant multiplies it by -1.

We also see from the determinantal form that the three states a, b, c *must* all be different, for otherwise two columns would be identical, and the determinant would be zero. This is just Pauli's Exclusion Principle: no two fermions can be in the same state. Although these determinantal wave functions (sometimes called Slater determinants) are only strictly correct for noninteracting fermions, they are a useful beginning in describing electrons in atoms (or in a metal), with the electron-electron repulsion approximated by a single-particle potential. For example, the Coulomb field in an atom, as seen by the outer electrons, is partially shielded by the inner electrons, and a suitable $V(r)$ can be constructed self-consistently, by computing the single-particle eigenstates and finding their associated charge densities.

Space and Spin Wave Functions

Suppose we have two electrons in some spin-independent potential $V(r)$ (for example in an atom). We know the two-electron wave function is antisymmetric. Now, the Hamiltonian has no spin-dependence, so we must be able to construct a set of common eigenstates of the Hamiltonian, the total spin, and the z- component of the total spin.

For two electrons, there are four basis states in the spin space. The eigenstates of S and S_z are the singlet state

$$\chi_S(s_1,s_2) \;=\; |S_{tot}=0,S_z=0\rangle \;=\; (1/\sqrt{2})(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{10.4.3}$$

and the triplet states

$$\chi_T^1(s_1,s_2) \;=\; |1,1\rangle \;=\; |\uparrow\uparrow\rangle,\;\; |1,0\rangle = \;(1/\sqrt{2})(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle),\;\; |1,-1\rangle = |\downarrow\downarrow\rangle \tag{10.4.4}$$

where the first arrow in the ket refers to the spin of particle 1, the second to particle 2.

It is evident by inspection that the singlet spin wave function is antisymmetric in the two particles, the triplet symmetric. The total wave function for the two electrons in a common eigenstate of $S,\ S_z$ and the Hamiltonian H has the form:

$$\Psi(\vec{r}_1,\ \vec{r}_2,s_1,s_2) = \psi(\vec{r}_1,\ \vec{r}_2)\chi(s_1,s_2) \tag{10.4.5}$$

and Ψ must be antisymmetric. It follows that a pair of electrons in the *singlet* spin state must have a *symmetric spatial wave function*, $\psi(\vec{r}_1,\ \vec{r}_2) = \psi(\vec{r}_2,\ \vec{r}_1)$, whereas electrons in the *triplet* state, that is, with their spins parallel, have an *antisymmetric* spatial wave function.

Dynamical Consequences of Symmetry

This overall antisymmetry requirement actually determines the magnetic properties of atoms. The electron’s magnetic moment is aligned with its spin, and *even though the spin variables do not appear in the Hamiltonian, the energy of the eigenstates depends on the relative spin orientation*. This arises from the electrostatic repulsion energy between the electrons. In the spatially antisymmetric state, the two electrons have zero probability of being at the same place, and are on average further apart than in the spatially symmetric state. Therefore, the electrostatic repulsion raises the energy of the spatially symmetric state above that of the spatially antisymmetric state. It follows that *the lower energy state has the spins pointing in the same direction*. This argument is still valid for more than two electrons, and leads to Hund’s rule for the magnetization of incompletely filled inner shells of electrons in transition metal atoms and rare earths: if the shell is half filled or less, all the spins point in the same direction. This is the first step in understanding ferromagnetism.

Another example of the importance of overall wave function antisymmetry for fermions is provided by the specific heat of hydrogen gas. This turns out to be heavily dependent on whether the two protons (spin one-half) in the H_2 molecule have their spins parallel or antiparallel, even though that alignment involves only a very tiny interaction energy. If the proton spins are antiparallel, that is to say in the singlet state, the molecule is called parahydrogen. The triplet state is called orthohydrogen. These two distinct gases are remarkably stable—in the absence of magnetic impurities, para–ortho transitions take weeks.

The actual energy of interaction of the proton spins is of course completely negligible in the specific heat. The important contributions to the specific heat are the usual kinetic energy term, and the rotational energy of the molecule. This is where the overall (space×spin) antisymmetric wave function for the protons plays a role. Recall that the parity of a state with rotational angular momentum l is $(-1)^l$. Therefore, parahydrogen, with an antisymmetric proton *spin* wave function, must have a symmetric proton *space* wave function, and so can only have *even* values of the rotational angular momentum. Orthohydrogen can only have odd values. The energy of the rotational level with angular momentum l is $E_l^{rot} = \hbar^2 l(l+1)/I$, so the two kinds of hydrogen gas have different sets of rotational energy levels, and consequently different specific heats.

Symmetry of Three-Electron Wave Functions

Things get trickier when we go to three electrons. There are now $2^3 = 8$ basis states in the spin space. Four of these are accounted for by the spin 3/2 state with all spins pointing in the same direction. This is evidently a symmetric state, so must be multiplied by an antisymmetric spatial wave function, a determinant. But the other four states are two pairs of total spin $1/2$ states. They are orthogonal to the symmetric spin 3/2 state, so they can’t be symmetric, but they can’t be antisymmetric either, since in each such state two of the spins must be pointing in the same direction! An example of such a state (following Baym, page 407) is

$$\chi(s_1,s_2,s_3) \;=\; |\uparrow_1\rangle(1/\sqrt{2})(|\uparrow_2\downarrow_3\rangle - |\downarrow_2\uparrow_3\rangle). \tag{10.4.6}$$

Evidently, this must be multiplied by a spatial wave function symmetric in 2 and 3, but to get a total wave function with *overall* antisymmetry it is necessary to add more terms:

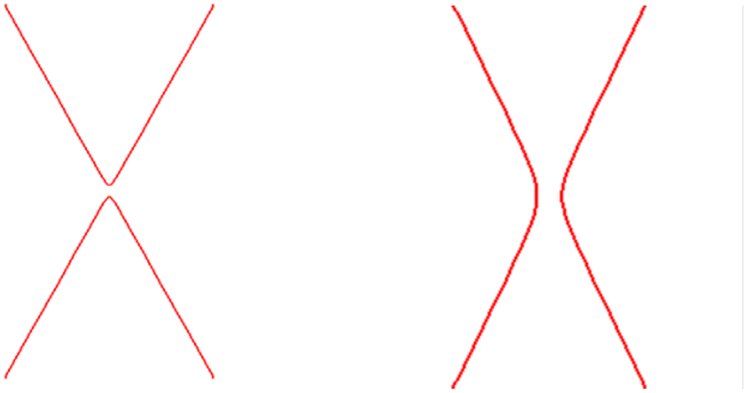
$$\Psi(1,2,3) = \chi(s_1,s_2,s_3)\psi(\vec{r}_1,\ \vec{r}_2,\vec{r}_3) + \chi(s_2,s_3,s_1)\psi(\vec{r}_2,\ \vec{r}_3,\vec{r}_1) + \chi(s_3,s_1,s_2)\psi(\vec{r}_3,\vec{r}_1,\ \vec{r}_2) \tag{10.4.7}$$

(from Baym). Requiring the spatial wave function $\psi(\vec{r}_1,\ \vec{r}_2,\vec{r}_3)$ to be symmetric in 2, 3 is sufficient to guarantee the overall antisymmetry of the total wave function Ψ. Particle enthusiasts might be interested to note that functions exactly like this arise in constructing the spin/flavor wave function for the proton in the quark model (Griffiths, *Introduction to Elementary Particles*, page 179).

For more than three electrons, similar considerations hold. The mixed symmetries of the spatial wave functions and the spin wave functions which together make a totally antisymmetric wave function are quite complex, and are described by Young diagrams (or tableaux). There is a simple introduction, including the generalization to SU(3), in Sakurai, section 6.5. See also §63 of Landau and Lifshitz.

Scattering of Identical Particles

As a preliminary exercise, consider the *classical* picture of scattering between two positively charged particles, for example α-particles, viewed in the center of mass frame. If an outgoing α is detected at an angle θ to the path of ingoing α #1, it could be #1 deflected through θ, or #2 deflected through $\pi-\theta$. (see figure). Classically, we could tell which one it was by watching the collision as it happened, and keeping track.



However, in a quantum mechanical scattering process, we cannot keep track of the particles unless we bombard them with photons having wavelength substantially less than the distance of closest approach. This is just like detecting an electron at a particular place when there are two electrons in a one dimensional box: the probability *amplitude* for finding an α coming out at angle θ to the ingoing direction of one of them is the sum of the amplitudes (*not* the sum of the probabilities!) for scattering through θ and $\pi-\theta$.

Writing the asymptotic scattering wave function in the standard form for scattering from a fixed target,

$$\psi(\vec{r}) \approx e^{ikz} + f(\theta)\frac{e^{ikr}}{r} \tag{10.4.8}$$

the two-particle wave function in the center of mass frame, in terms of the relative coordinate, is given by symmetrizing:

$$\psi(\vec{r}) \approx e^{ikz} + e^{-ikz} + (f(\theta)+f(\pi-\theta))\frac{e^{ikr}}{r}. \tag{10.4.9}$$

How does the particle symmetry affect the actual scattering rate at an angle θ? If the particles were distinguishable, the differential cross section would be

$$\left(\frac{d\sigma}{d\Omega}\right)_{distinguishable} = |f(\theta)|^2 + |f(\pi-\theta)|^2, \tag{10.4.10}$$

but quantum mechanically

$$\left(\frac{d\sigma}{d\Omega}\right) = |f(\theta)+f(\pi-\theta)|^2. \tag{10.4.11}$$

This makes a big difference! For example, for scattering through 90°, where $f(\theta)=f(\pi-\theta)$, the quantum mechanical scattering rate is *twice* the classical (distinguishable) prediction.

Furthermore, if we make the standard expansion of the scattering amplitude $f(\theta)$ in terms of partial waves,

$$f(\theta)=\sum_{l=0}^{\infty}(2l+1)a_l P_l(\cos\theta) \tag{10.4.12}$$

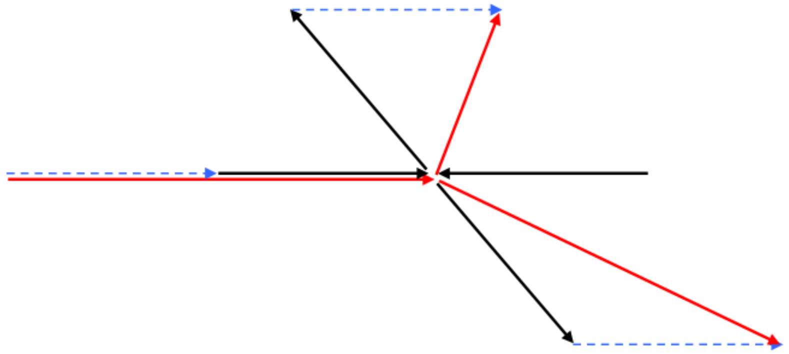
then

$$\begin{aligned} f(\theta)+f(\pi-\theta) &= \sum_{l=0}^{\infty}(2l+1)a_l(P_l(\cos\theta)+P_l(cos(\pi-\theta))) \\ &= \sum_{l=0}^{\infty}(2l+1)a_l(P_l(\cos\theta)+P_l(-\cos\theta)) \end{aligned} \tag{10.4.13}$$

and since $P_l(-x)=(-1)^l P_l(x)$ the scattering only takes place in even partial wave states. This is the same thing as saying that the overall wave function of two identical bosons is symmetric, so if they are in an eigenstates of total angular momentum, from $P_l(-x)=(-1)^l P_l(x)$ it has to be a state of even *l*.

For fermions in an antisymmetric spin state, such as proton-proton scattering with the two proton spins forming a singlet, the spatial wave function is symmetric, and the argument is the same as for the boson case above. For *parallel* spin protons, however, the spatial wave function has to be *antisymmetric*, and the scattering amplitude will then be $f(\theta)-f(\pi-\theta)$. In this case there is *zero* scattering at 90°!

Note that for (nonrelativistic) equal mass particles, the scattering angle in the center of mass frame is twice the scattering angle in the fixed target (lab) frame. This is easily seen in the diagram below. The four equal-length black arrows, two in, two out, forming an X, are the center of mass momenta. The lab momenta are given by adding the (same length) blue dotted arrow to each, reducing one of the ingoing momenta to zero, and giving the (red arrow) lab momenta (slightly displaced for clarity). The outgoing lab momenta are the diagonals of rhombi (equal-side parallelograms), hence at right angles and bisecting the center of mass angles of scattering.



CHAPTER OVERVIEW

11: Other Material

11.1: Schrödinger Spreadsheets
11.2: Homeworks, exams, etc.

11.1: Schrödinger Spreadsheets

11.2: Homeworks, exams, etc.

Index

A

alpha decay
  3.3: Energy-Time Uncertainty Principle
Annihilation operator
  3.4: The Simple Harmonic Oscillator
associated Laguerre polynomials
  4.6: The Hydrogen Atom

B

bosons
  5.1: Bosons and Fermions

C

Cauchy’s theorem
  2.4: Complex Variable, Stationary Path Integrals
Coherent States
  3.6: Coherent States
Creation operator
  3.4: The Simple Harmonic Oscillator
curl
  4.3: Note on Curvilinear Coordinates
curvature
  3.1: 1-D Schrödinger Equation - Example Systems
Curvilinear Coordinates
  4.3: Note on Curvilinear Coordinates

D

density matrix
  7: The Density Matrix
  7.1: The Density Matrix
differential cross section
  10.1: Scattering Theory
Div
  4.3: Note on Curvilinear Coordinates
double slit experiment
  1.4: The Uncertainty Principle

E

Exponential Operator Algebra
  3.9: Appendix- Some Exponential Operator Algebra

F

Fermi’s Golden Rule
  9.5: Time-Dependent Perturbation Theory
fermions
  5.1: Bosons and Fermions
Finite Depth Square Well
  3.1: 1-D Schrödinger Equation - Example Systems
Function spaces
  2.3: Function Spaces

G

Gibbs’ phenomenon
  2.1: Fourier Series and Integrals, the Dirac Function
Grad
  4.3: Note on Curvilinear Coordinates
Green’s function
  10.1: Scattering Theory

H

harmonic oscillator
  3.4: The Simple Harmonic Oscillator
Hermite polynomial
  3.4: The Simple Harmonic Oscillator
Hermitian
  3.2: General Uncertainty Principal

I

Infinite Depth Square Well
  3.1: 1-D Schrödinger Equation - Example Systems
interaction Hamiltonian
  9.6: The Photoelectric Effect in Hydrogen

L

Ladder operators
  3.4: The Simple Harmonic Oscillator
Laplacian operator
  4.3: Note on Curvilinear Coordinates
Legendre Functions
  4.4: Orbital Eigenfunctions in 3-D
Linear Stark Effect
  9.1: Time-Independent Perturbation Theory
Lorentz force law
  6.1: Charged Particle in a Magnetic Field

M

mixed states
  7.1: The Density Matrix

N

Neumann Functions
  10.2: More Scattering Theory - Partial Waves
number operator
  3.4: The Simple Harmonic Oscillator

P

parity
  3.1: 1-D Schrödinger Equation - Example Systems
Partial Waves
  10.2: More Scattering Theory - Partial Waves
Particle in Box
  5.1: Bosons and Fermions
Path integrals
  3.7: Path Integrals
  3.8: Path Integrals for the SHO
Pauli exclusion principle
  5.1: Bosons and Fermions
Peierls Transition
  9.2: The Peierls Transition - an Unexpected Insulator
phase space
  3.4: The Simple Harmonic Oscillator
photoelectric cross section
  9.6: The Photoelectric Effect in Hydrogen
Photoionization
  9.6: The Photoelectric Effect in Hydrogen
Poisson Brackets
  6.1: Charged Particle in a Magnetic Field
Principle of Least Action
  3.7: Path Integrals
Principle of Least Time
  3.7: Path Integrals
Propagators
  3.5: Propagators and Representations
Pure States
  7.1: The Density Matrix

Q

Quadratic Stark Effect
  9.1: Time-Independent Perturbation Theory
Quantizing Radiation
  9.7: Quantizing Radiation

R

Representations (Quantum)
  3.5: Propagators and Representations

S

S matrix
  10.3: Scattering Amplitudes, Bound States, Resonances
Saddlepoint
  2.4: Complex Variable, Stationary Path Integrals
scalar operator
  4.8: Tensor Operators
Scattering
  10: Scattering Theory
scattering length
  10.2: More Scattering Theory - Partial Waves
Schwartz’s inequality
  2.2: Linear Algebra
Spherical Bessel Functions
  10.2: More Scattering Theory - Partial Waves

T

Tensor Operators
  4.8: Tensor Operators
Translation Operator
  3.6: Coherent States
tunneling
  3.1: 1-D Schrödinger Equation - Example Systems
tunneling probability
  3.3: Energy-Time Uncertainty Principle
two photon absorption
  9.5: Time-Dependent Perturbation Theory

V

Van der Waals forces
  9.3: Van Der Waals Forces between Atoms
Variational Method
  8.1: Variational Methods

W

WKB Approximation
  8.2: The WKB Approximation
WKB Connection Formula
  8.3: Note on the WKB Connection Formula

Index

A

alpha decay
  3.3: Energy-Time Uncertainty Principle
Annihilation operator
  3.4: The Simple Harmonic Oscillator
associated Laguerre polynomials
  4.6: The Hydrogen Atom

B

bosons
  5.1: Bosons and Fermions

C

Cauchy's theorem
  2.4: Complex Variable, Stationary Path Integrals
Coherent States
  3.6: Coherent States
Creation operator
  3.4: The Simple Harmonic Oscillator
curl
  4.3: Note on Curvilinear Coordinates
curvature
  3.1: 1-D Schrödinger Equation - Example Systems
Curvilinear Coordinates
  4.3: Note on Curvilinear Coordinates

D

density matrix
  7: The Density Matrix
  7.1: The Density Matrix
differential cross section
  10.1: Scattering Theory
Div
  4.3: Note on Curvilinear Coordinates
double slit experiment
  1.4: The Uncertainty Principle

E

Exponential Operator Algebra
  3.9: Appendix- Some Exponential Operator Algebra

F

Fermi's Golden Rule
  9.5: Time-Dependent Perturbation Theory
fermions
  5.1: Bosons and Fermions
Finite Depth Square Well
  3.1: 1-D Schrödinger Equation - Example Systems
Function spaces
  2.3: Function Spaces

G

Gibbs' phenomenon
  2.1: Fourier Series and Integrals, the Dirac Function
Grad
  4.3: Note on Curvilinear Coordinates
Green's function
  10.1: Scattering Theory

H

harmonic oscillator
  3.4: The Simple Harmonic Oscillator
Hermite polynomial
  3.4: The Simple Harmonic Oscillator
Hermitian
  3.2: General Uncertainty Principal

I

Infinite Depth Square Well
  3.1: 1-D Schrödinger Equation - Example Systems
interaction Hamiltonian
  9.6: The Photoelectric Effect in Hydrogen

L

Ladder operators
  3.4: The Simple Harmonic Oscillator
Laplacian operator
  4.3: Note on Curvilinear Coordinates
Legendre Functions
  4.4: Orbital Eigenfunctions in 3-D
Linear Stark Effect
  9.1: Time-Independent Perturbation Theory
Lorentz force law
  6.1: Charged Particle in a Magnetic Field

M

mixed states
  7.1: The Density Matrix

N

Neumann Functions
  10.2: More Scattering Theory - Partial Waves
number operator
  3.4: The Simple Harmonic Oscillator

P

parity
  3.1: 1-D Schrödinger Equation - Example Systems
Partial Waves
  10.2: More Scattering Theory - Partial Waves
Particle in Box
  5.1: Bosons and Fermions
Path integrals
  3.7: Path Integrals
  3.8: Path Integrals for the SHO
Pauli exclusion principle
  5.1: Bosons and Fermions
Peierls Transition
  9.2: The Peierls Transition - an Unexpected Insulator
phase space
  3.4: The Simple Harmonic Oscillator
photoelectric cross section
  9.6: The Photoelectric Effect in Hydrogen
Photoionization
  9.6: The Photoelectric Effect in Hydrogen
Poisson Brackets
  6.1: Charged Particle in a Magnetic Field
Principle of Least Action
  3.7: Path Integrals
Principle of Least Time
  3.7: Path Integrals
Propagators
  3.5: Propagators and Representations
Pure States
  7.1: The Density Matrix

Q

Quadratic Stark Effect
  9.1: Time-Independent Perturbation Theory
Quantizing Radiation
  9.7: Quantizing Radiation

R

Representations (Quantum)
  3.5: Propagators and Representations

S

S matrix
  10.3: Scattering Amplitudes, Bound States, Resonances
Saddlepoint
  2.4: Complex Variable, Stationary Path Integrals
scalar operator
  4.8: Tensor Operators
Scattering
  10: Scattering Theory
scattering length
  10.2: More Scattering Theory - Partial Waves
Schwartz's inequality
  2.2: Linear Algebra
Spherical Bessel Functions
  10.2: More Scattering Theory - Partial Waves

T

Tensor Operators
  4.8: Tensor Operators
Translation Operator
  3.6: Coherent States
tunneling
  3.1: 1-D Schrödinger Equation - Example Systems
tunneling probability
  3.3: Energy-Time Uncertainty Principle
two photon absorption
  9.5: Time-Dependent Perturbation Theory

V

Van der Waals forces
  9.3: Van Der Waals Forces between Atoms
Variational Method
  8.1: Variational Methods

W

WKB Approximation
  8.2: The WKB Approximation
WKB Connection Formula
  8.3: Note on the WKB Connection Formula

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Quantum Mechanics (Fowler)

Webpages: 68

All licenses found:

- Undeclared: 100% (68 pages)

By Page

- Quantum Mechanics (Fowler) - *Undeclared*
    - Front Matter - *Undeclared*
        - TitlePage - *Undeclared*
        - InfoPage - *Undeclared*
        - Table of Contents - *Undeclared*
        - Licensing - *Undeclared*
    - 1: Reviewing Elementary Stuff - *Undeclared*
        - 1.1: Breakdown of Classical Mechanics - *Undeclared*
        - 1.2: Early Quantum Mechanics - *Undeclared*
        - 1.3: Wave Equations, Wavepackets and Superposition - *Undeclared*
        - 1.4: The Uncertainty Principle - *Undeclared*
        - 1.5: Electron in a Box - *Undeclared*
    - 2: Some Essential Math - *Undeclared*
        - 2.1: Fourier Series and Integrals, the Dirac Function - *Undeclared*
        - 2.2: Linear Algebra - *Undeclared*
        - 2.3: Function Spaces - *Undeclared*
        - 2.4: Complex Variable, Stationary Path Integrals - *Undeclared*
    - 3: Mostly 1-D Quantum Mechanics - *Undeclared*
        - 3.1: 1-D Schrödinger Equation - Example Systems - *Undeclared*
        - 3.2: General Uncertainty Principal - *Undeclared*
        - 3.3: Energy-Time Uncertainty Principle - *Undeclared*
        - 3.4: The Simple Harmonic Oscillator - *Undeclared*
        - 3.5: Propagators and Representations - *Undeclared*
        - 3.6: Coherent States - *Undeclared*
        - 3.7: Path Integrals - *Undeclared*
        - 3.8: Path Integrals for the SHO - *Undeclared*
        - 3.9: Appendix- Some Exponential Operator Algebra - *Undeclared*
    - 4: Angular Momentum, Spin and the Hydrogen Atom - *Undeclared*
        - 4.1: Angular Momentum Operator Algebra - *Undeclared*
        - 4.2: Orbital Eigenfunctions- 2-D Case - *Undeclared*
        - 4.3: Note on Curvilinear Coordinates - *Undeclared*
        - 4.4: Orbital Eigenfunctions in 3-D - *Undeclared*
        - 4.5: Spin - *Undeclared*
        - 4.6: The Hydrogen Atom - *Undeclared*
        - 4.7: Adding Angular Momenta - *Undeclared*
        - 4.8: Tensor Operators - *Undeclared*
    - 5: Interlude - The Nature of Electrons - *Undeclared*
        - 5.1: Bosons and Fermions - *Undeclared*
        - 5.2: Multielectron Atoms - *Undeclared*
    - 6: Charged Particle in Magnetic Field - *Undeclared*
        - 6.1: Charged Particle in a Magnetic Field - *Undeclared*
    - 7: The Density Matrix - *Undeclared*
        - 7.1: The Density Matrix - *Undeclared*
    - 8: Approximate Methods - *Undeclared*
        - 8.1: Variational Methods - *Undeclared*
        - 8.2: The WKB Approximation - *Undeclared*
        - 8.3: Note on the WKB Connection Formula - *Undeclared*
    - 9: Perturbation Theory - *Undeclared*
        - 9.1: Time-Independent Perturbation Theory - *Undeclared*
        - 9.2: The Peierls Transition - an Unexpected Insulator - *Undeclared*
        - 9.3: Van Der Waals Forces between Atoms - *Undeclared*
        - 9.4: The Interaction Representation - *Undeclared*
        - 9.5: Time-Dependent Perturbation Theory - *Undeclared*
        - 9.6: The Photoelectric Effect in Hydrogen - *Undeclared*
        - 9.7: Quantizing Radiation - *Undeclared*
    - 10: Scattering Theory - *Undeclared*
        - 10.1: Scattering Theory - *Undeclared*
        - 10.2: More Scattering Theory - Partial Waves - *Undeclared*
        - 10.3: Scattering Amplitudes, Bound States, Resonances - *Undeclared*
        - 10.4: Identical Particles- Symmetry and Scattering - *Undeclared*

- 11: Other Material - *Undeclared*
    - 11.1: Schrödinger Spreadsheets - *Undeclared*
    - 11.2: Homeworks, exams, etc. - *Undeclared*
- Back Matter - *Undeclared*
    - Index - *Undeclared*
    - Index - *Undeclared*
    - Glossary - *Undeclared*
    - Detailed Licensing - *Undeclared*