The Physics of Quantum Mechanics

Daniel F. Styer

The Physics of Quantum Mechanics

Daniel F. Styer
Schiffer Professor of Physics, Oberlin College

This book is in **draft** form — it is not polished or complete. It needs more problems, it needs sample problems. I appreciate your comments.

Instructions for living a life:
Pay attention.
Be astonished.
Tell about it.

— Mary Oliver, *Sometimes*

Dedicated to all my students: past, present, and future.

Contents

Synoptic Contents

Welcome

What is quantum mechanics and why should I care about it?

1. What Is Quantum Mechanics About?

Classical mechanics is wrong, but what is right? We explore, in the context of modern experiments with qubits, the atomic phenomena that quantum mechanics needs to explain.

2. Forging Mathematical Tools

We build a framework for quantum mechanics, using a mathematical tool called “amplitude”.

3. Refining Mathematical Tools

We build another mathematical tool, the operator, out of amplitude.

4. Formalism

Melding the physics with the mathematical tools.

5. Time Evolution

How do amplitudes change with time?

6. The Quantum Mechanics of Position

The framework, built to treat qubits, extends to treat continuous position as well.

7. Particle in an Infinite Square Well

Our first problem with a particle in one-dimensional space, showing the central role played by energy.

8. The Free Particle

Our second problem with a particle in one-dimensional space, confirming the central role played by energy.

9. Energy Eigenproblems

Since energy plays a central role, we examine it carefully. Can we generalize from these two examples to more general problems? We find that solving particular problems strengthens our conceptual understanding, and that conceptual understanding strengthens our skill in solving particular problems.

10. The Simple Harmonic Oscillator

A third example, one that appears throughout physics, from molecules to field theory.

11. Perturbation Theory

Most problems can't be solved exactly. But approximation schemes including perturbation theory can give superb results even in the absence of an exact solution.

12. More Dimensions, More Particles

We've whet our appetites with a single particle in one dimension. Now we move on to the main feast.

13. Angular Momentum

14. Central Force Problem and a First Look at Hydrogen

15. Identical Particles

This surprisingly subtle topic deserves a chapter of its own.

16. A First Look at Helium

17. Breather

Let's pause in our headlong rush to more realistic, more complex systems. What have we uncovered, what needs to be uncovered in the future?

18. Hydrogen

We apply our new knowledge to physical (rather than model) systems.

19. Helium

20. Atoms

21. Molecules

22. WKB: The Quasiclassical Approximation

23. The Interaction of Matter and Radiation

24. The Territory Ahead

What hasn't this book done?

Welcome

Why would anyone want to study a book titled *The Physics of Quantum Mechanics*?

Starting in the year 1900, physicists exploring the newly discovered atom found that the atomic world of electrons and protons is not just smaller than our familiar world of trees, balls, and automobiles, it is also fundamentally different in character. Objects in the atomic world obey different rules from those obeyed by a tossed ball or an orbiting planet. These atomic rules are so different from the familiar rules of everyday physics, so counterintuitive and unexpected, that it took more than 25 years of intense research to uncover them.

But it is really only since the year 1990 that physicists have come to appreciate that the rules of the atomic world (now called "quantum mechanics") are not just *different* from the everyday rules (now called "classical mechanics"). The atomic rules are also far *richer*. The atomic rules provide for phenomena like particle interference and entanglement that are simply absent from the everyday world. Every phenomenon of classical mechanics is also present in quantum mechanics, but the quantum world provides for many additional phenomena.

Here's an analogy: Some films are in black-and-white and some are in color. It does not malign any black-and-white film to say that a color film has more possibilities, more richness. In fact, black-and-white films are simply one category of color films, because black and white are both colors. Anyone moving from the world of only black-and-white to the world of color is opening up the door to a new world — a world ripe with new possibilities and new expression — without closing the door to the old world.

This same flood of richness and freshness comes from entering the quantum world. It is a difficult world to enter, because we humans have no experience, no intuition, no expectations about this world. Even our language, invented by people living in the everyday world, has no words for the new quantal phenomena — just as a language among a race of the color-blind would have no word for “red”.

Reading this book is not easy: it is like a color-blind student learning about color from a color-blind teacher. The book is just one long argument, building up the structure of a world that we can explore not through touch or through sight or through scent, but only through logic. Those willing to follow and to challenge the logic, to open their minds to a new world, will find themselves richly rewarded.

The place of quantum mechanics in nature

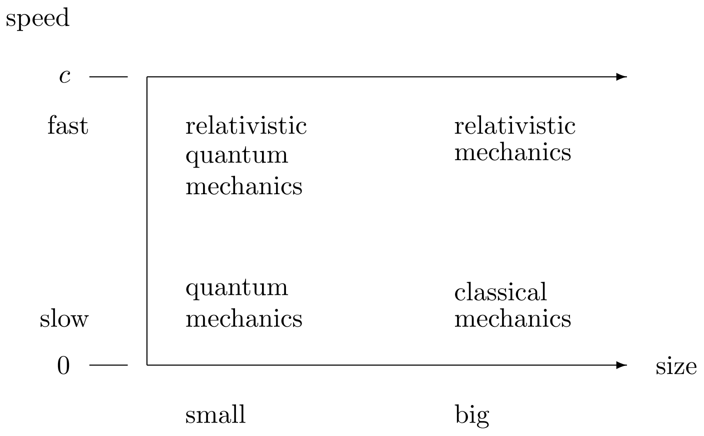
Quantum mechanics is the framework for describing and analyzing small things, like atoms and nuclei. Quantum mechanics also applies to big things, like baseballs and galaxies, but when applied to big things, certain approximations become legitimate: taken together, these are called the *classical approximation* to quantum mechanics, and the result is the familiar classical mechanics.

Quantum mechanics is not only less familiar and less intuitive than classical mechanics; it is also harder than classical mechanics. So whenever the classical approximation is sufficiently accurate, we would be foolish not to use it. This leads some to develop the misimpression that quantum mechanics applies to small things, while classical mechanics applies to big things. No. Quantum mechanics applies to all sizes, but classical mechanics is a good approximation to quantum mechanics when it is applied to big things.

For what size is the classical approximation good enough? That depends on the accuracy desired. The higher the accuracy demanded, the more situations will require full quantal treatment rather than approximate classical treatment. But as a rule of thumb, something as big as a DNA strand is almost always treated classically, not quantum mechanically.

This situation is analogous to the relationship between relativistic mechanics and classical mechanics. Relativity applies always, but classical mechanics is a good approximation to relativistic mechanics when applied to slow things (that is, with speeds much less than light speed c). The speed at which the classical approximation becomes legitimate depends upon the accuracy demanded, but as a rule of thumb particles moving less than a quarter of light speed are treated classically.

The difference between the quantal case and the relativistic case is that while relativistic mechanics is less familiar, less comforting, and less expected than classical mechanics, it is no more intricate than classical mechanics. Quantum mechanics, in contrast, is less familiar, less comforting, less expected, *and* more intricate than classical mechanics. This intricacy makes quantum mechanics harder than classical mechanics, yes, but also richer, more textured, more nuanced. Whether to curse or celebrate this intricacy is your choice.



Finally, is there a framework that applies to situations that are both fast and small? There is: it is called "relativistic quantum mechanics" and is closely related to "quantum field theory". Ordinary non-relativistic quantum mechanics is a good approximation for relativistic quantum mechanics when applied to slow things. Relativistic mechanics is a good approximation for relativistic quantum mechanics when applied to big things. And classical mechanics is a good approximation for relativistic quantum mechanics when applied to big, slow things.

What you can expect from this book

This book introduces quantum mechanics at the third- or fourth-year American undergraduate level. It assumes the reader knows about. . . .

This is a book about physics, not about mathematics. The word "physics" derives from the Greek word for "nature", so the emphasis lies in nature, not in the mathematics we use to describe nature. Thus the book starts with experiments about nature, then builds mathematical machinery to describe nature, then erects a formalism ("postulates"), and then moves on to applications, where the formalism is applied to nature and where insight into both nature and formalism is deepened.

The book never abandons its focus on nature. It provides a balanced, interwoven treatment of concepts, formalism, and applications so that each strand reinforces the other. (The three candles on the cover represent these three strands.) Without doubt, quantum mechanics is both beautiful and difficult. I have been pursuing quantum mechanics for more than fifty years — questioning, experimenting, calculating, simulating, reading, writing, pondering, proving, exploring, teaching — and I still find it shocking.

There are both "exercises" and "problems". The exercises are interwoven with the text, and provide checks to see whether you understand the material or are just skimming uncritically. Most are quick, and many can be performed in your head without benefit of pen and paper. You should do all the exercises as you come to them.

The problems are placed at the end of the pertinent section or chapter. There are many problems at many levels of difficulty, but no problem is there just for "make-work": each has a "moral to the story". Some problems are essential to the logical development of the subject: these are labeled (unsurprisingly) "essential". Other problems promote learning far better than simple reading can: these are labeled "recommended". A few problems, called "projects", are suggestions for open-ended explorations. Sample problems build both mathematical technique and physical insight.

Most physics textbooks face the quandary: Should exposition go from general to specific or vice versa? Richard Feynman asks this question in his book *Statistical Mechanics.* On the first page[1] he writes out a fundamental

---

[1] R.P. Feynman, *Statistical Mechanics: A Set of Lectures* (W.A. Benjamin, Reading, Massachusetts, 1972) page 1.

law, then writes "This fundamental law is the summit of statistical mechanics, and the entire subject is either the slide-down from this summit, as the principle is applied to various cases, or the climb-up to where the fundamental law is derived and the concepts of thermal equilibrium and temperature T clarified."

This book uses neither strategy: It begins with one specific system — the magnetic moment of a silver atom — and introduces the central quantities of amplitude and state and operator as they apply to that system. It then gives the general structure ("formalism") for quantum mechanics and, once that's in place, applies the general results to many and various systems.[2]

The book does not merely convey correct ideas, it also refutes misconceptions. Just to get started, I list the most important and most pernicious misconceptions about quantum mechanics: (a) An electron has a position but you don't know what it is. (b) The only states are energy states. (c) The wavefunction $\psi(\vec{r}, t)$ is "out there" in space and you could reach out and touch it if only your fingers were sufficiently sensitive.

The object of the biographical footnotes in this book is twofold: First, to present the briefest of outlines of the subject's historical development, lest anyone get the misimpression that quantum mechanics arose fully formed, like Aphrodite from sea foam. Second, to show that the founders of quantum mechanics were not inaccessible giants, but people with foibles and strengths, with interests both inside and outside of physics, just like you and me.

---

[2] As a child growing up on a farm, I became familiar, one by one, with many wildflowers and field crops. When I took a course on plant taxonomy in college, I learned a scheme that organized all of my familiarity into a structure of plant "families". It was easy for me to learn the characteristics of the *Caryophyllaceae* family, for example, because I already knew the wildflower Chickweed, a member of that family. Similarly for the *Rosaceae* and the Apple blossom. Once I knew the structure, it was easy for me to learn new species, not one-by-one, but by fitting them into that overarching structure. Other students in the class lacked my familiarity with individual flower species, so the general structure we all learned, which seemed to me natural and organic, seemed to them arbitrary and contrived. They were never able to fit new species into it. My intent in this book is to build your understanding of quantum mechanics in a similar pattern of organic growth.

Acknowledgments

I learned quantum mechanics from stellar teachers. My high school chemistry teacher Frank Dugan introduced me not only to quantum mechanics but to the precept that science involves hard, fulfilling work in addition to dreams and imagination. When I was an undergraduate, John Boccio helped mold my understanding of quantum mechanics, and also molded the shape of my life. In graduate school N. David Mermin, Vinay Ambegaokar, Neil Ashcroft, Michael Peskin, and Kurt Gottfried pushed me without mercy but pushed me in the direction of understanding and away from the mind-numbing attitude of "shut up and calculate". My debt to my thesis adviser, Michael Fisher, is incalculable. I've been inspired by research lectures from Tony Leggett, Jürg Fröhlich, Jennifer and Lincoln Chayes, Shelly Goldstein, and Chris Fuchs, among others.

I have taught quantum mechanics to thousands of students from the general audience level through advanced undergraduates. Their questions, confusions, triumphs, and despairs have infused my own understanding of the discipline. I cannot name them all, but I would be remiss if I did not thank my former students Paul Kimoto, Gail Welsh, K. Tabetha Hole, Gary Felder, Sarah Clemmens, Dahyeon Lee, Victor Wong, Noah Morris, Avay Subedi, and Shuran Zhu.

My scientific prose style was developed by Michael Fisher and N. David Mermin. In particular this book's structure of "first lay out the phenomena (chapter 1), then build mathematical tools to describe those phenomena" echos the structure of Fisher's 1964 essay "The Nature of Critical Points". I have also absorbed lessons in writing from John McPhee, Maurice Forrester, and Terry Tempest Williams. My teaching style has been influenced especially by Mark Heald, Tony French, Edwin Taylor, Arnold Arons, and Robert H. Romer.

Chapter 1

What is Quantum Mechanics About?

1.1 Quantization

We are used to things that vary continuously: An oven can take on any temperature, a recipe might call for any quantity of flour, a child can grow to a range of heights. If I told you that an oven might take on the temperature of 172.1 °C or 181.7 °C, but that a temperature of 173.8 °C was physically impossible, you would laugh in my face.

So you can imagine the surprise of physicists on 14 December 1900, when Max Planck announced that certain features of blackbody radiation (that is, of light in thermal equilibrium) could be explained by assuming that the energy of the light could *not* take on any value, but only certain discrete values. Specifically, Planck found that light of frequency ω could take on only the energies of

$$E = \hbar\omega(n + \tfrac{1}{2}), \qquad \text{where } n = 0, 1, 2, 3, \ldots, \tag{1.1}$$

and where the constant $\hbar$ (now called the "reduced Planck constant") is

$$\hbar = 1.054\,571\,817 \times 10^{-34} \text{ J s.} \tag{1.2}$$

(I use modern terminology and the current value for $\hbar$, rather than the terminology and value used by Planck in 1900.)

That is, light of frequency ω can have an energy of $3.5\,\hbar\omega$, and it can have an energy of $4.5\,\hbar\omega$, but it is physically impossible for this light to have an energy of $3.8\,\hbar\omega$. Any numerical quantity that can take on only discrete values like this is called "quantized". By contrast, a numerical quantity that can take on any value is called "continuous".

The photoelectric effect supplies additional evidence that the energy of light comes only in discrete values. And if the energy of light comes in discrete values, then it's a good guess that the energy of an atom comes in discrete values too. This good guess was confirmed through investigations of atomic spectra (where energy goes into or out of an atom via absorption or emission of light) and through the Franck–Hertz experiment (where energy goes into or out of an atom via collisions).

Furthermore, if the energy of an atom comes in discrete values, then it's a good guess that other properties of an atom — such as its magnetic moment — also take on only discrete values. The theme of this book is that these good guesses have all proved to be correct.

The story of Planck's[1] discovery is a fascinating one, but it's a difficult and elaborate story because it involves not just quantization, but also thermal equilibrium and electromagnetic radiation. The story of the discovery of atomic energy quantization is just as fascinating, but again fraught with intricacies. In an effort to remove the extraneous and dive deep to the heart of the matter, we focus on the magnetic moment of an atom. We will, to the extent possible, do a quantum-mechanical treatment of an atom's magnetic moment while maintaining a classical treatment of all other aspects — such as its energy and momentum and position. (In chapter 6, "The Quantum Mechanics of Position", we take up a quantum-mechanical treatment of position, momentum, and energy.)

1.1.1 *The Stern-Gerlach experiment*

An electric current flowing in a loop produces a magnetic moment, so it makes sense that the electron orbiting (or whatever it does) an atomic nucleus would produce a magnetic moment for that atom. And of course, it also makes sense that physicists would be itching to measure that magnetic moment.

It is not difficult to measure the magnetic moment of, say, a scout compass. Place the magnetic compass needle in a known magnetic field and measure the torque that acts to align the needle with the field. You

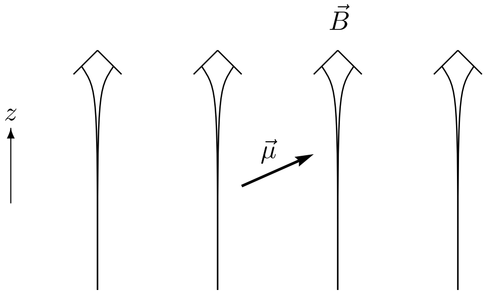
---

[1] Max Karl Ernst Ludwig Planck (1858–1947) was a German theoretical physicist particularly interested in thermodynamics and radiation. Concerning his greatest discovery, the introduction of quantization into physics, he wrote, "I can characterize the whole procedure as an act of desperation, since, by nature I am peaceable and opposed to doubtful adventures." [Letter from Planck to R.W. Wood, 7 October 1931, quoted in J. Mehra and H. Rechenberg, *The Historical Development of Quantum Theory* (Springer–Verlag, New York, 1982) volume 1, page 49.]

will need to measure an angle and you might need to look up a formula in your magnetism textbook, but there is no fundamental difficulty.

Measuring the magnetic moment of an atom is a different matter. You can't even see an atom, so you can't watch it twist in a magnetic field like a compass needle. Furthermore, because the atom is very small, you expect the associated magnetic moment to be very small, and hence very hard to measure. The technical difficulties are immense.

These difficulties must have deterred but certainly did not stop Otto Stern and Walter Gerlach.[2] They realized that the twisting of a magnetic moment in a *uniform* magnetic field could not be observed for atomic-sized magnets, and also that the moment would experience zero net force. But they also realized that a magnetic moment in a *non-uniform* magnetic field *would* experience a net force, and that this force could be used to measure the magnetic moment.



A classical magnetic moment in a non-uniform magnetic field.

A classical magnetic moment $\vec{\mu}$, situated in a magnetic field $\vec{B}$ that points in the z direction and increases in magnitude in the z direction, is subject to a force

$$\mu_z \frac{\partial B}{\partial z}, \tag{1.3}$$

where μ_z is the z-component of the magnetic moment or, in other words, the projection of $\vec{\mu}$ on the z axis. (If this is not obvious to you, then work problem 1.1, "Force on a classical magnetic moment", on page 11.)

---

[2] Otto Stern (1888–1969) was a Polish-German-Jewish physicist who made contributions to both theory and experiment. He left Germany for the United States in 1933 upon the Nazi ascension to power. Walter Gerlach (1889–1979) was a German experimental physicist. During the Second World War he led the physics section of the Reich Research Council and for a time directed the German effort to build a nuclear bomb.

Stern and Gerlach used this fact to measure the z-component of the magnetic moment of an atom. First, they heated silver in an electric "oven". The vaporized silver atoms emerged from a pinhole in one side of the oven, and then passed through a non-uniform magnetic field. At the far side of the field the atoms struck and stuck to a glass plate. The entire apparatus had to be sealed within a good vacuum, so that collisions with nitrogen molecules would not push the silver atoms around. The deflection of an atom away from straight-line motion is proportional to the magnetic force, and hence proportional to the projection μ_z. In this ingenious way, Stern and Gerlach could measure the z-component of the magnetic moment of an atom even though any single atom is invisible.

Before reading on, pause and think about what results you would expect from this experiment.

Here are the results that I expect: I expect that an atom which happens to enter the field with magnetic moment pointing straight up (in the z direction) will experience a large upward force. Hence it will move upward and stick high up on the glass-plate detector. I expect that an atom which happens to enter with magnetic moment pointing straight down (in the $-z$ direction) will experience a large downward force, and hence will stick far down on the glass plate. I expect that an atom entering with magnetic moment tilted upward, but not straight upward, will move upward but not as far up as the straight-up atoms, and the mirror image for an atom entering with magnetic moment tilted downward. I expect that an atom entering with horizontal magnetic moment will experience a net force of zero, so it will pass through the non-uniform field undeflected.

Furthermore, I expect that when a silver atom emerges from the oven source, its magnetic moment will be oriented randomly — as likely to point in one direction as in any other. There is only one way to point straight up, so I expect that very few atoms will stick high on the glass plate. There are many ways to point horizontally, so I expect many atoms to pass through undeflected. There is only one way to point straight down, so I expect very few atoms to stick far down on the glass plate.[3]

In summary, I expect that atoms would leave the magnetic field in any of a range of deflections: a very few with large positive deflection, more with a

---

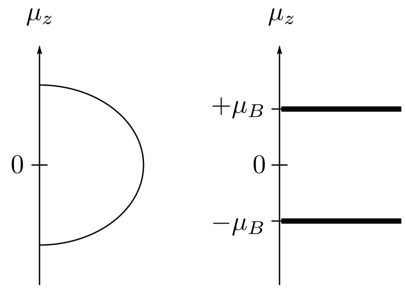
[3] To be specific, this reasoning suggests that the number of atoms with moment tilted at angle θ relative to the z direction is proportional to $\sin\theta$, where θ ranges from 0° to 180°. You might want to prove this to yourself, but we'll never use this result so don't feel compelled.

small positive deflection, a lot with no deflection, some with a small negative deflection, and a very few with large negative deflection. This continuity of deflections reflects a continuity of magnetic moment projections.

In fact, however, this is not what happens at all! The projection μ_z does not take on a continuous range of values. Instead, it is quantized and takes on only two values, one positive and one negative. Those two values are called $\mu_z = \pm\mu_B$ where μ_B, the so-called "Bohr magneton", has the measured value of

$$\mu_B = 9.274\,010\,078 \times 10^{-24} \text{ J/T}, \tag{1.4}$$

with an uncertainty of 3 in the last decimal digit.



Distribution of μ_z

Expected:

Actual:

The Stern-Gerlach experiment was initially performed with silver atoms but has been repeated with many other types of atoms. When nitrogen is used, the projection μ_z takes on one of the four quantized values of $+3\mu_B$, $+\mu_B$, $-\mu_B$, or $-3\mu_B$. When sulfur is used, it takes on one of the five quantized values of $+4\mu_B$, $+2\mu_B$, 0, $-2\mu_B$, and $-4\mu_B$. For no atom do the values of μ_z take on the broad continuum of my classical expectation. For all atoms, the projection μ_z is quantized.

Problems

1.1 **Force on a classical magnetic moment**

The force on a classical magnetic moment is most easily calculated using "magnetic charge fiction": Consider the magnetic moment to consist of two "magnetic charges" of magnitude $+m$ and $-m$, separated by the position vector $\vec{d}$ running from $-m$ to $+m$. The magnetic moment is then $\vec{\mu} = m\vec{d}$.

a. Use B_+ for the magnitude of the magnetic field at $+m$, and B_- for the magnitude of the magnetic field at $-m$. Show that the net force on the magnetic moment is in the z direction with magnitude $mB_+ - mB_-$.
b. Use d_z for the z-component of the vector $\vec{d}$. Show that to high accuracy

$$B_+ = B_- + \frac{\partial B}{\partial z} d_z.$$

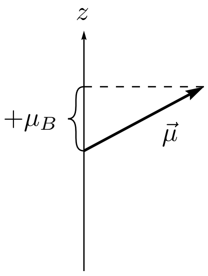
Surely, for distances of atomic scale, this accuracy is more than adequate.
c. Derive expression (1.3) for the force on a magnetic moment.

1.1.2 *The conundrum of projections*

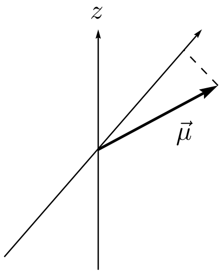
I would expect the projections μ_z of a silver atom to take on a continuous range of values. But in fact, these values are quantized: Whenever μ_z is measured, it turns out to be either $+\mu_B$ or $-\mu_B$, and never anything else. This is counterintuitive and unexpected, but we can live with the counterintuitive and unexpected — it happens all the time in politics.

However, this fact of quantization appears to result in a logical contradiction, because there are many possible axes upon which the magnetic moment can be projected. The figures on the next page make it clear that it is impossible for *any* vector to have a projection of either $\pm\mu_B$ on *all* axes!

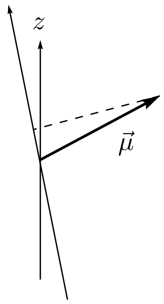
Because if the projection of $\vec{\mu}$ on the z axis is $+\mu_B$ …



…then the projection of $\vec{\mu}$ on this second axis must be more than $+\mu_B$ …

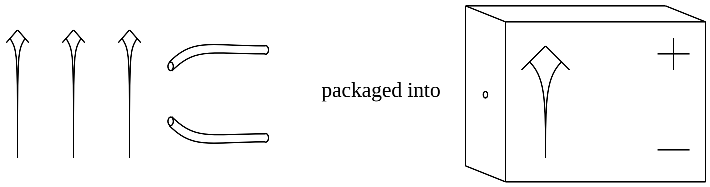


…while the projection of $\vec{\mu}$ on this third axis must be less than $+\mu_B$.

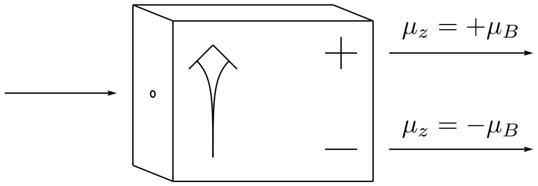


Whenever we measure the magnetic moment, projected onto any axis, the result is either $+\mu_B$ or $-\mu_B$. Yet is it impossible for the projection of any classical arrow on *all* axes to be either $+\mu_B$ or $-\mu_B$! This seeming contradiction is called “the conundrum of projections”. We can live with the counterintuitive, the unexpected, the strange, but we cannot live with a logical contradiction. How can we resolve it?

The resolution comes not from meditating on the question, but from experimenting about it. Let us actually measure the projection on one axis, and then on a second. To do this easily, we modify the Stern-Gerlach apparatus and package it into a box called a “Stern-Gerlach analyzer”. This box consists of a Stern-Gerlach apparatus followed by “pipes” that channel the outgoing atoms into horizontal paths.[4] This chapter treats only silver atoms, so we use analyzers with two exit ports.



An atom enters a vertical analyzer through the single hole on the left. If it exits through the upper hole on the right (the “+ port”) then the outgoing atom has $\mu_z = +\mu_B$. If it exits through the lower hole on the right (the “− port”) then the outgoing atom has $\mu_z = -\mu_B$.

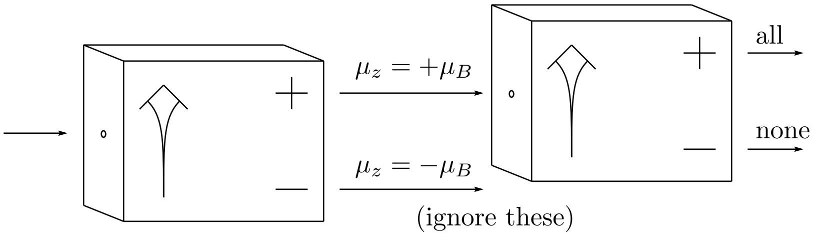


---

[4] In general, the “pipes” will manipulate the atoms through electromagnetic fields, not through touching. One way way to make such “pipes” is to insert a second Stern-Gerlach apparatus, oriented upside-down relative to the first. The atoms with $\mu_z = +\mu_B$, which had experienced an upward force in the first half, will experience an equal downward force in the second half, and the net impulse delivered will be zero. But whatever their manner of construction, the pipes must *not* change the magnetic moment of an atom passing through them.

1.1.3 *Two vertical analyzers*

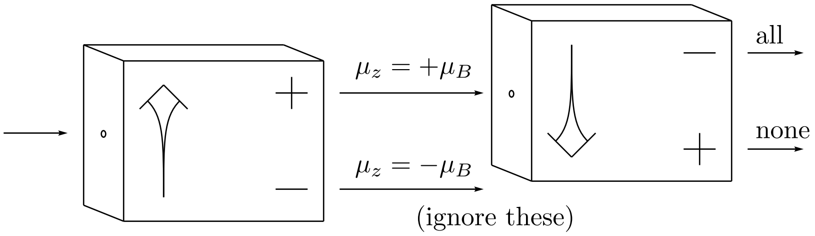
In order to check the operation of our analyzers, we do preliminary experiments. Atoms are fed into a vertical analyzer. Any atom exiting from the + port is then channeled into a second vertical analyzer. That atom exits from the + port of the second analyzer. This makes sense: the atom had $\mu_z = +\mu_B$ when exiting the first analyzer, and the second analyzer confirms that it has $\mu_z = +\mu_B$.



Furthermore, if an atom exiting from the − port of the first analyzer is channeled into a second vertical analyzer, then that atom exits from the − port of the second analyzer.

1.1.4 *One vertical and one upside-down analyzer*

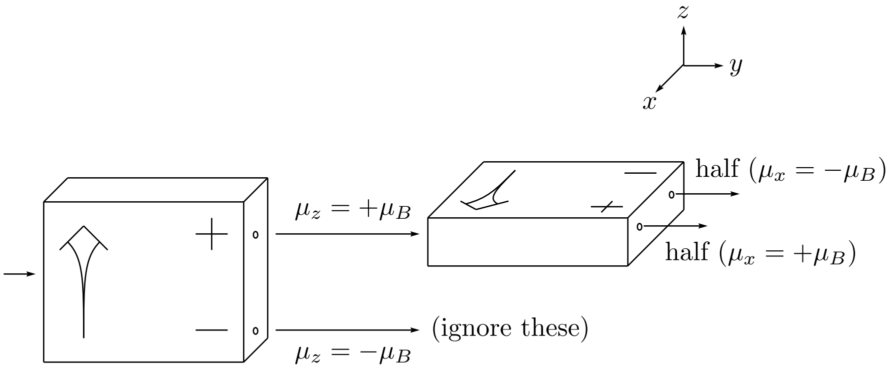
Atoms are fed into a vertical analyzer. Any atom exiting from the + port is then channeled into a second analyzer, but this analyzer is oriented *upside-down*. What happens? If the projection on an upward-pointing axis is $+\mu_B$ (that is, $\mu_z = +\mu_B$), then the projection on a downward-pointing axis is $-\mu_B$ (we write this as $\mu_{(-z)} = -\mu_B$). So I expect that these atoms will emerge from the − port of the second analyzer (which happens to be the higher port). And this is exactly what happens.



Similarly, if an atom exiting from the − port of the first analyzer is channeled into an upside-down analyzer, then that atom emerges from the + port of the second analyzer.

1.1.5 *One vertical and one horizontal analyzer*

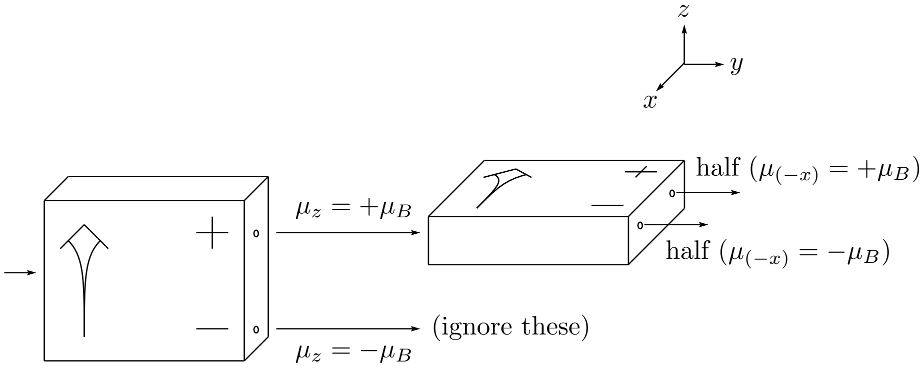
Atoms are fed into a vertical analyzer. Any atom exiting from the + port is then channeled into a second analyzer, but this analyzer is oriented *horizontally.* The second analyzer doesn't measure the projection μ_z, it measures the projection μ_x. What happens in this case? Experiment shows that the atoms emerge randomly: half from the + port, half from the − port.



This makes some sort of sense. If a classical magnetic moment were vertically oriented, it would have $\mu_x = 0$, and such a classical moment would go straight through a horizontal Stern-Gerlach analyzer. We've seen that atomic magnetic moments *never* go straight through. If you "want" to go straight but are forced to turn either left or right, the best you can do is turn left half the time and right half the time. (Don't take this paragraph literally…atoms have no personalities and they don't "want" anything. But it is a useful mnemonic.)

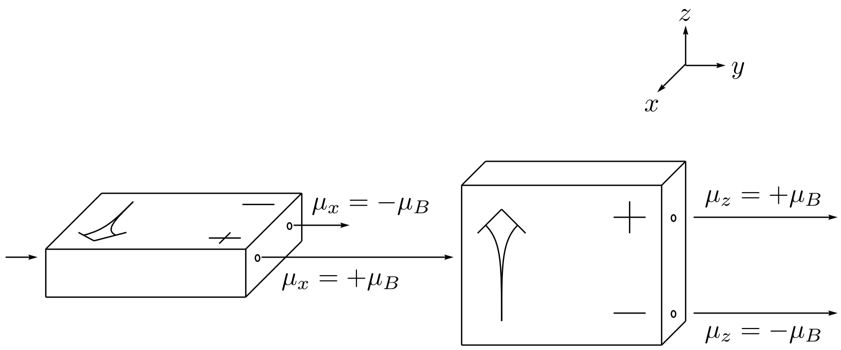
1.1.6 *One vertical and one backwards horizontal analyzer*

Perform the same experiment as above (section 1.1.5), except insert the horizontal analyzer in the opposite sense, so that it measures the projection on the negative x axis rather than the positive x axis. Again, half the atoms emerge from the $+$ port, and half emerge from the $-$ port.



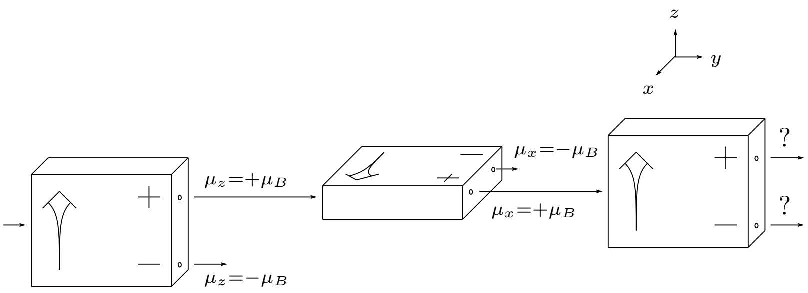
1.1.7 *One horizontal and one vertical analyzer*

A $+x$ analyzer followed by a $+z$ analyzer is the same apparatus as above (section 1.1.6), except that both analyzers are rotated as a unit by 90° about the y axis. So of course it has the same result: half the atoms emerge from the $+$ port, and half emerge from the $-$ port.



1.1.8 *Three analyzers*

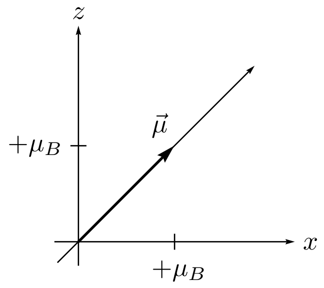
Atoms are fed into a vertical analyzer. Any atom exiting from the + port is then channeled into a horizontal analyzer. Half of these atoms exit from the + port of the horizontal analyzer (see section 1.1.5), and *these* atoms are channeled into a third analyzer, oriented vertically. What happens at the third analyzer?



There are two ways to think of this: (I) When the atom emerged from the + port of the first analyzer, it was determined to have $\mu_z = +\mu_B$. When that same atom emerged from the + port of the second analyzer, it was determined to have $\mu_x = +\mu_B$. Now we know two projections of the magnetic moment. When it enters the third analyzer, it still has $\mu_z = +\mu_B$, so it will emerge from the + port. (II) The last two analyzers in this sequence are a horizontal analyzer followed by a vertical analyzer, and from section 1.1.7 we know what happens in this case: a 50/50 split. That will happen in this case, too.

So, analysis (I) predicts that all the atoms entering the third analyzer will exit through the + port and none through the − port. Analysis (II) predicts that half the atoms will exit through the + port and half through the − port.

Experiment shows that analysis (II) gives the correct result. But what could possibly be wrong with analysis (I)? Let's go through line by line: "When the atom emerged from the + port of the first analyzer, it was determined to have $\mu_z = +\mu_B$." Nothing wrong here — this is what an analyzer does. "When that same atom emerged from the + port of the second analyzer, it was determined to have $\mu_x = +\mu_B$." Ditto. "Now we know two projections of the magnetic moment." This has got to be the problem. To underscore that problem, look at the figure below.



If an atom *did* have both $\mu_z = +\mu_B$ *and* $\mu_x = +\mu_B$, then the projection on an axis rotated 45° from the vertical would be $\mu_{45^\circ} = +\sqrt{2}\,\mu_B$. But the Stern-Gerlach experiment assures us that whenever μ_{45° is measured, the result is either $+\mu_B$ or $-\mu_B$, and never $+\sqrt{2}\,\mu_B$. In summary, *it is not possible for a moment to have a projection on both the z axis and on the x axis.* Passing to the fourth sentence of analysis (I) — "When the atom enters the third analyzer, it still has $\mu_z = +\mu_B$, so it will emerge from the + port" — we immediately see the problem. The atom emerging from the + port of the second analyzer does *not* have $\mu_z = +\mu_B$ — it doesn't have a projection on the z axis at all.

Because it's easy to fall into misconceptions, let me emphasize what I'm saying and what I'm not saying:

> I'm saying that if an atom has a value for μ_x, then it doesn't have a value for μ_z.
>
> I'm not saying that the atom has a value for μ_z but no one knows what it is.
>
> I'm not saying that the atom has a value for μ_z but that value is changing rapidly.
>
> I'm not saying that the atom has a value for μ_z but that value is changing unpredictably.
>
> I'm not saying that a random half of such atoms have the value $\mu_z = +\mu_B$ and the other half have the value $\mu_z = -\mu_B$.
>
> I'm not saying that the atom has a value for μ_z which will be disturbed upon measurement.

The atom with a value for μ_x *does not have a value for* μ_z in the same way that *love does not have a color.*

This is a new phenomenon, and it deserves a new name. That name is "indeterminacy". This is perhaps not the best name, because it might suggest, incorrectly, that an atom with a value for μ_x has a value for μ_z and we merely haven't yet determined what that value is. The English language was invented by people who didn't understand quantum mechanics, so it is not surprising that there are no perfectly appropriate names for quantum mechanical phenomena. This is a defect in our language, not a defect in quantum mechanics or in our understanding of quantum mechanics, and it is certainly not a defect in nature.[5]

How can a vector have a projection on one axis but not on another? It is the job of the rest of this book to answer that question, [6] but one thing is clear already: The visualization of an atomic magnetic moment as a classical arrow must be wrong.

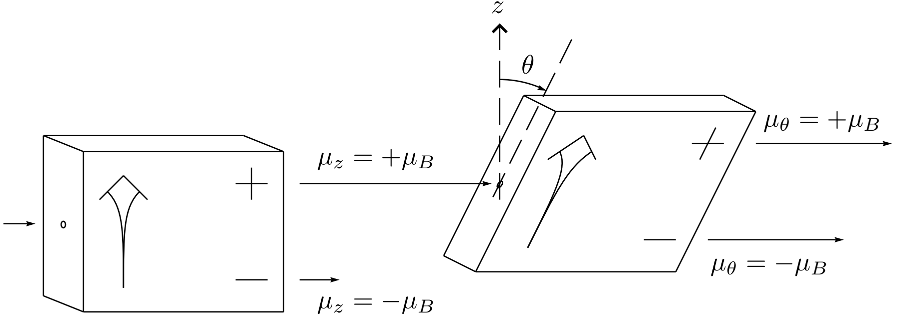
---

[5] In exactly the same manner, the name "orange" applies to light within the wavelength range 590–620 nm and the name"red" applies to light within the wavelength range 620–740 nm, but the English language has no word to distinguish the wavelength range 1590–1620 nm from the wavelength range 1620–1740 nm. This is not because optical light is "better" or "more deserving" than infrared light. It is due merely to the accident that our eyes detect optical light but not infrared light.

[6] Preview: In quantum mechanics, the magnetic moment is represented mathematically not by a vector but by a vector operator.

1.1.9 *The upshot*

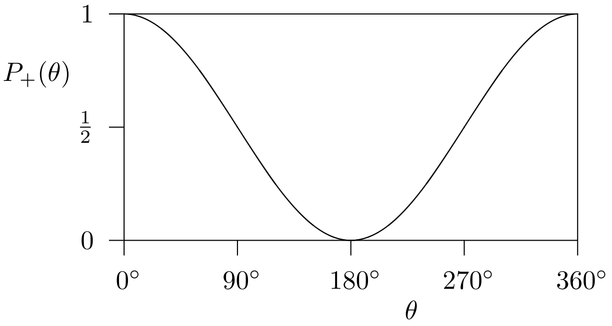
We escape from the conundrum of projections through probability. If an atom has $\mu_z = +\mu_B$, and if the projection on some other axis is measured, then the result cannot be predicted with certainty: we instead give probabilities for the various results. If the second analyzer is rotated by angle θ relative to the vertical, the probability of emerging from the + port of the second analyzer is called $P_+(\theta)$.



We already know some special values: from section 1.1.3, $P_+(0°) = 1$; from section 1.1.5, $P_+(90°) = \frac{1}{2}$; from section 1.1.4, $P_+(180°) = 0$; from section 1.1.6, $P_+(270°) = \frac{1}{2}$; from section 1.1.3, $P_+(360°) = 1$. It is not hard to guess the curve that interpolates between these values:

$$P_+(\theta) = \cos^2(\theta/2), \tag{1.5}$$

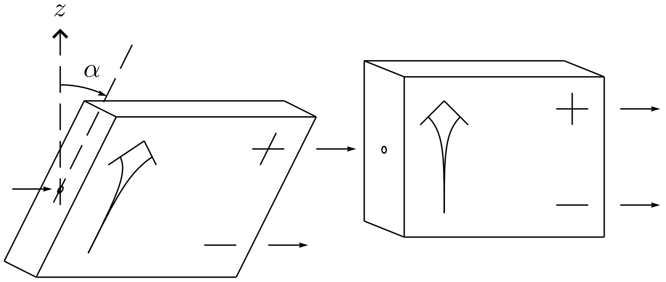
and experiment confirms this guess.



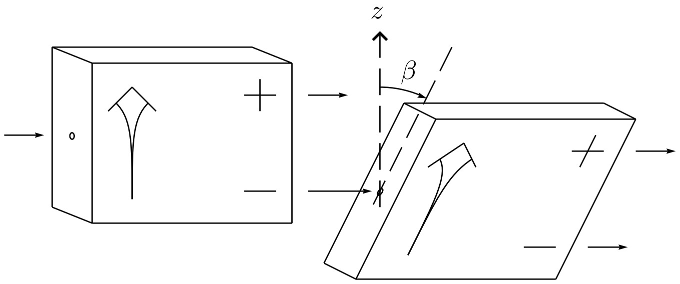
Problems

1.2 **Exit probabilities** (essential problem)

a. An analyzer is tilted from the vertical by angle α. An atom leaving its + port is channeled into a vertical analyzer. What is the probability that this atom emerges from the + port? The − port? (*Clue:* Use the "rotate as a unit" concept introduced in section 1.1.7.)

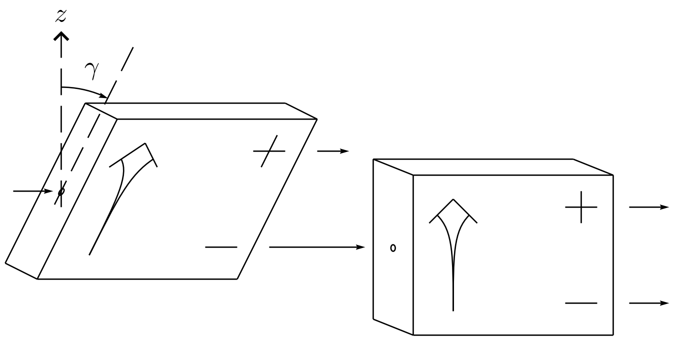


b. An atom exiting the − port of a vertical analyzer behaves exactly like one exiting the + port of an upside-down analyzer (see section 1.1.4). Such an atom is channeled into an analyzer tilted from the vertical by angle β. What is the probability that this atom emerges from the + port? The − port?



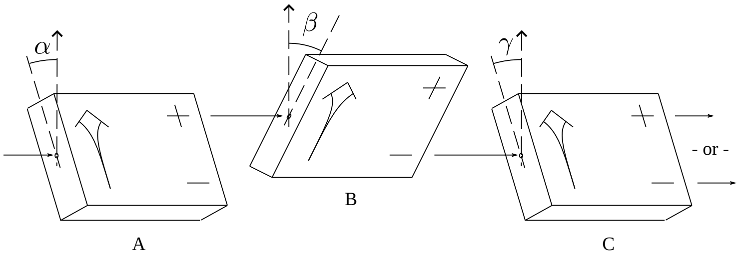
(Problem continues on next page.)

c. An analyzer is tilted from the vertical by angle γ. An atom leaving its $-$ port is channeled into a vertical analyzer. What is the probability that this atom emerges from the + port? The $-$ port?



1.3 Multiple analyzers

An atom with $\mu_z = +\mu_B$ is channeled through the following line of three Stern-Gerlach analyzers.



Find the probability that it emerges from (a) the $-$ port of analyzer A; (b) the + port of analyzer B; (c) the + port of analyzer C; (d) the $-$ port of analyzer C.

1.4 Properties of the $P_+(\theta)$ function

a. An atom exits the + port of a vertical analyzer; that is, it has $\mu_z = +\mu_B$. Argue that the probability of this atom exiting from the $-$ port of a θ analyzer is the same as the probability of it exiting from the + port of a $(180° - \theta)$ analyzer.

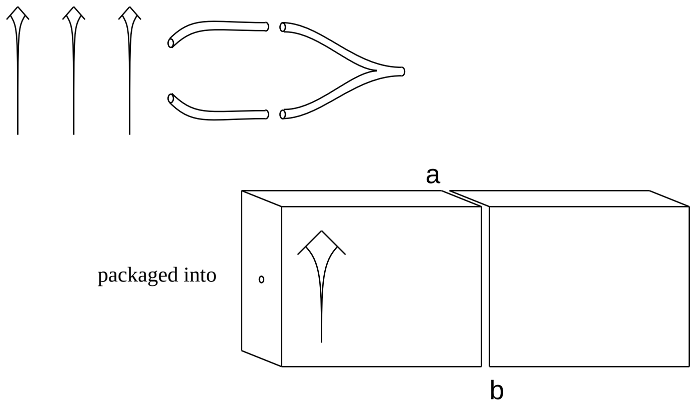
b. Conclude that the $P_+(\theta)$ function introduced in section 1.1.9 must satisfy

$$P_+(\theta) + P_+(180^\circ - \theta) = 1.$$

c. Does the experimental result (1.5) satisfy this condition?

1.2 Interference

There are more quantum mechanical phenomena to uncover. To support our exploration, we build a new experimental device called the "analyzer loop".[7] This is nothing but a Stern-Gerlach analyzer followed by "piping" that channels the two exit paths together again.[8]



The device must be constructed to high precision, so that there can be no way to distinguish whether the atom passed through by way of the top

---

[7] We build it in our minds. The experiments described in this section have never been performed exactly as described here, although researchers are getting close. [See Shimon Machluf, Yonathan Japha, and Ron Folman, "Coherent Stern–Gerlach momentum splitting on an atom chip" *Nature Communications* **4** (9 September 2013) 2424.] We know the results that would come from these experiments because conceptually parallel (but more complex!) experiments *have* been performed on photons, neutrons, atoms, and molecules.

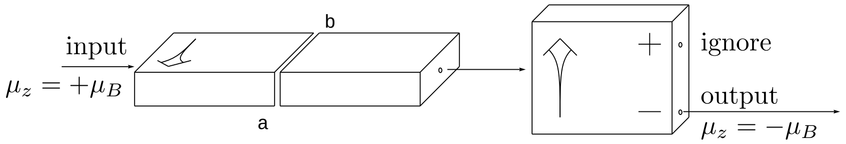
[8] If you followed the footnote on page 14, you will recall that these "pipes" manipulate atoms through electromagnetic fields, not through touching. One way to make them would be to insert two more Stern-Gerlach apparatuses, the first one upside-down and the second one rightside-up relative to the initial apparatus. But whatever the manner of their construction, the pipes must *not* change the magnetic moment of an atom passing through them.

path or the bottom path. For example, the two paths must have the same length: If the top path were longer, then an atom going through via the top path would take more time, and hence there *would* be a way to tell which way the atom passed through the analyzer loop.

In fact, the analyzer loop is constructed so precisely that it doesn't change the character of the atom passing through it. If the atom enters with $\mu_z = +\mu_B$, it exits with $\mu_z = +\mu_B$. If it enters with $\mu_x = -\mu_B$, it exits with $\mu_x = -\mu_B$. If it enters with $\mu_{17^\circ} = -\mu_B$, it exits with $\mu_{17^\circ} = -\mu_B$. It is hard to see why anyone would want to build such a device, because they're expensive (due to the precision demands), and they do absolutely nothing!

Once you made one, however, you could convert it into something useful. For example, you could insert a piece of metal blocking path `a`. In that case, all the atoms exiting would have taken path `b`, so (if the analyzer loop were oriented vertically) all would emerge with $\mu_z = -\mu_B$.

Using the analyzer loop, we set up the following apparatus: First, channel atoms with $\mu_z = +\mu_B$ into a horizontal analyzer loop.[9] Then, channel the atoms emerging from that analyzer loop into a vertical analyzer. Ignore atoms emerging from the + port of the vertical analyzer and look for atoms emerging from the − port.



We execute three experiments with this set-up: first we pass atoms through when path `a` is blocked, then when path `b` is blocked, finally when neither path is blocked.

1.2.1 *Path* `a` *blocked*

(1) Atoms enter the analyzer loop with $\mu_z = +\mu_B$.
(2) Half of them attempt path `a`, and end up impaled on the blockage.
(3) The other half take path `b`, and emerge from the analyzer loop with $\mu_x = -\mu_B$.

---

[9] To make sure that all of these atoms have $\mu_z = +\mu_B$, they are harvested from the + port of a vertical analyzer.

(4) Those atoms then enter the vertical analyzer. Similar to the result of section 1.1.7, half of these atoms emerge from the + port and are ignored. Half of them emerge from the − port and are counted.
(5) The overall probability of passing through the set-up is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

If you perform this experiment, you will find that this analysis is correct and that these results are indeed obtained.

1.2.2 *Path* b *blocked*

(1) Atoms enter the analyzer loop with $\mu_z = +\mu_B$.
(2) Half of them attempt path b, and end up impaled on the blockage.
(3) The other half take path a, and emerge from the analyzer loop with $\mu_x = +\mu_B$.
(4) Those atoms then enter the vertical analyzer. Exactly as in section 1.1.7, half of these atoms emerge from the + port and are ignored. Half of them emerge from the − port and are counted.
(5) The overall probability of passing through the set-up is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

Once again, experiment confirms these results.

1.2.3 *Neither path blocked*

Here, I have not just one, but two ways to analyze the experiment:

Analysis I:

(1) An atom passes through the set-up either via path b or via path a.
(2) From section 1.2.1, the probability of passing through via path b is $\frac{1}{4}$.
(3) From section 1.2.2, the probability of passing through via path a is $\frac{1}{4}$.
(4) Thus the probability of passing through the entire set-up is $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

Analysis II:

(1) Because "the analyzer loop is constructed so precisely that it doesn't change the character of the atom passing through it", the atom emerges from the analyzer loop with $\mu_z = +\mu_B$.
(2) When such atoms enter the vertical analyzer, all of them emerge through the + port. (See section 1.1.3.)

(3) Thus the probability of passing through the entire set-up is zero.

These two analyses cannot both be correct. Experiment confirms the result of analysis II, but what could possibly be wrong with analysis I? Item (2) is already confirmed through the experiment of section 1.2.1, item (3) is already confirmed through the experiment of section 1.2.2, and don't tell me that I made a mistake in the arithmetic of item (4). The only thing left is item (1): "An atom passes through the set-up either via path b or via path a." This simple, appealing, common-sense statement must be *wrong*!

Just a moment ago, the analyzer loop seemed like a waste of money and skill. But in fact, a horizontal analyzer loop is an extremely clever way of correlating the path through the analyzer loop with the value of μ_x: If the atom has $\mu_x = +\mu_B$, then it takes path a. If the atom has $\mu_x = -\mu_B$, then it takes path b. If the atom has $\mu_z = +\mu_B$, then it *doesn't have a value of* μ_x and hence it *doesn't take a path.*

Notice again what I'm saying: I'm not saying the atom takes one path or the other but we don't know which. I'm not saying the atom breaks into two pieces and each half traverses its own path. I'm saying the atom doesn't take a path. The $\mu_z = +\mu_B$ atoms within the horizontal analyzer loop *do not have a position* in the same sense that *love does not have a color.* If you think of an atom as a smaller, harder version of a classical marble, then you're visualizing the atom incorrectly.

Once again, our experiments have uncovered a phenomenon that doesn't happen in daily life, so there is no word for it in conventional language.[10] Sometimes people say that "the atom takes both paths", but that phrase does not really get to the heart of the new phenomenon. I have asked students to invent a new word to represent this new phenomenon, and my favorite of their many suggestions is "ambivate" — a combination of ambulate and ambivalent — as in "an atom with $\mu_z = +\mu_B$ ambivates through both paths of a horizontal analyzer loop". While this is a great word, it hasn't caught on. The conventional name for this phenomenon is "quantal interference".

---

[10] In exactly the same way, there was no need for the word "latitude" or the word "longitude" when it was thought that the Earth was flat. The discovery of the near-spherical character of the Earth forced our forebears to invent new words to represent these new concepts. Words do not determine reality; instead reality determines which words are worth inventing.

The name "quantal interference" comes from a (far-fetched) analogy with interference in wave optics. Recall that in the two-slit interference of light, there are some observation points that have a light intensity if light passes through slit a alone, and the same intensity if light passes through slit b alone, but zero intensity if light passes through both slits. This is called "destructive interference". There are other observation points that have a light intensity if the light passes through slit a alone, and the same intensity if light passes through slit b alone, but four times that intensity if light passes through both slits. This is called "constructive interference". But in fact the word "interference" is a poor name for this phenomenon as well. It's adapted from a football term, and football players never (or at least never intentionally) run "constructive interference".

One last word about language: The device that I've called the "analyzer loop" is more conventionally called an "interferometer". I didn't use that name at first because that would have given away the ending.

Back on page 8 I said that, to avoid unnecessary distraction, this chapter would "to the extent possible, do a quantum-mechanical treatment of an atom's magnetic moment while maintaining a classical treatment of all other aspects — such as its energy and momentum and position". You can see now why I put in that qualifier "to the extent possible": we have found that within an interferometer, a quantum-mechanical treatment of magnetic moment demands a quantum-mechanical treatment of position as well.

Exercise 1.A. Paradox?

a. The year is 1492, and you are discussing with a friend the radical idea that the earth is round. "This idea can't be correct," objects your friend, "because it contains a paradox. If it were true, then a traveler moving always due east would eventually arrive back at his starting point. Anyone can see that that's not possible!" Convince your friend that this paradox is not an internal inconsistency in the round-earth idea, but an inconsistency between the round-earth idea and the picture of the earth as a plane, a picture which your friend has internalized so thoroughly that he can't recognize it as an approximation rather than the absolute truth.

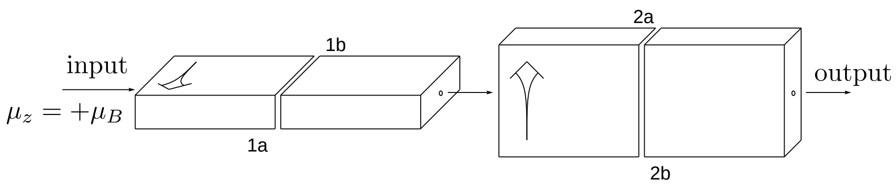
b. The year is 2092, and you are discussing with a friend the radical idea of quantal interference. "This idea can't be correct," objects your friend, "because it contains a paradox. If it were true, then an atom passing through branch a would have to know whether branch b were open or blocked. Anyone can see that that's not possible!" Convince your friend that this paradox is not an internal inconsistency in quantum mechanics, but an inconsistency between quantal ideas and the picture of an atom as a hard little marble that always has a definite position, a picture which your friend has internalized so thoroughly that he can't recognize it as an approximation rather than the absolute truth.

1.2.4 *Sample Problem: Constructive interference*

Consider the same set-up as on page 25, but now ignore atoms leaving the $-$ port of the vertical analyzer and consider as output atoms leaving the $+$ port. What is the probability of passing through the set-up when path a is blocked? When path b is blocked? When neither path is blocked?

Solution: $\frac{1}{4}$; $\frac{1}{4}$; 1. Because $\frac{1}{4}+\frac{1}{4}<1$, this is an example of constructive interference.

1.2.5 *Sample Problem: Two analyzer loops*



Atoms with $\mu_z = +\mu_B$ are channeled through a horizontal analyzer loop (number 1), then a vertical analyzer loop (number 2). If all paths are open, 100% of the incoming atoms exit from the output. What percentage of the incoming atoms leave from the output if the following paths are blocked?

(a) 2a
(b) 2b
(c) 1a
(d) 1b
(e) 1b and 2a
(f) 1a and 2b

Solution: Only two principles are needed to solve this problem: First, an atom leaving an unblocked analyzer loop leaves in the same condition it had when it entered. Second, an atom leaving an analyzer loop with one path blocked leaves in the condition specified by the path that it took, regardless of the condition it had when it entered. Use of these principles gives the solution in the table on the next page. Notice that in changing from situation (a) to situation (e), you add blockage, yet you increase the output!

paths blocked	input condition	path taken through # 1	intermediate condition	path taken through # 2	output condition	probability of input → output
none	$\mu_z = +\mu_B$	"both"	$\mu_z = +\mu_B$	a	$\mu_z = +\mu_B$	100%
2a	$\mu_z = +\mu_B$	"both"	$\mu_z = +\mu_B$	100% blocked at a	none	0%
2b	$\mu_z = +\mu_B$	"both"	$\mu_z = +\mu_B$	a	$\mu_z = +\mu_B$	100%
1a	$\mu_z = +\mu_B$	50% blocked at a 50% pass through b	$\mu_x = -\mu_B$	"both"	$\mu_x = -\mu_B$	50%
1b	$\mu_z = +\mu_B$	50% pass through a 50% blocked at b	$\mu_x = +\mu_B$	"both"	$\mu_x = +\mu_B$	50%
1b and 2a	$\mu_z = +\mu_B$	50% pass through a 50% blocked at b	$\mu_x = +\mu_B$	25% blocked at a 25% pass through b	$\mu_z = -\mu_B$	25%
1a and 2b	$\mu_z = +\mu_B$	50% blocked at a 50% pass through b	$\mu_x = -\mu_B$	25% pass through a 25% blocked at b	$\mu_z = +\mu_B$	25%

1.2.6 *Sample Problem: Find the flaw*

No one would write a computer program and call it finished without testing and debugging their first attempt. Yet some approach physics problem solving in exactly this way: they get to the equation that is “the solution”, stop, and then head off to bed for some well-earned sleep without investigating whether the solution makes sense. This is a loss, because the real fun and interest in a problem comes not from our cleverness in finding “the solution”, but from uncovering what that solution tells us about nature. To give you experience in this reflection step, I’ve designed “find the flaw” problems in which you don’t find the solution, you only test it. Here’s an example.

Find the flaw: Tilted analyzer loop
Four students — Aldo, Beth, Celine, and Denzel — work problem 1.5 presented on the next page. All find the same answer for part (a), namely zero, but for parts (b) and (c) they produce four different answers! Their candidate answers are:

	(b)	(c)
Aldo	$\cos^4(\theta/2)$	$\sin^4(\theta/2)$
Beth	$\frac{1}{4}\sin(\theta)$	$\frac{1}{4}\sin(\theta)$
Celine	$\frac{1}{4}\sqrt{2}\sin(\theta/2)$	$\frac{1}{4}\sqrt{2}\sin(\theta/2)$
Denzel	$\frac{1}{2}\sin^2(\theta)$	$\frac{1}{2}\sin^2(\theta)$

Without actually working the problem, provide simple reasons showing that all of these candidates must be wrong.

Solution: For the special case $\theta = 0°$ the correct answers for (b) and (c) are both 0. Aldo’s answer to (b) fails this test.

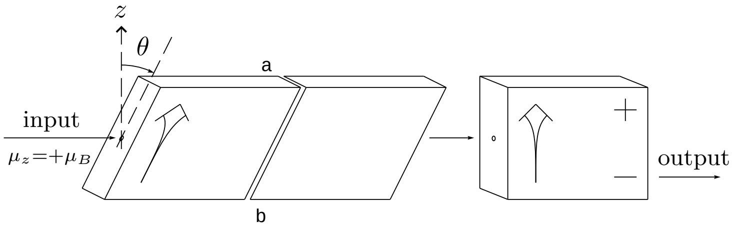
The special case $\theta = 90°$ was investigated in sections 1.2.1 and 1.2.2: in this case the answers for (b) and (c) are both $\frac{1}{4}$. Denzel’s answer fails this test.

Beth’s answer gives negative probabilities when $180° < \theta < 360°$. Bad idea!

The answer should not change when θ increases by 360°. Celine's answer fails this test. (For example, it gives the answer $+\frac{1}{4}$ when $\theta = 90^\circ$ and $-\frac{1}{4}$ when $\theta = 450^\circ$, despite the fact that 90° and 450° are the same angle.)

Problems

1.5 **Tilted analyzer loop** (recommended problem)

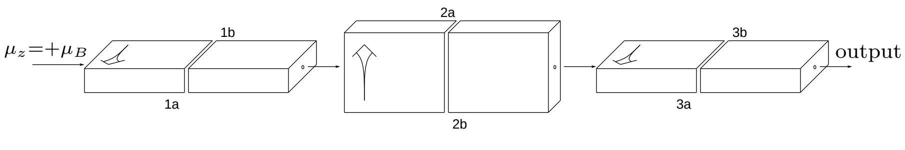


An atom with $\mu_z = +\mu_B$ enters the analyzer loop (interferometer) shown above, tilted at angle θ to the vertical. The outgoing atom enters a z-analyzer, and whatever comes out the − port is considered output. What is the probability for passage from input to output when:

a. Paths a and b are both open?
b. Path b is blocked?
c. Path a is blocked?

1.6 **Three analyzer loops** (recommended problem)

Atoms with $\mu_z = +\mu_B$ are channeled into a horizontal analyzer loop, followed by a vertical analyzer loop, followed by a horizontal analyzer loop.



If all paths are open, 100% of the incoming atoms exit from the output. What percent of the incoming atoms leave from the output if the following paths are blocked?

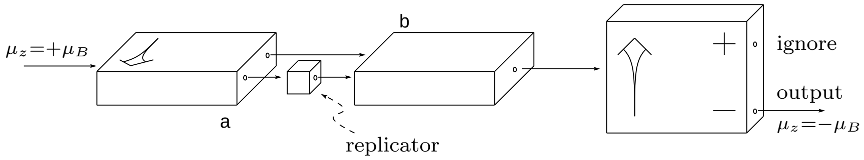
(a) 3a
(b) 3b
(c) 2a
(d) 2b
(e) 1b
(f) 2a and 3b
(g) 1b and 3b
(h) 1b and 3a
(i) 1b and 3a and 2a

(Note that in going from situation (h) to situation (i) you get *more* output from *increased* blockage.)

1.3 Aharonov-Bohm effect

We have seen how to sort silver atoms using a Stern-Gerlach analyzer, made of a non-uniform magnetic field. But how do atoms behave in a uniform magnetic field? In general, this is an elaborate question, (treated in section 5.4), and the answer will depend on the initial condition of the atom's magnetic moment, on the magnitude of the field, and on the amount of time that the atom spends in the field. But for one special case the answer, determined experimentally, is easy. If a silver atom is exposed to uniform magnetic field $\vec{B}$ for exactly the right amount of time (which turns out to be a time of $\pi\hbar/\mu_B B$), then the atom emerges with exactly the same magnetic condition it had initially: If it starts with $\mu_z = -\mu_B$, it ends with $\mu_z = -\mu_B$. If it starts with $\mu_x = +\mu_B$, it ends with $\mu_x = +\mu_B$. If it starts with $\mu_{29^\circ} = +\mu_B$, it ends with $\mu_{29^\circ} = +\mu_B$. Thus for atoms moving at a given speed, we can build a box containing a uniform magnetic field with just the right length so that any atom passing through it will spend just the right amount of time to emerge in the same condition it had when it entered. We call this box a "replicator".

If you play with one of these boxes you'll find that you can build any elaborate set-up of sources, detectors, blockages, and analyzers, and that inserting a replicator into any path will not affect the outcome of any experiment. But notice that this apparatus list does not include interferometers (our "analyzer loops")! Build the interference experiment of page 25. Do not block either path. Instead, slip a replicator into one of the two paths a or b — it doesn't matter which.



Without the replicator *no* atom emerges at output. But experiment shows that after inserting the replicator, *all* the atoms emerge at output.

How can this be? Didn't we just say of a replicator that "any atom passing through it will. . . emerge in the same condition it had when it entered"? Indeed we did, and indeed this is true. But an atom with $\mu_z = +\mu_B$ doesn't pass through path a or path b — it ambivates through both paths.

If the atom *did* take one path or the other, then the replicator *would* have no effect on the experimental results. The fact that it *does* have an effect is proof that the atom *doesn't* take one path or the other.

The fact[11] that one can perform this remarkable experiment was predicted theoretically (in a different context) by Walter Franz. He announced his result in Danzig (now Gdańsk, Poland) in May 1939, just months before the Nazi invasion of Poland, and his prediction was largely forgotten in the resulting chaos. The effect was rediscovered theoretically by Werner Ehrenberg and Raymond Siday in 1949, but they published their result under the opaque title of "The refractive index in electron optics and the principles of dynamics" and their prediction was also largely forgotten. The effect was rediscovered theoretically a third time by Yakir Aharonov and David Bohm in 1959, and this time it sparked enormous interest, both experimental and theoretical. The phenomenon is called today the "Aharonov-Bohm effect".

Problem

1.7 **Bomb-testing interferometer**[12] (recommended problem)
The Acme Bomb Company sells a bomb triggered by the presence of silver, and claims that the trigger is so sensitive that the bomb explodes when its trigger absorbs even a single silver atom. You have heard similar extravagant claims from other manufacturers, so you're suspicious.

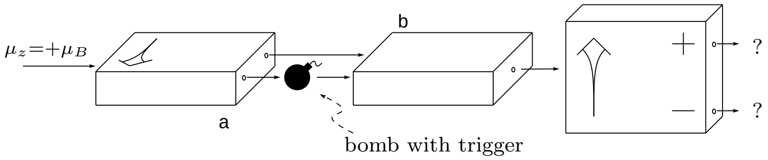
---

[11] See B.J. Hiley, "The early history of the Aharonov-Bohm effect" (17 April 2013) https://arxiv.org/abs/1304.4736.

[12] Avshalom C. Elitzur and Lev Vaidman, "Quantum mechanical interaction-free measurements" *Foundations of Physics* **23** (July 1993) 987–997.

You purchase a dozen bombs, then shoot individual silver atoms at each in turn. The first bomb tested explodes! The trigger worked as advertised, but now it's useless because it's blasted to pieces. The second bomb tested doesn't explode — the atom slips through a hole in the trigger. This confirms your suspicion that not *all* the triggers are as sensitive as claimed, so this bomb is useless to you as well. If you continue testing in this fashion, at the end all your good bombs will be blown up and you will be left with a stash of bad bombs.

So instead, you set up the test apparatus sketched here:



An atom with $\mu_z = +\mu_B$ enters the interferometer. If the bomb trigger has a hole, then the atom ambivates through both paths, arrives at the analyzer with $\mu_z = +\mu_B$, and exits the + port of the analyzer. But if the bomb trigger is good, then either (a) the atom takes path a and sets off the bomb, or else (b) the atom takes path b.

a. If the bomb trigger is good, what is the probability of option (a)? Of option (b)?
b. If option (b) happens, what kind of atom arrives at the analyzer? What is the probability of that atom exiting through the + port? The − port?

Conclusion: If the atom exits through the − port, then the bomb is good. If it exits through the + port then the bomb might be good or bad and further testing is required. But you *can* determine that the bomb trigger is good without blowing it up!

1.4 Light on the atoms

Our conclusion that, under some circumstances, the atom "does not have a position" is so dramatically counterintuitive that you might — no, you *should* — be tempted to test it experimentally. Set up the interference experiment on page 25, but instead of simply allowing atoms to pass through the interferometer, watch to see which path the atom takes through the set-up. To watch them, you need light. So set up the apparatus with lamps trained on the two paths a and b.

Send in one atom. There's a flash of light at path a.

Another atom. Flash of light at b.

Another atom. Flash at b again.

Then a, then a, then b.

You get the drift. Always the light appears at one path or the other. (In fact, the flashes come at random with probability $\frac{1}{2}$ for a flash at a and $\frac{1}{2}$ for a flash at b.) Never is there no flash. Never are there "two half flashes". The atom always has a position when passing through the interferometer. "So much", say the skeptics, "for this metaphysical nonsense about 'the atom takes both paths'."

But wait. Go back and look at the output of the vertical analyzer. When we ran the experiment with no light, the probability of coming out the $-$ port was 0. When we turn the lamps on, then the probability of coming out the $-$ port becomes $\frac{1}{2}$.

When the lamps are off, analysis II on page 26 is correct: the atoms ambivate through both paths, and the probability of exiting from the $-$ port is 0. When the lamps are on and a flash is seen at path a, then the atom *does* take path a, and now the analysis of section 1.2.2 on page 26 is correct: the probability of exiting from the $-$ port is $\frac{1}{2}$.

The process when the lamps are on is called "observation" or "measurement", and a lot of nonsense has come from the use of these two words. The important thing is whether the light is present or absent. Whether or not the flashes are "observed" by a person is irrelevant. To prove this to yourself, you may, instead of observing the flashes in person, record the flashes on video. If the lamps are on, the probability of exiting from the $-$ port is $\frac{1}{2}$. If the lamps are off, the probability of exiting from the $-$ port is 0. Now, after the experiment is performed, you may either destroy the video, or play it back to a human audience, or play it back to a feline audience. Surely, by this point it is too late to change the results at the exit port.

It's not just light. Any method you can dream up for determining the path taken will show that the atom takes just one path, but that method will also change the output probability from 0 to $\frac{1}{2}$. No person needs to actually read the results of this mechanism: as long as the mechanism is at work, as long as it is in principle possible to determine which path is taken, then one path is taken and no interference happens.

What happens if you train a lamp on path a but leave path b in the dark? In this case a flash means the atom has taken path a. No flash means the atom has taken path b. In both cases the probability of passage for the atom is $\frac{1}{2}$.

How can the atom taking path b "know" that the lamp at path a is turned on? The atom initially "sniffs out" *both* paths, like a fog creeping down two passageways. The atom that eventually does take path b in the dark started out attempting both paths, and that's how it "knows" the lamp at path a is on. This is called the "Renninger negative-result experiment".

It is not surprising that the presence or absence of light should affect an atom's motion: this happens even in classical mechanics. When an object absorbs or reflects light, that object experiences a force, so its motion is altered. For example, a baseball tossed upward in a gymnasium with the overhead lamps off attains a slightly greater height that an identical baseball experiencing an identical toss in the same gymnasium with the overhead lamps on, because the downward-directed light beams push the baseball downward. (This is the same "radiation pressure" that is responsible for the tails of comets. And of course, the effect occurs whenever the lamps are turned on: whether any person actually watches the illuminated baseball is irrelevant.) This effect is negligible for typical human-scale baseballs and tosses and lamps, but atoms are far smaller than baseballs and it is reasonable that the light should alter the motion of an atom more than it alters the motion of a baseball.

One last experiment: Look for the atoms with dim light. In this case, some of the atoms will pass through with a flash. But — because of the dimness — some atoms will pass through without any flash at all. For those atoms passing through with a flash, the probability for exiting the $-$ port is $\frac{1}{2}$. For those atoms passing through without a flash, the probability of exiting the $-$ port is 0.

1.5 Entanglement

I have claimed that an atom with $\mu_z = +\mu_B$ *doesn't have* a value of μ_x, and that when such an atom passes through a horizontal interferometer, it *doesn't have* a position. You might say to yourself, "These claims are so weird, so far from common sense, that I just can't accept them. I believe this atom *does* have a value of μ_x and *does* have a position, but something else very complicated is going on to make the atom *appear* to lack a μ_x and a position. I don't know what that complicated thing is, but just because I haven't yet thought it up yet doesn't mean that it doesn't exist."

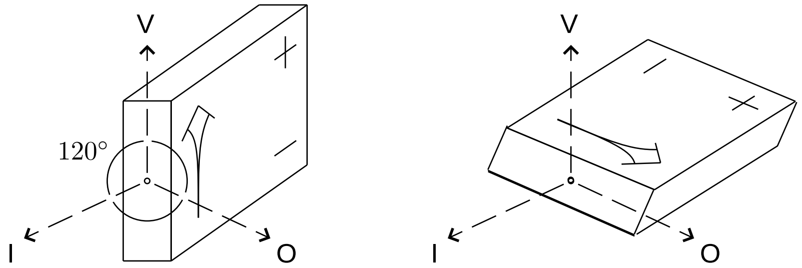
If you think this, you're in good company: Einstein[13] thought it too. This section introduces a new phenomenon of quantum mechanics, and shows that no local deterministic mechanism, no matter how complex or how fantastic, can give rise to all the results of quantum mechanics. Einstein was wrong.

1.5.1 *Flipping Stern-Gerlach analyzer*

A new piece of apparatus helps us uncover this new phenomenon of nature. Mount a Stern-Gerlach analyzer on a stand so that it can be oriented either vertically (0°), or tilted one-third of a circle clockwise (+120°), or tilted one-third of a circle counterclockwise (−120°). Call these three orientations V (for vertical), O (for out of the page), or I (for into the page). As an atom approaches the analyzer, select one of these three orientations at random, flip the analyzer to that orientation, and allow the atom to pass through as usual. As a new atom approaches, again select an orientation at random, flip the analyzer, and let the atom pass through. Repeat many times.

---

[13] Although Albert Einstein (1879–1955) is most famous for his work on relativity, he claimed that he had "thought a hundred times as much about the quantum problems as I have about general relativity theory." (Remark to Otto Stern, reported in Abraham Pais, *"Subtle is the Lord. . . ": The Science and the Life of Albert Einstein*, [Oxford University Press, Oxford, UK, 1982] page 9.) Concerning the importance of various traits in science (and in life) he wrote "I have no special talents. I am only passionately curious." (Letter to Carl Seelig, 11 March 1952, the Albert Einstein Archives 39-013.)



Flipping Stern-Gerlach analyzer. The arrows V*,* O*, and* I*, oriented* 120° *apart, all lie within the plane perpendicular to the atom's approach path.*

What happens if an atom with $\mu_z = +\mu_B$ enters a flipping analyzer? With probability $\frac{1}{3}$, the atom enters a vertical analyzer (orientation V), and in that case it exits the + port with probability 1. With probability $\frac{1}{3}$, the atom enters an out-of-the-page analyzer (orientation O), and in that case (see equation 1.5) it exits the + port with probability

$$\cos^2(120^\circ/2) = \tfrac{1}{4}.$$

With probability $\frac{1}{3}$, the atom enters an into-the-page analyzer (orientation I), and in that case it exits the + port with probability $\frac{1}{4}$. Thus the overall probability of this atom exiting through the + port is

$$\tfrac{1}{3} \times 1 + \tfrac{1}{3} \times \tfrac{1}{4} + \tfrac{1}{3} \times \tfrac{1}{4} = \tfrac{1}{2}. \tag{1.6}$$

A similar analysis shows that if an atom with $\mu_z = -\mu_B$ enters the flipping analyzer, it exits the + port with probability $\frac{1}{2}$.

You could repeat the analysis for an atom entering with $\mu_{(+120^\circ)} = +\mu_B$, but you don't need to. Because the three orientations are exactly one-third of a circle apart, rotational symmetry demands that an atom entering with $\mu_{(+120^\circ)} = +\mu_B$ behaves exactly as an atom entering with $\mu_z = +\mu_B$.

In conclusion, an atom entering in any of the six conditions $\mu_z = +\mu_B$, $\mu_z = -\mu_B$, $\mu_{(+120^\circ)} = +\mu_B$, $\mu_{(+120^\circ)} = -\mu_B$, $\mu_{(-120^\circ)} = +\mu_B$, or $\mu_{(-120^\circ)} = -\mu_B$ will exit through the + port with probability $\frac{1}{2}$.

1.5.2 *EPR source of atom pairs*

Up to now, our atoms have come from an oven. For the next experiments we need a special source[14] that expels two atoms at once, one moving to the left

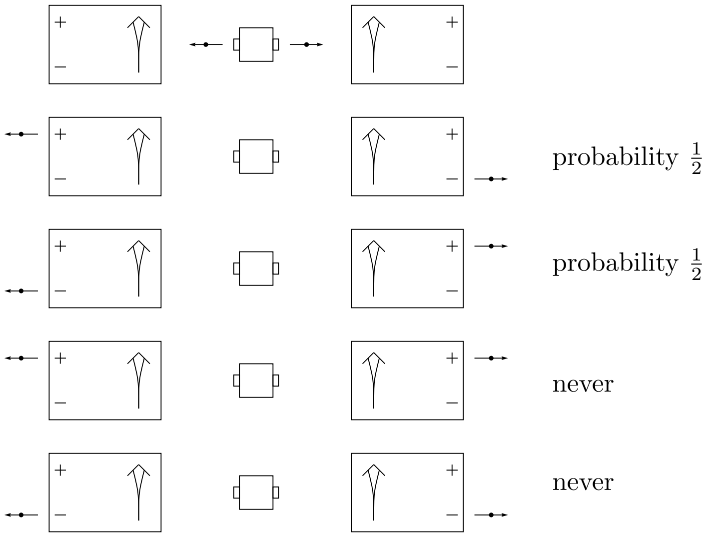
---

[14] The question of how to build this special source need not concern us at the moment: it is an experimental fact that such sources do exist. One way to make one would start with

and the other to the right. For the time being we call this an "EPR" source, which produces an atomic pair in an "EPR" condition. The letters come from the names of those who discovered this condition: Albert Einstein, Boris Podolsky, and Nathan Rosen. After investigating this condition we will develop a more descriptive name.

The following experiments investigate the EPR condition:

(1) Each atom encounters a vertical Stern-Gerlach analyzer. The experimental result: *the two atoms exit through opposite ports.* To be precise: with probability $\frac{1}{2}$, the left atom exits + and the right atom exits −, and with probability $\frac{1}{2}$, the left atom exits − and the right atom exits +, but it never happens that both atoms exit + or that both atoms exit −.



You might suppose that this is because for half the pairs, the left atom is generated with $\mu_z = +\mu_B$ while the right atom is generated with $\mu_z = -\mu_B$, while for the other half of the pairs, the left atom is generated with $\mu_z = -\mu_B$ while the right atom is generated with

---

a diatomic molecule with zero magnetic moment. Cause the molecule to disintegrate and eject the two daughter atoms in opposite directions. Because the initial molecule had zero magnetic moment, the pair of daughter atoms will have the properties of magnetic moment described. In fact, it's easier to build a source, not for a pair of atoms, but for a pair of photons using a process called spontaneous parametric down-conversion.

> $\mu_z = +\mu_B$. This supposition seems suspicious, because it singles out the z axis as special, but at this stage in our experimentation it's possible.

(2) Repeat the above experiment with horizontal Stern-Gerlach analyzers. The experimental result: Exactly the same as in experiment **(1)**! The two atoms always exit through opposite ports.

> Problem 1.9 on page 50 demonstrates that the results of this experiment rule out the supposition presented at the bottom of experiment **(1)**.

(3) Repeat the above experiment with the two Stern-Gerlach analyzers oriented at +120°, or with both oriented at −120°, or with both oriented at 57°, or for any other angle, as long as both have the same orientation. The experimental result: Exactly the same for any orientation!

(4) In an attempt to trick the atoms, we set the analyzers to vertical, then launch the pair of atoms, then (while the atoms are in flight) switch both analyzers to, say, 42°, and have the atoms encounter these analyzers both with switched orientation. The experimental result: Regardless of what the orientation is, and regardless of when that orientation is set, the two atoms always exit through opposite ports.

Here is one way to picture this situation: The pair of atoms has a total magnetic moment of zero. But whenever the projection of a single atom on any axis is measured, the result must be $+\mu_B$ or $-\mu_B$, never zero. The only way to insure that that total magnetic moment, projected on any axis, sums to zero is the way described above. Do not put too much weight on this picture: like the "wants to go straight" story of section 1.1.5 (page 16), this is a classical story that happens to give the correct result. The definitive answer to any question is always experiment, not any picture or story, however appealing it may be.

These four experiments show that it is impossible to describe the condition of the atoms through anything like "the left atom has $\mu_z = +\mu_B$, the right atom has $\mu_z = -\mu_B$". How *can* we describe the condition of the pair? This will require further experimentation. For now, we say it has an EPR condition.

1.5.3 *EPR atom pair encounters flipping Stern-Gerlach analyzers*

A pair of atoms leaves the EPR source, and each atom travels at the same speed to vertical analyzers located 100 meters away. The left atom exits the − port, the right atom exits the + port. When the pair is flying from source to analyzer, it's not correct to describe it as "the left atom has $\mu_z = -\mu_B$, the right atom has $\mu_z = +\mu_B$", but after the atoms leave their analyzers, then this is a correct description.

Now shift the left analyzer one meter closer to the source. The left atom encounters its analyzer before the right atom encounters its. Suppose the left atom exits the − port, while the right atom is still in flight toward its analyzer. We know that when the right atom eventually does encounter its vertical analyzer, it will exit the + port. Thus it is correct to describe the right atom as having "$\mu_z = +\mu_B$", even though that atom hasn't yet encountered its analyzer.

Replace the right vertical analyzer with a flipping Stern-Gerlach analyzer. (In the figure below, it is in orientation O, out of the page.) Suppose the left atom encounters its vertical analyzer and exits the − port. Through the reasoning of the previous paragraph, the right atom now has $\mu_z = +\mu_B$. We know that when such an atom encounters a flipping Stern-Gerlach analyzer, it exits the + port with probability $\frac{1}{2}$.

Similarly, if the left atom encounters its vertical analyzer and exits the + port, the right atom now has $\mu_z = -\mu_B$, and once it arrives at its flipping analyzer, it will exit the − port with probability $\frac{1}{2}$. Summarizing these two paragraphs: Regardless of which port the left atom exits, the right atom will exit the opposite port with probability $\frac{1}{2}$.

Now suppose that the left analyzer were not vertical, but instead in orientation I, tilted into the page by one-third of a circle. It's easy to see that, again, regardless of which port the left atom exits, the right atom will exit the opposite port with probability $\frac{1}{2}$.

Finally, suppose that the left analyzer is a flipping analyzer. Once again, the two atoms will exit from opposite ports with probability $\frac{1}{2}$.

The above analysis supposed that the left analyzer was one meter closer to the source than the right analyzer, but clearly it also works if the right analyzer is one meter closer to the source than the left analyzer. Or one centimeter. One suspects that the same result will hold even if the two analyzers are exactly equidistant from the source, and experiment bears out this suspicion.

In summary: Each atom from this EPR source enters a flipping Stern-Gerlach analyzer.

(A) The atoms exit from opposite ports with probability $\frac{1}{2}$.
(B) If the two analyzers happen to have the same orientation, the atoms exit from opposite ports.

This is the prediction of quantum mechanics, and experiment confirms this prediction.

1.5.4 *The prediction of local determinism*

Suppose you didn't know anything about quantum mechanics, and you were told the result that "if the two analyzers have the same orientation, the atoms exit from opposite ports." Could you explain it?

I am sure you could. In fact, there are two possible explanations: First, the communication explanation. The left atom enters its vertical analyzer, and notices that it's being pulled toward the + port. It calls up the right atom with its walkie-talkie and says "If your analyzer has orientation I or O then you might go either way, but if your analyzer has orientation V you've got to go to the − port!" This is a possible explanation, but it's not a *local* explanation. The two analyzers might be 200 meters apart, or they might be 200 light-years apart. In either case, the message would have to get from the left analyzer to the right analyzer instantaneously. The walkie-talkies would have to use not radio waves, which propagate at the speed of light, but some sort of not-yet-discovered "insta-rays". Physicists have always been skeptical of non-local explanations, and since the advent of relativity they have grown even more skeptical, so we set this explanation aside. Can you find a local explanation?

Again, I am sure you can. Suppose that when the atoms are launched, they have some sort of characteristic that specifies which exit port they will take when they arrive at their analyzer. This very reasonable supposition, called "determinism", pervades all of classical mechanics. It is similar to saying "If I stand atop a 131 meter cliff and toss a ball horizontally with speed 23.3 m/s, I can predict the angle with which the ball strikes the ground, even though that event will happen far away and long in the future." In the case of the ball, the resulting strike angle is encoded into the initial position and velocity. In the case of the atoms, it's not clear how the exit port will be encoded: perhaps through the orientation of its magnetic moment, perhaps in some other, more elaborate way. But the method of encoding is irrelevant: if local determinism holds, then something within the atom determines which exit port it will take when it reaches its analyzer.[15] I'll represent this "something" through a code like $(++-)$. The first symbol means that if the atom encounters an analyzer in orientation V, it will exit through the $+$ port. The second means that if it encounters an analyzer in orientation O, it will exit through the $+$ port. The third means that if it encounters an analyzer in orientation I, it will exit through the $-$ port. The only way to ensure that "if the two analyzers have the same orientation, the atoms exit from opposite ports" is to assume that when the two atoms separate from each other within the source, they have opposite codes. If the left atom has $(+-+)$, the right atom must have $(-+-)$. If the left atom has $(---)$, the right atom must have $(+++)$. This is the local deterministic scheme for explaining fact (B) that "if the two analyzers have the same orientation, the atoms exit from opposite ports".

But can this scheme explain fact (A)? Let's investigate. Consider first the case mentioned above: the left atom has $(+-+)$ and the right atom has $(-+-)$. These atoms will encounter analyzers set to any of $3^2 = 9$ possible pairs of orientations. We list them below, along with with exit ports taken by the atoms. (For example, the third line of the table considers a left analyzer in orientation V and a right analyzer in orientation I. The left atom has code $(+-+)$, and the first entry in that code determines that the left atom will exit from the V analyzer through the $+$ port. The right atom has code $(-+-)$, and the third entry in that code determines that the right atom will exit from the I analyzer through the $-$ port.)

---

[15] But remember that in quantum mechanics determinism does not hold. The information *can't* be encoded within the three projections of a classical magnetic moment vector, because at any one instant, the quantum magnetic moment vector has only one projection.

left port	left analyzer	right analyzer	right port	opposite?
+	V	V	−	yes
+	V	O	+	no
+	V	I	−	yes
−	O	V	−	no
−	O	O	+	yes
−	O	I	−	no
+	I	V	−	yes
+	I	O	+	no
+	I	I	−	yes

Each of the nine orientation pairs (VV, OI, etc.) are equally likely, five of the orientation pairs result in atoms exiting from opposite ports, so when atoms of this type emerge from the source, the probability of these atoms exiting from opposite ports is $\frac{5}{9}$.

What about a pair of atoms generated with different codes? Suppose the left atom has $(--+)$ so the right atom must have $(++-)$. If you perform the analysis again, you will find that the probability of atoms exiting from opposite ports is once again $\frac{5}{9}$.

Suppose the left atom has $(---)$, so the right atom must have $(+++)$. The probability of the atoms exiting from opposite ports is of course 1.

There are, in fact, just $2^3 = 8$ possible codes:

code for left atom	probability of exiting opposite
+ + +	1
− + +	5/9
+ − +	5/9
+ + −	5/9
+ − −	5/9
− + −	5/9
− − +	5/9
− − −	1

If the source makes left atoms of only type $(--+)$, then the probability of atoms exiting from opposite ports is $\frac{5}{9}$. If the source makes left atoms of only type $(+++)$, then the probability of atoms exiting from opposite ports is 1. If the source makes left atoms of type $(--+)$ half the time, and of type $(+++)$ half the time, then the probability of atoms exiting from opposite ports is halfway between $\frac{5}{9}$ and 1, namely $\frac{7}{9}$. But no matter how the source makes atoms, the probability of atoms exiting from opposite ports must be somewhere between $\frac{5}{9}$ and 1.

But experiment and quantum mechanics agree: That probability is actually $\frac{1}{2}$ — and $\frac{1}{2}$ is *not* between $\frac{5}{9}$ and 1. No local deterministic scheme — no matter how clever, or how elaborate, or how baroque — can give the result $\frac{1}{2}$. There is no "something within the atom that determines which exit port it will take when it reaches its analyzer". If the magnetic moment has a projection on axis V, then it *doesn't have* a projection on axis O or axis I.

There is a reason that Einstein, despite his many attempts, never produced a scheme that explained quantum mechanics in terms of some more fundamental, local and deterministic mechanism. It is not that Einstein wasn't clever. It is that *no such scheme exists.*

1.5.5 *The upshot*

This is a new phenomenon — one totally absent from classical physics — so it deserves a new name, something more descriptive than "EPR". Einstein called it "spooky action at a distance".[16] The phenomenon is spooky all right, but this phrase misses the central point that the phenomenon involves "correlations at a distance", whereas the word "action" suggests "cause-and-effect at a distance". Erwin Schrödinger coined the term "entanglement" for this phenomenon and said it was "not… *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought".[17] The world has followed Schrödinger and the phenomenon is today called entanglement. We will later investigate entanglement in more detail, but for now we will just call

---

[16] Letter from Einstein to Max Born, 3 March 1947, *The Born-Einstein Letters* (Macmillan, New York, 1971) translated by Irene Born.

[17] Erwin Schrödinger, "Discussion of probability relations between separated systems" *Mathematical Proceedings of the Cambridge Philosophical Society* **31** (October 1935) 555–563.

our EPR source a "source of entangled atom pairs" and describe the condition of the atom pair as "entangled".

The failure of local determinism described above is a special case of "Bell's Theorem", developed by John Bell[18] in 1964. The theorem has by now been tested experimentally numerous times in numerous contexts (various different angles; various distances between the analyzers; various sources of entangled pairs; various kinds of particles flying apart — gamma rays, or optical photons, or ions). In every test, quantum mechanics has been shown correct and local determinism wrong. What do we gain from these results?

First, they show that nature does not obey local determinism. To our minds, local determinism is common sense and any departure from it is weird. Thus whatever theory of quantum mechanics we eventually develop will be, to our eyes, weird. This will be a strength, not a defect, in the theory. The weirdness lies in nature, not in the theory used to describe nature.

Each of us feels a strong psychological tendency to reject the unfamiliar. In 1633, the Holy Office of the Inquisition found Galileo Galilei's idea that the Earth orbited the Sun so unfamiliar that they rejected it. The inquisitors put Galileo on trial and forced him to abjure his position. From the point of view of nature, the trial was irrelevant, Galileo's abjuration was irrelevant: the Earth orbits the Sun whether the Holy Office finds that fact comforting or not. It is our job as scientists to change our minds to fit nature; we do not change nature to fit our preconceptions. Don't make the inquisitors' mistake.

Second, the Bell's theorem result guides not just our calculations about nature but also our visualizations of nature, and even the very idea of what it means to "understand" nature. Lord Kelvin[19] framed the situation perfectly in his 1884 Baltimore lectures: "I never satisfy myself until I can

---

[18] John Stewart Bell (1928–1990), a Northern Irish physicist, worked principally in accelerator design, and his investigation of the foundations of quantum mechanics was something of a hobby. Concerning tests of his theorem, he remarked that "The reasonable thing just doesn't work." [Jeremy Bernstein, *Quantum Profiles* (Princeton University Press, Princeton, NJ, 1991) page 84.]

[19] William Thomson, the first Baron Kelvin (1824–1907), was an Irish mathematical physicist and engineer who worked in Scotland. He is best known today for establishing the thermodynamic temperature scale that bears his name, but he also made fundamental contributions to electromagnetism. He was knighted for his engineering work on the first transatlantic telegraph cable.

make a mechanical model of a thing. If I can make a mechanical model I can understand it. As long as I cannot make a mechanical model all the way through I cannot understand, and this is why I cannot get the electromagnetic theory."[20] If we take this as our meaning of "understand", then the experimental tests of Bell's theorem assure us that we will never be able to understand quantum mechanics.[21] What is to be done about this? There are only two choices. Either we can give up on understanding, or we can develop a new and more appropriate meaning for "understanding".

Max Born[22] argued for the first choice: "The ultimate origin of the difficulty lies in the fact (or philosophical principle) that we are compelled to use the words of common language when we wish to describe a phenomenon, not by logical or mathematical analysis, but by a picture appealing to the imagination. Common language has grown by everyday experience and can never surpass these limits."[23] Born felt that it was impossible to visualize or "understand" quantum mechanics: all you could do was grind through the "mathematical analysis".

Humans are visual animals, however, and I have found that when we are told not to visualize, we do so anyway. But we do so in an illicit and uncritical way. For example, many people visualize an atom passing through an interferometer as a small, hard, marble, with a definite position, despite the already-discovered fact that this visualization is untenable. Many people visualize a photon as a "ball of light" despite the fact that a photon (as conventionally defined) has a definite energy and hence can never have a position.

It *is* possible to develop a visualization and understanding of quantum mechanics. This *can't* be done by building a "mechanical model all the way through". It must be done through *both* analogy and contrast: atoms

---

[20] William Thomson, "Baltimore lectures on wave theory and molecular dynamics," in Robert Kargon and Peter Achinstein, editors, *Kelvin's Baltimore Lectures and Modern Theoretical Physics* (MIT Press, Cambridge, MA, 1987) page 206.

[21] The first time I studied quantum mechanics seriously, I wrote in the margin of my textbook "Good God they do it! But how?" I see now that I was looking for a mechanical mechanism undergirding quantum mechanics. It doesn't exist, but it's very natural for anyone to *want* it to exist.

[22] Max Born (1882–1970) was a German-Jewish theoretical physicist with a particular interest in optics. At the University of Göttingen in 1925 he directed Werner Heisenberg's research which resulted in the first formulation of quantum mechanics. His granddaughter, the British-born Australian actress and singer Olivia Newton-John, is famous for her 1981 hit song "Physical".

[23] Max Born, *Atomic Physics*, sixth edition (Hafner Press, New York, 1957) page 97.

behave in *some* ways like small hard marbles, in *some* ways like classical waves, and in *some* ways like a cloud or fog of probability. Atoms don't behave *exactly* like any of these things, but if you keep in mind both the analogy and its limitations, then you can develop a pretty good visualization and understanding.

And that brings us back to the name "entanglement". It's an important name for an important phenomenon, but it suggests that the two distant atoms are connected mechanically, through strings. They aren't. The two atoms are correlated — if the left comes out +, the right comes out −, and vice versa — but they aren't correlated because of some signal sent back and forth through either strings or walkie-talkies. Entanglement involves correlation without causality.

Problems

1.8 **An atom walks into an analyzer**
Execute the "similar analysis" mentioned in the sentence below equation (1.6).

1.9 **A supposition squashed** (essential problem)
If atoms were generated according to the supposition presented below experiment **(1)** on page 41, then would would happen when they encountered the two horizontal analyzers of experiment **(2)**?

1.10 **A probability found through local determinism**
Suppose that the codes postulated on page 45 did exist. Suppose also that a given source produces the various possible codes with these probabilities:

code for left atom	probability of making such a pair
$+++$	1/2
$++-$	1/4
$+--$	1/8
$--+$	1/8

If this given source were used in the experiment of section 1.5.3 with distant flipping Stern-Gerlach analyzers, what would be the probability of the two atoms exiting from opposite ports?

1.11 **A probability found through quantum mechanics**
In the test of Bell's inequality (the experiment of section 1.5.3), what is the probability given by quantum mechanics that, if the orientation settings are different, the two atoms exit from opposite ports?

1.6 Quantum cryptography

We've seen a lot of new phenomena, and the rest of this book is devoted to filling out our understanding of these phenomena and applying that understanding to various circumstances. But first, can we use them for anything?

We can. The sending of coded messages used to be the province of armies and spies and giant corporations, but today everyone does it. All transactions through automatic teller machines are coded. All Internet commerce is coded. This section describes a particular, highly reliable encoding scheme and then shows how quantal entanglement may someday be used to implement this scheme. (Quantum cryptography was used to securely transmit voting ballots cast in the Geneva canton of Switzerland during parliamentary elections held 21 October 2007. But it is not today in regular use anywhere.)

In this section I use names conventional in the field of coded messages (called cryptography). Alice and Bob wish to exchange private messages, but they know that Eve is eavesdropping on their communication. How can they encode their messages to maintain their privacy?

1.6.1 *The Vernam cipher*

The Vernam cipher or "one-time pad" technique is the only coding scheme proven to be absolutely unbreakable (if used correctly). It does not *rely* on the use of computers — it was invented by Gilbert Vernam in 1919 — but today it is mostly implemented using computers, so I'll describe it in that context.

Data are stored on computer disks through a series of magnetic patches on the disk that are magnetized either "up" or "down". An "up" patch is taken to represent `1`, and a "down" patch is taken to represent `0`. A string of seven patches is used to represent a character. For example, by a convention called ASCII, the letter “a” is represented through the sequence `1100001` (or, in terms of magnetizations, up, up, down, down, down, down, up). The letter “W” is represented through the sequence `1010111`. Any computer the world around will represent the message “What?” through the sequence

`1010111 1101000 1100001 1110100 0111111`

This sequence is called the “plaintext”.

But Alice doesn’t want a message recognizable by any computer the world around. She wants to send the message “What?” to Bob in such a way that Eve will *not* be able to read the message, even though Eve has eavesdropped on the message. Here is the scheme invented by Vernam: Before sending her message, Alice generates a string of random `0`s and `1`s just as long as the message she wants to send — in this case, $7 \times 5 = 35$ bits. She might do this by flipping 35 coins, or by flipping one coin 35 times. I’ve just done that, producing the random number

`0100110 0110011 1010110 1001100 1011100`

Then Alice gives Bob a copy of that random number – the “key”.

Instead of sending the plaintext, Alice modifies her plaintext into a coded “ciphertext” using the key. She writes down her plaintext and writes the key below it, then works through column by column. For each position, if the key is `0` the plaintext is left unchanged; but if the key is `1` the plaintext is reversed (from `0` to `1` or vice versa). For the first column, the key is `0`, so Alice *doesn’t* change the plaintext: the first character of ciphertext is the same as the first character of plaintext. For the second column, the key is `1`, so Alice *does* change the plaintext: the second character of ciphertext is the reverse of the second character of plaintext. Alice goes through all the columns, duplicating the plaintext where the key is `0` and reversing the plaintext where the key is `1`.

plaintext: `1010111 1101000 1100001 1110100 0111111`
key: `0100110 0110011 1010110 1001100 1011100`
ciphertext: `1110001 1011011 0110111 0111000 1100011`

Then, Alice sends out her ciphertext over open communication lines.

Now, the ciphertext that Bob (and Eve) receive translates to *some* message through the ASCII convention – in fact, it translates to "q[78c" — but because the key is random, the ciphertext is just as random. Bob deciphers Alice's message by carrying out the encoding process on the ciphertext, namely, duplicating the ciphertext where the key is `0` and reversing the ciphertext where the key is `1`. The result is the plaintext. Eve does not know the key, so she cannot produce the plaintext.

The whole scheme relies on the facts that the key is (1) random and (2) unknown to Eve. The very name "one-time pad" underscores that a key can only be used once and must then be discarded. If a single key is used for two messages, then the second key is not "random" — it is instead perfectly correlated with the first key. There are easy methods to break the code when a key is reused.

Generating random numbers is not easy, and the Vernam cipher demands keys as long as the messages transmitted. As recently as 1992, high-quality computer random-number generators were classified by the U.S. government as munitions, along with tanks and fighter planes, and their export from the country was prohibited.

And of course Eve must not know the key. So there must be some way for Alice to get the key to Bob securely. If they have some secure method for transmitting keys, why don't they just use that same secure method for sending their messages?

In common parlance, the word "random" can mean "unimportant, not worth considering" (as in "Joe made a random comment"). So it may seem remarkable that a major problem for government, the military, and commerce is the generation and distribution of randomness, but that is indeed the case.

1.6.2 *Quantum mechanics to the rescue*

Since quantum mechanics involves randomness, it seems uniquely positioned to solve this problem. Here's one scheme.

Alice and Bob set up a source of entangled atoms halfway between their two homes. Both of them erect vertical Stern-Gerlach analyzers to detect the atoms. If Alice's atom comes out $+$, she will interpret it as a `1`, if $-$, a `0`. Bob interprets his atoms in the opposite sense. Since the entangled atoms always exit from opposite ports, Alice and Bob end up with the same random number, which they use as a key for their Vernam-cipher communications over conventional telephone or computer lines.

This scheme will indeed produce and distribute copious, high-quality random numbers. But Eve can get at those same numbers through the following trick: She cuts open the atom pipe leading from the entangled source to Alice's home, and inserts a vertical interferometer.[24] She watches the atoms pass through her interferometer. If the atom takes path a, Eve knows that when Alice receives that same atom, it will exit from Eve's + port. If the atom takes path b, the opposite holds. Eve gets the key, Eve breaks the code.

It's worth looking at this eavesdropping in just a bit more detail. When the two atoms depart from their source, they are entangled. It is *not* true that, say, Alice's atom has $\mu_z = +\mu_B$ while Bob's atom has $\mu_z = -\mu_B$ — the pair of atoms is in the condition we've called "entangled", but the individual atoms themselves are not in *any* condition. However, after Eve sees the atom taking path a of her interferometer, then the two atoms are no longer entangled — now it *is* true that Alice's atom has the condition $\mu_z = +\mu_B$ while Bob's atom has the condition $\mu_z = -\mu_B$. The key received by Alice and Bob will be random whether or not Eve is listening in. To test for evesdropping, Alice and Bob must examine it in some other way.

Replace Alice and Bob's vertical analyzers with flipping Stern-Gerlach analyzers. After Alice receives her random sequence of pluses and minuses, encountering her random sequence of analyzer orientations, she sends both these sequences to Bob over an open communication line. (Eve will intercept this information but it won't do her any good, because she won't know the corresponding information for Bob.) Bob now knows both the results at his analyzer and the results at Alice's analyzer, so he can test to see whether the atom pairs were entangled. If he finds that they were, then Eve is not listening in. If he finds that they were not entangled, then he knows for certain that Eve *is* listening in, and they must not use their compromised key.

Is there some other way for Eve to tap the line? No! If the atom pairs pass the test for entanglement, then no one can know the values of their

---

[24] Inspired by James Bond, I always picture Eve as exotic beauty in a little black dress slinking to the back of an eastern European café to tap the diplomatic cable which conveniently runs there. But in point of fact Eve would be a computer.

μ_z projections because those projections don't exist! We have guaranteed that no one has intercepted the key by the interferometer method, or by any other method whatsoever.

Once Bob has tested for entanglement, he and Alice still have a lot of work to do. For a key they must use only those random numbers produced when their two analyzers happen to have the same orientations. There are detailed protocols specifying how Alice and Bob must exchange information about their analyzer orientations, in such a way that Eve can't uncover them. I won't describe these protocols because while they tell you how clever people are, they tell you nothing about how nature behaves. But you should take away that entanglement is not merely a phenomenon of nature: it is also a natural resource.

Problem

1.12 **All about Eve**
Suppose Eve uses a vertical interferometer to watch the atoms en route to Bob. Now the atom pairs are not entangled when they reach Alice and Bob. What then is the probability of the two atoms exiting from opposite ports? Compare to the probability when the atom pairs are entangled. What is the probability of the two atoms exiting from opposite ports when the two analyzers both have orientation O? Compare to the probability when the atom pairs are entangled.

1.7 What is a qubit?

We've devoted an entire chapter to the magnetic moment of a silver atom. Perhaps you find this inappropriate: do you really care so much about silver atoms? Yes you do, because the phenomena and principles we've established concerning the magnetic moment of a silver atom apply to a host of other systems: the polarization of a light photon, the hybridization of a benzene molecule, the position of the nitrogen atom within an ammonia molecule, the neutral kaon, and more. Such systems are called "two-state systems" or "spin-$\frac{1}{2}$ systems" or "qubit systems". The ideas we establish concerning the magnetic moment of a silver atom apply equally well to all these systems.

After developing these ideas in the next four chapters, we will (in chapter 6, "The Quantum Mechanics of Position") generalize them to continuous systems like the position of an electron.

Problem

1.13 **Questions** (recommended problem)
Answering questions is an important scientific skill and, like any skill, it is sharpened through practice. This book gives you plenty of opportunities to develop that skill. Asking questions is another important scientific skill.[25] To hone that skill, write down a list of questions you have about quantum mechanics at this point. Be brief and pointed: you will not be graded for number or for verbosity. In future problems, I will ask you to add to your list.

⟦For example, one of my questions would be: "Can entanglement be used to send a message from the left analyzer to the right analyzer?"⟧

---

[25] "The important thing is not to stop questioning," said Einstein. "Never lose a holy curiosity." [Interview by William Miller, "Death of a Genius", *Life* magazine, volume **38**, number 18 (2 May 1955) pages 61–64 on page 64.]

Chapter 2

Forging Mathematical Tools

When you walked into your introductory classical mechanics course, you were already familiar with the phenomena of introductory classical mechanics: flying balls, spinning wheels, colliding billiard balls. Your introductory mechanics textbook didn't need to introduce these things to you, but instead jumped right into describing these phenomena mathematically and explaining them in terms of more general principles.

The first chapter of this textbook made you familiar with the phenomena of quantum mechanics: quantization, interference, and entanglement — at least, insofar as these phenomena are manifest in the behavior of the magnetic moment of a silver atom. You are now, with respect to quantum mechanics, at the same level that you were, with respect to classical mechanics, when you walked into your introductory mechanics course. It is now our job to describe these quantal phenomena mathematically, to explain them in terms of more general principles, and (eventually) to investigate situations more complex than the magnetic moment of one or two silver atoms.

2.1 What is a quantal state?

We've been talking about the state of the silver atom's magnetic moment by saying things like "the projection of the magnetic moment on the z axis is $\mu_z = -\mu_B$" or "$\mu_x = +\mu_B$" or "$\mu_\theta = -\mu_B$". This notation is clumsy. First of all, it requires you to write down the same old μs time and time again. Second, the most important thing is the axis (z or x or θ), and the symbol for the axis is also the smallest and easiest to overlook.

P.A.M. Dirac[1] invented a notation that overcomes these faults. He looked at descriptions like

$$\mu_z = -\mu_B \quad \text{or} \quad \mu_x = +\mu_B \quad \text{or} \quad \mu_\theta = -\mu_B$$

and noted that the only difference from one expression to the other was the axis subscript and the sign in front of μ_B. Since the only thing that distinguishes one expression from another is $(z, -)$, or $(x, +)$, or $(\theta, -)$, Dirac thought, these should be the only things we need to write down. He denoted these three states as

$$|z-\rangle \quad \text{or} \quad |x+\rangle \quad \text{or} \quad |\theta-\rangle.$$

The placeholders $|\quad\rangle$ are simply ornaments to remind us that we're talking about quantal states, just as the arrow atop $\vec{r}$ is simply an ornament to remind us that we're talking about a vector. States expressed using this notation are sometimes called "kets".

Simply establishing a notation doesn't tell us much. Just as in classical mechanics, we say we know a state when we know all the information needed to describe the system now and to predict its future. In our universe the classical time evolution law is

$$\vec{F} = m\frac{d^2\vec{r}}{dt^2}$$

and so the state is specified by giving both a position $\vec{r}$ and a velocity $\vec{v}$. If nature had instead provided the time evolution law

$$\vec{F} = m\frac{d^3\vec{r}}{dt^3}$$

then the state would have been specified by giving a position $\vec{r}$, a velocity $\vec{v}$, and an acceleration $\vec{a}$. The specification of state is dictated by nature, not by humanity, so we can't know how to specify a state until we know the laws of physics governing that state. Since we don't yet know the laws of quantal physics, we can't yet know exactly how to specify a quantal state.

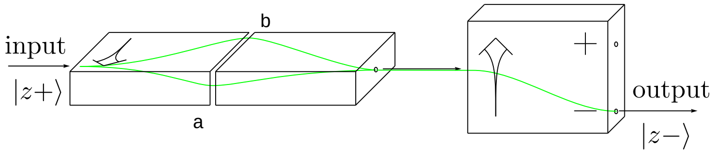
Classical intuition makes us suppose that, to specify the magnetic moment of a silver atom, we need to specify all three components μ_z, μ_x, and μ_y. We have already seen that nature precludes such a specification: if the magnetic moment has a value for μ_z, then it doesn't have a value for μ_x,

---

[1] The Englishman Paul Adrien Maurice Dirac (1902–1984) in 1928 formulated a relativistically correct quantum mechanical equation that turns out to describe the electron. In connection with this so-called Dirac equation, he predicted the existence of antimatter. Dirac was painfully shy and notoriously cryptic.

and it's absurd to demand a specification for something that doesn't exist. As we learn more and more quantum physics, we will learn better and better how to specify states. There will be surprises. But always keep in mind that (just as in classical mechanics) it is experiment, not philosophy or meditation, and certainly not common sense, that tells us how to specify states.

2.2 Amplitude



An atom initially in state $|z+\rangle$ ambivates through the apparatus above. We have already seen that, when the atom ambivates in darkness,

$$\begin{aligned}&\text{probability to go from input to output} \neq \\ &\quad\text{probability to go from input to output via path } \mathsf{a} \\ &\quad+ \text{probability to go from input to output via path } \mathsf{b}.\end{aligned} \tag{2.1}$$

On the other hand, it makes sense to associate some sort of "influence to go from input to output via path a" with the path through a and a corresponding "influence to go from input to output via path b" with the path through b. This postulated influence is called "probability amplitude" or just "amplitude".[2] Whatever amplitude is, its desired property is that

$$\begin{aligned}&\text{amplitude to go from input to output} = \\ &\quad\text{amplitude to go from input to output via path } \mathsf{a} \\ &\quad+ \text{amplitude to go from input to output via path } \mathsf{b}.\end{aligned} \tag{2.2}$$

For the moment, the very existence of amplitude is nothing but a hopeful surmise. Scientists cannot now and indeed never will be able to prove that the concept of amplitude applies to all situations. That's because new situations are being investigated every day, and perhaps tomorrow a new

---

[2] The name "amplitude" is a poor one, because it is also used for the maximum value of a sinusoidal signal — in the function $A\sin(\omega t)$, the symbol A represents the amplitude — and this sinusoidal signal "amplitude" has nothing to do with the quantal "amplitude". One of my students correctly suggested that a better name for quantal amplitude would be "proclivity". But it's too late now to change the word.

situation will be discovered that cannot be described in terms of amplitudes. But as of today, that hasn't happened.

The role of amplitude, whatever it may prove to be, is to calculate probabilities. We set forth…

Three desirable rules for amplitude

(1) *From amplitude to probability.* For every possible action there is an associated amplitude, such that

$$\text{probability for the action} = |\text{amplitude for the action}|^2.$$

(2) *Actions in series.* If an action takes place through several successive stages, the amplitude for that action is the product of the amplitudes for each stage.

(3) *Actions in parallel.* If an action could take place in several possible ways, the amplitude for that action is the sum of the amplitudes for each possibility.

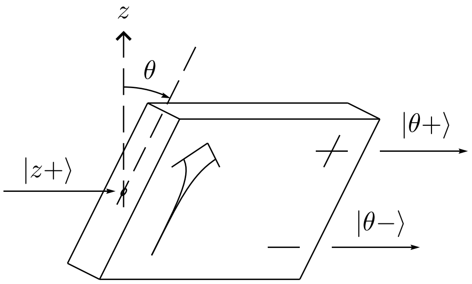
The first rule is a simple way to make sure that probabilities are always positive. The second rule is a natural generalization of the rule for probabilities in series — that if an action happens through several stages, the probability for the action as a whole is the product of the probabilities for each stage. And the third rule simply restates the "desired property" presented in equation (2.2).

We apply these rules to various situations that we've already encountered, beginning with the interference experiment sketched above. Recall the probabilities already established (first column in table):

	probability	|amplitude|	amplitude
go from input to output	0	0	0
go from input to output via path a	$\frac{1}{4}$	$\frac{1}{2}$	$+\frac{1}{2}$
go from input to output via path b	$\frac{1}{4}$	$\frac{1}{2}$	$-\frac{1}{2}$

If rule (1) is to hold, then the amplitude to go from input to output must also be 0, while the amplitude to go via a path must have magnitude $\frac{1}{2}$ (second column in table). According to rule (3), the two amplitudes to go via a and via b must sum to zero, so they cannot both be represented by positive numbers. Whatever mathematical entity is used to represent amplitude, it must enable two such entities, each with non-zero magnitude, to sum to zero. There are many such entities: real numbers, complex numbers, hypercomplex numbers, and vectors in three dimensions are all possibilities. For this particular interference experiment, it suffices to assign real numbers to amplitudes: the amplitude to go via path a is $+\frac{1}{2}$, and the amplitude to go via path b is $-\frac{1}{2}$. (Third column in table. The negative sign could have been assigned to path a rather than to path b: this choice is merely conventional.) For other interference experiments (see section 2.8), complex numbers are required. It turns out that, for all situations yet encountered, one can represent amplitude mathematically as a complex number. Once again, this reflects the results of experiment, not of philosophy or meditation.

The second situation we'll consider is a Stern-Gerlach analyzer.



The amplitude that an atom entering the θ-analyzer in state $|z+\rangle$ exits in state $|\theta+\rangle$ is called[3] $\langle\theta+|z+\rangle$. That phrase is a real mouthful, so the symbol $\langle\theta+|z+\rangle$ is pronounced "the amplitude that $|z+\rangle$ is in $|\theta+\rangle$", even though this briefer pronunciation leaves out the important role of the analyzer.[4] From rule (1), we know that

$$|\langle\theta+|z+\rangle|^2 = \cos^2(\theta/2) \tag{2.3}$$

$$|\langle\theta-|z+\rangle|^2 = \sin^2(\theta/2). \tag{2.4}$$

---

[3] The states appear in the symbol in the opposite sequence from their appearance in the description.

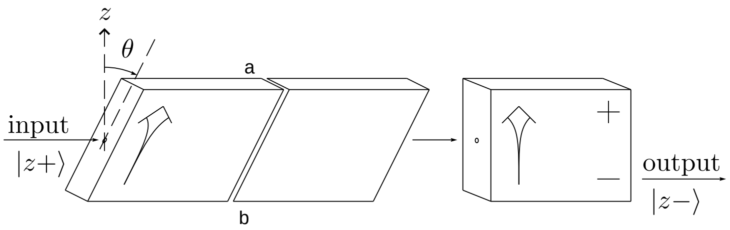
[4] The ultimate source of such problems is that the English language was invented by people who did not understand quantum mechanics, hence they never produced concise, accurate phrases to describe quantal phenomena. In the same way, the ancient phrase "search the four corners of the Earth" is still colorful and practical, and is used today even by those who know that the Earth doesn't have four corners.

You can also use rule (1), in connection with the experiments described in problem 1.2, "Exit probabilities" (on page 22) to determine that

$$
\begin{aligned}
|\langle z{+}|\theta{+}\rangle|^2 &= \cos^2(\theta/2)\\
|\langle z{-}|\theta{+}\rangle|^2 &= \sin^2(\theta/2)\\
|\langle \theta{+}|z{-}\rangle|^2 &= \sin^2(\theta/2)\\
|\langle \theta{-}|z{-}\rangle|^2 &= \cos^2(\theta/2)\\
|\langle z{+}|\theta{-}\rangle|^2 &= \sin^2(\theta/2)\\
|\langle z{-}|\theta{-}\rangle|^2 &= \cos^2(\theta/2).
\end{aligned}
$$

Clearly analyzer experiments like these determine the *magnitude* of an amplitude. No analyzer experiment can determine the *phase* of an amplitude. To determine phases, we must perform interference experiments.

So the third situation is an interference experiment.



Rule (2), actions in series, tells us that the amplitude to go from $|z{+}\rangle$ to $|z{-}\rangle$ via path a is the product of the amplitude to go from $|z{+}\rangle$ to $|\theta{+}\rangle$ times the amplitude to go from $|\theta{+}\rangle$ to $|z{-}\rangle$:

$$\text{amplitude to go via path } \mathsf{a} = \langle z{-}|\theta{+}\rangle\langle\theta{+}|z{+}\rangle.$$

Similarly

$$\text{amplitude to go via path } \mathsf{b} = \langle z{-}|\theta{-}\rangle\langle\theta{-}|z{+}\rangle.$$

And then rule (3), actions in parallel, tells us that the amplitude to go from $|z{+}\rangle$ to $|z{-}\rangle$ is the sum of the amplitude to go via path a and the amplitude to go via path b. In other words

$$\langle z{-}|z{+}\rangle = \langle z{-}|\theta{+}\rangle\langle\theta{+}|z{+}\rangle + \langle z{-}|\theta{-}\rangle\langle\theta{-}|z{+}\rangle. \tag{2.5}$$

We know the magnitude of each of these amplitudes from analyzer experiments:

amplitude	magnitude
$\langle z{-}\vert z{+}\rangle$	0
$\langle z{-}\vert\theta{+}\rangle$	$\vert\sin(\theta/2)\vert$
$\langle\theta{+}\vert z{+}\rangle$	$\vert\cos(\theta/2)\vert$
$\langle z{-}\vert\theta{-}\rangle$	$\vert\cos(\theta/2)\vert$
$\langle\theta{-}\vert z{+}\rangle$	$\vert\sin(\theta/2)\vert$

The task now is to assign phases to these magnitudes in such a way that equation (2.5) is satisfied. In doing so we are faced with an embarrassment of riches: there are *many* consistent ways to make this assignment. Here are two commonly used conventions:

amplitude	convention I	convention II
$\langle z{-}\vert z{+}\rangle$	0	0
$\langle z{-}\vert\theta{+}\rangle$	$\sin(\theta/2)$	$i\sin(\theta/2)$
$\langle\theta{+}\vert z{+}\rangle$	$\cos(\theta/2)$	$\cos(\theta/2)$
$\langle z{-}\vert\theta{-}\rangle$	$\cos(\theta/2)$	$\cos(\theta/2)$
$\langle\theta{-}\vert z{+}\rangle$	$-\sin(\theta/2)$	$-i\sin(\theta/2)$

There are two things to notice about these amplitude assignments. First, one normally assigns values to physical quantities by experiment, or by calculation, but not "by convention". Second, both of these conventions show unexpected behaviors: Because the angle 0° is the same as the angle 360°, one would expect that $\langle 0^\circ{+}\vert z{+}\rangle$ would equal $\langle 360^\circ{+}\vert z{+}\rangle$, whereas in fact the first amplitude is $+1$ and the second is -1. Because the state $\vert 180^\circ{-}\rangle$ (that is, $\vert\theta{-}\rangle$ with $\theta = 180^\circ$) is the same as the state $\vert z{+}\rangle$, one would expect that $\langle 180^\circ{-}\vert z{+}\rangle = 1$, whereas in fact $\langle 180^\circ{-}\vert z{+}\rangle$ is either -1 or $-i$, depending on convention. These two observations underscore the fact that amplitude is a mathematical tool that enables us to calculate physically observable quantities, like probabilities. It is not itself a physical entity. No experiment measures amplitude. Amplitude is not "out there, physically present in space" in the way that, say, a nitrogen molecule is.

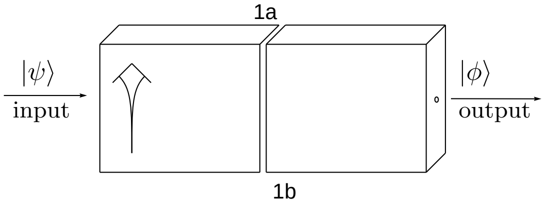
A good analogy is that an amplitude convention is like a language. Any language is a human convention: there is no intrinsic connection between a physical horse and the English word "horse", or the German word "pferd", or the Swahili word “farasi”. The fact that language is pure human convention, and that there are multiple conventions for the name of a horse, doesn’t mean that language is unimportant: on the contrary language is an immensely powerful tool. And the fact that language is pure human convention doesn’t mean that you can’t develop intuition about language: on the contrary if you know the meaning of “arachnid” and the meaning of “phobia”, then your intuition for English tells you that “arachnophobia” means fear of spiders. Exactly the same is true for amplitude: it is a powerful tool, and with practice you can develop intuition for it.

When I introduced the phenomenon of quantal interference on page 27, I said that there was no word or phrase in the English language that accurately represents what’s going on: It’s flat-out wrong to say “the atom takes path a” and it’s flat-out wrong to say “the atom takes path b”. It gives a wrong impression to say “the atom takes no path” or “the atom takes both paths”. I introduced the phrase “the atom ambivates through the two paths of the interferometer”. Now we have a technically correct way of describing the phenomenon: “the atom has an amplitude to take path a and an amplitude to take path b”.

Here’s another warning about language: If an atom in state $|\psi\rangle$ enters a vertical analyzer, the amplitude for it to exit from the + port is $\langle z+|\psi\rangle$. (And of course the amplitude for it exit from the − port is $\langle z-|\psi\rangle$.) This is often stated “If the atom is in state $|\psi\rangle$, the amplitude of it being in state $|z+\rangle$ is $\langle z+|\psi\rangle$.” This is an acceptable shorthand for the full explanation, which requires thinking about an analyzer experiment, even though the shorthand never mentions the analyzer. But never say “If the atom is in state $|\psi\rangle$, the probability of it being in state $|z+\rangle$ is $|\langle z+|\psi\rangle|^2$.” This gives the distinct and incorrect impression that before entering the analyzer, the atom was either in state $|z+\rangle$ or in state $|z-\rangle$, and you just didn’t know which it was. Instead, say “If an atom in state $|\psi\rangle$ enters a vertical analyzer, the probability of exiting from the + port in state $|z+\rangle$ is $|\langle z+|\psi\rangle|^2$.”

2.2.1 *Sample Problem: Two paths*

Find an equation similar to equation (2.5) representing the amplitude to start in state $|\psi\rangle$ at input, ambivate through a vertical interferometer, and end in state $|\phi\rangle$ at output.



Solution: Because of rule (2), actions in series, the amplitude for the atom to take the top path is the product

$$\langle\phi|z+\rangle\langle z+|\psi\rangle.$$

Similarly the amplitude for it to take the bottom path is

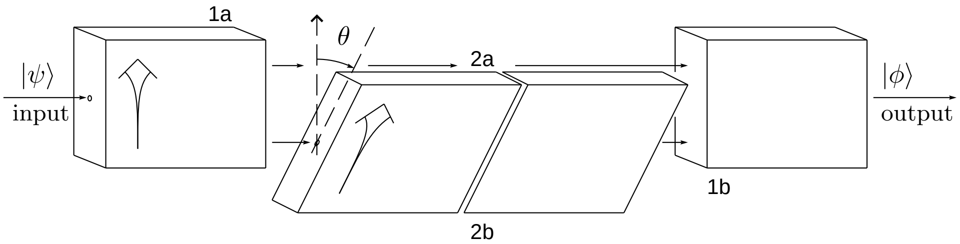
$$\langle\phi|z-\rangle\langle z-|\psi\rangle.$$

Because of rule (3), actions in parallel, the amplitude for it to ambivate through both paths is the sum of these two, and we conclude that

$$\langle\phi|\psi\rangle = \langle\phi|z+\rangle\langle z+|\psi\rangle + \langle\phi|z-\rangle\langle z-|\psi\rangle. \tag{2.6}$$

2.2.2 *Sample Problem: Three paths*

Stretch apart a vertical interferometer, so that the recombining rear end is far from the splitting front end, and insert a θ interferometer into the bottom path. Now there are *three* paths from input to output. Find an equation similar to equation (2.5) representing the amplitude to start in state $|\psi\rangle$ at input and end in state $|\phi\rangle$ at output.



Solution:

$$\begin{aligned}\langle\phi|\psi\rangle = {} & \langle\phi|z+\rangle\langle z+|\psi\rangle \\ & + \langle\phi|z-\rangle\langle z-|\theta+\rangle\langle\theta+|z-\rangle\langle z-|\psi\rangle \\ & + \langle\phi|z-\rangle\langle z-|\theta-\rangle\langle\theta-|z-\rangle\langle z-|\psi\rangle\end{aligned} \tag{2.7}$$

Problems

2.1 **Talking about interference**
An atom in state $|\psi\rangle$ ambivates through a vertical analyzer. We say, appropriately, that "the atom has an amplitude to take the top path and an amplitude to take the bottom path". Find expressions for those two amplitudes and describe, in ten sentences or fewer, why it is *not* appropriate to say "the atom has probability $|\langle z+|\psi\rangle|^2$ to take the top path and probability $|\langle z-|\psi\rangle|^2$ to take the bottom path".

2.2 **Other conventions**
Two conventions for assigning amplitudes are given in the table on page 63. Show that if $\langle z-|\theta+\rangle$ and $\langle z-|\theta-\rangle$ are multiplied by phase factor $e^{i\alpha}$, and if $\langle z+|\theta+\rangle$ and $\langle z+|\theta-\rangle$ are multiplied by phase factor $e^{i\beta}$ (where α and β are both real), then the resulting amplitudes are just as good as the original (for either convention I or convention II).

2.3 **Peculiarities of amplitude**
Page 63 pointed out some of the peculiarities of amplitude; this problem points out another. Since the angle θ is the same as the angle $360° + \theta$, one would expect that $\langle\theta+|z+\rangle$ would equal $\langle(360° + \theta)+|z+\rangle$. Show, using either of the conventions given in the table on page 63, that this expectation is false. What is instead correct?

2.3 Reversal-conjugation relation

Working with amplitudes is made easier through the theorem that the amplitude to go from state $|\psi\rangle$ to state $|\phi\rangle$ and the amplitude to go in the opposite direction are related through complex conjugation:

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*. \tag{2.8}$$

The proof below works for states of the magnetic moment of a silver atom — the kind of states we've worked with so far — but in fact the result holds for any quantal system.

The proof relies on three facts: First, the probability for one state to be analyzed into another depends only on the magnitude of the angle between the incoming magnetic moment and the analyzer, and not on the sense of that angle. (An atom in state $|z+\rangle$ has the same probability of leaving the + port of an analyzer whether it is rotated 17° clockwise or 17° counterclockwise.) Thus

$$|\langle\phi|\psi\rangle|^2 = |\langle\psi|\phi\rangle|^2. \tag{2.9}$$

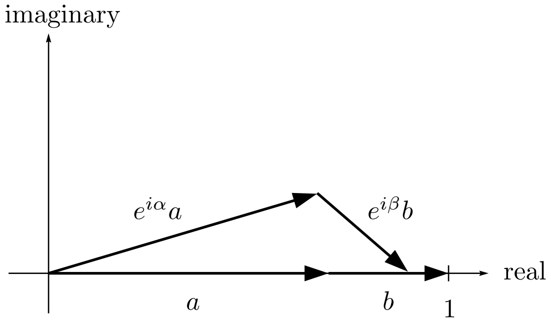
Second, an atom exits an interferometer in the same state in which it entered, so

$$\langle\phi|\psi\rangle = \langle\phi|\theta{+}\rangle\langle\theta{+}|\psi\rangle + \langle\phi|\theta{-}\rangle\langle\theta{-}|\psi\rangle. \tag{2.10}$$

Third, an atom entering an analyzer comes out somewhere, so

$$1 = |\langle\theta{+}|\psi\rangle|^2 + |\langle\theta{-}|\psi\rangle|^2. \tag{2.11}$$

The proof also relies on a mathematical result called "the triangle inequality for complex numbers": If a and b are real numbers with $a+b=1$, and in addition $e^{i\alpha}a + e^{i\beta}b = 1$, with α and β real, then $\alpha = \beta = 0$. You can find very general, very abstract, proofs of the triangle inequality, but the complex plane sketch below encapsulates the idea:



From the first fact (2.9), the two complex numbers $\langle\phi|\psi\rangle$ and $\langle\psi|\phi\rangle$ have the same magnitude, so they differ only in phase. Write this statement as

$$\langle\phi|\psi\rangle = e^{i\delta}\langle\psi|\phi\rangle^* \tag{2.12}$$

where the phase δ is a real number that might depend on the states $|\phi\rangle$ and $|\psi\rangle$. Apply this general result first to the particular state $|\phi\rangle = |\theta+\rangle$:

$$\langle\theta+|\psi\rangle = e^{i\delta_+}\langle\psi|\theta+\rangle^*, \tag{2.13}$$

and then to the particular state $|\phi\rangle = |\theta-\rangle$:

$$\langle\theta-|\psi\rangle = e^{i\delta_-}\langle\psi|\theta-\rangle^*, \tag{2.14}$$

where the two real numbers δ_+ and δ_- might be different. Our objective is to prove that $\delta_+ = \delta_- = 0$.

Apply the second fact (2.10) with $|\phi\rangle = |\psi\rangle$, giving

$$\begin{aligned}1 &= \langle\psi|\theta+\rangle\langle\theta+|\psi\rangle + \langle\psi|\theta-\rangle\langle\theta-|\psi\rangle \\ &= e^{i\delta_+}\langle\psi|\theta+\rangle\langle\psi|\theta+\rangle^* + e^{i\delta_-}\langle\psi|\theta-\rangle\langle\psi|\theta-\rangle^* \\ &= e^{i\delta_+}|\langle\psi|\theta+\rangle|^2 + e^{i\delta_-}|\langle\psi|\theta-\rangle|^2 \\ &= e^{i\delta_+}|\langle\theta+|\psi\rangle|^2 + e^{i\delta_-}|\langle\theta-|\psi\rangle|^2.\end{aligned} \tag{2.15}$$

Compare this result to the third fact (2.11)

$$1 = |\langle\theta+|\psi\rangle|^2 + |\langle\theta-|\psi\rangle|^2 \tag{2.16}$$

and use the triangle inequality with $a = |\langle\theta+|\psi\rangle|^2$ and $b = |\langle\theta-|\psi\rangle|^2$. The two phases δ_+ and δ_- must vanish, so the "reversal-conjugation relation" is proven.

2.4 Establishing a phase convention

Although there are multiple alternative phase conventions for amplitudes (see problem 2.2 on page 66), we will from now on use only phase convention I from page 63:

$$\begin{aligned}\langle z+|\theta+\rangle &= \cos(\theta/2) \\ \langle z-|\theta+\rangle &= \sin(\theta/2) \\ \langle z+|\theta-\rangle &= -\sin(\theta/2) \\ \langle z-|\theta-\rangle &= \cos(\theta/2)\end{aligned} \tag{2.17}$$

In particular, for $\theta = 90°$ we have

$$\begin{aligned}\langle z+|x+\rangle &= 1/\sqrt{2} \\ \langle z-|x+\rangle &= 1/\sqrt{2} \\ \langle z+|x-\rangle &= -1/\sqrt{2} \\ \langle z-|x-\rangle &= 1/\sqrt{2}\end{aligned} \tag{2.18}$$

This convention has a desirable special case for $\theta = 0^\circ$, namely

$$\begin{aligned}\langle z{+}|\theta{+}\rangle &= 1\\ \langle z{-}|\theta{+}\rangle &= 0\\ \langle z{+}|\theta{-}\rangle &= 0\\ \langle z{-}|\theta{-}\rangle &= 1\end{aligned} \tag{2.19}$$

but an unexpected special case for $\theta = 360^\circ$, namely

$$\begin{aligned}\langle z{+}|\theta{+}\rangle &= -1\\ \langle z{-}|\theta{+}\rangle &= 0\\ \langle z{+}|\theta{-}\rangle &= 0\\ \langle z{-}|\theta{-}\rangle &= -1\end{aligned} \tag{2.20}$$

This is perplexing, given that the angle $\theta = 0^\circ$ is the same as the angle $\theta = 360^\circ$! Any convention will have similar perplexing cases. Such perplexities underscore the fact that amplitudes are important mathematical tools used to calculate probabilities, but are not “physically real”.

Given these amplitudes, we can use the interference result (2.6) to calculate any amplitude of interest:

$$\begin{aligned}\langle\phi|\psi\rangle &= \langle\phi|z{+}\rangle\langle z{+}|\psi\rangle + \langle\phi|z{-}\rangle\langle z{-}|\psi\rangle\\ &= \langle z{+}|\phi\rangle^*\langle z{+}|\psi\rangle + \langle z{-}|\phi\rangle^*\langle z{-}|\psi\rangle\end{aligned} \tag{2.21}$$

where in the last line we have used the reversal-conjugation relation (2.8).

Problems

2.4 **Other conventions, other peculiarities**
Write what this section would have been had we adopted convention II rather than convention I from page 63. In addition, evaluate the four amplitudes of equation (2.17) for $\theta = +180^\circ$ and $\theta = -180^\circ$.

2.5 **Finding amplitudes** (recommended problem)
Using the interference idea embodied in equation (2.21), calculate the amplitudes $\langle\theta{+}|54^\circ{+}\rangle$ and $\langle\theta{-}|54^\circ{+}\rangle$ as a function of θ. Do these amplitudes have the values you expect for $\theta = 54^\circ$? For $\theta = 234^\circ$? Plot $\langle\theta{+}|54^\circ{+}\rangle$ for θ from 0° to 360°. Compare the result for $\theta = 0^\circ$ and $\theta = 360^\circ$.

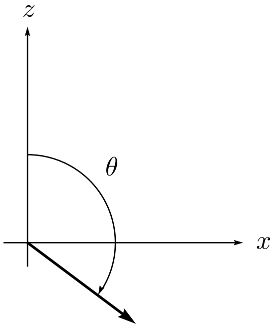
2.6 **Rotations**
Use the interference idea embodied in equation (2.21) to show that

$$\begin{aligned}\langle x{+}|\theta{+}\rangle &= \tfrac{1}{\sqrt{2}}[\cos(\theta/2) + \sin(\theta/2)]\\ \langle x{-}|\theta{+}\rangle &= -\tfrac{1}{\sqrt{2}}[\cos(\theta/2) - \sin(\theta/2)]\\ \langle x{+}|\theta{-}\rangle &= \tfrac{1}{\sqrt{2}}[\cos(\theta/2) - \sin(\theta/2)]\\ \langle x{-}|\theta{-}\rangle &= \tfrac{1}{\sqrt{2}}[\cos(\theta/2) + \sin(\theta/2)]\end{aligned} \tag{2.22}$$

If and only if you enjoy trigonometric identities, you should then show that these results can be written equivalently as

$$\begin{aligned}\langle x+|\theta+\rangle &= \cos((\theta-90^\circ)/2)\\ \langle x-|\theta+\rangle &= \sin((\theta-90^\circ)/2)\\ \langle x+|\theta-\rangle &= -\sin((\theta-90^\circ)/2)\\ \langle x-|\theta-\rangle &= \cos((\theta-90^\circ)/2)\end{aligned} \tag{2.23}$$

This makes perfect geometric sense, as the angle relative to the x axis is 90° less than the angle relative to the z axis:



2.5 How can I specify a quantal state?

We introduced the Dirac notation for quantal states on page 58, but haven't yet fleshed out that notation by specifying a state mathematically. Start with an analogy:

2.5.1 *How can I specify a position vector?*

We are so used to writing down the position vector $\vec{r}$ that we rarely stop to ask ourselves what it means. But the plain fact is that whenever we measure a length (say, with a meter stick) we find not a vector, but a single number! Experiments measure never the vector $\vec{r}$ but always a scalar — the dot product between $\vec{r}$ and some other vector, call it $\vec{s}$ for "some other".

If we know the dot product between $\vec{r}$ and every vector $\vec{s}$, then we know everything there is to know about $\vec{r}$. Does this mean that to specify $\vec{r}$, we must keep a list of all possible dot products $\vec{s}\cdot\vec{r}$? Of course not. . . such a list would be infinitely long!

You know that if you write $\vec{r}$ in terms of an orthonormal basis $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$, namely

$$\vec{r} = r_x\hat{\imath} + r_y\hat{\jmath} + r_z\hat{k} \tag{2.24}$$

where $r_x = \hat{\imath}\cdot\vec{r}$, $r_y = \hat{\jmath}\cdot\vec{r}$, and $r_z = \hat{k}\cdot\vec{r}$, then you've specified the vector. Why? Because if you know the triplet (r_x, r_y, r_z) and the triplet (s_x, s_y, s_z), then you can easily find the desired dot product

$$\vec{s}\cdot\vec{r} = \begin{pmatrix} s_x & s_y & s_z \end{pmatrix} \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} = s_x r_x + s_y r_y + s_z r_z. \tag{2.25}$$

It's a lot more compact to specify the vector through three dot products — namely $\hat{\imath}\cdot\vec{r}$, $\hat{\jmath}\cdot\vec{r}$, and $\hat{k}\cdot\vec{r}$ — from which you can readily calculate an infinite number of desired dot products, than it is to list all infinity dot products themselves!

2.5.2 *How can I specify a quantal state?*

Like the position vector $\vec{r}$, the quantal state $|\psi\rangle$ cannot by itself be measured. But if we determine (through some combination of analyzer experiments, interference experiments, and convention) the amplitude $\langle\sigma|\psi\rangle$ for every possible state $|\sigma\rangle$, then we know everything there is to know about $|\psi\rangle$. Is there some compact way of specifying the state, or do we have to keep an infinitely long list of all these amplitudes?

This nut is cracked through the interference experiment result

$$\langle\sigma|\psi\rangle = \langle\sigma|\theta+\rangle\langle\theta+|\psi\rangle + \langle\sigma|\theta-\rangle\langle\theta-|\psi\rangle, \tag{2.26}$$

which simply says, in symbols, that the atom exits an interferometer in the same state in which it entered (see equation 2.10). It gets hard to keep track of all these symbols, so I'll introduce the names

$$\begin{aligned} \langle\theta+|\psi\rangle &= \psi_+ \\ \langle\theta-|\psi\rangle &= \psi_- \end{aligned}$$

and

$$\begin{aligned} \langle\theta+|\sigma\rangle &= \sigma_+ \\ \langle\theta-|\sigma\rangle &= \sigma_-. \end{aligned}$$

From the reversal-conjugation relation, this means

$$\begin{aligned}\langle\sigma|\theta+\rangle &= \sigma_+^*\\ \langle\sigma|\theta-\rangle &= \sigma_-^*.\end{aligned}$$

In terms of these symbols, the interference result (2.26) is

$$\langle\sigma|\psi\rangle = \sigma_+^*\psi_+ + \sigma_-^*\psi_- = \begin{pmatrix}\sigma_+^* & \sigma_-^*\end{pmatrix}\begin{pmatrix}\psi_+\\ \psi_-\end{pmatrix}. \tag{2.27}$$

And this is our shortcut! By keeping track of only two amplitudes, ψ_+ and ψ_-, for each state, we can readily calculate any amplitude desired. We don't have to keep an infinitely long list of amplitudes.

This dot product result for computing amplitude is so useful and so convenient that sometimes people say the amplitude *is* a dot product. No. The amplitude reflects analyzer experiments, plus interference experiments, plus convention. The dot product is a powerful mathematical tool for computing amplitudes. (A parallel situation: There are many ways to find the latitude and longitude coordinates for a point on the Earth's surface, but the easiest is to use a GPS device. Some people are so enamored of this ease that they call the latitude and longitude the "GPS coordinates". But in fact the coordinates were established long before the Global Positioning System was built.)

2.5.3 *What is a basis?*

For vectors in three-dimensional space, an orthonormal basis[5] such as $\{\hat{i}, \hat{j}, \hat{k}\}$ is a set of three vectors of unit magnitude perpendicular to each other. As we've seen, the importance of a basis is that every vector $\vec{r}$ can be represented as a sum over these basis vectors,

$$\vec{r} = r_x\hat{i} + r_y\hat{j} + r_z\hat{k},$$

and hence any vector $\vec{r}$ can be conveniently represented through the triplet

$$\begin{pmatrix} r_x\\ r_y\\ r_z\end{pmatrix} = \begin{pmatrix}\hat{i}\cdot\vec{r}\\ \hat{j}\cdot\vec{r}\\ \hat{k}\cdot\vec{r}\end{pmatrix}.$$

For quantal states, we've seen that a set of two states such as $\{|\theta+\rangle, |\theta-\rangle\}$ plays a similar role, so it too is called a basis. For the magnetic

---

[5] The plural of "basis" is "bases", pronounced "base-ease".

moment of a silver atom, two states $|a\rangle$ and $|b\rangle$ constitute a basis whenever $\langle a|b\rangle = 0$, and the analyzer experiment of section 1.1.4 shows that the states $|\theta+\rangle$ and $|\theta-\rangle$ certainly satisfy this requirement. In the basis $\{|a\rangle, |b\rangle\}$ an arbitrary state $|\psi\rangle$ can be conveniently represented through the pair of amplitudes

$$\begin{pmatrix} \langle a|\psi\rangle \\ \langle b|\psi\rangle \end{pmatrix}.$$

2.5.4 *Hilbert space*

We have learned to express a physical state as a mathematical entity — namely, using the $\{|a\rangle, |b\rangle\}$ basis, the state $|\psi\rangle$ is represented as a column matrix of amplitudes

$$\begin{pmatrix} \langle a|\psi\rangle \\ \langle b|\psi\rangle \end{pmatrix}.$$

This mathematical entity is called a "state vector in Hilbert[6] space".

For example, in the basis $\{|z+\rangle, |z-\rangle\}$ the state $|\theta+\rangle$ is represented by

$$\begin{pmatrix} \langle z+|\theta+\rangle \\ \langle z-|\theta+\rangle \end{pmatrix} = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}. \tag{2.28}$$

Exercise 2.A. What is the representation of the state $|\theta-\rangle$ in this basis?

In contrast, in the basis $\{|x+\rangle, |x-\rangle\}$ that same state $|\theta+\rangle$ is represented (in light of equation 2.22) by the different column matrix

$$\begin{pmatrix} \langle x+|\theta+\rangle \\ \langle x-|\theta+\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}[\cos(\theta/2) + \sin(\theta/2)] \\ -\frac{1}{\sqrt{2}}[\cos(\theta/2) - \sin(\theta/2)] \end{pmatrix}. \tag{2.29}$$

Write down the interference experiment result twice

$$\begin{aligned} \langle a|\psi\rangle &= \langle a|z+\rangle\langle z+|\psi\rangle + \langle a|z-\rangle\langle z-|\psi\rangle \\ \langle b|\psi\rangle &= \langle b|z+\rangle\langle z+|\psi\rangle + \langle b|z-\rangle\langle z-|\psi\rangle \end{aligned}$$

and then write these two equations as one using column matrix notation

$$\begin{pmatrix} \langle a|\psi\rangle \\ \langle b|\psi\rangle \end{pmatrix} = \begin{pmatrix} \langle a|z+\rangle \\ \langle b|z+\rangle \end{pmatrix} \langle z+|\psi\rangle + \begin{pmatrix} \langle a|z-\rangle \\ \langle b|z-\rangle \end{pmatrix} \langle z-|\psi\rangle.$$

---

[6] The German mathematician David Hilbert (1862–1943) made contributions to functional analysis, geometry, mathematical physics, and other areas. He formalized and extended the concept of a vector space. Hilbert and Albert Einstein raced to uncover the field equations of general relativity, but Einstein beat Hilbert by a matter of weeks.

Notice the column matrix representations of states $|\psi\rangle$, $|z+\rangle$, and $|z-\rangle$, and write this equation as

$$|\psi\rangle = |z+\rangle\langle z+|\psi\rangle + |z-\rangle\langle z-|\psi\rangle. \tag{2.30}$$

And now we have a new thing under the sun. We never talk about adding together two classical states, nor multiplying them by numbers, but this equation gives us the meaning of such state addition in quantum mechanics. This is a new mathematical tool, it deserves a new name, and that name is "superposition". Superposition[7] is the mathematical reflection of the physical phenomenon of interference, and the equation (2.30) corresponds the sentence: "When an atom in state $|\psi\rangle$ ambivates through a vertical interferometer, it has amplitude $\langle z+|\psi\rangle$ of taking path a and amplitude $\langle z-|\psi\rangle$ of taking path b; its state is a superposition of the state of an atom taking path a and the state of an atom taking path b."

Superposition is not familiar from daily life or from classical mechanics, but there is a story[8] that increases understanding: "A medieval European traveler returns home from a journey to India, and describes a rhinoceros as a sort of cross between a dragon and a unicorn." In this story the rhinoceros, an animal that is *not* familiar but that *does* exist, is described as intermediate (a "sort of cross") between two fantasy animals (the dragon and the unicorn) that *are* familiar (to the medieval European) but that *do not* exist.

Similarly, an atom in state $|z+\rangle$ ambivates through both paths of a horizontal interferometer. This action is *not* familiar but *does* happen, and it is characterized as a superposition (a "sort of cross") between two actions ("taking path a" and "taking path b") that *are* familiar (to all of us steeped in the classical approximation) but that *do not* happen.

In principle, any calculation performed using the Hilbert space representation of states could be performed by considering suitable, cleverly designed analyzer and interference experiments. But it's a lot easier to use the abstract Hilbert space machinery. (Similarly, any result in electrostatics could be found using Coulomb's Law, but it's a lot easier to use the abstract electric field and electric potential. Any calculation involving vectors

---

[7] Classical particles do not exhibit superposition, but classical waves do. This is the meaning behind the cryptic statement "in quantum mechanics, an electron behaves somewhat like a particle and somewhat like a wave" or the even more cryptic phrase "wave-particle duality".

[8] Invented by John D. Roberts, but first published in Robert T. Morrison and Robert N. Boyd, *Organic Chemistry*, second edition (Allyn & Bacon, Boston, 1966) page 318.

could be performed graphically, but it's a lot easier to use abstract components. Any addition or subtraction of whole numbers could be performed by counting out marbles, but it's a lot easier to use abstract mathematical tools like carrying and borrowing.)

2.5.5 *Peculiarities of state vectors*

Because state vectors are built from amplitudes, and amplitudes have peculiarities (see pages 63 and 69), it is natural that state vectors have similar peculiarities. For example, since the angle θ is the same as the angle $\theta + 360°$, I would expect that the state vector $|\theta+\rangle$ would be the same as the state vector $|(\theta + 360°)+\rangle$.

But in fact, in the $\{|z+\rangle, |z-\rangle\}$ basis, the state $|\theta+\rangle$ is represented by

$$\begin{pmatrix} \langle z+|\theta+\rangle \\ \langle z-|\theta+\rangle \end{pmatrix} = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}, \tag{2.31}$$

so the state $|(\theta + 360°)+\rangle$ is represented by

$$\begin{aligned} \begin{pmatrix} \langle z+|(\theta+360°)+\rangle \\ \langle z-|(\theta+360°)+\rangle \end{pmatrix} &= \begin{pmatrix} \cos((\theta+360°)/2) \\ \sin((\theta+360°)/2) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta/2+180°) \\ \sin(\theta/2+180°) \end{pmatrix} = \begin{pmatrix} -\cos(\theta/2) \\ -\sin(\theta/2) \end{pmatrix}. \end{aligned} \tag{2.32}$$

So in fact $|\theta+\rangle = -|(\theta + 360°)+\rangle$. Bizarre!

This bizarreness is one facet of a general rule: If you multiply any state vector by a complex number with magnitude unity — a number such as -1, or i, or $\frac{1}{\sqrt{2}}(-1+i)$, or $e^{2.7i}$ — a so-called "complex unit" or "phase factor" — then you get a different state vector that represents the same state. This fact is called "global phase freedom" — you are free to set the overall phase of your state vector for your own convenience. This general rule applies only for multiplying both elements of the state vector by the same complex unit: if you multiply the two elements with *different* complex units, you *will* obtain a vector representing a different state (see problem 2.8 on page 78).

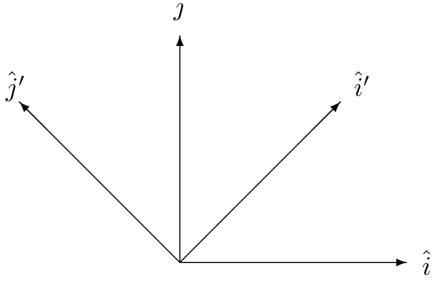
2.5.6 *Names for position vectors*

The vector $\vec{r}$ is specified in the basis $\{\hat{i}, \hat{j}, \hat{k}\}$ by the three components

$$\begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} = \begin{pmatrix} \hat{i}\cdot\vec{r} \\ \hat{j}\cdot\vec{r} \\ \hat{k}\cdot\vec{r} \end{pmatrix}.$$

Because this component specification is so convenient, it is sometimes said that the vector $\vec{r}$ is not just specified, but is *equal* to this triplet of numbers. That's false.

Think of the vector $\vec{r} = 5\hat{\imath} + 5\hat{\jmath}$. It is represented in the basis $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ by the triplet $(5, 5, 0)$. But this is not the only basis that exists. In the basis $\{\hat{\imath}' = (\hat{\imath}+\hat{\jmath})/\sqrt{2}, \hat{\jmath}' = (-\hat{\imath}+\hat{\jmath})/\sqrt{2}, \hat{k}\}$, that same vector is represented by the triplet $(5\sqrt{2}, 0, 0)$. If we had said that $\vec{r} = (5, 5, 0)$ and that $\vec{r} = (5\sqrt{2}, 0, 0)$, then we would be forced to conclude that $5 = 5\sqrt{2}$ and that $5 = 0$!



To specify a position vector $\vec{r}$, we use the components of $\vec{r}$ in a particular basis, usually denoted (r_x, r_y, r_z). We often write "$\vec{r} = (r_x, r_y, r_z)$" but in fact that's not exactly correct. The vector $\vec{r}$ represents a position — it is independent of basis. The row matrix (r_x, r_y, r_z) represents the components of that position vector in a particular basis — it is the "name" of the position in a particular basis. Instead of using an equals sign $=$ we use the symbol $\doteq$ to mean "represented by in a particular basis", as in "$\vec{r} \doteq (5, 5, 0)$" meaning "the vector $\vec{r} = 5\hat{\imath} + 5\hat{\jmath}$ is represented by the triplet $(5, 5, 0)$ in the basis $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$".

Vectors are physical things: a caveman throwing a spear at a mammoth was performing addition of position vectors, even though the caveman didn't understand basis vectors or Cartesian coordinates. The concept of "position" was known to cavemen who did not have any concept of "basis".

2.5.7 *Names for quantal states*

We've been specifying a state like $|\psi\rangle = |17°+\rangle$ by stating the axis upon which the projection of $\vec{\mu}$ is definite and equal to $+\mu_B$ — in this case, the axis tilted 17° from the vertical.

Another way to specify a state $|\psi\rangle$ would be to give the amplitude that $|\psi\rangle$ is in any possible state: that is, to list $\langle\theta{+}|\psi\rangle$ and $\langle\theta{-}|\psi\rangle$ for all values of θ: $0^\circ \leq \theta < 360^\circ$. One of those amplitudes (in this case $\langle 17^\circ{+}|\psi\rangle$) will have value 1, and finding this one amplitude would give us back the information in the specification $|17^\circ{+}\rangle$. In some ways this is a more convenient specification because we don't have to look up amplitudes: they're right there in the list. On the other hand it is an awful lot of information to have to carry around.

The Hilbert space approach is a third way to specify a state that combines the brevity of the first way with the convenience of the second way. Instead of listing the amplitude $\langle\sigma|\psi\rangle$ for *every* state $|\sigma\rangle$ we list only the two amplitudes $\langle a|\psi\rangle$ and $\langle b|\phi\rangle$ for the elements $\{|a\rangle, |b\rangle\}$ of a basis. We've already seen (equation 2.27) how quantal interference then allows us to readily calculate any amplitude.

Just as we said "the position vector $\vec{r}$ is represented in the basis $\{\hat{i}, \hat{j}, \hat{k}\}$ as $(1, 1, 0)$" or

$$\vec{r} \doteq (1, 1, 0),$$

so we say "the quantal state $|\psi\rangle$ is represented in the basis $\{|z{+}\rangle, |z{-}\rangle\}$ as

$$|\psi\rangle \doteq \begin{pmatrix} \langle z{+}|\psi\rangle \\ \langle z{-}|\psi\rangle \end{pmatrix}."$$

When you learned how to add position vectors, you learned to add them both geometrically (by setting them tail to head and drawing a vector from the first tail to the last head) and through components. The same holds for adding quantal states: You can add them physically, through interference experiments, or through components.

The equation

$$\vec{r} = \hat{i}r_x + \hat{j}r_y + \hat{k}r_z = \hat{i}(\hat{i}\cdot\vec{r}) + \hat{j}(\hat{j}\cdot\vec{r}) + \hat{k}(\hat{k}\cdot\vec{r})$$

for geometrical vectors is useful and familiar. The parallel equation

$$|\psi\rangle = |z{+}\rangle\langle z{+}|\psi\rangle + |z{-}\rangle\langle z{-}|\psi\rangle.$$

for state vectors is just as useful and will soon be just as familiar.

Problems

2.7 **Superposition and interference** (recommended problem)
On page 74 I wrote that "When an atom ambivates through an interferometer, its state is a superposition of the state of an atom taking path `a` and the state of an atom taking path `b`."

a. Write down a superposition equation reflecting this sentence for the interference experiment sketched on page 59.
b. Do the same for the interference experiment sketched on page 62.

2.8 **Representations** (recommended problem)
In the $\{|z+\rangle, |z-\rangle\}$ basis the state $|\psi\rangle$ is represented by

$$\begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}.$$

(In other words, $\psi_+ = \langle z+|\psi\rangle$ and $\psi_- = \langle z-|\psi\rangle$.)

a. If ψ_+ and ψ_- are both real, show that there is one and only one axis upon which the projection of $\vec{\mu}$ has a definite, positive value, and find the angle between that axis and the z axis in terms of ψ_+ and ψ_-.
b. What would change if you multiplied both ψ_+ and ψ_- by the same phase factor (complex unit)?
c. What would change if you multiplied ψ_+ and ψ_- by different phase factors?

This problem invites the question "What if the ratio of ψ_+/ψ_- is not pure real?" When you study more quantum mechanics, you will find that in this case the axis upon which the projection of $\vec{\mu}$ has a definite, positive value is not in the x-z plane, but instead has a component in the y direction as well.

2.9 **Addition of states**
Some students in your class wonder "What does it mean to 'add two quantal states'? You never add two classical states." For their benefit you decide to write four sentences interpreting the equation

$$|\psi\rangle = a|z+\rangle + b|z-\rangle \tag{2.33}$$

describing why you *can* add quantal states but *can't* add classical states. Your four sentences should include a formula for the amplitude a in terms of the states $|\psi\rangle$ and $|z+\rangle$.

2.10 **Names of six states, in two bases**

Write down the representations (the "names") of the states $|z+\rangle$, $|z-\rangle$, $|x+\rangle$, $|x-\rangle$, $|\theta+\rangle$, and $|\theta-\rangle$ in (a) the basis $\{|z+\rangle, |z-\rangle\}$ and in (b) the basis $\{|x+\rangle, |x-\rangle\}$.

2.11 **More peculiarities of states**

Because a vector pointing down at angle θ is the same as a vector pointing up at angle $\theta - 180°$, I would expect that $|\theta-\rangle = |(\theta - 180°)+\rangle$. Show that this expectation is false by uncovering the true relation between these two state vectors.

2.12 **Alternative approach to superposition**

We have said on page 71 that "if we determine the amplitude $\langle\sigma|\psi\rangle$ for every possible state $|\sigma\rangle$, then we know everything there is to know about $|\psi\rangle$." So, for example, if two particular states $|\psi_1\rangle$ and $|\psi_2\rangle$ have the same amplitudes $\langle\sigma|\psi_1\rangle = \langle\sigma|\psi_2\rangle$ for every state $|\sigma\rangle$, then the two states must be the same: $|\psi_1\rangle = |\psi_2\rangle$. In short, we can just erase the leading $\langle\sigma|$s from both sides.

Apply this idea to a more elaborate equation like the interference result

$$\langle\sigma|\psi\rangle = \langle\sigma|\theta+\rangle\langle\theta+|\psi\rangle + \langle\sigma|\theta-\rangle\langle\theta-|\psi\rangle, \tag{2.34}$$

and compare your conclusion to the superposition result (2.30).

2.13 **When does superposition generate a state?**

If the state vectors $|\phi\rangle$ and $|\chi\rangle$ represent quantal states, show that

$$|\psi\rangle = a|\phi\rangle + b|\chi\rangle.$$

represents a physical state provided that

$$|a|^2 + |b|^2 + 2\,\Re e\{a^* b\langle\phi|\chi\rangle\} = 1.$$

2.14 **Translation matrix**

(This problem requires background knowledge in the mathematics of matrix multiplication.)

Suppose that the representation of $|\psi\rangle$ in the basis $\{|z+\rangle, |z-\rangle\}$ is

$$\begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \begin{pmatrix} \langle z+|\psi\rangle \\ \langle z-|\psi\rangle \end{pmatrix}.$$

The representation of $|\psi\rangle$ in the basis $\{|\theta+\rangle, |\theta-\rangle\}$ is just as good, and we call it

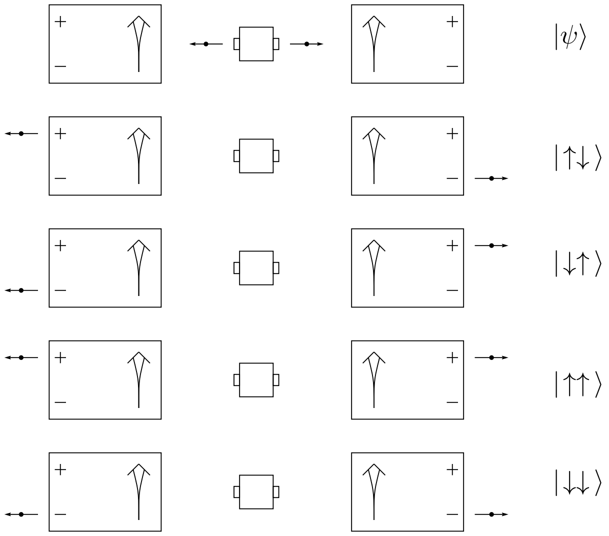
$$\begin{pmatrix} \psi'_+ \\ \psi'_- \end{pmatrix} = \begin{pmatrix} \langle\theta+|\psi\rangle \\ \langle\theta-|\psi\rangle \end{pmatrix}.$$

Show that you can "translate" between these two representations using the matrix multiplication

$$\begin{pmatrix} \psi'_+ \\ \psi'_- \end{pmatrix} = \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}.$$

2.6 States for entangled systems

In the Einstein-Podolsky-Rosen experiment **(1)** on page 41, with two vertical analyzers, the initial state is represented by $|\psi\rangle$, and various possible final states are represented by $|\uparrow\downarrow\rangle$ and so forth, as shown below. (In this section all analyzers will be vertical, so we adopt the oft-used convention that writes $|z+\rangle$ as $|\uparrow\rangle$ and $|z-\rangle$ as $|\downarrow\rangle$.)



The experimental results tell us that

$$\begin{aligned} |\langle\uparrow\downarrow|\psi\rangle|^2 &= \tfrac{1}{2} \\ |\langle\downarrow\uparrow|\psi\rangle|^2 &= \tfrac{1}{2} \\ |\langle\uparrow\uparrow|\psi\rangle|^2 &= 0 \\ |\langle\downarrow\downarrow|\psi\rangle|^2 &= 0. \end{aligned} \qquad (2.35)$$

Additional analysis (sketched in problem 15.14, "Normalization of singlet spin state") is needed to assign phases to these amplitudes. The results are

$$\begin{aligned} \langle\uparrow\downarrow|\psi\rangle &= +\tfrac{1}{\sqrt{2}} \\ \langle\downarrow\uparrow|\psi\rangle &= -\tfrac{1}{\sqrt{2}} \\ \langle\uparrow\uparrow|\psi\rangle &= 0 \\ \langle\downarrow\downarrow|\psi\rangle &= 0. \end{aligned} \qquad (2.36)$$

Using the generalization of equation (2.30) for a four-state basis, these results tell us that

$$\begin{aligned}|\psi\rangle &= |{\uparrow\downarrow}\rangle\langle{\uparrow\downarrow}|\psi\rangle + |{\downarrow\uparrow}\rangle\langle{\downarrow\uparrow}|\psi\rangle + |{\uparrow\uparrow}\rangle\langle{\uparrow\uparrow}|\psi\rangle + |{\downarrow\downarrow}\rangle\langle{\downarrow\downarrow}|\psi\rangle \\ &= \tfrac{1}{\sqrt{2}}(|{\uparrow\downarrow}\rangle - |{\downarrow\uparrow}\rangle).\end{aligned} \tag{2.37}$$

A simple derivation, with profound implications.

2.6.1 *State pertains to system, not to atom*

In this entangled situation there is no such thing as an "amplitude for the right atom to exit from the + port," because the probability for the right atom to exit from the + port depends on whether the left atom exits the + or the − port. The pair of atoms has a state, but the right atom by itself *doesn't have a state*, in the same way that an atom passing through an interferometer *doesn't have a position* and that love *doesn't have a color*.[9]

Leonard Susskind[10] puts it this way: If entangled states existed in auto mechanics as well as quantum mechanics, then an auto mechanic might tell you "I know everything about your car but ... I can't tell you anything about any of its parts."

---

[9] We noted on page 47 that Erwin Schrödinger came up with the *name* entanglement in 1935. But the *concept* of entanglement was expressed quite plainly in 1928 by Hermann Weyl, writing that if "two physical systems $\mathfrak{a}$ and $\mathfrak{b}$ are compounded to form a total system $\mathfrak{c}$ ... [then] *if the state of* $\mathfrak{a}$ *and the state of* $\mathfrak{b}$ *are known, the state of* $\mathfrak{c}$ *is in general not uniquely specified* .... In this significant sense quantum theory subscribes to the view that '*the whole is greater than the sum of its parts.*'" Hermann Weyl, *Gruppentheorie und Quantenmechanik* (S. Hirzel, Leipzig, 1928) pages 79–80. [Translated by H.P. Robertson as *The Theory of Groups and Quantum Mechanics* (Methuen and Company, London, 1931) pages 91–93. Translation reprinted by Dover Publications, New York, 1950.] Italics in original.

[10] Leonard Susskind and Art Friedman, *Quantum Mechanics: The Theoretical Minimum* (Basic Books, New York, 2014) page xii.

2.6.2 *"Collapse of the state vector"*

Set up this EPR experiment with the left analyzer 100 kilometers from the source, and the right analyzer 101 kilometers from the source. As soon as the left atom comes out of its − port, then it is known that the right atom will come out if its + port. The system is no longer in the entangled state $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$; instead the left atom is in state $|\downarrow\rangle$ and the right atom is in state $|\uparrow\rangle$. The state of the right atom has changed (some say it has "collapsed") despite the fact that it is 200 kilometers from the left analyzer that did the state changing!

This fact disturbs those who hold the misconception that states are physical things located out in space like nitrogen molecules, because it seems that information about state has made an instantaneous jump across 200 kilometers. In fact no information has been transferred from left to right: true, Alice at the left interferometer knows that the right atom will exit the + port 201 kilometers away, but Bob at the right interferometer *doesn't* have this information and won't unless she tells him in some conventional, light-speed-or-slower fashion.[11]

If Alice could in some magical way manipulate her atom to ensure that it would exit the − port, then she *could* send a message instantaneously. But Alice does not possess magic, so she cannot manipulate the left-bound atom in this way. Neither Alice, nor Bob, *nor even the left-bound atom itself* knows from which port it will exit. Neither Alice, nor Bob, *nor even the left-bound atom itself* can influence from which port it will exit.[12]

---

[11] If you are familiar with gauges in electrodynamics, you will find quantal state similar to the Coulomb gauge. In the Coulomb gauge, the electric potential at a point in space changes the instant that any charged particle moves, regardless of how far away that charged particle is. This does not imply that information moves instantly, because electric potential by itself is not measurable. The same applies for quantal state.

[12] There is a phenomenon with the unfortunate name of "quantum teleportation" that permits information to travel from one location to another location far away. The name suggests that the information travels instantaneously, but in fact it travels at the speed of light or slower. See Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels" *Physical Review Letters* **70** (29 March 1993) 1895–1899.

2.6.3 *Measurement and entanglement*

Back in section 1.4, "Light on the atoms" (page 36), we discussed the character of "observation" or "measurment" in quantum mechanics. Let's bring our new machinery concerning quantal states to bear on this situation.

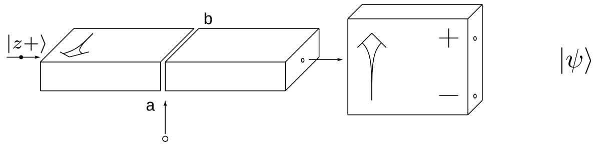
The figure on the next page shows, in the top panel, a potential measurement about to happen. An atom (represented by a black dot) in state $|z+\rangle$ approaches a horizontal interferometer at the same time that a photon (represented by a white dot) approaches path a of that interferometer.

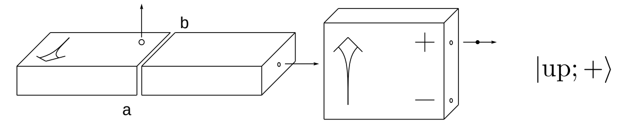
We employ a simplified model in which the photon either misses the atom, in which case it continues undeflected upward, or else the photon interacts with the atom, in which case it is deflected outward from the page. In this model there are four possible outcomes, shown in the bottom four panels of the figure.

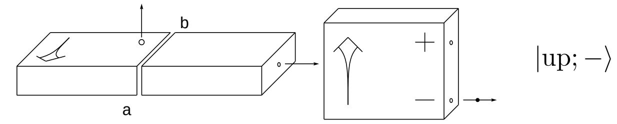
After this potential measurement, the system of photon plus atom is in an entangled state: the states shown on the right must list *both* the condition of the photon ("up" or "out") *and* the condition of the atom (+ or −).

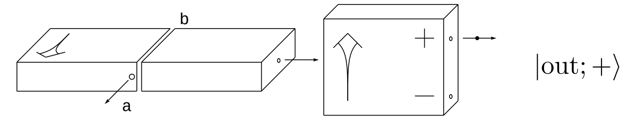
If the photon misses the atom, then the atom must emerge from the + port of the analyzer: there is zero probability that the system has final state $|\text{up}; -\rangle$. But if the photon interacts with the atom, then the atom might emerge from either port: there is non-zero probability that the system has final state $|\text{out}; -\rangle$. These two states are exactly the same as far as the atom is concerned; they differ only in the position of the photon.

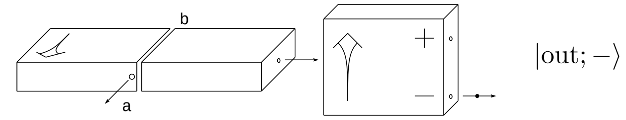
If we focus only on the atom, we would say that something strange has happened (a "measurement" at path a) that enabled the atom to emerge from the − port which (in the absence of "measurement") that atom would never do. But if we focus on the entire system of photon plus atom, then it is an issue of entanglement, not of measurement.











Problem

2.15 Amplitudes for "Measurement and entanglement"
Suppose that, in the "simplified model" for measurement and entanglement, the probability for photon deflection is $\frac{1}{5}$. Find the four probabilities $|\langle \text{up}; +|\psi\rangle|^2$, $|\langle \text{up}; -|\psi\rangle|^2$, $|\langle \text{out}; +|\psi\rangle|^2$, and $|\langle \text{out}; -|\psi\rangle|^2$.

2.7 What is a qubit?

At the end of the last chapter (on page 55) we listed several so-called "two-state systems" or "spin-$\frac{1}{2}$ systems" or "qubit systems". You might have found these terms strange: There are an infinite number of states for the magnetic moment of a silver atom: $|z+\rangle$, $|1°+\rangle$, $|2°+\rangle$, and so forth. Where does the name "two-state system" come from? You now see the answer: it's short for "two-basis-state system".

The term "spin" originated in the 1920s when it was thought that an electron was a classical charged rigid sphere that created a magnetic moment through spinning about an axis. A residual of that history is that people still call[13] the state $|z+\rangle$ by the name "spin up" and by the symbol $|\uparrow\rangle$, and the state $|z-\rangle$ by "spin down" and $|\downarrow\rangle$. (Sometimes the association is made in the opposite way.) Meanwhile the state $|x+\rangle$ is given the name "spin sideways" and the symbol $|\rightarrow\rangle$.

Today, two-basis-state systems are more often called "qubit" systems from the term used in quantum information processing. In a classical computer, like the ones we use today, a bit of information can be represented physically by a patch of magnetic material on a disk: the patch magnetized "up" is interpreted as a 1, the patch magnetized "down" is interpreted as a 0. Those are the only two possibilities. In a quantum computer, a qubit of information can be represented physically by the magnetic moment of a silver atom: the atom in state $|z+\rangle$ is interpreted as $|1\rangle$, the atom in state $|z-\rangle$ is interpreted as $|0\rangle$. But the atom might be in any (normalized) superposition $a|1\rangle + b|0\rangle$, so rather than two possibilities there are an infinite number.

Furthermore, qubits can interfere with and become entangled with other qubits, options that are simply unavailable to classical bits. With more states, and more ways to interact, quantum computers can only be faster than classical computers, and even as I write these possibilities are being explored.

In today's state of technology, quantum computers are hard to build, and they may never live up to their promise. But maybe they will.

---

[13] The very most precise and pedantic people restrict the term "spin" to elementary particles, such as electrons and neutrinos. For composite systems like the silver atom they speak instead of "the total angular momentum $\vec{J}$ of the silver atom in its ground state, projected on a given axis, and divided by $\hbar$." For me, the payoff in precision is not worth the penalty in polysyllables.

Chapters 1 and 2 have focused on two-basis-state systems, but of course nature provides other systems as well. For example, the magnetic moment of a nitrogen atom (mentioned on page 11) is a "four-basis-state" system, where one basis is

$$|z;+2\rangle, \quad |z;+1\rangle, \quad |z;-1\rangle, \quad |z;-2\rangle. \tag{2.38}$$

And chapter 6 shifts our focus to a system with an infinite number of basis states.

2.8 Photon polarization

This book develops the principles of quantum mechanics using a particular system, the magnetic moment of a silver atom, which has two basis states. Another system with two basis states is polarized light. I do not use this system mainly because photons are less familiar than atoms. These problems develop the quantum mechanics of photon polarization much as the text developed the quantum mechanics of magnetic moment.

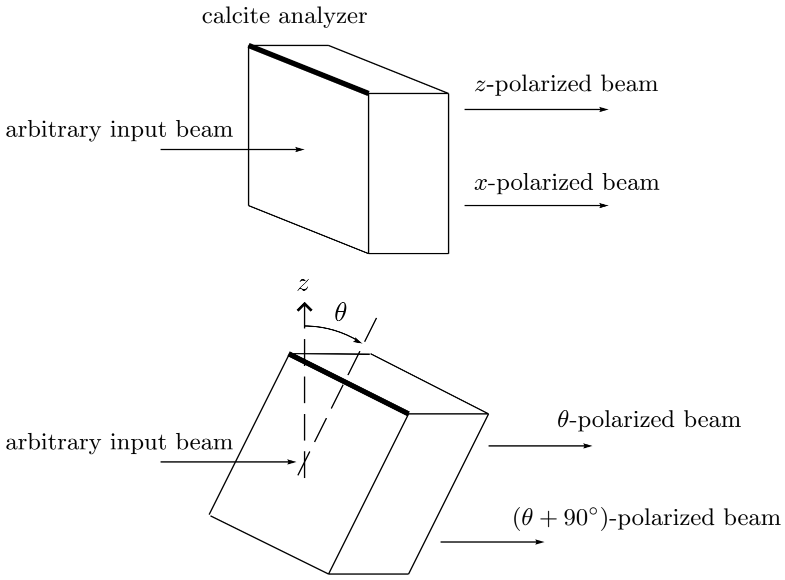
One cautionary note: There is always a tendency to view the photon as a little bundle of electric and magnetic fields, a "wave packet" made up of these familiar vectors. This view is completely incorrect. In quantum electrodynamics, in fact, the electric field is a classical macroscopic quantity that takes on meaning only when a large number of photons are present.

2.16 **Classical description of polarized light**

When a beam of unpolarized light passes through an ideal polarizing sheet, the emerging beam is of lower intensity and is "polarized", that is, the electric field vector undulates but points only parallel or antiparallel to the polarizing axis of that sheet. When a beam of vertically polarized light (a "z-polarized beam") is passed through an ideal polarizing sheet with polarizing axis oriented at an angle θ to the vertical, the beam is reduced in intensity and emerges with electric field undulating parallel to that sheet's polarizing axis (a "θ-polarized beam"). The sheet performs these feats by absorbing any component of electric field perpendicular to its polarizing axis. Recall that the intensity of a light beam is proportional to the square of the maximum value of the undulating electric field. Show that if the incoming z-polarized beam has intensity I_0, then the outgoing θ-polarized beam has intensity $I_0 \cos^2\theta$. Show that this expression gives the expected results when θ is 0°, 90°, 180° or 270°.

2.17 Quantal description of polarized light: Analyzers

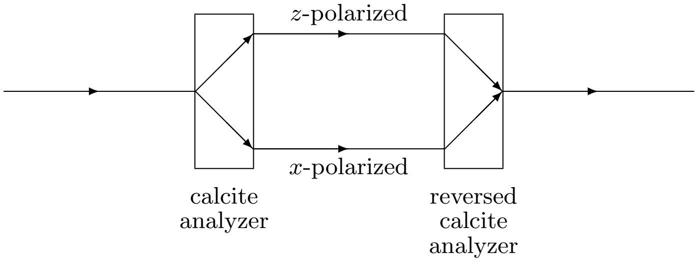
In quantum mechanics, a photon state is described by three quantities: energy, direction of motion, and polarization. We ignore the first two quantities. There are an infinite number of possible polarization states: each photon in a z-polarized beam is in the $|z\rangle$ state, each photon in a θ-polarized beam ($0^\circ \leq \theta < 180^\circ$) is in the $|\theta\rangle$ state, etc. In the quantum description, when a photon in state $|z\rangle$ encounters a polarizing sheet oriented at angle θ to the vertical, then either it is absorbed (with probability $\sin^2\theta$) or else it emerges as a photon in state $|\theta\rangle$ (with probability $\cos^2\theta$). A polarizing sheet is thus not an analyzer: whereas an analyzer would split the incident beam into two (or more) beams, the polarizing sheet absorbs one of the beams that an analyzer would emit. An analyzer can instead be constructed out of any material that exhibits double refraction, such as a calcite crystal:



What are the probabilities $|\langle z|\theta\rangle|^2$, $|\langle z|\theta + 90^\circ\rangle|^2$?

2.18 Interference

As usual, two analyzers, one inserted backwards, make up an analyzer loop.



Invent a series of experiments that demonstrates quantum interference. (I used input photons in state $|z\rangle$, passed through an analyzer loop rotated at angle θ to the vertical, followed by a vertical analyzer. But you might develop some other arrangement.) Show that the results of these experiments, plus the results of problem 2.17, are consistent with the amplitudes

$$\begin{array}{ll} \langle z|\theta\rangle = \cos\theta & \langle z|\theta+90^\circ\rangle = -\sin\theta \\ \langle x|\theta\rangle = \sin\theta & \langle x|\theta+90^\circ\rangle = \cos\theta. \end{array} \tag{2.39}$$

2.19 Circular polarization

Just as it is possible to analyze any light beam into z- and x-polarized beams, or into θ- and $(\theta+90^\circ)$-polarized beams, so it is possible to analyze any beam into right- and left-circularly polarized beams. (There is also "elliptically polarized light", that interpolates smoothly between circular and linear polarization.) Classical optics shows that any linearly polarized beam splits half-and-half into right- and left-circularly polarized light when so analyzed.

Quantum mechanics maintains that right- and left-circularly polarized beams are made up of photons in the $|R\rangle$ and $|L\rangle$ states, respectively. The amplitudes thus have magnitudes

$$\begin{array}{l} |\langle R|\ell p\rangle| = 1/\sqrt{2} \\ |\langle L|\ell p\rangle| = 1/\sqrt{2} \end{array} \tag{2.40}$$

where $|\ell p\rangle$ is any linearly polarized state. An RL analyzer loop is described through the equation

$$\langle\theta|R\rangle\langle R|z\rangle + \langle\theta|L\rangle\langle L|z\rangle = \langle\theta|z\rangle = \cos\theta. \tag{2.41}$$

Show that no real valued amplitudes can satisfy both relations (2.40) and (2.41), but that the complex values

$$\begin{array}{ll} \langle L|\theta\rangle = e^{i\theta}/\sqrt{2} & \langle L|z\rangle = 1/\sqrt{2} \\ \langle R|\theta\rangle = e^{-i\theta}/\sqrt{2} & \langle R|z\rangle = 1/\sqrt{2} \end{array} \tag{2.42}$$

are satisfactory!

Problems

2.20 **Analysis of a poetic sentence**

The poet Christian Wiman writes[14] that "If quantum entanglement is true, if related particles react in similar or opposite ways even when separated by tremendous distances, then it is obvious that the whole world is alive and communicating in ways we do not fully understand."

Critique this sentence.

2.21 **Questions** (recommended problem)

Update your list of quantum mechanics questions that you started at problem 1.13 on page 56. Write down new questions and, if you have uncovered answers to any of your old questions, write them down briefly.

⟦For example, one of my questions would be: "I'd like to see a proof that the global phase freedom mentioned on page 75, which obviously changes the amplitudes computed, does not change any experimentally accessible result."⟧

---

[14] *My Bright Abyss* (Farrar, Straus and Giroux, New York, 2013) page 35. See also pages 51–52.

Chapter 3

Refining Mathematical Tools

3.1 Products and operators

3.1.1 *The inner product*

We started with an expression for amplitude $\langle\phi|\psi\rangle$. At equation (2.27) we learned how to divorce this one complex number, involving both $|\phi\rangle$ and $|\psi\rangle$, into a dot product between a row matrix concerning $|\phi\rangle$ alone and a column matrix concerning $|\psi\rangle$ alone:

$$\langle\phi|\psi\rangle = \begin{pmatrix} \phi_+^* & \phi_-^* \end{pmatrix} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \phi_+^*\psi_+ + \phi_-^*\psi_-. \tag{3.1}$$

We say that the 2×1 column matrix

$$\begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} \text{ represents the state } |\psi\rangle \text{ as a “ket”.} \tag{3.2}$$

In a parallel development for the left-hand side, we say that the 1×2 row matrix

$$\begin{pmatrix} \phi_+^* & \phi_-^* \end{pmatrix} \text{ represents the state } \langle\phi| \text{ as a “bra”.} \tag{3.3}$$

In this context, the dot product is called an “inner product” or a “bracket”.

Exercise 3.A. In a certain basis, the states $|\psi\rangle$ and $|\phi\rangle$ are represented by

$$|\psi\rangle \doteq \tfrac{1}{5}\begin{pmatrix} -3 \\ 4i \end{pmatrix} \qquad |\phi\rangle \doteq \tfrac{1}{7}\begin{pmatrix} 2+3i \\ 6 \end{pmatrix}.$$

What is the inner product $\langle\psi|\phi\rangle$? What is $\langle\phi|\psi\rangle$?
Answers: $\langle\psi|\phi\rangle = -\frac{1}{35}(6+33i)$, $\langle\phi|\psi\rangle = -\frac{1}{35}(6-33i)$.

Exercise 3.B. Suppose $|\chi'\rangle = e^{i\theta}|\chi\rangle$. What is $\langle\chi'|$ in terms of $\langle\chi|$?

3.1.2 *The outer product*

Let's go back to our equation that represents the interference experiment: For any states $|\phi\rangle$ and $|\psi\rangle$, and for any pair of basis states $|a\rangle$ and $|b\rangle$,

$$\langle\phi|\psi\rangle = \langle\phi|a\rangle\langle a|\psi\rangle + \langle\phi|b\rangle\langle b|\psi\rangle.$$

Now effect the divorce of each amplitude into an inner product of states:

$$\langle\phi|\psi\rangle = \langle\phi| \Big\{ |a\rangle\langle a| + |b\rangle\langle b| \Big\} |\psi\rangle. \tag{3.4}$$

Our question: What's that thing between curly brackets?

In any particular basis, $|a\rangle$ is represented by a 2×1 column matrix, while $\langle a|$ is represented by a 1×2 row matrix. Thus the product $|a\rangle\langle a|$ is represented by a 2×2 square matrix. Similarly for $|b\rangle\langle b|$. Thus, in any particular basis, the thing between curly brackets is represented by a 2×2 matrix.

If this confuses you, then think of it this way. If

$$|\alpha\rangle \doteq \begin{pmatrix} \alpha_a \\ \alpha_b \end{pmatrix} \qquad \text{and} \qquad |\beta\rangle \doteq \begin{pmatrix} \beta_a \\ \beta_b \end{pmatrix},$$

then

$$\langle\alpha| \doteq \begin{pmatrix} \alpha_a^* & \alpha_b^* \end{pmatrix} \qquad \text{and} \qquad \langle\beta| \doteq \begin{pmatrix} \beta_a^* & \beta_b^* \end{pmatrix}.$$

The "inner product" is the 1×1 matrix

$$\langle\alpha|\beta\rangle = \begin{pmatrix} \alpha_a^* & \alpha_b^* \end{pmatrix} \begin{pmatrix} \beta_a \\ \beta_b \end{pmatrix} = \alpha_a^*\beta_a + \alpha_b^*\beta_b,$$

while the "outer product" is represented by the 2×2 matrix

$$|\alpha\rangle\langle\beta| \doteq \begin{pmatrix} \alpha_a \\ \alpha_b \end{pmatrix} \begin{pmatrix} \beta_a^* & \beta_b^* \end{pmatrix} = \begin{pmatrix} \alpha_a\beta_a^* & \alpha_a\beta_b^* \\ \alpha_b\beta_a^* & \alpha_b\beta_b^* \end{pmatrix}.$$

A piece of terminology: The outer product $|\alpha\rangle\langle\beta|$ is called an "operator" and the square matrix that represents it in a particular basis is called a "matrix". The two terms are often used interchangeably, but if you care to make the distinction then this is how to make it. It's conventional to symbolize operators with hats, like $\hat{A}$.

Exercise 3.C. In a certain basis, the states $|\psi\rangle$ and $|\phi\rangle$ are represented by

$$|\psi\rangle \doteq \tfrac{1}{5}\begin{pmatrix} -3 \\ 4i \end{pmatrix} \qquad |\phi\rangle \doteq \tfrac{1}{7}\begin{pmatrix} 2+3i \\ 6 \end{pmatrix}.$$

What is the outer product $|\psi\rangle\langle\phi|$? What is $|\phi\rangle\langle\psi|$?

Answers:

$$|\psi\rangle\langle\phi| = -\tfrac{1}{35}\begin{pmatrix} (6-9i) & 18 \\ (-12-8i) & -24i \end{pmatrix}, \quad |\phi\rangle\langle\psi| = -\tfrac{1}{35}\begin{pmatrix} (6+9i) & (-12+8i) \\ 18 & 24i \end{pmatrix}.$$

Exercise 3.D. Suppose $|\chi'\rangle = e^{i\theta}|\chi\rangle$. What is $|\chi'\rangle\langle\chi'|$ in terms of $|\chi\rangle\langle\chi|$?

With these ideas in place, we see what's inside the curly brackets of expression (3.4) — it's the identity operator

$$\hat{1} = |a\rangle\langle a| + |b\rangle\langle b|,$$

and this holds true for any basis $\{|a\rangle, |b\rangle\}$.

We check this out two ways. First, in the basis $\{|z+\rangle, |z-\rangle\}$, we find the representation for the operator

$$|z+\rangle\langle z+| + |z-\rangle\langle z-|.$$

Remember that in this basis

$$|z+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{while} \qquad |z-\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

so

$$|z+\rangle\langle z+| \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.5}$$

Meanwhile

$$|z-\rangle\langle z-| \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.6}$$

Thus

$$|z+\rangle\langle z+| + |z-\rangle\langle z-| \doteq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Yes! As required, this combination is the identity matrix, which is of course the representation of the identity operator.

For our second check, in the basis $\{|z+\rangle, |z-\rangle\}$ we find the representation for the operator

$$|\theta+\rangle\langle\theta+| + |\theta-\rangle\langle\theta-|.$$

Remember (equation 2.28) that in this basis

$$|\theta+\rangle \doteq \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} \qquad \text{while} \qquad |\theta-\rangle \doteq \begin{pmatrix} -\sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix},$$

so

$$\begin{aligned} |\theta+\rangle\langle\theta+| &\doteq \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\theta/2) & \cos(\theta/2)\sin(\theta/2) \\ \sin(\theta/2)\cos(\theta/2) & \sin^2(\theta/2) \end{pmatrix}. \end{aligned} \tag{3.7}$$

Meanwhile

$$|\theta-\rangle\langle\theta-| \doteq \begin{pmatrix} -\sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix} \begin{pmatrix} -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix} \tag{3.8}$$
$$= \begin{pmatrix} \sin^2(\theta/2) & -\sin(\theta/2)\cos(\theta/2) \\ -\cos(\theta/2)\sin(\theta/2) & \cos^2(\theta/2) \end{pmatrix}.$$

(As a check, notice that when $\theta = 0$, equation (3.7) reduces to equation (3.5), and equation (3.8) reduces to equation (3.6).) Thus

$$\begin{aligned} |\theta+\rangle\langle\theta+| + |\theta-\rangle\langle\theta-| &\doteq \begin{pmatrix} \cos^2(\theta/2) & \cos(\theta/2)\sin(\theta/2) \\ \sin(\theta/2)\cos(\theta/2) & \sin^2(\theta/2) \end{pmatrix} \\ &\quad + \begin{pmatrix} \sin^2(\theta/2) & -\sin(\theta/2)\cos(\theta/2) \\ -\cos(\theta/2)\sin(\theta/2) & \cos^2(\theta/2) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Yes! Once again this combination is the identity matrix.

3.1.3 *Silver atom in a magnetic field*

A silver atom is exposed to a magnetic field B in the x-direction for a time t. What is its state at the end of this time? The answer is not obvious, but here it is. If the atom starts off in state $|x+\rangle$, it becomes $e^{i\gamma}|x+\rangle$, where $\gamma = \mu_B Bt/\hbar$. If it starts off in state $|x-\rangle$, it becomes $e^{-i\gamma}|x-\rangle$. If it starts off in some arbitrary state $|\psi_0\rangle$, it becomes

$$|\psi_t\rangle = e^{i\gamma}|x+\rangle\langle x+|\psi_0\rangle + e^{-i\gamma}|x-\rangle\langle x-|\psi_0\rangle. \tag{3.9}$$

We interpret this equation according to the three amplitude rules on page 60. The atom has two paths forward in time: the first path is as $|x+\rangle$ and the second path is as $|x-\rangle$. (These paths are not physically separate in space, but they are two separate paths nevertheless.) If it takes the first path, the amplitude of moving forward in time is $e^{i\gamma}$. So the first term in equation (3.9) is the amplitude $\langle x+|\psi_0\rangle$ that the first path is taken, times the amplitude $e^{i\gamma}$ of moving forward in time along that path. (Multiply amplitudes in series.) The second term has a similar interpretation, and as usual we sum the amplitudes for the two parallel paths.

We can again effect the divorce and write expression (3.9) as

$$|\psi_t\rangle = \left\{ e^{i\gamma}|x+\rangle\langle x+| + e^{-i\gamma}|x-\rangle\langle x-| \right\} |\psi_0\rangle, \tag{3.10}$$

and we can write the expression in curly brackets as the "time evolution operator"

$$\hat{U} = e^{i\gamma}|x+\rangle\langle x+| + e^{-i\gamma}|x-\rangle\langle x-|. \tag{3.11}$$

The time evolution operator has nothing to do with the initial or final states.

Exercise 3.E. What happens at the special time $t_S = \pi\hbar/\mu_B B$? At twice that time?

Exercise 3.F. Write the matrix representing operator $\hat{U}$ in the $\{|z+\rangle, |z-\rangle\}$ basis.

3.2 Measurement

What happens when an atom in state $|\psi\rangle$ passes through a θ-analyzer? Or, what is the same thing, what happens when an atom in state $|\psi\rangle$ is measured to find the projection of $\vec{\mu}$ on the θ axis? (We call the projection of $\vec{\mu}$ on the θ axis μ_θ.)

The atom enters the analyzer in state $|\psi\rangle$. It has two possible fates:

- It emerges from the + port, in which case the atom has been measured to have $\mu_\theta = +\mu_B$, and it emerges in state $|\theta+\rangle$. This happens with probability $|\langle\theta+|\psi\rangle|^2$.
- It emerges from the − port, in which case the atom has been measured to have $\mu_\theta = -\mu_B$, and it emerges in state $|\theta-\rangle$. This happens with probability $|\langle\theta-|\psi\rangle|^2$.

What is the mean[1] value of μ_θ?

$$\begin{aligned}\langle\mu_\theta\rangle &= (+\mu_B)|\langle\theta+|\psi\rangle|^2 + (-\mu_B)|\langle\theta-|\psi\rangle|^2 \\ &= (+\mu_B)\langle\theta+|\psi\rangle^*\langle\theta+|\psi\rangle + (-\mu_B)\langle\theta-|\psi\rangle^*\langle\theta-|\psi\rangle \\ &= (+\mu_B)\langle\psi|\theta+\rangle\langle\theta+|\psi\rangle + (-\mu_B)\langle\psi|\theta-\rangle\langle\theta-|\psi\rangle \\ &= \langle\psi|\left\{(+\mu_B)|\theta+\rangle\langle\theta+| + (-\mu_B)|\theta-\rangle\langle\theta-|\right\}|\psi\rangle\end{aligned}$$

---

[1] The "mean value" is also called the "average value" and sometimes the "expected value" or the "expectation value". The latter name is particularly poor. If you toss a die, the mean value of the number facing up is 3.5. Yet no one expects to toss a die and find the number 3.5 facing up!

In the last line we have again effected the divorce — writing amplitudes in terms of inner products between states. The part in curly brackets is again independent of the state.

Given the last line, it makes sense to define an operator associated with the measurement of μ_θ, namely

$$\hat{\mu}_\theta = (+\mu_B)|\theta+\rangle\langle\theta+| + (-\mu_B)|\theta-\rangle\langle\theta-|, \tag{3.12}$$

so that if the atom is in state $|\psi\rangle$ and the value of μ_θ is measured, then the mean value of the measurement is

$$\langle\mu_\theta\rangle = \langle\psi|\hat{\mu}_\theta|\psi\rangle. \tag{3.13}$$

Notice what we've done here: To find the mean value of μ_θ for a particular atom, we've split up the problem into an operator $\hat{\mu}_\theta$ involving only the measuring device and a state $|\psi\rangle$ involving only the atomic state.

And notice what we have *not* done here. The operator $\hat{\mu}_\theta$ does not act upon the state of the atom going into the analyzer to produce the state of the atom going out of the analyzer: In fact that output state is unknown. That is how the time evolution operator (3.11) behaves, but it is *not* how the measurement operator (3.12) behaves.

3.2.1 *Sample Problem: Matrix representation of $\hat{\mu}_\theta$*

What is the matrix representation of $\hat{\mu}_\theta$ in the basis $\{|z+\rangle, |z-\rangle\}$? Evaluate for the special cases $\theta = 0$, $\theta = 90°$, and $\theta = 180°$.

We have already found representations for the outer product $|\theta+\rangle\langle\theta+|$ at equation (3.7) and for the outer product $|\theta-\rangle\langle\theta-|$ at equation (3.8). Using these expressions

$$\begin{aligned}
\hat{\mu}_\theta &= (+\mu_B)|\theta+\rangle\langle\theta+| + (-\mu_B)|\theta-\rangle\langle\theta-| \\
&\doteq (+\mu_B)\begin{pmatrix} \cos^2(\theta/2) & \cos(\theta/2)\sin(\theta/2) \\ \sin(\theta/2)\cos(\theta/2) & \sin^2(\theta/2) \end{pmatrix} \\
&\qquad +(-\mu_B)\begin{pmatrix} \sin^2(\theta/2) & -\sin(\theta/2)\cos(\theta/2) \\ -\cos(\theta/2)\sin(\theta/2) & \cos^2(\theta/2) \end{pmatrix} \\
&= \mu_B \begin{pmatrix} \cos^2(\theta/2) - \sin^2(\theta/2) & 2\cos(\theta/2)\sin(\theta/2) \\ 2\cos(\theta/2)\sin(\theta/2) & \sin^2(\theta/2) - \cos^2(\theta/2) \end{pmatrix} \\
&= \mu_B \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \qquad (3.14)
\end{aligned}$$

where in the last line I have used the trigonometric half-angle formulas that everyone memorized in high school and then forgot. (I forgot them too, but I know where to look them up.)

In particular, using the values $\theta = 0$, $\theta = 90°$, and $\theta = 180°$,

$$\hat{\mu}_z \doteq \mu_B \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \hat{\mu}_x \doteq \mu_B \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\mu}_{(-z)} \doteq \mu_B \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \qquad (3.15)$$

Furthermore

$$\hat{\mu}_\theta = \cos\theta\, \hat{\mu}_z + \sin\theta\, \hat{\mu}_x.$$

Which is convenient because the unit vector $\hat{r}$ in the direction of θ is

$$\hat{r} = \cos\theta\, \hat{k} + \sin\theta\, \hat{\imath}.$$

(In the first equation, a hat represents an operator. In the second, it represents a unit vector.)

So, knowing the operator associated with a measurement, we can easily find the resulting mean value for any given state when measured. But we often want to know more than the mean. We want to know also the standard deviation. Indeed we would like to know everything about the measurement: the possible results, the probability of each result, the state the system will be in after the measurement is performed. Surprisingly, all this information is wrapped up within the measurement operator as well.

We know that there are only two states that have a definite value of μ_θ, namely $|\theta+\rangle$ and $|\theta-\rangle$. How do these states behave when acted upon by the operator $\hat{\mu}_\theta$?

$$\begin{aligned}\hat{\mu}_\theta|\theta+\rangle &= \Big\{(+\mu_B)|\theta+\rangle\langle\theta+| + (-\mu_B)|\theta-\rangle\langle\theta-|\Big\}|\theta+\rangle\\ &= (+\mu_B)|\theta+\rangle\langle\theta+|\theta+\rangle + (-\mu_B)|\theta-\rangle\langle\theta-|\theta+\rangle\\ &= (+\mu_B)|\theta+\rangle(1) + (-\mu_B)|\theta-\rangle(0)\\ &= (+\mu_B)|\theta+\rangle\end{aligned}$$

In other words, when the operator $\hat{\mu}_\theta$ acts upon the state $|\theta+\rangle$, the result is $(+\mu_B)$ times that same state $|\theta+\rangle$ — and $(+\mu_B)$ is exactly the result that we would always obtain if we measured μ_θ for an atom in state $|\theta+\rangle$! A parallel result holds for $|\theta-\rangle$.

To convince you of how rare this phenomena is, let me apply the operator $\hat{\mu}_\theta$ to some other state, say $|z+\rangle$. The result is

$$\begin{aligned}\hat{\mu}_\theta|z+\rangle &= \Big\{(+\mu_B)|\theta+\rangle\langle\theta+| + (-\mu_B)|\theta-\rangle\langle\theta-|\Big\}|z+\rangle\\ &= (+\mu_B)|\theta+\rangle\langle\theta+|z+\rangle + (-\mu_B)|\theta-\rangle\langle\theta-|z+\rangle\\ &= (+\mu_B)|\theta+\rangle(\cos(\theta/2)) + (-\mu_B)|\theta-\rangle(-\sin(\theta/2)).\end{aligned}$$

But

$$\begin{aligned}|\theta+\rangle &= |z+\rangle\langle z+|\theta+\rangle + |z-\rangle\langle z-|\theta+\rangle = |z+\rangle(\quad\cos(\theta/2)) + |z-\rangle(\sin(\theta/2))\\ |\theta-\rangle &= |z+\rangle\langle z+|\theta-\rangle + |z-\rangle\langle z-|\theta-\rangle = |z+\rangle(-\sin(\theta/2)) + |z-\rangle(\cos(\theta/2)),\end{aligned}$$

so

$$\begin{aligned}\hat{\mu}_\theta|z+\rangle &= (+\mu_B)|\theta+\rangle(\cos(\theta/2)) + (-\mu_B)|\theta-\rangle(-\sin(\theta/2))\\ &= \mu_B\left[|z+\rangle(\cos^2(\theta/2) - \sin^2(\theta/2)) + |z-\rangle(2\cos(\theta/2)\sin(\theta/2))\right]\\ &= \mu_B\left[|z+\rangle\cos\theta + |z-\rangle\sin\theta\right],\end{aligned}$$

where in the last line I have again used the half-remembered half-angle formulas.

The upshot is that most of the time, $\hat{\mu}_\theta$ acting upon $|z+\rangle$ does *not* produce a number times $|z+\rangle$ — most of the time it produces some combination of $|z+\rangle$ and $|z-\rangle$. In fact the only case in which $\hat{\mu}_\theta$ acting upon $|z+\rangle$ produces a number times $|z+\rangle$ is when $\sin\theta = 0$, that is when $\theta = 0$ or when $\theta = 180°$.

The states when $\hat{\mu}_\theta$ acting upon $|\psi\rangle$ produces a number times the original state $|\psi\rangle$ are rare: they are called "eigenstates". The associated numbers are called "eigenvalues". We have found the two eigenstates of $\hat{\mu}_\theta$: they are $|\theta+\rangle$ with eigenvalue $+\mu_B$ and $|\theta-\rangle$ with eigenvalue $-\mu_B$.

$$\begin{aligned} \hat{\mu}_\theta|\theta+\rangle &= (+\mu_B)|\theta+\rangle \qquad \text{eigenstate } |\theta+\rangle \text{ with eigenvalue } +\mu_B \\ \hat{\mu}_\theta|\theta-\rangle &= (-\mu_B)|\theta-\rangle \qquad \text{eigenstate } |\theta-\rangle \text{ with eigenvalue } -\mu_B \end{aligned}$$

The eigenstates are the states with definite values of μ_θ. And the eigenvalues are those values!

The German word *eigen* derives from the same root as the English word "own", as in "my own state". It means "associated with" "peculiar to" or "belonging to". The eigenstate $|\theta-\rangle$ is the state "belonging to" a θ projection of value $-\mu_B$.

Summary: The quantum theory of measurement

This summarizes the quantum theory of measurement as applied to the measurement of $\vec{\mu}$ projected onto the unit vector in the direction of θ:

The operator $\hat{\mu}_\theta$ has two eigenstates which constitute a complete and orthonormal basis:

state $|\theta+\rangle$ with eigenvalue $+\mu_B$
state $|\theta-\rangle$ with eigenvalue $-\mu_B$

(a) If you measure μ_θ of an atom in an eigenstate of $\hat{\mu}_\theta$, then the number measured will be the corresponding eigenvalue, and the atom will remain in that eigenstate.

(b) If you measure μ_θ of an atom in an arbitrary state $|\psi\rangle$, then the number measured will be one of the two eigenvalues of $\hat{\mu}_\theta$: It will be $+\mu_B$ with probability $|\langle\theta+|\psi\rangle|^2$, it will be $-\mu_B$ with probability $|\langle\theta-|\psi\rangle|^2$. If the value measured was $+\mu_B$, then the atom will leave in state $|\theta+\rangle$, if the value measured was $-\mu_B$, then the atom will leave in state $|\theta-\rangle$.

Exercise 3.G. Show that part **(a)** of the summary follows from **(b)**.

3.3 Are states and operators "real"?

This is a philosophical question for which there's no specific meaning and hence no specific answer. But in my opinion, states and operators are mathematical tools that enable us to efficiently and accurately calculate the probabilities that can be found through repeated analyzer experiments, interference experiments, and indeed all experiments.[2] They are not "real".

Indeed, it is possible to formulate quantum mechanics in such a way that probabilities and amplitudes are found without using the mathematical tools of "state" and "operator" at all. Richard Feynman[3] and Albert Hibbs do just this in their 1965 book *Quantum Mechanics and Path Integrals.* States and operators do not make an appearance until deep into their book, and even when they do appear they are not essential. In my opinion, this Feynman "sum over histories" formulation is the most intuitively appealing approach to quantum mechanics. There is, however, a price to be paid for this appeal: it's very difficult to work problems in the Feynman formulation.

3.4 Lightning linear algebra

Linear algebra provides many of the mathematical tools used in quantum mechanics. This section will scan through and summarize linear algebra to drive home the main points…it won't attempt to prove things or to develop the theory in the most elegant form using the smallest number of assumptions.

---

[2] For more extensive treatment, see N. David Mermin, "What's bad about this habit?" *Physics Today* **62** (5) (May 2009) 8–9, and the discussion about this essay in *Physics Today* **62** (9) (September 2009) 10–15.

[3] Richard Feynman (1918–1988) was an American theoretical physicist of unconventional outlook, exuberance, and style. He invented a practical technique for calculations in quantum electrodynamics, developed a model for weak decay, and wrote forcefully that "For a successful technology, reality must take precedence over public relations, for Nature cannot be fooled." [*What Do You Care What Other People Think?* (W.W. Norton, New York, 1988) page 237.]

3.4.1 *What is a vector?*

A “scalar” is either a real number (x) or a complex number (z).

A “vector” will be notated either as $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, or as $\vec{r}$ (particularly for vectors that are arrows), or as $|\psi\rangle$, $|\phi\rangle$, $|\chi\rangle$ (particularly for state vectors in quantum mechanics).

In addition, there must be some rule for multiplying a vector by a scalar and a rule for adding vectors, so that $\mathbf{a} + z\mathbf{b}$ is a vector.

I won’t define “vector” any more than I defined “number”. But I will give some examples:

arrows in 2- or 3- or N-dimensional space
n-tuples, with real entries or with complex entries
polynomials
functions
$n \times m$ matrices
functions that are “square-integrable” (a set called “L^2”)

A “square-integrable” function is a function $f(x)$ of a single real variable, either real- or complex-valued, such that the integral

$$\int_{-\infty}^{+\infty} |f(x)|^2 \, dx$$

is finite.

> ⟦This section describes the linear algebra concept of “vector” as developed by Giuseppe Peano in 1888 and generalized by David Hilbert and Erhard Schmidt in 1908. A different mathematical concept, which unfortunately uses the same name “vector”, is more in line with the idea of “vector as arrow” and readily generalizes to tensors. This different concept was developed by Gregorio Ricci-Curbastro and Tullio Levi-Civita in 1900. A polynomial is a vector in the first sense but not in the second. When you read in any math book about “vectors”, be sure you know which of the two different concepts is meant.⟧

3.4.2 *Inner product*

The "inner product" is a function from the ordered pairs of vectors to the scalars,

$$\text{IP}(\mathbf{a}, \mathbf{b}) = \text{ a real or complex number,} \tag{3.16}$$

that satisfies

$$\text{IP}(\mathbf{a}, \mathbf{b} + \mathbf{c}) = \text{IP}(\mathbf{a}, \mathbf{b}) + \text{IP}(\mathbf{a}, \mathbf{c}) \tag{3.17}$$

$$\text{IP}(\mathbf{a}, z\mathbf{b}) = z\,\text{IP}(\mathbf{a}, \mathbf{b}) \tag{3.18}$$

$$\text{IP}(\mathbf{a}, \mathbf{b}) = [\text{IP}(\mathbf{b}, \mathbf{a})]^* \tag{3.19}$$

$$\text{IP}(\mathbf{a}, \mathbf{a}) > 0 \quad \text{ unless } \mathbf{a} = \mathbf{0}. \tag{3.20}$$

It follows from equation (3.19) that $\text{IP}(\mathbf{a}, \mathbf{a})$ is real. Equation (3.20) demands also that it's positive.

Why is there a complex conjugation in equation (3.19)? Why not just demand that $\text{IP}(\mathbf{a}, \mathbf{b}) = \text{IP}(\mathbf{b}, \mathbf{a})$? The complex conjugation is needed for consistency with (3.20). If it weren't there, then

$$\text{IP}(i\mathbf{a}, i\mathbf{a}) = (i \cdot i)\text{IP}(\mathbf{a}, \mathbf{a}) = -\text{IP}(\mathbf{a}, \mathbf{a}) < 0.$$

Notation: $\text{IP}(\mathbf{a}, \mathbf{b}) = (\mathbf{a}, \mathbf{b}) = \mathbf{a}\cdot\mathbf{b}$, $\text{IP}(|\phi\rangle, |\psi\rangle) = \langle\phi|\psi\rangle$.

Definition: The "norm" of $|\psi\rangle$ is $\sqrt{\langle\psi|\psi\rangle}$.

Examples of inner products: For arrows in 3-dimensional space,

$$\vec{a} \cdot \vec{b} = (\text{length of } \vec{a})(\text{length of } \vec{b})(\text{cosine of the angle between } \vec{a} \text{ and } \vec{b}). \tag{3.21}$$

For n-tuples $\mathbf{a} = (a_1, a_2, \ldots a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots b_n)$,

$$\mathbf{a}\cdot\mathbf{b} = a_1^* b_1 + a_2^* b_2 + \cdots + a_n^* b_n. \tag{3.22}$$

For functions $\phi(x)$ and $\psi(x)$ in L^2, the inner product is

$$(\phi(x), \psi(x)) = \int_{-\infty}^{+\infty} \phi^*(x)\psi(x)\,dx. \tag{3.23}$$

Exercise 3.H. Show that the three "examples of inner products" listed above satisfy the four defining characteristics of the inner product given in equations (3.17) through (3.20).

One consequence of the definition of inner product is that

$$|\langle\phi|\psi\rangle| \leq \sqrt{\langle\phi|\phi\rangle}\,\sqrt{\langle\psi|\psi\rangle}. \tag{3.24}$$

This is called the "Schwarz inequality".

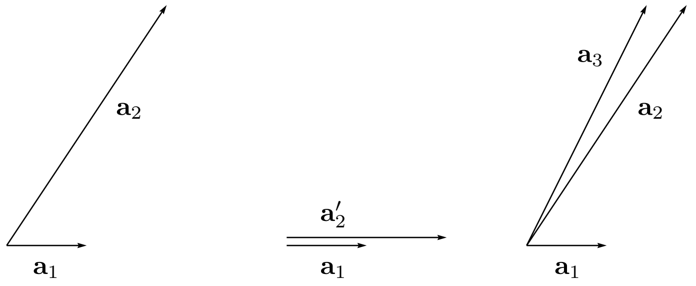
Exercise 3.I. Interpret the Schwarz inequality for position vectors in three-dimensional space.

Exercise 3.J. Prove the Schwarz inequality for any kind of vector by defining $|\chi\rangle = \langle\phi|\psi\rangle\,|\phi\rangle - \langle\phi|\phi\rangle\,|\psi\rangle$ and then using the fact that the norm of $|\chi\rangle$ is nonnegative.

3.4.3 *Building new vectors from old*

Given some vectors, say $\mathbf{a}_1$ and $\mathbf{a}_2$, what vectors can you build from them using scalar multiplication and vector addition?

Example: arrows in the plane.



(a) (b) (c)

In (a), any arrow in the plane can be built out of $\mathbf{a}_1$ and $\mathbf{a}_2$. In other words, any arrow in the plane can be written in the form $\mathbf{r} = r_1\mathbf{a}_1 + r_2\mathbf{a}_2$. We say that "the set $\{\mathbf{a}_1, \mathbf{a}_2\}$ spans the plane".

In (b), we cannot build the whole plane from $\mathbf{a}_1$ and $\mathbf{a}_2'$. These two vectors do not span the plane.

In (c), the set $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ spans the plane, but the set is redundant: you don't *need* all three. You can build $\mathbf{a}_3$ from $\mathbf{a}_1$ and $\mathbf{a}_2$: $\mathbf{a}_3 = \mathbf{a}_2 - \frac{1}{2}\mathbf{a}_1$, so anything that can be built from $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ can also be built from $\{\mathbf{a}_1, \mathbf{a}_2\}$.

A set is said to be "linearly independent" when you can't build any member of the set out of the other members. The set $\{\mathbf{a}_1, \mathbf{a}_2\}$ is linearly independent, the set $\{\mathbf{a}_1, \mathbf{a}_3\}$ is linearly independent, the set $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ is not.

So any arrow $\mathbf{r}$ in the plane has a unique representation in terms of $\{\mathbf{a}_1, \mathbf{a}_2\}$ but not in terms of $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$. For example,

$$\begin{aligned}\mathbf{r} = 2\mathbf{a}_3 &= -1\mathbf{a}_1 + 2\mathbf{a}_2 + 0\mathbf{a}_3 \\ &= \quad 0\mathbf{a}_1 + 0\mathbf{a}_2 + 2\mathbf{a}_3\end{aligned}$$

A spanning set of linearly independent vectors is called a "basis". A basis is a minimum set of building blocks from which any vector you want can be constructed. In any given basis, there is a unique representation for an arbitrary vector. It's easy to see that all bases have the same number of members, and this number is called the dimensionality, N.

The easiest basis to work with is an "orthonormal basis": A basis $\{|1\rangle, |2\rangle, \ldots, |N\rangle\}$ is orthonormal if

$$\langle n|m\rangle = \delta_{n,m}. \tag{3.25}$$

The symbol on the right-hand side is called the "Kronecker[4] delta":

$$\delta_{n,m} \equiv \begin{cases} 1 \text{ for } n = m \\ 0 \text{ for } n \neq m \end{cases}. \tag{3.26}$$

For any basis an arbitrary vector $|\psi\rangle$ can be written

$$|\psi\rangle = \psi_1|1\rangle + \psi_2|2\rangle + \cdots + \psi_N|N\rangle = \sum_{n=1}^{N} \psi_n|n\rangle, \tag{3.27}$$

but for many bases it's hard to find the coefficients ψ_n. For an orthonormal basis, however, it's easy. Take the inner product of basis member $|m\rangle$ with $|\psi\rangle$, giving

$$\langle m|\psi\rangle = \sum_{n=1}^{N} \psi_n \langle m|n\rangle = \sum_{n=1}^{N} \psi_n \delta_{m,n} = \psi_m. \tag{3.28}$$

Thus the expansion (3.27) is

$$|\psi\rangle = \sum_{n=1}^{N} |n\rangle\langle n|\psi\rangle. \tag{3.29}$$

---

[4] Leopold Kronecker (1823–1891), German mathematician. After earning his Ph.D. he spent a decade managing a farm, which made him financially comfortable enough that he could pursue mathematics research for the rest of his life as a private scholar without university position.

You have seen this formula in the context of arrows. For example, using two-dimensional arrows with the orthonormal basis $\{\hat{\imath}, \hat{\jmath}\}$, you know that

$$\vec{r} = x\,\hat{\imath} + y\,\hat{\jmath},$$

where

$$x = \hat{\imath} \cdot \vec{r} \quad \text{and} \quad y = \hat{\jmath} \cdot \vec{r}.$$

Thus

$$\vec{r} = \hat{\imath}\,(\hat{\imath} \cdot \vec{r}) + \hat{\jmath}\,(\hat{\jmath} \cdot \vec{r}),$$

which is just an instance of the more general expression (3.29).

3.4.4 *Representations*

Any vector $|\psi\rangle$ is completely specified by the N numbers $\psi_1, \psi_2, \ldots \psi_N$ (that is, the N numbers $\langle n|\psi\rangle$). We say that in the basis $\{|1\rangle, |2\rangle, \ldots, |N\rangle\}$, the vector $|\psi\rangle$ is "represented by" the column matrix

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} = \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \\ \vdots \\ \langle N|\psi\rangle \end{pmatrix}. \tag{3.30}$$

It is very easy to manipulate vectors through their representations, so representations are used often. So often, that some people go overboard and say that the vector $|\psi\rangle$ is *equal* to this column matrix. This is false. The matrix representation is a name for the vector, but is not equal to the vector — much as the word "tree" is a name for a tree, but is not the same as a tree. The symbol for "is represented by" is $\doteq$, so we write

$$|\psi\rangle \doteq \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} = \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \\ \vdots \\ \langle N|\psi\rangle \end{pmatrix}. \tag{3.31}$$

What can we do with representations? Here's a way to connect an inner product, which is defined solely through the list of properties (3.17)–(3.20), to a formula in terms of representations.

$$\begin{aligned}
&\langle\phi|\psi\rangle && [\![\text{ using (3.29) } \ldots]\!] \\
&= \langle\phi| \left\{ \sum_n |n\rangle\langle n|\psi\rangle \right\} && [\![\text{ using (3.17) } \ldots]\!] \\
&= \sum_n \langle\phi|n\rangle\langle n|\psi\rangle && [\![\text{ using (3.19) } \ldots]\!] \\
&= \sum_n \phi_n^* \psi_n \\
&= (\phi_1^* \quad \phi_2^* \quad \cdots \quad \phi_N^*) \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}
\end{aligned}$$

We will sometimes say that $\langle\phi|$ is the "dual vector" to $|\phi\rangle$ and is represented by the row matrix

$$(\phi_1^* \quad \phi_2^* \quad \cdots \quad \phi_N^*). \tag{3.32}$$

Transformation of representations

In the orthonormal basis $\{|1\rangle, |2\rangle, \ldots, |N\rangle\}$, the vector $|\psi\rangle$ is represented by an N-tuple

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}. \tag{3.33}$$

But in the different orthonormal basis $\{|1'\rangle, |2'\rangle, \ldots, |N'\rangle\}$, the vector $|\psi\rangle$ is represented by the different N-tuple

$$\begin{pmatrix} \psi'_1 \\ \psi'_2 \\ \vdots \\ \psi'_N \end{pmatrix}. \tag{3.34}$$

How are these two representations related?

$$\begin{aligned}
\psi'_n &= \langle n'|\psi\rangle \\
&= \langle n'| \left\{ \sum_m |m\rangle\langle m|\psi\rangle \right\} \\
&= \sum_m \langle n'|m\rangle\langle m|\psi\rangle
\end{aligned}$$

so

$$\begin{pmatrix} \psi'_1 \\ \psi'_2 \\ \vdots \\ \psi'_N \end{pmatrix} = \begin{pmatrix} \langle 1'|1\rangle & \langle 1'|2\rangle & \cdots & \langle 1'|N\rangle \\ \langle 2'|1\rangle & \langle 2'|2\rangle & \cdots & \langle 2'|N\rangle \\ \vdots & & & \vdots \\ \langle N'|1\rangle & \langle N'|2\rangle & \cdots & \langle N'|N\rangle \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}. \tag{3.35}$$

Exercise 3.K. Two orthonormal bases. A two-state system has orthonormal basis $\{|a_1\rangle, |a_2\rangle\}$. Show that the set

$$\begin{aligned} |b_1\rangle &= \cos\phi\,|a_1\rangle + \sin\phi\,|a_2\rangle \\ |b_2\rangle &= \mp\sin\phi\,|a_1\rangle \pm \cos\phi\,|a_2\rangle, \end{aligned}$$

where ϕ is any real number whatsoever, is also orthonormal.

3.4.5 *Operators*

In introductory calculus, a function f is a rule that associates with each number x another number $y = f(x)$. The concept of "function" can be extended to vectors, but it is traditional to call such functions "operators": an operator $\hat{A}$ is a rule that associates with each vector $|\psi\rangle$ another vector $|\phi\rangle$:

$$|\phi\rangle = \hat{A}|\psi\rangle. \tag{3.36}$$

We have seen that one may multiply a vector by a scalar or add two vectors. Are there similar operations for operators? There are. The product of scalar c times operator $\hat{A}$ is the operator $(c\hat{A})$ where

$$(c\hat{A})|\psi\rangle = c(\hat{A}|\psi\rangle). \tag{3.37}$$

The sum of two operators is defined through

$$(\hat{A}+\hat{B})|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle. \tag{3.38}$$

Furthermore, the product of two operators is defined as the action of the two operators successively:

$$(\hat{A}\hat{B})|\psi\rangle = \hat{A}(\hat{B}|\psi\rangle). \tag{3.39}$$

It is *not* necessarily true that the product $\hat{A}\hat{B}$ is the same as the product $\hat{B}\hat{A}$. If it *is* true then the two operators are said to "commute". That is, two operators $\hat{A}$ and $\hat{B}$ *commute* if and only if

$$\hat{A}\hat{B}|\psi\rangle = \hat{B}\hat{A}|\psi\rangle \tag{3.40}$$

for *every* vector $|\psi\rangle$.

Exercise 3.L. Examples of operators.
Take as vectors, functions of the real variable x: $|\psi\rangle = \psi(x)$.
Operator $\hat{A}$ is multiplication by x: $\hat{A}|\psi\rangle = x\psi(x)$.
Operator $\hat{B}$ is multiplication by x^2: $\hat{B}|\psi\rangle = x^2\psi(x)$.
Operator $\hat{C}$ is differentiation: $\hat{C}|\psi\rangle = d\psi(x)/dx$.
Show that operators $\hat{A}$ and $\hat{B}$ commute, but that $\hat{A}$ and $\hat{C}$ do not.
Do operators $\hat{B}$ and $\hat{C}$ commute?

Definition: The operator $\hat{A}\hat{B} - \hat{B}\hat{A}$ is called "the commutator of $\hat{A}$ and $\hat{B}$" and represented by $[\hat{A}, \hat{B}]$.

An operator $\hat{A}$ is said to be "linear" if, for all vectors $|\psi\rangle$ and $|\phi\rangle$, and for all scalars c_1 and c_2,

$$\hat{A}(c_1|\psi\rangle + c_2|\phi\rangle) = c_1\hat{A}|\psi\rangle + c_2\hat{A}|\phi\rangle. \tag{3.41}$$

It is remarkable[5] that nearly all operators of interest in quantum mechanics are linear.

Exercise 3.M. Take as vectors functions of the variable x: $|\psi\rangle = \psi(x)$. Show that the operator "d/dx" is linear but that the operator "log" is not.

Exercise 3.N. Show that if $\hat{A}$ and $\hat{B}$ are linear, then so are $c_1\hat{A} + c_2\hat{B}$ and $\hat{A}\hat{B}$.

If you know how $\hat{A}$ acts upon each member of a basis set $\{|1\rangle, |2\rangle, \ldots, |N\rangle\}$, then you know everything there is to know about $\hat{A}$, because for any vector $|\psi\rangle$

$$\hat{A}|\psi\rangle = \hat{A}\left\{\sum_n \psi_n|n\rangle\right\} = \sum_n \psi_n\hat{A}|n\rangle, \tag{3.42}$$

and the vectors $\hat{A}|n\rangle$ are known.

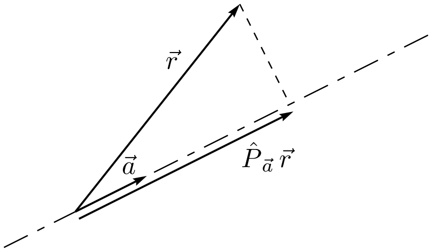
Examples of linear operators:

- The identity operator: $\hat{1}|\psi\rangle = |\psi\rangle$.

---

[5] "The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift, which we neither understand nor deserve." — Eugene Wigner [*Communications on Pure and Applied Mathematics* **13** (1960) 1–14].

- Rotations in the plane. (Linear because the sum of the rotated arrows is the same as the rotation of the summed arrows.)
- The "projection operator" $\hat{P}_{\vec{a}}$, defined in terms of some fixed vector $\vec{a}$ as
$$\hat{P}_{\vec{a}}\,\vec{r} = \left(\vec{a}\cdot\vec{r}\right)\vec{a} \tag{3.43}$$
This is often used for vectors $\vec{a}$ of norm 1, in which case, for arrows in space, it looks like:



- More generally, for any fixed $\vec{a}$ and $\vec{b}$, the operator
$$\hat{S}\,\vec{r} = \left(\vec{b}\cdot\vec{r}\right)\vec{a} \tag{3.44}$$
is linear.

The examples illustrate that the action of an operator can be quite complex indeed — differentiation, integration, and exponentiation are all operators. But sometimes there are special cases of simplicity lurking within the general complexity. If the effect of an operator on some particular vector $|\chi\rangle$ is simply to multiply that vector by a constant number,
$$\hat{A}|\chi\rangle = \lambda|\chi\rangle, \tag{3.45}$$
then that particular vector is called an "eigenvector" of $\hat{A}$, and the number λ is called an "eigenvalue".

Exercise 3.O. Take as vectors functions of the variable x.

a. If the operator is differentiation, d/dx, show that the function e^{ax} is an eigenvector (with what eigenvalue?) but that the function $\cos(kx)$ is not.
b. If the operator is double differentiation, d^2/dx^2, show that the functions e^{ax} and $\cos(kx)$ are eigenvectors (with what eigenvalues?).

c. If the operator is $x \cdot d/dx$, show that the function x^n (with $n \geq 1$) is an eigenvector (with what eigenvalue?).

The German word *eigen* means (see page 99) "associated with". As concerns the differentiation operator d/dx, the function e^{3x} is "associated with" 3, the function e^{4x} is "associated with" 4, but the function $e^{3x} + e^{4x}$ is not "associated with" any number — it is not an eigenfunction of the differentiation operator.

Operator functions

If $\hat{A}$ is an operator, can we assign a meaning to $\cos \hat{A}$ or to $\exp \hat{A}$?

Let's start with some simple functions. The operator $\hat{A}^2$ simply means $\hat{A}$ applied twice. The operator $c\hat{A}^3$ means $\hat{A}$ applied three times, then multiplied by the scalar c. We can similarly define $c\hat{A}^n$.

Exercise 3.P. An operator squared.
Take as vectors functions of the variable x: $|\psi\rangle = \psi(x)$. If the operator $\hat{A}$ is "$\pm ic\, d/dx$", what is the operator $\hat{A}^2$?

Exercise 3.Q. Eigenproblem for functions of operators, I.
The operator $\hat{A}$ has eigenvectors $|a_i\rangle$ and eigenvalues a_i. Show that the operator $c\hat{A}^n$, where $n = 0, 1, 2, 3, \ldots$, has the same eigenvectors $|a_i\rangle$ and eigenvalues ca_i^n. (*Clue:* Establish the result for $n = 0$ and $n = 1$, then use mathematical induction.)

Now, if $f(x)$ is a real function that can be represented by a power series (Taylor) expansion,

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \qquad c_n \text{ real}, \tag{3.46}$$

then we *define* the function of an operator as

$$f(\hat{A}) = \sum_{n=0}^{\infty} c_n \hat{A}^n \qquad c_n \text{ real}, \tag{3.47}$$

Exercise 3.R. Eigenproblem for functions of operators, II.
The operator $\hat{A}$ has eigenvectors $|a_i\rangle$ and eigenvalues a_i. Show that the operator $f(\hat{A})$ has the same eigenvectors $|a_i\rangle$ and eigenvalues $f(a_i)$. (*Clue:* Use the previous exercise.)

Outer products

Recall the operator of equation (3.44):

$$\hat{S}\,\vec{r} = \left(\vec{b}\cdot\vec{r}\right)\vec{a}.$$

In quantum mechanical notation, this is

$$\hat{S}|\psi\rangle = |a\rangle\langle b|\psi\rangle, \tag{3.48}$$

The operator $\hat{S}$ is written as $|a\rangle\langle b|$ and called "the outer product of $|a\rangle$ and $|b\rangle$". This means neither more nor less than the defining equation (3.48).

For any orthonormal basis $\{|1\rangle, |2\rangle, \ldots, |N\rangle\}$, consider the operator

$$\hat{T} \equiv |1\rangle\langle 1| + |2\rangle\langle 2| + \cdots + |N\rangle\langle N|. \tag{3.49}$$

The effect of this operator on an arbitrary vector $|\psi\rangle$ is given in equation (3.29), which shows that $\hat{T}|\psi\rangle = |\psi\rangle$ for any $|\psi\rangle$. Hence the remarkable equation

$$\hat{1} = \sum_n |n\rangle\langle n|. \tag{3.50}$$

This might look like magic, but in means nothing more than equation (3.29): that a vector may be resolved into its components. The operator of equation (3.50) simply represents the act of chopping a vector into its components and reassembling them. It is the mathematical representation of an analyzer loop!

Representations of linear operators

A linear operator can be represented in a given basis by an $N \times N$ matrix. If

$$|\phi\rangle = \hat{A}|\psi\rangle, \tag{3.51}$$

then

$$\begin{aligned}
\langle n|\phi\rangle &= \langle n|\hat{A}|\psi\rangle \\
&= \langle n|\hat{A}\hat{1}|\psi\rangle \\
&= \langle n|\hat{A}\left\{\sum_m |m\rangle\langle m|\right\}|\psi\rangle \\
&= \sum_m \langle n|\hat{A}|m\rangle\langle m|\psi\rangle,
\end{aligned} \tag{3.52}$$

or, in matrix form,

$$\begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{pmatrix} = \begin{pmatrix} \langle 1|\hat{A}|1\rangle & \langle 1|\hat{A}|2\rangle & \cdots & \langle 1|\hat{A}|N\rangle \\ \langle 2|\hat{A}|1\rangle & \langle 2|\hat{A}|2\rangle & \cdots & \langle 2|\hat{A}|N\rangle \\ \vdots & & & \vdots \\ \langle N|\hat{A}|1\rangle & \langle N|\hat{A}|2\rangle & \cdots & \langle N|\hat{A}|N\rangle \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}. \tag{3.53}$$

The matrix M that represents operator $\hat{A}$ in this particular basis has elements $\mathsf{M}_{n,m} = \langle n|\hat{A}|m\rangle$.

In a different basis, the same operator $\hat{A}$ will be represented by a different matrix. You can figure out for yourself how to transform the matrix representation of an operator in one basis into the matrix representation of that operator in a second basis. But it's not all that important to do so. Usually you work in the abstract operator notation until you've figured out the easiest basis to work with, and then work in only that basis.

Unitary operators

If the norm of $\hat{U}|\psi\rangle$ equals the norm of $|\psi\rangle$ for all $|\psi\rangle$, then $\hat{U}$ should be called "norm preserving" but in fact is called "unitary". The rotation operator is unitary.

Hermitian conjugate

For every operator $\hat{A}$ there is a unique operator $\hat{A}^\dagger$, the "Hermitian[6] conjugate" (or "Hermitian adjoint") of $\hat{A}$ such that

$$\langle\phi|\hat{A}^\dagger|\psi\rangle = \langle\psi|\hat{A}|\phi\rangle^* \tag{3.54}$$

for all vectors $|\psi\rangle$ and $|\phi\rangle$. If the matrix elements for $\hat{A}$ are $\mathsf{M}_{n,m}$, then the matrix elements for $\hat{A}^\dagger$ are $\mathsf{K}_{n,m} = \mathsf{M}^*_{m,n}$.

Hermitian operators

An operator $\hat{A}$ is said to be "Hermitian" when, for all vectors $|\psi\rangle$ and $|\phi\rangle$,

$$\langle\phi|\hat{A}|\psi\rangle = \langle\psi|\hat{A}|\phi\rangle^*. \tag{3.55}$$

---

[6] Charles Hermite (1822-1901), French mathematician who contributed to number theory, orthogonal polynomials, elliptic functions, quadratic forms, and linear algebra. Teacher of Hadamard and Poincaré, father-in-law of Picard.

For such an operator, $\hat{A}^{\dagger} = \hat{A}$. Matrix representations of Hermitian operators have $\mathsf{M}_{n,m} = \mathsf{M}^*_{m,n}$.

Think about the very simple operator that is multiplication by a constant: $\hat{A}|\psi\rangle = c|\psi\rangle$. Then $\langle\phi|\hat{A}|\psi\rangle = c\langle\phi|\psi\rangle$ while $\langle\psi|\hat{A}|\phi\rangle = c\langle\psi|\phi\rangle$, so $\langle\psi|\hat{A}|\phi\rangle^* = c^*\langle\psi|\phi\rangle^* = c^*\langle\phi|\psi\rangle$. The operator $\hat{A}$ is Hermitian if and only if the constant c is real.

Exercise 3.S. Show that if $\hat{A}$ is a linear operator and $(\mathbf{a}, \hat{A}\mathbf{a})$ is real for all vectors $\mathbf{a}$, then $\hat{A}$ is Hermitian. (*Clue:* Employ the hypothesis with $\mathbf{a} = \mathbf{b} + \mathbf{c}$ and $\mathbf{a} = \mathbf{b} + i\mathbf{c}$.

Exercise 3.T. Show that any operator of the form

$$\hat{A} = c_a|a\rangle\langle a| + c_b|b\rangle\langle b| + \cdots + c_z|z\rangle\langle z|,$$

where the c_n are real constants, is Hermitian.

Exercise 3.U. Show that, when $\hat{A}$ and $\hat{B}$ are Hermitian: (a) $c_1\hat{A} + c_2\hat{B}$ is Hermitian if c_1 and c_2 are real, and (b) $\hat{A}\hat{B}$ is Hermitian if $\hat{A}$ and $\hat{B}$ commute.

Hermitian operators are important in quantum mechanics because if an operator is to correspond to an observable, then that operator must be Hermitian.

Theorem: Hermitian operator eigenproblem.
If $\hat{H}$ is Hermitian, then: (a) All of its eigenvalues are real. (b) There is an orthonormal basis consisting of eigenvectors of $\hat{H}$.

Corollaries: If the orthonormal basis mentioned in (b) is $\{|1\rangle, |2\rangle, \ldots, |N\rangle\}$, and $\hat{H}|n\rangle = \lambda_n|n\rangle$, then

$$\hat{H} = \lambda_1|1\rangle\langle 1| + \lambda_2|2\rangle\langle 2| + \cdots + \lambda_N|N\rangle\langle N|. \tag{3.56}$$

The matrix representation of $\hat{H}$ in this basis is diagonal:

$$\hat{H} \doteq \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{pmatrix}. \tag{3.57}$$

Exercise 3.V. You know from the above theorem that if an operator is Hermitian then all of its eigenvalues are real. Show that the converse is false by producing a counterexample. (*Clue:* Try a 2×2 upper triangular matrix.)

Exercise 3.W. Suppose $\hat{A}$ is a Hermitian operator with eigenvectors $|\alpha\rangle$ and $|\beta\rangle$ corresponding to eigenvalues α and β. Show that if $\alpha \neq \beta$, then $|\alpha\rangle$ and $|\beta\rangle$ are orthogonal ($\langle\alpha|\beta\rangle = 0$). (*Clue:* Compare $(\alpha, \hat{A}\beta)$ with $(\hat{A}\alpha, \beta)$, using the fact that α and β are real.)

3.4.6 *Diagonalizing the matrix representing a Hermitian operator*

We will often have occasion (see for example page 154) to find the orthonormal basis of eigenvectors guaranteed to exist by the theorem on Hermitian operator eigenproblems.

For example, the matrix

$$\begin{pmatrix} 7 & i6 \\ -i6 & 2 \end{pmatrix} \tag{3.58}$$

represents, in some given basis, a Hermitian operator. We know this is true because if you transpose the matrix and conjugate each element, you come back to the original matrix (that is, $\mathsf{M}_{n,m} = \mathsf{M}^*_{m,n}$ for all elements of the matrix). An eigenvector of that Hermitian operator, represented in the same basis, satisfies

$$\begin{pmatrix} 7 & i6 \\ -i6 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \tag{3.59}$$

where λ is the eigenvalue. But can we find the three unknowns x, y, and λ? At first glance it seems hopeless, because there are three unknowns and only two equations.

The puzzle is unlocked through this key. The matrix equation is

$$\mathsf{M}\underset{\sim}{x} = \lambda\underset{\sim}{x} = \lambda\mathsf{I}\underset{\sim}{x}, \tag{3.60}$$

where M stands for the square matrix, $\underset{\sim}{x}$ stands for the unknown column matrix representing the eigenvector, and I stands for the square identity matrix. This is equivalent to

$$\Big[\mathsf{M} - \lambda\mathsf{I}\Big]\,\underset{\sim}{x} = 0. \tag{3.61}$$

We can effortlessly find one solution, namely $\underset{\sim}{x} = \mathsf{0}$, but this solution is *not* the desired eigenvector. In fact, if the matrix $\mathsf{M} - \lambda\mathsf{I}$ is invertible, that's the only solution, namely

$$\underset{\sim}{x} = \left[\mathsf{M} - \lambda\mathsf{I}\right]^{-1}\mathsf{0} = \mathsf{0}.$$

So if there is to be an eigenvector, the matrix $\mathsf{M} - \lambda\mathsf{I}$ must be non-invertible. You might recall that a non-invertible matix has determinant zero, so we must have

$$\det|\mathsf{M} - \lambda\mathsf{I}| = 0. \tag{3.62}$$

And this is the key that unlocks the puzzle. This equation involves only the eigenvalues, not the eigenvectors. So we use it to find the eigenvalues, and once we know them we look for the eigenvectors.

Let's apply this strategy to our matrix (3.58):

$$\begin{aligned}0 &= \det\begin{vmatrix} 7-\lambda & i6 \\ -i6 & 2-\lambda \end{vmatrix} \\ &= (7-\lambda)(2-\lambda) - (i6)(-i6) \\ &= \lambda^2 - 9\lambda - 22 \\ \lambda &= \tfrac{1}{2}\left[9 \pm \sqrt{9^2 - 4\cdot(-22)}\right] \\ &= -2 \text{ or } 11.\end{aligned}$$

Now we know the two eigenvalues! As promised by the theorem on Hermitian operator eigenproblems, they are both real.

The next step is to find eigenvectors: I'll start with the eigenvector associated with eigenvalue −2, and leave it as an exercise to find the one associated with 11. Going back to equation (3.59), we search for x and y such that

$$\begin{pmatrix} 7 & i6 \\ -i6 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = -2\begin{pmatrix} x \\ y \end{pmatrix}. \tag{3.63}$$

This one matrix equation stands for two equations, namely

$$\begin{aligned} 7x + i6y &= -2x \\ -i6x + 2y &= -2y \end{aligned}$$

or

$$\begin{aligned} 9x + i6y &= 0 \\ -i6x + 4y &= 0 \end{aligned}$$

or

$$\begin{aligned} 3x + i2y &= 0 \\ -i3x + 2y &= 0. \end{aligned} \tag{3.64}$$

Perhaps your heart skips a beat at this point, because the two equations are not independent! The second equation is just $-i$ times the first. This is a feature, not a bug. It simply reflects the fact that an eigenvector, multiplied by a number, is again an eigenvector with the same eigenvalue. In other words, any vector of the form

$$\begin{pmatrix} x \\ i\frac{3}{2}x \end{pmatrix}, \tag{3.65}$$

for any real or complex value of x, is an eigenvector.

Which of this abundance of riches should we choose? I like to use eigenvectors that are normalized, that is eigenvectors for which

$$\begin{pmatrix} x^* & -i\frac{3}{2}x^* \end{pmatrix} \begin{pmatrix} x \\ i\frac{3}{2}x \end{pmatrix} = 1.$$

This says that

$$|x|^2 + \tfrac{9}{4}|x|^2 = 1 \quad \text{or} \quad |x| = \tfrac{2}{\sqrt{13}}.$$

This still leaves us with an infinite number of choices. We could pick

$$x = \tfrac{2}{\sqrt{13}}, \text{ or } x = -\tfrac{2}{\sqrt{13}}, \text{ or } x = i\tfrac{2}{\sqrt{13}}, \text{ or even } x = (i+1)\sqrt{\tfrac{2}{13}},$$

but I like to keep it simple and straightforward (KISS), so I'll pick the first choice and say that the eigenvector, represented in the basis we've been using throughout, is

$$\tfrac{1}{\sqrt{13}} \begin{pmatrix} 2 \\ i3 \end{pmatrix}. \tag{3.66}$$

Exercise 3.X. Verify that the column matrix (3.66) indeed represents an eigenvector of (3.58) with eigenvalue -2.

Exercise 3.Y. The other eigenvector. Show that an eigenvector of (3.58) with eigenvalue 11 is

$$\tfrac{1}{\sqrt{13}} \begin{pmatrix} i3 \\ 2 \end{pmatrix}. \tag{3.67}$$

Exercise 3.Z. Verify that, as guaranteed by the theorem on Hermitian operator eigenproblems, eigenvectors (3.66) and (3.67) are orthogonal.

In the original basis, our Hermitian operator is represented by the matrix (3.58). In the new orthonormal basis consisting of vectors (3.66) and (3.67), our operator is represented by the different matrix

$$\begin{pmatrix} -2 & 0 \\ 0 & 11 \end{pmatrix}. \tag{3.68}$$

To diagonalize a matrix M, representing a Hermitian operator:

1. In initial basis, the matrix representation of $\hat{A}$ is the $N \times N$ matrix M. The eigenvectors of $\hat{A}$ satisfy $\hat{A}|e_n\rangle = \lambda_n|e_n\rangle$.

2. Find N eigenvalues by solving the N^{th} order polynomial equation

$$\det|\mathsf{M} - \lambda\mathsf{I}| = 0.$$

3. Find the representation e_n of the eigenvector $|e_n\rangle$ by solving N simultaneous linear equations

$$\mathsf{M}\mathsf{e}_n = \lambda_n \mathsf{e}_n.$$

In this equation, M is an $N \times N$ matrix, e_n is an $N \times 1$ matrix (the N unknowns), and λ_n is a known number (determined in step 2).

4. In the basis $\{|e_1\rangle, |e_2\rangle, \ldots, |e_N\rangle\}$, the matrix representation of $\hat{A}$ is diagonal

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{pmatrix}.$$

This algorithm is appropriate for analytical work but poor (unstable) for numerical work. Instead, use the "Jacobi[7] algorithm", which finds the eigenvalues and eigenvectors simultaneously.

The process described in this section is called "diagonalizing the matrix", which can give the unfortunate and incorrect impression that the process involves changing the operator. No. It changes the basis in which

---

[7] Carl Jacobi (1804–1851), prolific German-Jewish mathematician. A measure of his accomplishments is that his name appears in this book three times despite the fact that he died 49 years before quantum mechanics was discovered.

the operator is represented, so it changes the matrix representation, but it does not change the operator itself.

The mathematical tool of matrix diagonalization is used throughout science and engineering: it finds principal rotation axes for rigid body motion, normal modes for molecular vibrations, normal modes for bridge vibrations, and makes transfer matrices useful in statistical mechanics. The tool you learn to use here will help you many times over all your life.

3.5 Extras

Change of basis

Suppose the two amplitudes $\langle z+|\psi\rangle$ and $\langle z-|\psi\rangle$ are known. Then we can easily find the amplitudes $\langle \theta+|\psi\rangle$ and $\langle \theta-|\psi\rangle$, for any value of θ, through

$$\begin{aligned}\langle \theta+|\psi\rangle &= \langle \theta+|z+\rangle\langle z+|\psi\rangle + \langle \theta+|z-\rangle\langle z-|\psi\rangle\\ \langle \theta-|\psi\rangle &= \langle \theta-|z+\rangle\langle z+|\psi\rangle + \langle \theta-|z-\rangle\langle z-|\psi\rangle\end{aligned}$$

These two equations might seem arcane, but in fact each one just represents the interference experiment performed with a vertical analyzer: The state $|\psi\rangle$ is unaltered if the atom travels through the two branches of a vertical interferometer, that is via the upper $z+$ branch *and* the lower $z-$ branch. And if the state is unaltered then the amplitude to go to state $|\theta+\rangle$ is of course also unaltered.

The pair of equations is most conveniently written as a matrix equation

$$\begin{pmatrix}\langle \theta+|\psi\rangle\\ \langle \theta-|\psi\rangle\end{pmatrix} = \begin{pmatrix}\langle \theta+|z+\rangle & \langle \theta+|z-\rangle\\ \langle \theta-|z+\rangle & \langle \theta-|z-\rangle\end{pmatrix}\begin{pmatrix}\langle z+|\psi\rangle\\ \langle z-|\psi\rangle\end{pmatrix}.$$

The 2×1 column matrix on the right side is called the representation of state $|\psi\rangle$ in the basis $\{|z+\rangle, |z-\rangle\}$. The 2×1 column matrix on the left side is called the representation of state $|\psi\rangle$ in the basis $\{|\theta+\rangle, |\theta-\rangle\}$. The square 2×2 matrix is independent of the state $|\psi\rangle$, and depends only on the geometrical relationship between the initial basis $\{|z+\rangle, |z-\rangle\}$ and the final basis $\{|\theta+\rangle, |\theta-\rangle\}$:

$$\begin{pmatrix}\langle \theta+|z+\rangle & \langle \theta+|z-\rangle\\ \langle \theta-|z+\rangle & \langle \theta-|z-\rangle\end{pmatrix} = \begin{pmatrix}\cos(\theta/2) & \sin(\theta/2)\\ -\sin(\theta/2) & \cos(\theta/2)\end{pmatrix}.$$

Terms concerning quantum states

For atoms in state $|z+\rangle$, the probability of measuring μ_θ and finding $\mu_\theta = +\mu_B$ is $\cos^2(\theta/2)$. We say "The projection probability from $|z+\rangle$ to $|\theta+\rangle$ is $\cos^2(\theta/2)$." This situation is frequently, but incorrectly, described as "The probability that an atom in state $|z+\rangle$ is in state $|\theta+\rangle$ is $\cos^2(\theta/2)$."

If the projection probability from $|A\rangle$ to $|B\rangle$ is zero, and vice versa, the two states are *orthogonal.* (For example, $|z+\rangle$ and $|z-\rangle$ are orthogonal, whereas $|z+\rangle$ and $|x-\rangle$ are not.)

Given a set of states $\{|A\rangle, |B\rangle, \ldots, |N\rangle\}$, this set is said to be *complete* if an atom in *any* state is analyzed into one state of this set. In other words, it is complete if

$$\sum_{i=A}^{N}(\text{projection probability from any given state to } |i\rangle) = 1.$$

(For example, the set $\{|\theta+\rangle, |\theta-\rangle\}$ is complete.)

General definition of basis

We say that a set of states $\{|a\rangle, |b\rangle, \ldots, |n\rangle\}$ is a *basis* if both of the following apply:

- An atom in *any* state is analyzed into one member of this set. That is, for any state $|\psi\rangle$
$$|\langle a|\psi\rangle|^2 + |\langle b|\psi\rangle|^2 + \cdots + |\langle n|\psi\rangle|^2 = 1. \tag{3.69}$$
- There is zero amplitude for one member to be another member. That is
$$\begin{aligned} &\langle a|b\rangle = 0, \langle a|c\rangle = 0, \ldots, \langle a|n\rangle = 0, \\ &\langle b|c\rangle = 0, \ldots, \langle b|n\rangle = 0, \\ &\text{etc.} \end{aligned} \tag{3.70}$$

For example, the set $\{|\theta+\rangle, |\theta-\rangle\}$ is a basis for any value of θ. The set $\{|z+\rangle, |x-\rangle\}$ is not a basis.

Problems

3.1 Change of basis
The set $\{|a\rangle, |b\rangle\}$ is an orthonormal basis.

a. Show that the set $\{|a'\rangle, |b'\rangle\}$, where

$$|a'\rangle = +\cos\phi|a\rangle + \sin\phi|b\rangle$$
$$|b'\rangle = -\sin\phi|a\rangle + \cos\phi|b\rangle$$

is also an orthonormal basis. (The angle ϕ is simply a parameter — it has no physical significance.)

b. Write down the transformation matrix from the $\{|a\rangle, |b\rangle\}$ basis representation to the $\{|a'\rangle, |b'\rangle\}$ basis representation.

(If you suspect a change of basis is going to help you, but you're not sure how or why, this change often works, so it's a good one to try first. You can adjust ϕ to any parameter you want, but it's been my experience that it is most often helpful when $\phi = 45°$.)

3.2 **Change of representation, I**

If the set $\{|a\rangle, |b\rangle\}$ is an orthonormal basis, then the set $\{|a'\rangle, |b'\rangle\}$, where $|a'\rangle = |b\rangle$ and $|b'\rangle = |a\rangle$ is also an orthonormal basis — it's just a reordering of the original basis states. Find the transformation matrix. If state $|\psi\rangle$ is represented in the $\{|a\rangle, |b\rangle\}$ basis as

$$\begin{pmatrix} \psi_a \\ \psi_b \end{pmatrix},$$

then how is this state represented in the $\{|a'\rangle, |b'\rangle\}$ basis?

3.3 **Change of representation, II**

Same as the previous problem, but use $|a'\rangle = i|a\rangle$ and $|b'\rangle = -i|b\rangle$.

3.4 **Inner product**

You know that the inner product between two position unit vectors is the cosine of the angle between them. What is the inner product between the states $|z+\rangle$ and $|\theta+\rangle$? Does the geometrical interpretation hold?

3.5 **Outer product**

Using the $\{|z+\rangle, |z-\rangle\}$ basis representations

$$|\psi\rangle \doteq \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} \qquad |\phi\rangle \doteq \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix}$$

$$|\theta+\rangle \doteq \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} \qquad |\theta-\rangle \doteq \begin{pmatrix} -\sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix},$$

write representations for $|\theta+\rangle\langle\theta+|$ and $|\theta-\rangle\langle\theta-|$, then for $\langle\phi|\theta+\rangle\langle\theta+|\psi\rangle$ and $\langle\phi|\theta-\rangle\langle\theta-|\psi\rangle$, and finally verify that

$$\langle\phi|\psi\rangle = \langle\phi|\theta+\rangle\langle\theta+|\psi\rangle + \langle\phi|\theta-\rangle\langle\theta-|\psi\rangle.$$

3.6 **Measurement operator**

Write the representation of the $\hat{\mu}_\theta$ operator

$$\hat{\mu}_\theta = (+\mu_B)|\theta+\rangle\langle\theta+| + (-\mu_B)|\theta-\rangle\langle\theta-|$$

in the $\{|z+\rangle, |z-\rangle\}$ basis. Using this representation, verify that $|\theta+\rangle$ and $|\theta-\rangle$ are eigenvectors.

3.7 **The trace**

The *trace* of $N \times N$ matrix A (with components a_{ij}) is defined as the sum of its diagonal elements, that is

$$\mathrm{tr}\{A\} = \sum_{i=1}^{N} a_{ii}.$$

Show that $\mathrm{tr}\{AB\} = \mathrm{tr}\{BA\}$, and hence that $\mathrm{tr}\{ABCD\} = \mathrm{tr}\{DABC\} = \mathrm{tr}\{CDAB\}$, etc. (the so-called "cyclic invariance" of the trace). However, show that $\mathrm{tr}\{ABC\}$ does not generally equal $\mathrm{tr}\{CBA\}$ by constructing a counterexample. (All matrices are square.)

3.8 **The outer product**

Any two complex N-tuples can be multiplied to form an $N \times N$ matrix as follows: (The star represents complex conjugation.)

$$x = (x_1 \; x_2 \ldots x_N)$$

$$y = (y_1 \; y_2 \ldots y_N)$$

$$x \otimes y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} (y_1^* \; y_2^* \ldots y_N^*) = \begin{pmatrix} x_1 y_1^* & x_1 y_2^* & \ldots & x_1 y_N^* \\ x_2 y_1^* & x_2 y_2^* & \ldots & x_2 y_N^* \\ & \vdots & & \\ x_N y_1^* & x_N y_2^* & \ldots & x_N y_N^* \end{pmatrix}.$$

This so-called "outer product" is quite different from the familiar "dot product" or "inner product"

$$x \cdot y = (x_1^* \; x_2^* \ldots x_N^*) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = x_1^* \; y_1 + x_2^* \; y_2 + \cdots + x_N^* \; y_N.$$

Write a formula for the $i, \; j$ component of $x \otimes y$ and use it to show that the trace of an outer product is $\mathrm{tr}\{y \otimes x\} = x \cdot y$.

3.9 Pauli matrix algebra

Three important matrices are the *Pauli matrices*:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(We will call them σ_1, σ_2, σ_3, but others call them σ_x, σ_y, σ_z.)

a. Show that the four matrices $\{I, \sigma_1, \sigma_2, \sigma_3\}$, where

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

constitute a basis for the set of 2×2 matrices, by showing that any matrix

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

can be written as

$$A = z_0 I + z_1\sigma_1 + z_2\sigma_2 + z_3\sigma_3.$$

Produce formulas for the z_i in terms of the a_{ij}.

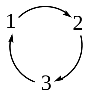
b. Show that

i. $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = I^2 = I$

ii. $\sigma_i\sigma_j = -\sigma_j\sigma_i$ for $i \neq j$

iii. $\sigma_1\sigma_2 = i\sigma_3$ (a)
$\sigma_2\sigma_3 = i\sigma_1$ (b)
$\sigma_3\sigma_1 = i\sigma_2$ (c)

Note: Equations (b) and (c) are called "cyclic permutations" of equation (a), because in each equation, the indices go in the order



and differ only by starting at different points on the circular "merry-go-round."

c. Show that for any complex numbers c_1, c_2, c_3,

$$(c_1\sigma_1 + c_2\sigma_2 + c_3\sigma_3)^2 = (c_1^2 + c_2^2 + c_3^2)I.$$

3.10 Diagonalizing the Pauli matrices

Find the eigenvalues and corresponding (normalized) eigenvectors for all three Pauli matrices.

3.11 Exponentiation of Pauli matrices

Define exponentiation of matrices through

$$e^M = \sum_{n=0}^{\infty} \frac{M^n}{n!}.$$

a. Show that

$$e^{z\sigma_i} = \cosh(z)I + \sinh(z)\sigma_i \quad \text{for } i = 1, 2, 3.$$

(*Clue:* Look up the series expansions of sinh and cosh.)

b. Show that

$$e^{(\sigma_1+\sigma_3)} = \cosh(\sqrt{2})I + \frac{\sinh(\sqrt{2})}{\sqrt{2}}(\sigma_1 + \sigma_3).$$

c. Prove that $e^{\sigma_1} e^{\sigma_3} \neq e^{(\sigma_1+\sigma_3)}$.

3.12 Unitary operators

Show that all the eigenvalues of a unitary operator have magnitude unity.

3.13 Commutator algebra

Prove that

$$\begin{aligned}
[\hat{A}, b\hat{B} + c\hat{C}] &= b[\hat{A}, \hat{B}] + c[\hat{A}, \hat{C}] \\
[a\hat{A} + b\hat{B}, \hat{C}] &= a[\hat{A}, \hat{C}] + b[\hat{B}, \hat{C}] \\
[\hat{A}, \hat{B}\hat{C}] &= \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C} \\
[\hat{A}\hat{B}, \hat{C}] &= \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B} \\
[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{B}, [\hat{C}, \hat{A}]] + [\hat{C}, [\hat{A}, \hat{B}]] &= 0 \quad \text{(the “Jacobi identity”).}
\end{aligned}$$

3.14 Questions (recommended problem)

Update your list of quantum mechanics questions that you started at problem 1.13 on page 56. Write down new questions and, if you have uncovered answers to any of your old questions, write them down briefly.

Chapter 4

Formalism

The previous three chapters described the experiments and reasoning that stand behind our current understanding of quantum mechanics. Some of it was rigorous, some of it was suggestive. Some of it was robust, some of it was mere analogy. Some of it was applicable to any quantum system, some of it was particular to the magnetic moment of a silver atom. This chapter sets forth in four rigorous statements (sometimes called "postulates") the things physicists hold to be true throughout non-relativistic quantum mechanics so that you'll know it straight, rather than get mixed up with the experiments and motivations and plausibility arguments.

A little confusion is a good thing — Niels Bohr[1] claimed that "those who are not shocked when they first come across quantum theory cannot possibly have understood it" — but these four statements should become firm and sharp in your mind.[2]

---

[1] Danish physicist (1885–1962), fond of revolutionary ideas. In 1913 he was the first to apply the ideas of the "old quantum theory" to atoms. In 1924 and again in 1929 he suggested that the law of energy conservation be abandoned, but both suggestions proved to be on the wrong track. Father of six children, all boys, one of whom won the Nobel Prize in Physics and another of whom played in the 1948 Danish Olympic field hockey team. This quote from Bohr was recalled by Werner Heisenberg in *Physics and Beyond* (Harper and Row, New York, 1971) page 206.

[2] This section owes a debt of gratitude to Daniel T. Gillespie, *A Quantum Mechanics Primer* (International Textbook Company, Scranton, Pennsylvania, 1970). This is the first serious book on quantum mechanics I ever read, and it mesmerized me. So I owe a personal debt of gratitude to Dr. Gillespie as well.

4.1 The quantal state

In classical mechanics, the state of the system is given by a few numbers that can be found by observation. For example, if the system is a single particle, it is the position $\vec{r}$ and the momentum $\vec{p}$. If the system is a magnetic moment, it is the three components of the moment vector $\vec{\mu}$. The system is specified by stating these so-called “observables”.

In quantum mechanics, there is a sharp distinction between state and observables. Concerning state, we have:

> **1. State.** The physical state of any system corresponds to a Hilbert space vector $|\psi\rangle$ with unit norm, and every Hilbert space vector with unit norm corresponds to a physical state. Two Hilbert space vectors that differ only by an overall scalar factor of magnitude one correspond to the same physical state. For example the vector $|\psi\rangle$ and the vector $c|\psi\rangle$, where c is any constant complex number with $|c|^2 = 1$, correspond to the same state. The most accurate statement is that the vector $|\psi\rangle$ is the “vector representing (or associated with) the state of the system”, but that is quite a mouthful so $|\psi\rangle$ is more frequently called the “state vector” of the system; the state of the system is said to “be represented by $|\psi\rangle$” or “the system is in state $|\psi\rangle$”. Anything in principle knowable about the state can be learned from the state vector $|\psi\rangle$.

The precise mathematical form taken by $|\psi\rangle$ depends upon the system under study. We have seen that for the magnetic moment of a silver atom $|\psi\rangle$ is a vector in a two-dimensional Hilbert space. For the magnetic moment of a nitrogen atom $|\psi\rangle$ is a vector in a four-dimensional Hilbert space (see page 11). In future explorations we will find the form taken by $|\psi\rangle$ for a single spinless particle ambivating in one dimension (equation 6.8), for a single particle with spin ambivating in one dimension (equation 12.8), for two spinless particles ambivating in three dimensions (equation 12.29), and more. This chapter focuses on the properties of state vector without regard to the specific system under study.

Exercise 4.A. If the state vector $|\psi\rangle$ has unit norm ($\langle\psi|\psi\rangle = 1$) and the complex number c has unit magnitude ($|c|^2 = 1$) show that the state $|\phi\rangle = c|\psi\rangle$ also has unit norm.

Exercise 4.B. Show that these two statements are equivalent:
(1) "The vector $|\psi\rangle$ and the vector $c|\psi\rangle$, where c is any constant complex number with $|c|^2 = 1$, correspond to the same state."
(2) "The vector $|\psi\rangle$ and the vector $e^{i\delta}|\psi\rangle$, for any real constant value of δ, correspond to the same state."

4.2 Observables

Statement 1 about "state" says that "Anything knowable about the state can be learned from the state vector $|\psi\rangle$" but doesn't say how to go about finding those knowable things. This section starts to answer that need by discussing quantal observables.

In quantum mechanics as in classical mechanics, an observable is something that can be found through a measurement of the system. If the system is a magnetic moment, for example, then the x-, y-, and z-components of the moment vector $\vec{\mu}$ are all observables. If the system is a single particle, the y-component of position $\vec{r}$, and the z-component of momentum $\vec{p}$ are observables. Any function of position and momentum, the most important of which is the energy, is an observable. The "measurement" of an observable is a physical process which, when performed on the system, yields a real number called the "value of the observable". This book treats only "ideal" measurements in which there is no experimental uncertainty.

> **2. Observables.** For each physical observable, there corresponds in the Hilbert space a linear Hermitian operator $\hat{A}$. This operator possesses a complete, orthonormal set of eigenvectors $|a_1\rangle$, $|a_2\rangle$, $|a_3\rangle$, ... with corresponding real eigenvalues a_1, a_2, a_3, ... such that
>
> $$\hat{A}|a_n\rangle = a_n|a_n\rangle \qquad n = 1, 2, 3, \ldots. \tag{4.1}$$
>
> Whenever this observable is measured, the result will be one of the eigenvalues a_1, a_2, a_3, ...

It might happen that two or more of the eigenvaues are the same: For example it could be that $a_4 = a_5$, despite the fact that $|a_4\rangle \neq |a_5\rangle$. Then this happens the eigenvalues are said to be "degenerate", a nasty name for an intriguing phenomenon.

4.3 Measurement

Statement 1 about “state” says that “Anything knowable about the state can be learned from the state vector $|\psi\rangle$.” Statement 2 about “observables” adds that whenever any observable is measured, the result will be one of the eigenvalues of the corresponding operator. But how can we learn which of those eigenvalues will be measured?

The answer goes back to the measurement process. Measurement is a physical process in which the system under study (such as the silver atom in section 2.6.3) becomes entangled with some other system — the measuring system — that probes the system under study (the photon in section 2.6.3). The full system consists of the system under study plus the measuring system To keep full information of the full system, we would have to keep track of both the silver atom and the photon for all times in the future.

But in most cases we don’t need full information, and don’t want to keep track of both the system under study and the measuring system. Instead we want to focus on just the system under study and, after it has done its job, ignore the measuring system. In those circumstances we use this statement:

> **3. Measurement.** If a system is in state $|\psi\rangle$ and the observable corresponding to operator $\hat{A}$ is measured, then the probability that the measurement will produce the result a_n is $|\langle a_n|\psi\rangle|^2$.

This is just our old friend amplitude made rigorous, precise, and more general.

Exercise 4.C. In order to interpret $|\langle a_n|\psi\rangle|^2$ as a probability, as claimed above, it must be true that

$$0 \leq |\langle a_n|\psi\rangle|^2 \leq 1.$$

Show that this is indeed correct. [*Clue:* Use the Schwarz inequality (3.24).]

Exercise 4.D. A system is in state $|\psi\rangle$. Show that the *mean value* for a measurement of $\hat{A}$ is

$$\langle \hat{A} \rangle = \langle \psi|\hat{A}|\psi\rangle. \tag{4.2}$$

4. Change of state upon measurement. If a system is in state $|\psi\rangle$ and the observable corresponding to operator $\hat{A}$ is measured producing the result a_n, then after that measurement the system is no longer in state $|\psi\rangle$, instead it is in an eigenstate of $\hat{A}$ with eigenvalue a_n.

This statement reflects the repeated measurement experiments of section 1.1.2: Before a measurement of μ_x, a silver atom in state $|z+\rangle$ *does not have* a value of μ_x. But after μ_x is measured and found to be, say, $-\mu_B$, then it *does have* a value of μ_x and is in state $|x-\rangle$.

Exercise 4.E. Measurement example. For a particular two-state system, two observables correspond to the operators $\hat{A}$ and $\hat{B}$. The eigenvectors of $\hat{A}$ are $|a_1\rangle$ and $|a_2\rangle$. The eigenvectors of $\hat{B}$ are

$$\begin{aligned} |b_1\rangle &= \tfrac{4}{5}|a_1\rangle + \tfrac{3}{5}|a_2\rangle \\ |b_2\rangle &= -\tfrac{3}{5}|a_1\rangle + \tfrac{4}{5}|a_2\rangle \end{aligned}$$

a. Show that if $\{|a_n\rangle\}$ is orthonormal (that is, $\langle a_n|a_m\rangle = \delta_{n,m}$), then $\{|b_n\rangle\}$ is orthonormal too.
b. Write equations for $\{|a_n\rangle\}$ in terms of $\{|b_n\rangle\}$.
c. The observable corresponding to $\hat{A}$ is measured giving result a_1. Then $\hat{B}$ is measured, then $\hat{A}$ is measured again. What is the probability that the final measurement finds the value of a_1? Of a_2? Do your two answers sum to 1 (as they must)?

4.3.1 *A quantitative measure of indeterminacy*

"The outcome of an experiment cannot, in general, be predicted. But the probabilities of various outcome *can* be calculated."

This does not mean we must give up all hope for prediction: For example one can readily calculate the mean (average) result of a measurement of $\hat{A}$ if the system is in state $|\psi\rangle$:

$$\langle\hat{A}\rangle = \sum_n |\langle a_n|\psi\rangle|^2 a_n = \sum_n \langle\psi|a_n\rangle a_n \langle a_n|\psi\rangle = \langle\psi|\hat{A}\psi\rangle. \qquad (4.3)$$

Exercise 4.F. A silver atom in state $|z+\rangle$ enters a horizontal analyzer and the value of μ_x is measured. What is the mean value $\langle\mu_x\rangle$? Do you expect that any single measurement will ever result in this mean value?

Since not all measurements will be result in the same outcome, it is important to know not only the mean but also the range or spread of possible outcomes. The traditional "root mean square" measure of spread is

$$\Delta A = \sqrt{\left\langle (\hat{A} - \langle \hat{A} \rangle)^2 \right\rangle}. \tag{4.4}$$

The quantity ΔA is called "the indeterminacy of $\hat{A}$".[3]

Exercise 4.G. Show that another expression for ΔA is

$$\Delta A^2 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2. \tag{4.5}$$

Exercise 4.H. A silver atom in state $|z+\rangle$ enters a horizontal analyzer and the value of μ_x is measured. What is the mean value $\langle \mu_x \rangle$? What is the indeterminacy $\Delta\mu_x$?

Exercise 4.I. A silver atom in state $|z+\rangle$ enters an analyzer tilted by 60° from the vertical and the value of μ_{60° is measured. What is the mean value $\langle \mu_{60^\circ} \rangle$? What is the indeterminacy $\Delta\mu_{60^\circ}$?

4.3.2 *Measurement of two observables*

Two observables are called "compatible" (or "simultaneously measurable") if, when you measure one, then measure the other, then measure the first again, you are guaranteed of getting the same result in the third measurement that you got in the first. (These measurements are so close in time that the system state does not change appreciably between measurements one and two, nor between measurements two and three.)

We have seen examples of compatibility and incompatibility in the realm of magnetic moments: Suppose you measure μ_z, then $\mu_{(-z)}$, then μ_z again. If the result of the first measurement is $+\mu_B$, then the result of the second will be $-\mu_B$, and the result of the third is guaranteed to be $+\mu_B$ again.

Exercise 4.J. What will happen if the result of the first measurement is $-\mu_B$?

---

[3] It is sometimes called "the uncertainty of $\hat{A}$" but this name is inappropriate. It's like saying "I am uncertain about the color of love", suggesting that love *does indeed* have a color but I'm just not certain what that color is.

The observables μ_z and $\mu_{(-z)}$ are compatible.

But suppose you measure μ_z, then μ_x, then μ_z again. If the result of the first measurement is $+\mu_B$, then the result of the second might be either $+\mu_B$ or $-\mu_B$, and the result of the third has probability $\frac{1}{2}$ of being $+\mu_B$ and probability $\frac{1}{2}$ of being $-\mu_B$. The observables μ_z and μ_x are incompatible.

In classical mechanics, all observables are compatible. The existence of incompatibility is one of the most remarkable facets of quantum mechanics. The following theorem is useful and interesting in its own right, and its proof shows statements 3 and 4 in action.

The Compatibility Theorem.
Two observables have corresponding operators $\hat{A}$ and $\hat{B}$. Then any one of the following sentences implies the other two:

(1) The two observables are compatible.
(2) The two operators $\hat{A}$ and $\hat{B}$ possess a common eigenbasis.
(3) The two operators $\hat{A}$ and $\hat{B}$ commute.

Proof: We shall prove the theorem only for the case that all the eigenvalues of $\hat{A}$ and of $\hat{B}$ are nondegenerate. The theorem is true even without this condition, but the proof is more intricate and less insightful.[4] We will show that sentence (1) implies sentence (2), and vice versa, then that sentence (2) implies sentence (3), and vice versa. It immediately follows that sentences (1) and (3) imply each other.

(1) *implies* (2): Statement 2 says that the first measurement will yield some eigenvalue of $\hat{A}$, say the value a_5. At the end of the second measurement the system must, by statement 4, be in some eigenstate of $\hat{B}$, perhaps $|b_7\rangle$. Now, by the definition of compatibility, the third measurement is guaranteed to yield value a_5. Our assumption of nondegeneracy insists that the only state so guaranteed is $|a_5\rangle$. Thus the state $|b_7\rangle$ is the same as the state $|a_5\rangle$. This argument can be repeated for eigenvalues a_1, for a_{12}, for any eigenvalue of $\hat{A}$: Any eigenvector of $\hat{A}$ must also be an eigenvector of $\hat{B}$. We have shown that the eigenbasis for $\hat{A}$ is also an eigenbasis for $\hat{B}$, which

---

[4] A complete proof is given in F. Mandl, *Quantum Mechanics* (Wiley, Chichester, UK, 1992) section 3.1.

is sentence (2). We can renumber the eigenvalues and eigenvectors so that there is some basis $\{|\phi_n\rangle\}$ such that

$$\hat{A}|\phi_n\rangle = a_n|\phi_n\rangle \quad \text{and} \quad \hat{B}|\phi_n\rangle = b_n|\phi_n\rangle \quad \text{for all } n = 1, 2, 3, \ldots. \tag{4.6}$$

In our example, one member of this basis is $|a_5\rangle = |b_7\rangle$ which we might call, say, $|\phi_3\rangle$, so that we renumber a_5 to a_3 and renumber b_7 to b_3.

(2) *implies* (1): The first measurement yields some eigenvalue of $\hat{A}$ and, by statement 4, leaves the system in some member $|\phi_n\rangle$ of the common eigenbasis. The second measurement yields the eigenvalue of $\hat{B}$ associated with $|\phi_n\rangle$ but leaves the system in that same state $|\phi_n\rangle$. So the third measurement will yield the same result as the first, which is the definition of compatible.

(2) *implies* (3): Consider some member of the common eigenbasis $|\phi_n\rangle$. We have

$$\begin{aligned}\hat{A}\hat{B}|\phi_n\rangle &= \hat{A}b_n|\phi_n\rangle = b_n\hat{A}|\phi_n\rangle = b_n a_n|\phi_n\rangle \\ \hat{B}\hat{A}|\phi_n\rangle &= \hat{B}a_n|\phi_n\rangle = a_n\hat{B}|\phi_n\rangle = a_n b_n|\phi_n\rangle\end{aligned}$$

whence

$$(\hat{A}\hat{B} - \hat{B}\hat{A})|\phi_n\rangle = [\hat{A}, \hat{B}]|\phi_n\rangle = 0. \tag{4.7}$$

But for sentence (3) to be true, we must show that

$$[\hat{A}, \hat{B}]|\psi\rangle = 0 \tag{4.8}$$

for *all* state vectors $|\psi\rangle$, not only for vectors within the common eigenbasis. Any vector $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_n \psi_n|\phi_n\rangle \tag{4.9}$$

where the expansion coefficients are $\psi_n = \langle\phi_n|\psi\rangle$ (completeness). Applying commutator $[\hat{A}, \hat{B}]$ to the expansion (4.9) results in the needed equation (4.8).

(3) *implies* (2): Given that $\hat{A}$ and $\hat{B}$ commute, then for any eigenvector $|a_n\rangle$ of $\hat{A}$, we have

$$\hat{A}\hat{B}|a_n\rangle = \hat{B}\hat{A}|a_n\rangle = \hat{B}a_n|a_n\rangle = a_n\hat{B}|a_n\rangle \tag{4.10}$$

whence the vector $\hat{B}|a_n\rangle$ is an eigenvector of $\hat{A}$ with eigenvalue a_n. Our assumption of nondegeneracy insists that all such vectors are proportional to $|a_n\rangle$, so

$$\hat{B}|a_n\rangle = C|a_n\rangle. \tag{4.11}$$

That is, $|a_n\rangle$ is an eigenvector of $\hat{B}$ with eigenvalue C. Every eigenvector of $\hat{A}$ is also an eigenvector of $\hat{B}$, So the two operators $\hat{A}$ and $\hat{B}$ possess a common eigenbasis. ▮

Exercise 4.K. We have established that the observables μ_z and $\mu_{(-z)}$ are compatible, whereas μ_z and μ_x are incompatible. Use result (3.15) to verify the compatibility theorem.

4.3.3 *The Heisenberg Indeterminacy Principle*

If two observables (one corresponding to $\hat{A}$ and the other to $\hat{B}$) are compatible, then we can legitimately say that some states have a value for both observables. But if they are incompatible, then no state has a value for both observables: if the system is in state $|a_6\rangle$, then asking for the value of observable $\hat{B}$ is like asking "What is the color of love?" Can we say anything quantitative in this situation? Remarkably, we can.

Theorem: Heisenberg[5] Indeterminacy Principle.
If two observables correspond to operators $\hat{A}$ and $\hat{B}$, and the commutator of those two operators is a scalar complex number d:

$$[\hat{A}, \hat{B}] = d, \tag{4.12}$$

then in any state the indeterminacies satisfy

$$\Delta A\, \Delta B \geq \tfrac{1}{2}|d|. \tag{4.13}$$

Proof: We will actually prove the "generalized indeterminacy relation" that in any state $|\psi\rangle$, the indeterminacies satisfy

$$\Delta A\, \Delta B \geq \tfrac{1}{2}|\langle\psi|[\hat{A}, \hat{B}]|\psi\rangle|. \tag{4.14}$$

The Heisenberg result follows immediately.

In this proof, it will prove convenient to write the inner product

$$\langle\psi_1|\psi_2\rangle \quad \text{as} \quad (\psi_1, \psi_2).$$

Recall that

$$(\Delta A)^2 = \langle(\hat{A} - \langle\hat{A}\rangle)^2\rangle \quad \text{and that} \quad (\Delta B)^2 = \langle(\hat{B} - \langle\hat{B}\rangle)^2\rangle,$$

which inspires us to define the new operators

$$\hat{A}' \equiv \hat{A} - \langle\hat{A}\rangle \quad \text{and} \quad \hat{B}' \equiv \hat{B} - \langle\hat{B}\rangle.$$

It is easy to show that $\hat{A}'$ and $\hat{B}'$ are Hermitian operators; that the commutator $[\hat{A}, \hat{B}] = [\hat{A}', \hat{B}']$; and that $(\Delta A)^2 = (\hat{A}'\psi, \hat{A}'\psi)$.

---

[5] Biographical information on Werner Heisenberg appears on page 209.

Exercise 4.L. Prove these three statements.

With this background, investigate the right-hand side of the generalized indeterminacy relation by writing

$$\begin{aligned}(\psi, [\hat{A}, \hat{B}]\psi) &= (\psi, [\hat{A}', \hat{B}']\psi)\\ &= (\psi, \hat{A}'\hat{B}'\psi) - (\psi, \hat{B}'\hat{A}'\psi)\\ &= (\hat{A}'\psi, \hat{B}'\psi) - (\hat{B}'\psi, \hat{A}'\psi)\\ &= (\hat{A}'\psi, \hat{B}'\psi) - (\hat{A}'\psi, \hat{B}'\psi)^*\\ &= 2i\,\Im m\Big\{(\hat{A}'\psi, \hat{B}'\psi)\Big\}.\end{aligned}$$

Taking the magnitude of both sides

$$|(\psi, [\hat{A}, \hat{B}]\psi)| = 2\left|\Im m\Big\{(\hat{A}'\psi, \hat{B}'\psi)\Big\}\right|.$$

The magnitude of the imaginary part of a complex number is always less than or equal to the magnitude of the complex number, so

$$|(\psi, [\hat{A}, \hat{B}]\psi)| \le 2\,|(\hat{A}'\psi, \hat{B}'\psi)|.$$

Apply the Schwarz inequality, using $\hat{A}'|\psi\rangle$ for the $|\phi\rangle$ in equation (3.24) and $\hat{B}'|\psi\rangle$ for the $|\psi\rangle$ there, to obtain

$$|(\hat{A}'\psi, \hat{B}'\psi)| \le \sqrt{(\hat{A}'\psi, \hat{A}'\psi)}\cdot\sqrt{(\hat{B}'\psi, \hat{B}'\psi)}.$$

Put the last two inequalities together to find

$$|(\psi, [\hat{A}, \hat{B}]\psi)| \le 2\sqrt{(\hat{A}'\psi, \hat{A}'\psi)}\cdot\sqrt{(\hat{B}'\psi, \hat{B}'\psi)}$$

or

$$|(\psi, [\hat{A}, \hat{B}]\psi)| \le 2\,\Delta A\,\Delta B,$$

which is the desired generalized indeterminacy relation. ▮

Exercise 4.M. A silver atom is in state $|z+\rangle$. Verify the generalized indeterminacy relation (4.14) using $\hat{A} = \hat{\mu}_z$, $\hat{B} = \hat{\mu}_x$.

Exercise 4.N. A silver atom is in state $|z+\rangle$. Verify the generalized indeterminacy relation (4.14) using $\hat{A} = \hat{\mu}_{60^\circ}$, $\hat{B} = \hat{\mu}_x$. [*Clue:* Use equation (3.14), and the results of exercises 4.H and 4.I.]

Exercise 4.O. Words matter.

To say “the color of love is uncertain” suggests that love has a color, but the speaker is not sure what that color is. To say “the color of love is indeterminate” is slightly better. But we’re really going here into territory where we’ve been before: there is no word in English that represents exactly a phenomenon in quantum mechanics. Can you invent a better word?

4.4 The role of formalism

We started off trying to follow the behavior of a silver atom as it passed through various magnetic fields, and we ended up with an elaborate mathematical structure of state vectors, Hilbert space, operators, and eigenstates. This is a good time to step back and focus, not on the formalism, but on what the formalism is good for: what it does, what it doesn't do, and why we should care. We do so by looking at a different mathematical formalism for a more familiar physical problem.

Here's the physical problem: Suppose I count out 178 marbles and put them in an empty bucket. Then I count out 252 more marbles and put them in the same bucket. How many marbles are in the bucket?

There are a number of ways to solve this problem. First, by experiment: One could actually count out and place the marbles, and then count the number of marbles in the bucket at the end of the process. Second, by addition using Arabic numerals, using the rules for addition of three-digit numbers ("carrying") that we all learned in elementary school. Third, by the trick of writing

$$178 + 252 = 180 + 250 = 430$$

which reduces the problem to two-digit addition. Fourth, by converting from Arabic numerals in base 10 (decimal) to Arabic numerals in base 8 (octal) and adding the octal numerals:

$$178_{\text{(dec)}} + 252_{\text{(dec)}} = 262_{\text{(oct)}} + 374_{\text{(oct)}} = 656_{\text{(oct)}} = 430_{\text{(dec)}}.$$

Fifth, by converting to Roman numerals and adding them using the Roman addition rules that are simple and direct, but that you probably didn't learn in elementary school. Sixth, by converting to Mayan numerals and adding them using rules that are, to you, even less familiar. If you think about it, you'll come up with other methods.

The formal processes of Arabic numeral addition, Roman numeral addition, and Mayan numeral addition are interesting only because they give the same result as the experimental method of counting out marbles. These formal, mathematical processes matter only because they reflect something about the physical world. (It's clear that addition using decimal Arabic numerals is considerably easier — and cheaper — than actually doing the experiment. If you were trained in octal or Roman or Mayan numerals, then you'd also find executing those algorithms easier than doing the experiment.)

Does the algorithm of "carrying" tell us anything about addition? For example, does it help us understand what's going on when we count out the total number of marbles in the bucket at the end of the experiment? I would answer "no". The algorithm of carrying tells us not about addition, but about how we represent numbers using Arabic numerals with decimal positional notation ("place value"). The "carry digits" are a convenient mathematical tool to help calculate the total number of marbles in the bucket. The amount of carrying involved differs depending upon whether the addition is performed in decimal or in octal. It is absurd to think that one could look into the bucket and identify which marbles were involved in the carry and which were not! Nevertheless, you can and should develop an intuition about whether or not a carry will be needed when performing a sum. Indeed, when we wrote $178 + 252$ as $180 + 250$, we did so precisely to avoid a carry.

There are many ways to find the sum of two integers. These different methods differ in ease of use, in familiarity, in concreteness, in ability to generalize to negative, fractional, and imaginary numbers. So you might prefer one method to another. But you can't say that one method is right and another is wrong: the significance of the various methods is, in fact, that they all produce the same answer, and that that answer is the same as the number of marbles in the bucket at the end of the process.

As with marbles in a bucket, so with classical mechanics. You know several formalisms — several algorithms — for solving problems in classical mechanics: the Newtonian formalism, the Lagrangian formalism, the Hamiltonian formalism, Poisson brackets, etc. These formal, mathematical, algorithmic processes are significant only because they reflect something about the physical world.

The mathematical manipulations involved in solving a particular problem using Newton's force-oriented method differ dramatically from the mathematical manipulations involved in solving that same problem using Hamilton's energy-oriented method, but the two answers will always be the same. Just as one can convert integers from a representation as decimal Arabic numerals to a representation as octal Arabic numerals, or as Roman numerals, or as Mayan numerals, so one can add any constant to a Hamiltonian and obtain a different Hamiltonian that is just as good as the original.

Poisson brackets don't actually exist out in nature — you can never perform an experiment to measure the numerical value of a Poisson bracket — but they are convenient mathematical tools that help us calculate the values of positions that we *can* measure.

Although Lagrangians, Hamiltonians, and Poisson brackets are features of the algorithm, not features of nature, it is nevertheless possible to develop intuition concerning Lagrangians, Hamiltonians, and Poisson brackets. You might call this "physical intuition" or you might call it "mathematical intuition" or "algorithmic intuition". Regardless of what you call it, it's a valuable thing to learn.

These different methods for solving classical problems differ in ease of use, in familiarity, in concreteness, in ability to generalize to relativistic and quantal situations. So you might prefer one method to another. But you can't say that one method is right and another is wrong: the significance of the various methods is, in fact, that they all produce the same answer, and that that answer is the same as the classical behavior exhibited by the system in question.

As with marbles in a bucket, and as with classical mechanics, so with quantum mechanics. This chapter has developed an elegant and perhaps formidable formal apparatus representing quantal states as vectors in Hilbert space and experiments as operators in Hilbert space. This is not the only way of solving problems in quantum mechanics: One could go back to the fundamental rules for combining amplitudes in series and in parallel (page 60), just as one could go back to solving arithmetic problems by throwing marbles into a bucket. Or one could develop more elaborate and more formal ways to solve quantum mechanics problems, just as one could use the Lagrangian or Hamiltonian formulations in classical mechanics. This book will not treat these alternative formulations of quantum mechanics: the path integral formulation (Feynman), the phase space formulation (Wigner), the density matrix formulation (for an introduction, see section 4.5), the variational formulation, the pilot wave formulation (de Broglie-Bohm), or any of the others. But be assured that these alternative formulations exist, and their existence proves that kets and operators are features of the algorithmic tools we use to solve quantum mechanical problems, not features of nature.[6]

---

[6] Felix Bloch recounts a telling story in "Reminiscences of Heisenberg and the early days of quantum mechanics" [*Physics Today* **29**(12) (December 1976) 23–27]. Heisenberg and Bloch "were on a walk and somehow began to talk about space. I had just read Weyl's

The mathematical manipulations involved in solving a particular problem using the Hilbert space formalism differ dramatically from the mathematical manipulations involved in solving that same problem using the rules for combining amplitudes in series and in parallel, but the two answers will always be the same. In almost all cases the Hilbert space formalism is far easier to apply, and that's why we use it. We use it so often that we can fall into the trap of thinking that kets and operators are features of nature, not features of an algorithm. But remember that just as one can convert integers from a representation as decimal Arabic numerals to a representation as octal Arabic numerals, or as Roman numerals, or as Mayan numerals, so one can multiply any state vector by a constant of magnitude unity to obtain a different state vector that is just as good as the original. State vectors don't actually exist out in nature — you can never perform an experiment to measure the numerical value of a state vector (or even of an amplitude) — but they are convenient mathematical tools that help us calculate the values of probabilities that we *can* measure.

Many students, faced with the formidable mathematical formalism of quantum mechanics, fall into the trap of despair. "How can nature possibly be so sophisticated and formal?" This is the same trap as wondering "How can marbles know the algorithm for carrying in the addition of decimal Arabic numerals?" Nature doesn't know anything about Hilbert space, just as marbles don't know anything about carrying. The fact that the formalism of quantum mechanics is more sophisticated than the formalism of addition, or the formalism of classical mechanics, simply reflects the two facts (noted briefly on page 3, to be explored further on page 179) that quantum mechanics is far removed from common sense, and that quantum mechanics is stupendously rich.

4.5 The density matrix

4.1 Definition

A system is in quantum state $|\psi\rangle$. Define the operator

$$\hat{\rho} = |\psi\rangle\langle\psi|,$$

---

book *Space, Time and Matter*, and under its influence was proud to declare that space was simply the field of linear operations. 'Nonsense,' said Heisenberg, 'space is blue and birds fly through it.' This may sound naive, but I knew him well enough by that time to fully understand the rebuke. What he meant was that it was dangerous for a physicist to describe Nature in terms of idealized abstractions too far removed from the evidence of actual observation."

called the *density matrix*, recall the definition of the trace function from problem 3.7, and show that the mean value of the observable associated with operator $\hat{A}$ in $|\psi\rangle$ is

$$\mathrm{tr}\{\hat{\rho}\hat{A}\}.$$

4.2 **Statistical mechanics**
Frequently physicists don't know exactly which quantum state their system is in. (For example, silver atoms coming out of an oven are in states of definite $\vec{\mu}$ projection, but there is no way to know which state any given atom is in.) In this case there are two different sources of measurement uncertainty: first, we don't know what state they system is in (statistical uncertainty, due to our ignorance) and second, even if we did know, we couldn't predict the result of every measurement (quantum indeterminacy, due to the way the world works). The density matrix formalism neatly handles both sources of uncertainty at once.

If the system could be in any of the states $|a\rangle$, $|b\rangle, \ldots,$ $|i\rangle, \ldots$ (not necessarily a basis set), and if it has probability p_i of being in state $|i\rangle$, then the density matrix

$$\hat{\rho} = \sum_i p_i |i\rangle\langle i|$$

is associated with the system. Show that the mean value of the observable associated with $\hat{A}$ is still given by

$$\mathrm{tr}\{\hat{\rho}\hat{A}\}.$$

4.3 **Trace of the density matrix**
Show that $\mathrm{tr}\{\hat{\rho}\} = 1$. (This can be either a long and tedious proof, or a short and insightful one.)

Problems

4.4 **Anticommutators**
The "anticommutator" of two operators $\hat{A}$ and $\hat{B}$ is defined as

$$\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{4.15}$$

Apply the techniques used in the proof of the generalized indeterminacy relation (4.14) to anticommutators instead of commutators to prove that

$$\Delta A \, \Delta B \geq \Re e\Big\{\langle\hat{A}\hat{B}\rangle - \langle\hat{A}\rangle\langle\hat{B}\rangle\Big\}. \tag{4.16}$$

4.5 **Questions** (recommended problem)
Update your list of quantum mechanics questions that you started at problem 1.13 on page 56. Write down new questions and, if you have uncovered answers to any of your old questions, write them down briefly.

Chapter 5

Time Evolution

5.1 Operator for time evolution

If quantum mechanics is to have a classical limit, then quantal states have to change with time. We write this time dependence explicitly as

$$|\psi(t)\rangle. \tag{5.1}$$

We seek the equations that govern this time evolution, the ones parallel to the classical time evolution equations, be they the Newtonian equations

$$\vec{F} = m\vec{a} \tag{5.2}$$

or the Lagrange equations

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0 \tag{5.3}$$

or the Hamilton equations

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i, \quad \frac{\partial H}{\partial p_i} = \dot{q}_i. \tag{5.4}$$

Assume the existence of some "time evolution operator" $\hat{U}(\Delta t)$ such that

$$|\psi(t+\Delta t)\rangle = \hat{U}(\Delta t)|\psi(t)\rangle. \tag{5.5}$$

You might think that this statement is so general that we haven't assumed anything — we've just said that things are going to change with time. In fact we've made a big assumption: just by our notation we've assumed that the time-evolution operator $\hat{U}$ is linear, independent of the state $|\psi\rangle$ that's evolving. That is, we've assumed that the *same* operator will time-evolve *any* different state. (The operator will, of course, depend on which system is evolving in time: the number of particles involved, their interactions, their masses, the value of the magnetic field in which they move, and so forth.)

By virtue of the meaning of time, we expect the operator $\hat{U}(\Delta t)$ to have these four properties:

(1) $\hat{U}(\Delta t)$ is unitary.
(2) $\hat{U}(\Delta t_2)\hat{U}(\Delta t_1) = \hat{U}(\Delta t_2 + \Delta t_1)$.
(3) $\hat{U}(\Delta t)$ is dimensionless.
(4) $\hat{U}(0) = \hat{1}$.

And it's also reasonable[1] to assume that the time-evolution operator can be expanded in a Taylor series:

$$\hat{U}(\Delta t) = \hat{U}(0) + \hat{A}\Delta t + \hat{B}(\Delta t)^2 + \cdots . \tag{5.6}$$

We know that $\hat{U}(0) = \hat{1}$, and we'll write the quadratic and higher-order terms as $\hat{B}(\Delta t)^2 + \cdots = \mathcal{O}(\Delta t^2)$ ... which is read "terms of order Δt^2 and higher" or just as "terms of order Δt^2". Finally, we'll write $\hat{A}$ in a funny way so that

$$\hat{U}(\Delta t) = \hat{1} - \frac{i}{\hbar}\hat{H}\Delta t + \mathcal{O}(\Delta t^2). \tag{5.7}$$

I could just say, "we define $\hat{H} = i\hbar\hat{A}$" but that just shunts aside the important question — why is this a useful definition? There are two reasons: First, the operator $\hat{H}$ turns out to be Hermitian. (We will prove this in this section.) Second, because it's Hermitian, it can represent a measured quantity. When we investigate the classical limit in section 6.9.4, we will see that it corresponds to the classical energy. For now, you should just verify for yourself that it has the correct *dimensions* for energy.

---

[1] You are familiar with expanding a function $f(\Delta t)$ in a Taylor series. Is it really legitimate to expand an operator $\hat{U}(\Delta t)$ in a Taylor series? How do you define the derivative of an operator? A limit involving operators? The magnitude of an operator? For what values of Δt does this series converge?

These are fascinating questions but they are questions about mathematics, not about nature. In fact the Taylor series for operators is perfectly legitimate but proving that legitimacy is a difficult task that would take us too far afield. If you are interested in such questions — or indeed any question concerning any facet of mathematical physics — I recommend the magisterial four-volume work *Methods of Modern Mathematical Physics* by Michael Reed and Barry Simon (Academic Press, New York, 1972–1978).

Theoretical physics is a branch of physics; it answers questions about nature. *Mathematical physics* is a branch of mathematics; it answers questions about structure. I find both fields fascinating and refuse to denigrate either, but this book is about physics, not mathematics.

The energy operator is called “the Hamiltonian” and represented by the letter $\hat{H}$ in honor of William Rowan Hamilton,[2] who first pointed out the central role that energy can play in time evolution in the formal theory of classical mechanics.

Theorem: The operator $\hat{H}$ defined in equation (5.7) is Hermitian.

Proof: The proof uses the fact that the norm of $|\psi(t+\Delta t)\rangle$ equals the norm of $|\psi(t)\rangle$:

$$|\psi(t+\Delta t)\rangle = |\psi(t)\rangle - \frac{i}{\hbar}\Delta t \underbrace{\hat{H}|\psi(t)\rangle}_{\equiv |\psi_H(t)\rangle} + \mathcal{O}(\Delta t^2). \tag{5.8}$$

Thus

$$\begin{aligned}
&\langle\psi(t+\Delta t)|\psi(t+\Delta t)\rangle \\
&= \left[\langle\psi(t)| + \frac{i}{\hbar}\Delta t\langle\psi_H(t)| + \mathcal{O}(\Delta t^2)\right]\left[|\psi(t)\rangle - \frac{i}{\hbar}\Delta t|\psi_H(t)\rangle + \mathcal{O}(\Delta t^2)\right] \\
&= \langle\psi(t)|\psi(t)\rangle + \frac{i}{\hbar}\Delta t\left[\langle\psi_H(t)|\psi(t)\rangle - \langle\psi(t)|\psi_H(t)\rangle\right] + \mathcal{O}(\Delta t^2) \\
1 &= 1 + \frac{i}{\hbar}\Delta t\left[\langle\psi(t)|\psi_H(t)\rangle^* - \langle\psi(t)|\psi_H(t)\rangle\right] + \mathcal{O}(\Delta t^2) \\
0 &= \frac{i}{\hbar}\Delta t\left[\langle\psi(t)|\hat{H}|\psi(t)\rangle^* - \langle\psi(t)|\hat{H}|\psi(t)\rangle\right] + \mathcal{O}(\Delta t^2).
\end{aligned} \tag{5.9}$$

This equation has to hold *for all values of* Δt, so the quantity in square brackets must vanish![3] That is,

$$\langle\psi(t)|\hat{H}|\psi(t)\rangle^* = \langle\psi(t)|\hat{H}|\psi(t)\rangle \tag{5.10}$$

for *all* vectors $|\psi(t)\rangle$. It follows from exercise 3.S on page 113 that operator $\hat{H}$ is Hermitian. ▮

We have written the time-evolution equation as

$$|\psi(t+\Delta t)\rangle = |\psi(t)\rangle - \frac{i}{\hbar}\Delta t\hat{H}|\psi(t)\rangle + \mathcal{O}(\Delta t^2). \tag{5.11}$$

---

[2] Hamilton (1805–1865) made important contributions to mathematics, optics, classical mechanics, and astronomy. At the age of 22 years, while still an undergraduate, he was appointed professor of astronomy at his university and the Royal Astronomer of Ireland. As far as I have been able to determine, he was not related to the American founding father Alexander Hamilton.

[3] If I said that $0 = ax + bx^2$, then solutions would be $x = 0$ and $x = -a/b$. But if I said that $0 = ax + bx^2$ holds *for all values of* x, then I would instead conclude that $a = 0$ and $b = 0$.

Rearrangement gives

$$\frac{|\psi(t+\Delta t)\rangle - |\psi(t)\rangle}{\Delta t} = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle + \mathcal{O}(\Delta t). \tag{5.12}$$

In the limit $\Delta t \to 0$, this gives

$$\boxed{\frac{d|\psi(t)\rangle}{dt} = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle}, \tag{5.13}$$

an important result known as the Schrödinger[4] equation!

5.2 Energy eigenstates are stationary states

Because the Hamiltonian operator $\hat{H}$ is Hermitian, its eigenvalues are real. We say that energy eigenstate $|e_n\rangle$ has energy eigenvalue e_n when

$$\hat{H}|e_n\rangle = e_n|e_n\rangle. \tag{5.14}$$

If a system starts in an energy state represented by $|e_n\rangle$, it remains in that state forever — we call it a "stationary state". This section argues informally that energy eigenstates are stationary states. A formal proof is given by the theorem "Formal solution of the Schrödinger equation" on page 163. The informal argument of this section provides less rigor but more insight than that formal proof.

> **Result:** If $|\psi(0)\rangle = (\text{number})|e_n\rangle$, then $|\psi(t)\rangle = (\text{number})'|e_n\rangle$, where both numbers have magnitude unity.

Argument: Start at time $t = 0$ and step forward a small time Δt:

$$\begin{aligned}
\frac{\Delta|\psi\rangle}{\Delta t} &\approx -\frac{i}{\hbar}\hat{H}|\psi(0)\rangle \\
&= -\frac{i}{\hbar}\hat{H}(\text{number})|e_n\rangle \\
&= -\frac{i}{\hbar}e_n(\text{number})|e_n\rangle \\
&= (\text{stuff})|e_n\rangle. \\
\Delta|\psi\rangle &= (\text{stuff})\Delta t|e_n\rangle.
\end{aligned}$$

---

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

That is, the change in the state vector is parallel to the initial state vector, so the new state vector $|\psi(\Delta t)\rangle = |\psi(0)\rangle + \Delta|\psi\rangle$ is again parallel to the initial state vector, and all three vectors are parallel to $|e_n\rangle$. Repeat for as many time steps as needed.

The vector $|\psi(\Delta t)\rangle$ is not only parallel to the vector $|\psi(0)\rangle$, but it also has the same norm. (Namely unity.) This can't happen for regular position vectors multiplied by real numbers. The only way to multiply a vector by a number, and get a different vector with the same norm, is to multiply by a complex number.

Theory meets reality

We now have a theorem stating that if the system starts off in an energy eigenstate, it remains in that state forever. Yet you know that if, say, a hydrogen atom starts off in its fifth excited state, it *does not* stay in that state forever: instead it quickly decays to the ground state.[5] So what's up?

The answer is that if the Hamiltonian in equation (5.13) were exact, then the atom *would* stay in that stationary state forever. But real atoms are subject to collisions and radiation, so any Hamiltonian we write down is *not* exactly correct. Phenomena like collisions and radiation, unaccounted for in the Hamiltonian (5.13), cause the atom to fall into its ground state.

Because collisions and radiation are small effects, an atom starting off in the fifth excited state stays in that stationary state for a "long" time — but that means long relative to typical atomic times, such as the characteristic time 10^{-17} seconds generated at problem **??.??** on page **??**. If you study more quantum mechanics,[6] you will find that a typical atomic excited state lifetime is 10^{-9} seconds. So the excited state lifetime is very short by human standards, but very long by atomic standards. (To say "very long" is an understatement: it is 100 million times longer; by contrast the Earth has completed only 66 million orbits since the demise of the dinosaurs.)

The decay is "quick" on a human time scale, but very slow on an atomic time scale, because the model Hamiltonian is not the exact Hamiltonian, but a very close approximation.

---

[5] The energy eigenstate with lowest energy eigenvalue has a special name: the ground state.

[6] See for example David J. Griffiths and Darrell F. Schroeter, *Introduction to Quantum Mechanics*, third edition (Cambridge University Press, Cambridge, UK, 2018) section 11.3.2, "The Lifetime of an Excited State".

To check these claims, you can work with hydrogen in a very dilute gas, so that collisions are rare. At first glance you would think that you could never remove the atom from the electromagnetic field, but in fact excited atoms in electromagnetc resonant cavities can have altered lifetimes.[7]

5.3 Working with the Schrödinger equation

Quantal states evolve according to the Schrödinger time-evolution equation

$$\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle. \tag{5.15}$$

We have shown that the linear operator $\hat{H}$ is Hermitian and has the dimensions of energy. I've stated that we are going to show, when we discuss the classical limit, that the operator $\hat{H}$ corresponds to energy, and this justifies the name "Hamiltonian operator". That's still not much knowledge! This is just as it was in classical mechanics: Time evolution is governed by $\vec{F} = m\vec{a}$, but this doesn't help you until you know what forces are acting. Similarly, in quantum mechanics the Schrödinger equation is true but doesn't help us until we know how to find the Hamiltonian operator.

We find the Hamiltonian operator in quantum mechanics in the same way that we find the force function in classical mechanics: by appeal to experiment, to special cases, to thinking about the system and putting the pieces together. It's a creative task to stitch together the hints that we know to find a Hamiltonian. Sometimes in this book I'll be able to guide you down this creative path. Sometimes, as in great art, the creative process came through a stroke of genius that can only be admired and not explained.

5.3.1 *Representations of the Schrödinger equation*

As usual, we become familiar with states through their components, that is through their representations in a particular basis:

$$|\psi(t)\rangle = \sum_n \psi_n |n\rangle. \tag{5.16}$$

---

[7] See Serge Haroche and Daniel Kleppner, "Cavity Quantum Electrodynamics" *Physics Today* **42** (1) (January 1989) 24–30 and Serge Haroche and Jean-Michel Raimond, "Cavity Quantum Electrodynamics" *Scientific American* **268** (4) (April 1993) 54–62.

We know that $|\psi(t)\rangle$ changes with time on the left-hand side, so something has to change with time on the right-hand side. Which is it, the expansion coefficients ψ_n or the basis states $|n\rangle$? The choice has nothing to do with nature — it is purely formal. All our experimental results will depend on $|\psi(t)\rangle$, and whether we ascribe the time evolution to the expansion coefficients or to the basis states is merely a matter of convenience. There are three common conventions, called "pictures": In the "Schrödinger picture", the expansion coefficients change with time while the basis states don't. In the "Heisenberg picture" the reverse is true. In the "interaction picture" both expansion coefficients and basis states change with time.

time constant	time dependent	name
$\{\lvert n\rangle\}$	$\psi_n(t)$	Schrödinger picture
ψ_n	$\{\lvert n(t)\rangle\}$	Heisenberg picture
nothing	$\psi_n(t)$, $\{\lvert n(t)\rangle\}$	interaction picture

This book will use the Schrödinger picture, but be aware that this is mere convention.

In the Schrödinger picture, the expansion coefficients $\langle n|\psi(t)\rangle = \psi_n(t)$ change in time according to

$$\frac{d}{dt}\langle n|\psi(t)\rangle = -\frac{i}{\hbar}\langle n|\hat{H}|\psi(t)\rangle = -\frac{i}{\hbar}\sum_m \langle n|\hat{H}|m\rangle\langle m|\psi(t)\rangle, \tag{5.17}$$

or, in other words, according to

$$\frac{d\psi_n(t)}{dt} = -\frac{i}{\hbar}\sum_m H_{n,m}\psi_m(t) \quad \text{where, recall} \quad H_{n,m} = H^*_{m,n}. \tag{5.18}$$

5.3.2 *A system with one basis state*

Consider a system with one basis state — say, a motionless hydrogen atom in its electronic ground state, which we call $|1\rangle$. Then

$$|\psi(t)\rangle = \psi_1(t)|1\rangle$$

If the initial state happens to be

$$|\psi(0)\rangle = |1\rangle,$$

then the time evolution problem is

$$\begin{aligned} \text{Initial condition: } & \psi_1(0) = 1 \\ \text{Differential equation: } & \frac{d\psi_1(t)}{dt} = -\frac{i}{\hbar}E_g\psi_1(t), \end{aligned}$$

where $E_g = \langle 1|\hat{H}|1\rangle$ is the energy of the ground state.

The solution is straightforward:

$$\psi_1(t) = 1e^{-(i/\hbar)E_g t}$$

or, in other words,

$$|\psi(t)\rangle = e^{-(i/\hbar)E_g t}|1\rangle. \tag{5.19}$$

Because two state vectors that differ only in phase represent the same state, the *state* doesn't change even though the *coefficient* $\psi_1(t)$ does change with time. The system stays always in the ground state.

When I was in high school, my chemistry teacher said that "an atom is a pulsating blob of probability". He was thinking of this equation, with the expansion coefficient $\psi_1(t)$ changing in time as

$$e^{-(i/\hbar)E_g t} = \cos((E_g/\hbar)t) - i\sin((E_g/\hbar)t). \tag{5.20}$$

On one hand you know that this function "pulsates" — that is, changes in time periodically with period $2\pi\hbar/E_g$. On the other hand you know also that this function represents an irrelevant overall phase — for example, it has no effect on any probability at all. My high school chemistry teacher was going overboard in ascribing physical reality to the mathematical tools we use to describe reality.

Exercise 5.A. Change energy zero. You know the energy zero is purely conventional so changing the energy zero shouldn't change anything in the physics. And indeed it changes only the phase, which is also purely conventional. In the words of my high school chemistry teacher this changes the "pulsation" rate — but it doesn't change anything about the behavior of the hydrogen atom.

5.4 A system with two basis states: The silver atom

Consider a system with two basis states — say, a silver atom in a uniform vertical magnetic field. Take the two basis states to be

$$|1\rangle = |z+\rangle \quad \text{and} \quad |2\rangle = |z-\rangle. \tag{5.21}$$

It's very easy to write down the differential equation

$$\frac{d}{dt}\begin{pmatrix}\psi_1(t)\\ \psi_2(t)\end{pmatrix} = -\frac{i}{\hbar}\begin{pmatrix}H_{1,1} & H_{1,2}\\ H_{2,1} & H_{2,2}\end{pmatrix}\begin{pmatrix}\psi_1(t)\\ \psi_2(t)\end{pmatrix} \tag{5.22}$$

but it's much harder to see what the elements in the Hamiltonian matrix should be — that is, it's hard to guess the Hamiltonian operator.

The classical energy for this system is

$$U = -\vec{\mu} \cdot \vec{B} = -\mu_z B. \tag{5.23}$$

Our guess for the quantum Hamiltonian is simply to change quantities into operators

$$\hat{H} = -\hat{\mu}_z B \tag{5.24}$$

where

$$\hat{\mu}_z = (+\mu_B)|z+\rangle\langle z+| + (-\mu_B)|z-\rangle\langle z-| \tag{5.25}$$

is the quantum mechanical operator corresponding to the observable μ_z (see equation 3.12). In this equation B is not an operator but simply a number, the magnitude of the classical magnetic field in which the silver atom is immersed. You might think that we should quantize the magnetic field as well as the atomic magnetic moment, and indeed a full quantum-mechanical treatment would have to include the quantum theory of electricity and magnetism. That's a task for later. For now, we'll accept the Hamiltonian (5.24) as a reasonable starting point, and indeed it turns out to describe this system to high accuracy, although not perfectly.[8]

It is an easy exercise to show that in the basis

$$\{|z+\rangle, |z-\rangle\} = \{|1\rangle, |2\rangle\},$$

the Hamiltonian operator (5.24) is represented by the matrix

$$\begin{pmatrix} H_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{pmatrix} = \begin{pmatrix} -\mu_B B & 0 \\ 0 & +\mu_B B \end{pmatrix}. \tag{5.26}$$

Thus the differential equations (5.22) become

$$\frac{d}{dt}\begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix} = -\frac{i}{\hbar}\begin{pmatrix} -\mu_B B & 0 \\ 0 & +\mu_B B \end{pmatrix}\begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix} \tag{5.27}$$

or

$$\begin{aligned} \frac{d\psi_1(t)}{dt} &= -\frac{i}{\hbar}(-\mu_B B)\psi_1(t) \\ \frac{d\psi_2(t)}{dt} &= -\frac{i}{\hbar}(+\mu_B B)\psi_2(t). \end{aligned}$$

---

[8] If you want perfection, you'll need to look at some discipline other than science.

The solutions are straightforward:

$$\begin{aligned}\psi_1(t) &= \psi_1(0)e^{-(i/\hbar)(-\mu_B B)t}\\ \psi_2(t) &= \psi_2(0)e^{-(i/\hbar)(+\mu_B B)t}.\end{aligned}$$

TALK about initial state $|z+\rangle$. Stationary state.

Suppose the initial state is

$$|x+\rangle = |z+\rangle\langle z+|x+\rangle + |z-\rangle\langle z-|x+\rangle = |z+\rangle\tfrac{1}{\sqrt{2}} + |z-\rangle\tfrac{1}{\sqrt{2}},$$

where we have used the amplitude conventions of equation (2.18). Then

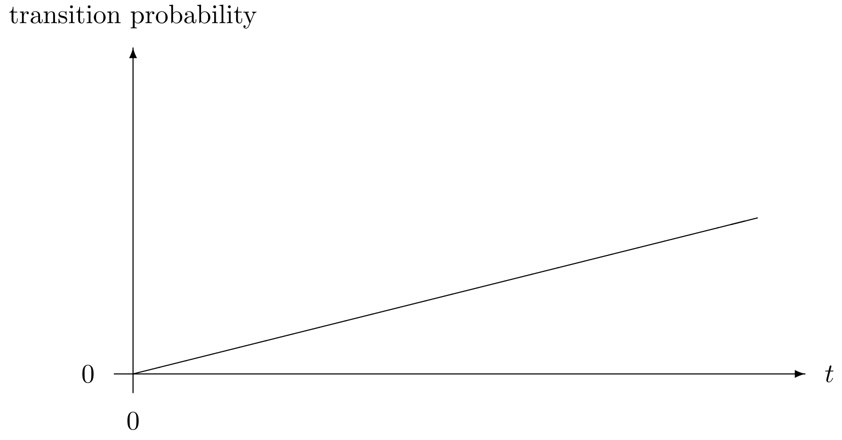
$$\psi_1(0) = \tfrac{1}{\sqrt{2}} \qquad \psi_2(0) = \tfrac{1}{\sqrt{2}}$$

so

$$|\psi(t)\rangle = \tfrac{1}{\sqrt{2}}e^{-(i/\hbar)(-\mu_B B)t}|z+\rangle + \tfrac{1}{\sqrt{2}}e^{-(i/\hbar)(+\mu_B B)t}|z-\rangle.$$

So the atom is produced in state $|x+\rangle$, then is exposed to a vertical magnetic field for time t, and ends up in the state mentioned above. If we now measure μ_x, what is the probability that it has changed from $+\mu_B$ to $-\mu_B$?

Before doing any calculation, I like to make a guess. My personal **expectation** is that the magnetic field induces a transition from $|x+\rangle$ to $|x-\rangle$, so the more time an atom spends in the field, the more likely it is to make the transition.



With the guess out of the way, let's do the **calculation**. The probability of transitioning from $|x+\rangle$ to $|x-\rangle$ is the square of the amplitude

$$\begin{aligned}\langle x - |\psi(t)\rangle &= \tfrac{1}{\sqrt{2}} e^{-(i/\hbar)(-\mu_B B)t} \langle x - |z+\rangle + \tfrac{1}{\sqrt{2}} e^{-(i/\hbar)(+\mu_B B)t} \langle x - |z-\rangle \\ &= \tfrac{1}{\sqrt{2}} e^{-(i/\hbar)(-\mu_B B)t} \left(-\tfrac{1}{\sqrt{2}}\right) + \tfrac{1}{\sqrt{2}} e^{-(i/\hbar)(+\mu_B B)t} \left(\tfrac{1}{\sqrt{2}}\right) \\ &= \frac{1}{2}\left[-e^{-(i/\hbar)(-\mu_B B)t} + e^{-(i/\hbar)(+\mu_B B)t}\right] \\ &= \frac{1}{2}\left[-2i \sin((1/\hbar)(\mu_B B)t)\right] \\ &= -i \sin\left(\frac{\mu_B B}{\hbar} t\right).\end{aligned}$$

The probability is

$$|\langle x - |\psi(t)\rangle|^2 = \sin^2\left(\frac{\mu_B B}{\hbar} t\right) \tag{5.28}$$

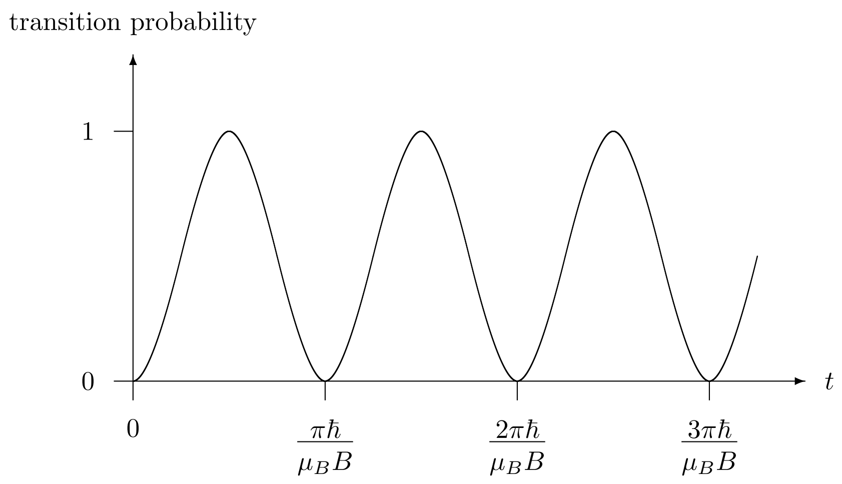
which starts at zero when $t = 0$, then goes up to 1, then goes back down to zero, with an oscillation period of

$$\frac{\pi\hbar}{\mu_B B}.$$

Reflection. The transition probability result, graphed below as a function of time, shows oscillatory behavior called "Rabi[9] flopping". This is the beat at the heart of an atomic clock.

---

[9] Isidor Isaac Rabi (1898–1988), Polish-Jewish-American physicist. He won the Nobel Prize for his discovery of nuclear magnetic resonance, but he contributed to the invention of the laser and of the atomic clock as well. His fascinating life cannot be summarized in a few sentences: I recommend John Rigden's biography *Rabi: Scientist and Citizen* (Basic Books, New York, 1987).



I have made bad guesses in my life, but none worse than the difference between my expectation graphed on page 150 and the real behavior graphed above. It's as if, while hammering a nail into a board, the first few strikes drive the nail deeper and deeper into the board, but additional strikes make the nail come out of the board. And one strike (at time $\pi\hbar/\mu_B B$) makes the nail pop out of the board altogether! Is there any way to account for this bizarre result other than shrugging that "It comes out of the math"?

There is. This is a form of interference[10] where the particle moves not from point to point through two possible slits, but from spin state to spin state with two possible intermediate states. The initial state is $|x+\rangle$ and the final state is $|x+\rangle$. The two possible intermediates are $|x-\rangle$ and $|x+\rangle$. There is an amplitude to go from $|x+\rangle$ to $|x+\rangle$ via $|x-\rangle$, and an amplitude to go from $|x+\rangle$ to $|x+\rangle$ by staying in $|x+\rangle$. At time $\frac{1}{2}\pi\hbar/|\mu_B B$ those two amplitudes interfere destructively so there is a small probability of ending up in $|x+\rangle$ and hence a large probability of ending up in $|x-\rangle$. At time $\pi\hbar/|\mu_B B$ those two amplitudes interfere constructively so there is a large probability of ending up in $|x+\rangle$ and hence a large probability of ending up in $|x-\rangle$.

---

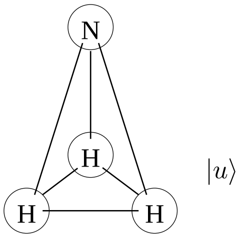
[10] This point of view is expounded by R.P. Feynman and A.R. Hibbs in section 6-5 of *Quantum Mechanics and Path Integrals*, emended edition (Dover Publications, Mineola, NY, 2010).

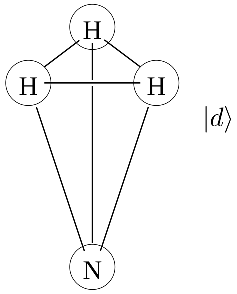
Problem

5.1 Some problem where initial state is $|\theta+\rangle$ and final is $|\phi+\rangle$ or similar.

5.5 Another two-state system: The ammonia molecule

Another system with two basis states is the ammonia molecule NH_3. If we ignore translation and rotation, and assume that the molecule is rigid,[11] then there are still two possible states for the molecule: state $|u\rangle$ with the nitrogen atom pointing up, and state $|d\rangle$ with the nitrogen atom pointing down. These are states of definite position for the nitrogen atom, but not states of definite energy (stationary states) because there is some amplitude for the nitrogen atom to tunnel from the "up" position to the "down" position. That is, if you start with the atom in state $|u\rangle$, then some time later it might be in state $|d\rangle$, because the nitrogen atom tunneled through the plane of hydrogen atoms.





What is the implication of such tunneling for the Hamiltonian matrix? The matrix we dealt with in equation (5.26) was diagonal, and hence the two differential equations split up ("decoupled") into one involving $\psi_1(t)$ and another involving $\psi_2(t)$. These were independent: If a system started

---

[11] That is, ignore vibration. These approximations seem, at first glance, to be absurd. They are in fact excellent approximations, because the tunneling is independent of translation, rotation, or vibration.

out in the state $|1\rangle$ (i.e. $\psi_1(t) = e^{-(i/\hbar)H_{1,1}t}$, $\psi_2(t) = 0$), then it stayed there forever. We've just said that this is *not* true for the ammonia molecule, so the Hamiltonian matrix must *not* be diagonal.

The Hamiltonian matrix in the $\{|u\rangle, |d\rangle\}$ basis has the form

$$\begin{pmatrix} H_{u,u} & H_{u,d} \\ H_{d,u} & H_{d,d} \end{pmatrix} = \begin{pmatrix} E & Ae^{i\phi} \\ Ae^{-i\phi} & E \end{pmatrix}. \tag{5.29}$$

The two off-diagonal elements must be complex conjugates of each other because the matrix is Hermitian. It's reasonable that the two on-diagonal elements are equal because the states $|u\rangle$ and $|d\rangle$ are mirror images and hence $\langle u|\hat{H}|u\rangle = \langle d|\hat{H}|d\rangle$. The term $Ae^{i\phi}$ is related to a tunneling amplitude. (SAY MORE HERE.) The term $Ae^{i\phi}$ implies that a molecule starting with the nitrogen atom up (state $|u\rangle$) will not stay that way forever. At some time it might "tunnel" to the down position (state $|d\rangle$).

For this Hamiltonian, the Schrödinger equation is

$$\frac{d}{dt}\begin{pmatrix} \psi_u(t) \\ \psi_d(t) \end{pmatrix} = -\frac{i}{\hbar}\begin{pmatrix} E & Ae^{i\phi} \\ Ae^{-i\phi} & E \end{pmatrix}\begin{pmatrix} \psi_u(t) \\ \psi_d(t) \end{pmatrix} \tag{5.30}$$

or

$$\begin{aligned} \frac{d\psi_u(t)}{dt} &= -\frac{i}{\hbar}\left[E\psi_u(t) + Ae^{i\phi}\psi_d(t)\right] \\ \frac{d\psi_d(t)}{dt} &= -\frac{i}{\hbar}\left[Ae^{i\phi}\psi_u(t) + E\psi_d(t)\right]. \end{aligned}$$

It's hard to see how to approach solving this pair of differential equations. The differential equation for $\psi_u(t)$ involves the unknown function $\psi_d(t)$, while the differential equation for $\psi_d(t)$ involves the unknown function $\psi_u(t)$. We were able to solve the differential equations (5.27) with ease precisely because they didn't involve such "crosstalk".

And this observation suggests a path forward: While the equations hard to solve in this initial basis, they would be easy to solve in a basis where the matrix is diagonal. So, following the four-step procedure on page 117, we search for a basis that diagonalizes the matrix.

---

1. The Hamiltonian is represented in the initial basis $\{|u\rangle, |d\rangle\}$ by

$$\mathsf{M} = \begin{pmatrix} E & Ae^{i\phi} \\ Ae^{-i\phi} & E \end{pmatrix}$$

2. Find the eigenvalues.

$$\begin{aligned}
\det\begin{vmatrix} E-\lambda & Ae^{i\phi} \\ Ae^{-i\phi} & E-\lambda \end{vmatrix} &= 0 \\
(E-\lambda)^2 - A^2 &= 0 \\
(E-\lambda)^2 &= A^2 \\
E-\lambda &= \pm A \\
\lambda &= E \pm A \\
\lambda_1 &= E - A \qquad (5.31) \\
\lambda_2 &= E + A \qquad (5.32)
\end{aligned}$$

As required by the theorem on Hermitian eigenproblems (page 113), the eigenvalues are real.

3. Find the eigenvectors.

We start with the eigenvector for $\lambda_1 = E - A$:

$$\begin{aligned}
\mathsf{M}\mathsf{e}_1 &= \lambda_1 \mathsf{e}_1 \\
(\mathsf{M} - \lambda_1 \mathsf{I})\mathsf{e}_1 &= 0 \\
\begin{pmatrix} E-\lambda_1 & Ae^{i\phi} \\ Ae^{-i\phi} & E-\lambda_1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} A & Ae^{i\phi} \\ Ae^{-i\phi} & A \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 1 & e^{i\phi} \\ e^{-i\phi} & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
x + e^{i\phi} y &= 0 \\
e^{-i\phi} x + y &= 0
\end{aligned}$$

As always (compare equation 3.64) these two equations are not independent! The second is $e^{-i\phi}$ times the first. The solution is $y = -e^{-i\phi}x$, so for any value of x

$$\mathsf{e}_1 = \begin{pmatrix} x \\ -e^{-i\phi}x \end{pmatrix}$$

represents an eigenvector.

Although I could choose any value of x that I wanted, it is most convenient to work with normalized eigenvectors, for which

$$\begin{aligned}
|x|^2 + |y|^2 &= 1 \\
|x|^2 + |-e^{-i\phi}x|^2 &= 1 \\
2|x|^2 &= 1
\end{aligned}$$

This equation has many solutions. I could pick

$$x = \frac{1}{\sqrt{2}} \quad \text{or} \quad x = -\frac{1}{\sqrt{2}} \quad \text{or} \quad x = \frac{i}{\sqrt{2}} \quad \text{or} \quad x = \frac{1+i}{2}$$

but there's no advantage to picking a solution with all sorts of unneeded symbols. So I choose the first possibility and write

$$\mathsf{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -e^{-i\phi} \end{pmatrix}.$$

This is the representation of $|e_1\rangle$ in the basis $\{|u\rangle, |d\rangle\}$.

Exercise 5.B. Show that an eigenvector associated with $\lambda_2 = E + A$ is

$$|e_2\rangle \doteq \mathsf{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{-i\phi} \end{pmatrix}. \tag{5.33}$$

Exercise 5.C. Verify that $\langle e_1|e_2\rangle = 0$, as required by the theorem on Hermitian eigenproblems (page 113).

In summary,

$$\begin{aligned} |e_1\rangle &= \tfrac{1}{\sqrt{2}} \left[|u\rangle - e^{-i\phi}|d\rangle\right] \\ |e_2\rangle &= \tfrac{1}{\sqrt{2}} \left[|u\rangle + e^{-i\phi}|d\rangle\right]. \end{aligned} \tag{5.34}$$

Exercise 5.D. Show that $\{|e_1\rangle, |e_2\rangle\}$ constitute a spanning set by building $|u\rangle$ and $|d\rangle$ out of $|e_1\rangle$ and $|e_2\rangle$.
(*Answer:* $|u\rangle = \frac{1}{\sqrt{2}}(|e_1\rangle + |e_2\rangle)$, $|d\rangle = \frac{1}{\sqrt{2}}e^{i\phi}(-|e_1\rangle + |e_2\rangle)$.)

4. In the basis $\{|e_1\rangle, |e_2\rangle\}$, the matrix representation of the Hamiltonian is

$$\begin{pmatrix} E-A & 0 \\ 0 & E+A \end{pmatrix}.$$

In the press of solving our immediate problem, it's easy to miss that we've reached a milestone here. We started our journey into quantum mechanics with the phenomenon of quantization. Continued exploration uncovered the phenomena of interference and entanglement. Attempting to describe these three phenomena we invented the tool of amplitude, and we have only now developed the mathematical machinery to the extent that that machinery can predict quantization: It predicts that the energy cannot take on any old value, but only the values $E-A$ and $E+A$. Having recognized this milestone, we continue with our immediate problem and see how to use it.

It's now straightforward to solve the differential equations. Using the notation

$$|\psi(t)\rangle = \bar{\psi}_1(t)|e_1\rangle + \bar{\psi}_2(t)|e_2\rangle,$$

the time evolution differential equations are

$$\begin{aligned}\frac{d\bar{\psi}_1(t)}{dt} &= -\frac{i}{\hbar}(E-A)\bar{\psi}_1(t)\\ \frac{d\bar{\psi}_2(t)}{dt} &= -\frac{i}{\hbar}(E+A)\bar{\psi}_2(t)\end{aligned}$$

with the immediate solutions

$$\begin{aligned}\bar{\psi}_1(t) &= \bar{\psi}_1(0)e^{-(i/\hbar)(E-A)t}\\ \bar{\psi}_2(t) &= \bar{\psi}_2(0)e^{-(i/\hbar)(E+A)t}.\end{aligned}$$

Thus

$$|\psi(t)\rangle = e^{-(i/\hbar)Et}\left[e^{-(i/\hbar)(-A)t}\bar{\psi}_1(0)|e_1\rangle + e^{-(i/\hbar)(+A)t}\bar{\psi}_2(0)|e_2\rangle\right]. \quad (5.35)$$

(I am surprised that this time evolution result — and indeed the result of any possible experiment — is independent of the phase ϕ of the off-diagonal element of the Hamiltonian. This surprise is explained in problem 5.11.)

Let's try out this general solution for a particular initial condition. Suppose the nitrogen atom starts out "up" — that is,

$$|\psi(0)\rangle = |u\rangle, \quad (5.36)$$

and we ask for the probability of finding it "down" — that is, $|\langle d|\psi(t)\rangle|^2$.

The initial expansion coefficients in the $\{|e_1\rangle, |e_2\rangle\}$ basis are (see equations 5.34)

$$\begin{aligned}\bar{\psi}_1(0) &= \langle e_1|\psi(0)\rangle = \langle e_1|u\rangle = \tfrac{1}{\sqrt{2}}\\ \bar{\psi}_2(0) &= \langle e_2|\psi(0)\rangle = \langle e_2|u\rangle = \tfrac{1}{\sqrt{2}}\end{aligned}$$

so

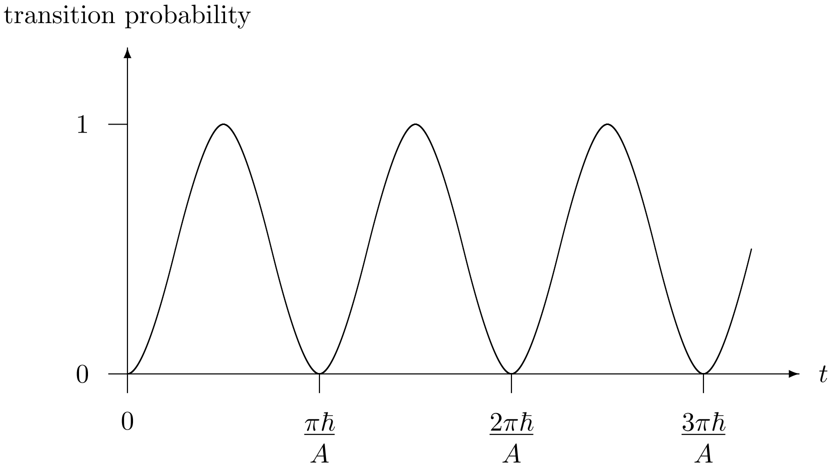
$$|\psi(t)\rangle = \tfrac{1}{\sqrt{2}}e^{-(i/\hbar)Et}\left[e^{+(i/\hbar)At}|e_1\rangle + e^{-(i/\hbar)At}|e_2\rangle\right].$$

The amplitude to find the nitrogen atom "down" is

$$\begin{aligned}
\langle d|\psi(t)\rangle &= \tfrac{1}{\sqrt{2}}e^{-(i/\hbar)Et}\left[e^{+(i/\hbar)At}\langle d|e_1\rangle + e^{-(i/\hbar)At}\langle d|e_2\rangle\right] \\
&= \tfrac{1}{\sqrt{2}}e^{-(i/\hbar)Et}\left[e^{+(i/\hbar)At}\left(-\tfrac{1}{\sqrt{2}}e^{-i\phi}\right) + e^{-(i/\hbar)At}\left(\tfrac{1}{\sqrt{2}}e^{-i\phi}\right)\right] \\
&= \tfrac{1}{2}e^{-i\phi}e^{-(i/\hbar)Et}\left[-e^{+(i/\hbar)At} + e^{-(i/\hbar)At}\right] \\
&= \tfrac{1}{2}e^{-i\phi}e^{-(i/\hbar)Et}\left[-2i\sin\left((1/\hbar)At\right)\right] \\
&= -ie^{-i\phi}e^{-(i/\hbar)Et}\sin\left(\frac{A}{\hbar}t\right)
\end{aligned}$$

and thus the probability of finding the nitrogen atom "down" is

$$|\langle d|\psi(t)\rangle|^2 = \sin^2\left(\frac{A}{\hbar}t\right). \tag{5.37}$$



This oscillation has period

$$\frac{\pi\hbar}{A} = \frac{2\pi\hbar}{\Delta E}$$

where ΔE represents the energy splitting between the two energy eigenvalues, $E + A$ and $E - A$.

This oscillation is at the heart of the MASER (Microwave Amplification by Simulated Emission of Radiation).

Reflection

In one sense we have solved the problem, using the mathematical trick of matrix diagonalization to produce solutions that at first glance (below equation 5.30) seemed beyond reach. But we should not stop there. In his book *Mathematics in Action*, O. Graham Sutton writes that "A technique succeeds in mathematical physics, not by a clever trick, or a happy accident, but because it expresses some aspect of a physical truth." What aspect of physical truth is exposed through the technique of matrix diagonalization?

What are these states we've been dealing with like?

- States $|u\rangle$ and $|d\rangle$ have definite positions for the nitrogen atom, namely "up" or "down". But they don't have definite energies. These states are sketched on page 153.
- States $|e_1\rangle$ and $|e_2\rangle$ have definite energies, namely $E - A$ or $E + A$. But they don't have definite positions for the nitrogen atom. They can't be sketched using classical ink. (For a molecule in this state the nitrogen atom is like a silver atom ambivating through "both branches" of an interferometer — the atom doesn't have a position.)

The mathematical technique of matrix diagonalization has led us to the physical truth of energy states. Most states *don't have an energy* and most states *aren't stationary states*. But if a state *does* have an energy, then it *is* a stationary state. In such states the nitrogen atom does not have a position. And in states where the nitrogen atom does have a position, the state does not have an energy.

There is a medical condition called prosopagnosia. People with this condition cannot recognize faces. This does not mean that those with prosopagnosia cannot recognize their friends — instead they use other ways to identify people, such as relying on voice, or clothing, or height.

All of us suffer from prosopagnosia with respect to the sates $|e_1\rangle$ and $|e_2\rangle$. This does not mean we cannot recognize those states, it just means we must rely on non-pictorial recognition. We must recognize them through their energies, not through the positions of their nitrogen atoms. The neurologist Oliver Sacks wrote an accurate and sympathetic account of a patient with prosopagnosia in his 1985 essay *The Man Who Mistook His Wife for a Hat*. Reading this essay might make you more sympathetic to your own prosopagnosia with respect to the state $|e_1\rangle$.

Exercise 5.E. Back when we discussed quatal interference, we said things like "An atom in state $|\psi\rangle$ ambivating through a vertical interfeometer doesn't take either path: instead it has amplitude $\langle z+|\psi\rangle$ to take path a and amplitude $\langle z-|\psi\rangle$ to take path b. For example, an atom in state $|\theta-\rangle$ has amplitude $\langle z+|\theta-\rangle = -\sin(\theta/2)$ to take path a."

Write a parallel statement by filling in the missing words from "For an ammonia molecule in state $|\psi\rangle$ the nitrogen atom doesn't have a position: instead .... For example, a molecule in state $|e_2\rangle$ has amplitude ... to be up."

Problems

5.2 Probability of no change

In equation (5.37) we found the probability that the nitrogen atom began in the "up" position (equation 5.36) and finished in the "down" position. Find the amplitude and the probability that the nitrogen atom will finish in the "up" position, and verify that these two probabilities sum to 1.

5.3 Tunneling for small times

Equation (5.35) solves the time evolution problem completely, for all time. But it doesn't give a lot of insight into what's "really going on". This problem provides some of that missing insight.

a. When the time involved is short, we can approximate time evolution through

$$|\psi(\Delta t)\rangle = \left[\hat{1} - \frac{i}{\hbar}\hat{H}\Delta t + \cdots\right]|\psi(0)\rangle. \tag{5.38}$$

Show that this equation, represented in the $\{|u\rangle, |d\rangle\}$ basis, is

$$\begin{pmatrix} \psi_u(\Delta t) \\ \psi_d(\Delta t) \end{pmatrix} \approx \begin{pmatrix} 1-(i/\hbar)E\Delta t & -(i/\hbar)Ae^{i\phi}\Delta t \\ -(i/\hbar)Ae^{-i\phi}\Delta t & 1-(i/\hbar)E\Delta t \end{pmatrix} \begin{pmatrix} \psi_u(0) \\ \psi_d(0) \end{pmatrix}. \tag{5.39}$$

b. Express the initial condition $|\psi(0)\rangle = |u\rangle$, used above at equation (5.36), in the $\{|u\rangle, |d\rangle\}$ basis, and show that, for small times,

$$\begin{pmatrix} \psi_u(\Delta t) \\ \psi_d(\Delta t) \end{pmatrix} \approx \begin{pmatrix} 1-(i/\hbar)E\Delta t \\ -(i/\hbar)Ae^{i\phi}\Delta t \end{pmatrix}. \tag{5.40}$$

c. This shows that the system starts with amplitude 1 for being in state $|u\rangle$, but that amplitude "seeps" (or "diffuses" or "hops") from $|u\rangle$ into $|d\rangle$. In fact, the amplitude to be found in $|d\rangle$ after a small time Δt has passed is $-(i/\hbar)Ae^{i\phi}\Delta t$. What is the probability of being found in $|d\rangle$? What is the condition for a "small" time?

d. Show that the same probability results from approximating result (5.37) for small times.

In a normal diffusion process – such as diffusion of blue dye from one water cell into an adjacent water cell – the dye spreads out uniformly and then net diffusion stops. But in this quantal amplitude diffusion, the amplitude is complex-valued. As such, the diffusion of more amplitude into the second cell can result, through destructive interference, in a decreased amplitude in the second cell. This interference gives rise to the oscillatory behavior demonstrated in equation (5.37).

e. While this approach does indeed provide a lot of insight, it also raises a puzzle. What, according to equation (5.40), is the probability of being found in the initial state $|u\rangle$ after a short time has passed? Conclude that the total probability is greater than 1! We will resolved this paradox in problem 11.1.

5.4 **Ammonia molecule: position of nitrogen atom**

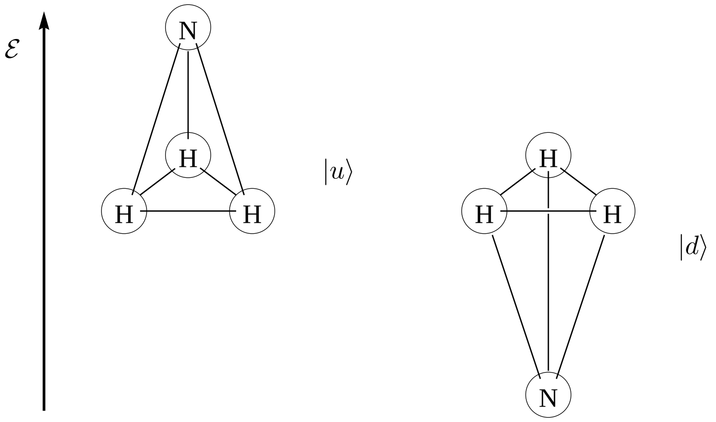
In state $|u\rangle$, the nitrogen atom is positioned a distance s above the plane of three hydrogen atoms; in state $|d\rangle$ it is positioned the same distance below. The position of the nitrogen atom is thus represented (compare equation 3.12) by the operator

$$\hat{z}_N = (+s)|u\rangle\langle u| + (-s)|d\rangle\langle d|. \tag{5.41}$$

Write the matrix representation of the $\hat{z}_N$ operator in the basis $\{|u\rangle, |d\rangle\}$ and in the basis $\{|e_1\rangle, |e_2\rangle\}$. What is the commutator $[\hat{z}_N, \hat{H}]$?

5.5 **Ammonia molecule in an electric field**

Place an ammonia molecule into an external electric field $\mathcal{E}$ perpendicular to the plane of hydrogen atoms.



Now the states $|u\rangle$ and $|d\rangle$ are no longer symmetric, so we can no longer assume that $\langle u|\hat{H}|u\rangle = \langle d|\hat{H}|d\rangle$. Indeed, the proper matrix representation of $\hat{H}$ in the $\{|u\rangle, |d\rangle\}$ basis is

$$\begin{pmatrix} E + p\mathcal{E} & Ae^{i\phi} \\ Ae^{-i\phi} & E - p\mathcal{E} \end{pmatrix},$$

where p is interpreted as the molecular dipole moment. Find the eigenvalues of $\hat{H}$. Check against the results (5.32) that apply when $\mathcal{E} = 0$.

5.6 **Project: Ammonia molecule in an electric field**

5.6 Formal properties of time evolution; Conservation laws

When not subject to "observation", quantal states evolve according to the Schrödinger time-evolution equation

$$\frac{d|\psi(t)\rangle}{dt} = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle. \tag{5.42}$$

The Hamiltonian operator $\hat{H}$ is Hermitian, with eigenvectors $\{|e_n\rangle\}$ and eigenvalues e_n:

$$\hat{H}|e_n\rangle = e_n|e_n\rangle. \tag{5.43}$$

These are called the "energy eigenstates" or "states of definite energy" or the "stationary states".

Theorem I: Formal solution of the Schrödinger equation.

$$\text{If } |\psi(0)\rangle = \sum_n \psi_n(0)|e_n\rangle, \text{ then } |\psi(t)\rangle = \sum_n \psi_n(0)e^{-(i/\hbar)e_n t}|e_n\rangle. \tag{5.44}$$

Proof: In component form, the Schrödinger equation is

$$\frac{d\psi_n(t)}{dt} = -\frac{i}{\hbar}\sum_m H_{n,m}\psi_m(t).$$

In the energy eigenbasis,

$$H_{n,m} = \left\{ \begin{array}{cc} e_n & n = m \\ 0 & n \neq m \end{array} \right\} = e_n \delta_{n,m}.$$

Thus

$$\frac{d\psi_n(t)}{dt} = -\frac{i}{\hbar}\sum_m e_n \delta_{n,m}\psi_m(t) = -\frac{i}{\hbar}e_n\psi_n(t)$$

and

$$\psi_n(t) = \psi_n(0)e^{-(i/\hbar)e_n t}. \qquad \blacksquare$$

So, this is how states change with time! But we can't measure states. How do things that we *can* observe change with time? We will first find how mean values change with time, then look at "the whole shebang" – not just the mean, but the full distribution.

Theorem II: Time evolution of means.

$$\frac{d\langle \hat{A}\rangle}{dt} = -\frac{i}{\hbar}\langle[\hat{A},\hat{H}]\rangle. \tag{5.45}$$

Proof: (Using mathematical notation for inner products.)

$$\begin{aligned}
\frac{d}{dt}\langle \hat{A}\rangle &= \frac{d}{dt}\Big(\psi(t), \hat{A}\psi(t)\Big) \\
&= \left(\frac{d\psi(t)}{dt}, \hat{A}\psi(t)\right) + \left(\psi(t), \hat{A}\frac{d\psi(t)}{dt}\right) \\
&= \left(-\frac{i}{\hbar}\hat{H}\psi(t), \hat{A}\psi(t)\right) + \left(\psi(t), \hat{A}\left[-\frac{i}{\hbar}\hat{H}\psi(t)\right]\right) \\
&\qquad [\![\text{use the fact that } \hat{H} \text{ is Hermitian}]\!] \\
&= \frac{i}{\hbar}\Big(\psi(t), \hat{H}\hat{A}\psi(t)\Big) - \frac{i}{\hbar}\Big(\psi(t), \hat{A}\hat{H}\psi(t)\Big) \\
&= -\frac{i}{\hbar}\Big(\psi(t), [\hat{A}\hat{H} - \hat{H}\hat{A}]\psi(t)\Big) \\
&= -\frac{i}{\hbar}\langle[\hat{A},\hat{H}]\rangle \qquad \blacksquare
\end{aligned}$$

Corollary: If $\hat{A}$ commutes with $\hat{H}$, then $\langle \hat{A} \rangle$ is constant.

However, just because the mean of a measurement doesn't change with time doesn't necessarily mean that *nothing* about the measurement changes with time. To fully specify the results of a measurement, you must also list the possible results, the eigenvalues a_n, and the probability of getting that result, namely $|\langle a_n|\psi(t)\rangle|^2$. The eigenvalues a_n are time constant, but how do the probabilities change with time?

Theorem III: Time evolution of projection probabilities.
If $|\phi\rangle$ is a time-independent state and $\hat{P}_\phi = |\phi\rangle\langle\phi|$ is its associated outer product, then

$$\frac{d}{dt}|\langle\phi|\psi(t)\rangle|^2 = -\frac{i}{\hbar}\langle[\hat{P}_\phi, \hat{H}]\rangle. \tag{5.46}$$

Proof:

$$\begin{aligned}\frac{d}{dt}|\langle\phi|\psi(t)\rangle|^2 &= \frac{d}{dt}\left[\langle\phi|\psi(t)\rangle\langle\phi|\psi(t)\rangle^*\right]\\ &= \left[\langle\phi|\frac{d}{dt}|\psi(t)\rangle\right]\langle\phi|\psi(t)\rangle^* + \langle\phi|\psi(t)\rangle\left[\langle\phi|\frac{d}{dt}|\psi(t)\rangle\right]^*\end{aligned}$$

But $\langle\phi|\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}\langle\phi|\hat{H}|\psi(t)\rangle$, so

$$\begin{aligned}\frac{d}{dt}|\langle\phi|\psi(t)\rangle|^2 &= -\frac{i}{\hbar}\left[\langle\phi|\hat{H}|\psi(t)\rangle\langle\phi|\psi(t)\rangle^* - \langle\phi|\psi(t)\rangle\langle\phi|\hat{H}|\psi(t)\rangle^*\right]\\ &= -\frac{i}{\hbar}\left[\langle\psi(t)|\phi\rangle\langle\phi|\hat{H}|\psi(t)\rangle - \langle\psi(t)|\hat{H}|\phi\rangle\langle\phi|\psi(t)\rangle\right]\\ &= -\frac{i}{\hbar}\left[\langle\psi(t)|\left\{|\phi\rangle\langle\phi|\hat{H} - \hat{H}|\phi\rangle\langle\phi|\right\}|\psi(t)\rangle\right]\\ &= -\frac{i}{\hbar}\langle\psi(t)|[\hat{P}_\phi, \hat{H}]|\psi(t)\rangle \qquad \blacksquare\end{aligned}$$

Lemma: Suppose $\hat{A}$ and $\hat{B}$ are commuting Hermitian operators. If $|a\rangle$ is an eigenvector of $\hat{A}$ and $\hat{P}_a = |a\rangle\langle a|$, then $[\hat{P}_a, \hat{B}] = 0$.

Proof of lemma: From the compatibility theorem (page 131), there is an eigenbasis $\{|b_n\rangle\}$ of $\hat{B}$ with $|b_1\rangle = |a\rangle$. Write $\hat{B}$ in diagonal form as

$$\hat{B} = \sum_n b_n|b_n\rangle\langle b_n|.$$

Then

$$\hat{B}|b_1\rangle\langle b_1| = \sum_n b_n|b_n\rangle\langle b_n|b_1\rangle\langle b_1| = \sum_n b_n|b_n\rangle\delta_{n,1}\langle b_1| = b_1|b_1\rangle\langle b_1|$$

while

$$|b_1\rangle\langle b_1|\hat{B} = \sum_n |b_1\rangle\langle b_1|b_n\rangle\langle b_n|b_n = \sum_n |b_1\rangle\delta_{1,n}\langle b_n|b_n = b_1|b_1\rangle\langle b_1|. \qquad \blacksquare$$

> *Corollary:* If $\hat{A}$ commutes with $\hat{H}$, then nothing about the measurement of $\hat{A}$ changes with time.
>
> *Definition:* The observable associated with such an operator is said to be "conserved".

Note that all these results apply to time evolution uninterrupted by measurements.

5.7 The neutral K meson

You know that elementary particles are characterized by their mass and charge, but that two particles of identical mass and charge can still behave differently. Physicists have invented characteristics such as "strangeness" and "charm" to label (not explain!) these differences. For example, the difference between the electrically neutral K meson K^0 and its antiparticle the $\bar{K}^0$ is described by attributing a strangeness of +1 to the K^0 and of −1 to the $\bar{K}^0$.

Most elementary particles are completely distinct from their antiparticles: an electron never turns into a positron! Such a change is prohibited by charge conservation. However this prohibition does not extend to the neutral K meson precisely because it is neutral. In fact, there is a time-dependent amplitude for a K^0 to turn into a $\bar{K}^0$. We say that the K^0 and the $\bar{K}^0$ are the two basis states for a two-state system. This two-state system has an observable strangeness, represented by an operator, and we have a K^0 when the system is in an eigenstate of strangeness with eigenvalue +1, and a $\bar{K}^0$ when the system is in an eigenstate of strangeness with eigenvalue −1. When the system is in other states it does not have a definite value of strangeness, and cannot be said to be "a K^0" or "a $\bar{K}^0$". The two strangeness eigenstates are denoted $|K^0\rangle$ and $|\bar{K}^0\rangle$.

5.7 **Strangeness**

Write an outer product expression for the strangeness operator $\hat{S}$, and find its matrix representation in the $\{|K^0\rangle,\ |\bar{K}^0\rangle\}$ basis. Note that this matrix is just the Pauli matrix σ_3.

5.8 **Charge Parity**

Define an operator $\widehat{CP}$ that turns one strangeness eigenstate into the other:

$$\widehat{CP}|K^0\rangle = |\bar{K}^0\rangle, \quad \widehat{CP}|\bar{K}^0\rangle = |K^0\rangle.$$

(CP stands for "charge parity", although that's not important here.) Write an outer product expression and a matrix representation (in the $\{|K^0\rangle,\ |\bar{K}^0\rangle\}$ basis) for the $\widehat{CP}$ operator. What is the connection between this matrix and the Pauli matrices? Show that the normalized eigenstates of CP are

$$\begin{aligned} |K_U\rangle &= \frac{1}{\sqrt{2}}\,(|K^0\rangle + |\bar{K}^0\rangle), \\ |K_S\rangle &= \frac{1}{\sqrt{2}}\,(|K^0\rangle - |\bar{K}^0\rangle). \end{aligned}$$

(The U and S stand for unstable and stable, but that's again irrelevant because we'll ignore K meson decay.)

5.9 **The Hamiltonian**

The time evolution of a neutral K meson is governed by the "weak interaction" Hamiltonian

$$\hat{H} = e\hat{1} + f\widehat{CP}.$$

(There is no way for you to derive this. I'm just telling you.) Show that the numbers e and f must be real.

5.10 **Time evolution**

Neutral K mesons are produced in states of definite strangeness because they are produced by the "strong interaction" Hamiltonian that conserves strangeness. Suppose one is produced at time $t = 0$ in state $|K^0\rangle$. Solve the Schrödinger equation to find its state for all time afterwards. Why is it easier to solve this problem using $|K_U\rangle$, $|K_S\rangle$ vectors rather than $|K^0\rangle$, $|\bar{K}^0\rangle$ vectors? Calculate and plot the probability of finding the meson in state $|K^0\rangle$ as a function of time.

⟦The neutral K meson system is extraordinarily interesting. I have oversimplified by ignoring decay. More complete treatments can be found in

Ashok Das & Adrian Melissinos, *Quantum Mechanics* (Gordon and Breach, New York, 1986) pages 172–173; R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, volume III (Addison-Wesley, Reading, Massachusetts, 1965) pages 11-12–20; Gordon Baym, *Lectures on Quantum Mechanics* (W.A. Benjamin, Reading, Massachusetts, 1969), pages 38–45; and Harry J. Lipkin, *Quantum Mechanics: New Approaches to Selected Topics* (North-Holland, Amsterdam, 1986) chapter 7.⟧

Problems

5.11 **The most general two-state Hamiltonian**

We've seen a number of two-state systems by now: the spin states of a spin-$\frac{1}{2}$ atom, the polarization states of a photon, the CP states of a neutral K-meson. ⟦For more two-state systems, see R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, volume III (Addison-Wesley, Reading, Massachusetts, 1965) chapters 9, 10, and 11.⟧ This problem investigates the most general possible Hamiltonian for any two-state system.

Because the Hamiltonian must be Hermitian, it must be represented by a matrix of the form

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

where a and b are real, but $c = |c|e^{i\gamma}$ might be complex. Thus the Hamiltonian is specified through four real numbers: a, b, magnitude $|c|$, and phase γ. This seems at first glance to be the most general Hamiltonian.

But remember that states can be modified by an arbitrary overall phase. If the initial basis is $\{|1\rangle, |2\rangle\}$, show that in the new basis $\{|1\rangle, |2'\rangle\}$, where $|2'\rangle = e^{-i\gamma}|2\rangle$, the Hamiltonian is represented by the matrix

$$\begin{pmatrix} a & |c| \\ |c| & b \end{pmatrix}$$

which is pure real and which is specified through only three real numbers.

5.12 **Questions** (recommended problem)

Update your list of quantum mechanics questions that you started at problem 1.13 on page 56. Write down new questions and, if you have uncovered answers to any of your old questions, write them down briefly.

Chapter 6

The Quantum Mechanics of Position

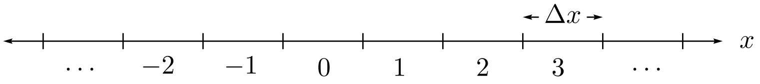
6.1 One particle in one dimension

Very early in this book (on page 8) we said we'd begin by treating only the magnetic moment of the atom quantum mechanically, and that once we got some grounding on the physical concepts and mathematical tools of quantum mechanics in this situation, we'd move on to the quantal treatment of other properties of the atom — such as its position, its momentum, and its energy. This was a very good thing that allowed us to uncover the phenomena of quantum mechanics — quantization, interference, and entanglement — to develop mathematical tools that describe those phenomena, to investigate time evolution, and to work on practical devices like atomic clocks, MASERs, and cryptosystems.

All good things must come to an end, but in this case we're ending one good thing to come onto an even better thing, namely the quantum mechanics of a continuous system. The system we'll pick first is a particle in one dimension. For the time being we'll ignore the atom's magnetic moment and internal constitution, and focus only on its position. Later in the book ⟦put in specific reference⟧ we'll treat both position and magnetic moment together.

Course-grained description

A single point particle ambivates in one dimension. We start off with a course-grained description of the particle's position: we divide the line into an infinite number of bins, each of width Δx. (We will later take the limit as the bin width vanishes and the number of bins grows to compensate.)



If we ask "In which bin is the particle positioned?" the answer might be "It's not in any of them. The particle doesn't have a position." Not all states have definite positions. On the other hand, there are *some* states that *do* have definite positions. If the particle has a position within bin 5 then we say that it is in state $|5\rangle$.

The set of states $\{|n\rangle\}$ with $n = 0, \pm 1, \pm 2, \pm 3, \ldots$ constitutes a basis, because the set is:

- Orthonormal. If the particle is in one bin, then it's not in any of the others. The mathematical expression of this property is
$$\langle n|m\rangle = \delta_{n,m}. \tag{6.1}$$
- Complete. If the particle does have a position, then it has a position within one of the bins. The mathematical expression of this property is
$$\sum_{n=-\infty}^{\infty} |n\rangle\langle n| = \hat{1}. \tag{6.2}$$

If the particle has no position, then its state $|\psi\rangle$ is a superposition of basis states
$$|\psi\rangle = \sum_{n=-\infty}^{\infty} \psi_n |n\rangle \tag{6.3}$$
where
$$\psi_n = \langle n|\psi\rangle \quad \text{so} \quad \sum_{n=-\infty}^{\infty} |\psi_n|^2 = 1. \tag{6.4}$$

The quantity $|\psi_5|^2$ is the probability that, if the position of the particle is measured (perhaps by shining a light down the one-dimensional axis), the particle will be found within bin 5. We should always say

"$|\psi_5|^2$ is the probability of finding the particle in bin 5", because the word “finding” suggests the whole story: Right now the particle has no position, but after you measure the position then it will have a position, and the probability that this position falls within bin 5 is $|\psi_5|^2$. This phrase is totally accurate but it’s a real mouthful. Instead one frequently hears

“$|\psi_5|^2$ is the probability that the particle is in bin 5”.

This is technically wrong. Before the position measurement, when the particle is in state $|\psi\rangle$, the particle *doesn’t have* a position. It has no probability of being in bin 5, or bin 6, or any other bin, just as love doesn’t have probability 0.5 of being red, 0.3 of being green, and 0.2 of being blue. Love doesn’t have a color, and the particle in state $|\psi\rangle$ doesn’t have a position.

Because the second, inaccurate, phrase is shorter than the first, correct, phrase, it is often used despite its falseness. You may use it too, as long as you don’t believe it.

Similarly, the most accurate statement is

“ψ_5 is the amplitude for finding the particle in bin 5”,

but you will frequently hear the brief and inaccurate

“ψ_5 is the amplitude that the particle is in bin 5”

instead.

Successively finer-grained descriptions

Suppose we want a more accurate description of the particle’s position properties. We can get it using a smaller value for the bin width Δx. Still more accurate descriptions come from using still smaller values of Δx. Ultimately we can produce a sequence of ever smaller bins homing in on the position of interest, say x_0. For all values of Δx, I will call the bin straddling x_0 by the name “bin k”. The relevant question seems at first to be: “What is the limit

$$\lim_{\Delta x \to 0} |\psi_k|^2 \text{ ?”}$$

In fact, this is not an interesting question. The answer is "zero". For example: Suppose you are presented with a narrow strip of lawn, 1000 meters long, which contains seven four-leaf clovers, scattered over the lawn at random. The probability of finding a four-leaf clover within a 2-meter wide bin is

$$\frac{7}{1000\text{ m}}(2\text{ m}) = 0.014.$$

The probability of finding a four-leaf clover within a 1-meter wide bin is

$$\frac{7}{1000\text{ m}}(1\text{ m}) = 0.007.$$

The probability of finding a four-leaf clover within a 1-millimeter wide bin is

$$\frac{7}{1000\text{ m}}(0.001\text{ m}) = 0.000007.$$

As the bin width goes to zero, the probability goes to zero as well.

As with clover, so with quantal probability. The interesting question concerns not the bin probability, which always goes to zero as the bins shrink to zero, but the *probability density*, that is, the probability of finding the particle per length.

Exercise 6.A. What is the probability density (including units) for finding a four-leaf clover in the strip of lawn described?

The probability per length of finding the particle at x_0, called the probability density at x_0, is the finite quantity

$$\lim_{\Delta x \to 0} \frac{|\psi_k|^2}{\Delta x}. \tag{6.5}$$

(Remember that the limit goes through a sequence of bins k, every one of which straddles the target point x_0.) In this expression both the numerator and denominator go to zero, but they approach zero in such a way that the ratio is finite. In other words, for small values of Δx, we have

$$|\psi_k|^2 \approx (\text{constant})\Delta x, \tag{6.6}$$

where that constant is the probability density for finding the particle at point x_0.

We need to understand both bin probabilities and bin amplitudes. Probabilities give the results for measurement experiments, but amplitudes give the results for both interference and measurement experiments. What does equation (6.6) say about bin amplitudes? It says that for small values of Δx

$$\psi_k \approx (\text{constant})' \sqrt{\Delta x} \tag{6.7}$$

whence the limit

$$\lim_{\Delta x \to 0} \frac{\psi_k}{\sqrt{\Delta x}}$$

exists. This limit defines the quantity, a function of x_0,

$$\lim_{\Delta x \to 0} \frac{\psi_k}{\sqrt{\Delta x}} = \psi(x_0). \tag{6.8}$$

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

The wavefunction evaluated at x_0 is sometimes called "the amplitude for the particle to have position x_0", but that's not exactly correct, because an amplitude squared is a probability whereas a wavefunction squared is a probability density. Instead this phrase is just shorthand for the more accurate phrase "$\psi(x_0)\sqrt{\Delta x}$ is the amplitude for finding the particle in an interval of short length Δx straddling position x_0, when the position is measured".

Working with wavefunction

When we were working with discrete systems, we said that the inner product could be calculated through

$$\langle \phi | \psi \rangle = \sum_n \phi_n^* \psi_n.$$

How does this pull over into continuous systems?

For any particular stage in the sequence of ever-smaller bins, the inner product is calculated through

$$\langle \phi | \psi \rangle = \sum_{i=-\infty}^{\infty} \phi_i^* \psi_i.$$

Prepare to take the limit $\Delta x \to 0$ by writing

$$\langle \phi|\psi\rangle = \sum_{i=-\infty}^{\infty} \frac{\phi_i^*}{\sqrt{\Delta x}} \frac{\psi_i}{\sqrt{\Delta x}} \Delta x.$$

Then

$$\langle \phi|\psi\rangle = \lim_{\Delta x \to 0} \sum_{i=-\infty}^{\infty} \frac{\phi_i^*}{\sqrt{\Delta x}} \frac{\psi_i}{\sqrt{\Delta x}} \Delta x = \int_{-\infty}^{+\infty} \phi^*(x)\psi(x)\, dx.$$

Exercise 6.B. What is the normalization condition for a wavefunction?

Basis states

When we went through the process of looking at finer and finer course-grainings, that is, taking $\Delta x \to 0$ and letting the number of bins increase correspondingly, we were not changing the physical state of the particle. Instead, we were just obtaining more and more accurate descriptions of that state. How? By using a larger and larger[1] basis! The sequence of intervals implies a sequence of basis states $|k\rangle$. What is the limit of that sequence?

One way to approach this question is to look at the sequence

$$\lim_{\Delta x \to 0} \psi_k = \lim_{\Delta x \to 0} \langle k|\psi\rangle = \left[\lim_{\Delta x \to 0} \langle k|\right] |\psi\rangle. \tag{6.9}$$

(Where, in the last step, we have acknowledged that in the sequence of finer-grained approximations involves changing the basis states $|k\rangle$, not the state of the particle $|\psi\rangle$.) This approach is not helpful because the limit always vanishes.

More useful is to look at the sequence

$$\lim_{\Delta x \to 0} \frac{\psi_k}{\sqrt{\Delta x}} = \lim_{\Delta x \to 0} \frac{\langle k|\psi\rangle}{\sqrt{\Delta x}} = \left[\lim_{\Delta x \to 0} \frac{\langle k|}{\sqrt{\Delta x}}\right] |\psi\rangle = \psi(x_0). \tag{6.10}$$

This sequence motivates the definition of the "position basis state"

$$|x_0\rangle = \lim_{\Delta x \to 0} \frac{|k\rangle}{\sqrt{\Delta x}}. \tag{6.11}$$

---

[1] You might object that the basis was not really getting bigger — it started out with an infinite number of bins and at each stage in the process always has an infinite number of bins. I will reply that in some sense it has a "larger infinity" than it started with. If you want to make this sense rigorous and precise, take a mathematics course on transfinite numbers.

This new entity $|x_0\rangle$ is not quite the same thing as the basis states like $|k\rangle$ that we've seen up to now, just as $\psi(x_0)$ is not quite the same thing as an amplitude. For example, $|k\rangle$ is dimensionless while $|x_0\rangle$ has the dimensions of $1/\sqrt{[\text{length}]}$. Mathematicians call the entity $|x_0\rangle$ not a "basis state" but a "rigged basis state". The word "rigged" carries the nautical connotation — a rigged ship is one outfitted for sailing and ready to move into action — and not the unsavory connotation — a rigged election is an unfair one. These are fascinating mathematical questions[2] but this is not a mathematics book, so we won't make a big fuss over the distinction.

Completeness relation for continuous basis states:

$$\hat{1} = \sum_{i=-\infty}^{\infty} |i\rangle\langle i| = \lim_{\Delta x \to 0} \sum_{i=-\infty}^{\infty} \frac{|i\rangle}{\sqrt{\Delta x}} \frac{\langle i|}{\sqrt{\Delta x}} \Delta x = \int_{-\infty}^{+\infty} |x\rangle\langle x|\, dx. \tag{6.12}$$

Orthogonality relation for continuous basis states:

$$\begin{aligned} \langle i|j\rangle &= \delta_{i,j} \\ \langle x|y\rangle &= 0 \qquad \text{when } x \neq y \\ \langle x|x\rangle &= \lim_{\Delta x \to 0} \frac{\langle i|i\rangle}{\Delta x} = \lim_{\Delta x \to 0} \frac{1}{\Delta x} = \infty \\ \langle x|y\rangle &= \delta(x-y). \end{aligned} \tag{6.13}$$

Just as the wavefunction is related to an amplitude but is not a true amplitude, and a rigged basis state $|x\rangle$ is related to a basis state but is not a true basis state, so the inner product result $\delta(x-y)$, the Dirac delta function, is related to a function but is not a true function. Mathematicians call it a a "Schwartz distribution". The Dirac delta function is discussed in Appendix B.

---

[2] See Rafael de la Madrid, "The role of the rigged Hilbert space in quantum mechanics" *European Journal of Physics* **26** (2005) 287–312.

Comparison of discrete and continuous basis states

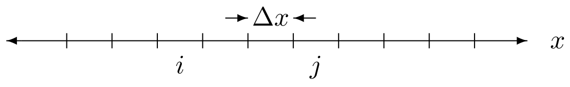
Discrete	**Continuous**
basis states $\vert n\rangle$; dimensionless	basis states $\vert x\rangle$; dimensions $\frac{1}{\sqrt{\text{length}}}$
$\psi_n = \langle n \vert \psi\rangle$	$\psi(x) = \langle x \vert \psi\rangle$
ψ_n is dimensionless	$\psi(x)$ has dimensions $\frac{1}{\sqrt{\text{length}}}$
$\sum_n \lvert\psi_n\rvert^2 = 1$	$\int_{-\infty}^{+\infty} \lvert\psi(x)\rvert^2 \, dx = 1$
$\langle n \vert m\rangle = \delta_{n,m}$	$\langle x \vert y\rangle = \delta(x-y)$
$\langle \phi \vert \psi\rangle = \sum_n \phi_n^* \psi_n$	$\langle \phi \vert \psi\rangle = \int_{-\infty}^{+\infty} \phi^*(x)\psi(x)\, dx$
$\sum_n \vert n\rangle\langle n\vert = \hat{1}$	$\int_{-\infty}^{+\infty} \vert x\rangle\langle x\vert \, dx = \hat{1}$

Exercise 6.C. Show that $\langle\phi|\psi\rangle = \int_{-\infty}^{+\infty} \phi^*(x)\psi(x)\,dx$ using the relation $\langle\phi|\psi\rangle = \langle\phi|\hat{1}|\psi\rangle$.

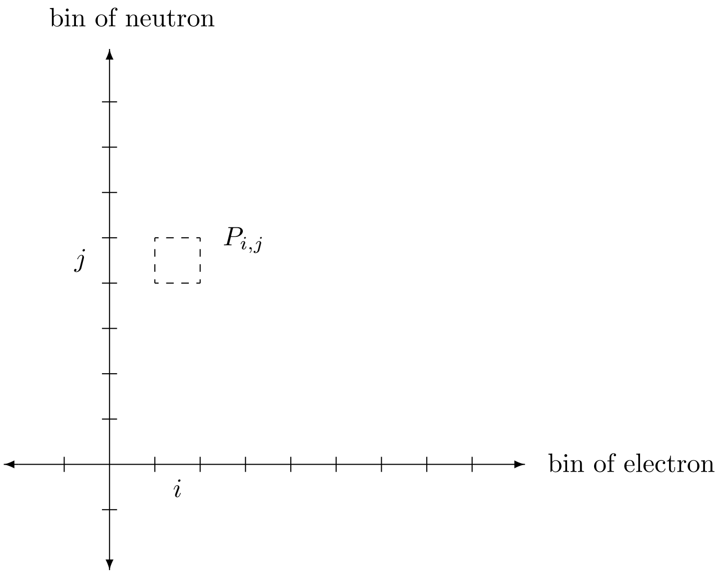
6.2 Two particles in one or three dimensions

Having discussed one particle in one dimension, we ask about two particles in one dimension.

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



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



To produce a probability density, we must divide the bin probability $P_{i,j}$ by $(\Delta x)^2$ (the area of the box above), and then take the limit as $\Delta x \to 0$, resulting in

$$\frac{P_{i,j}}{(\Delta x)^2} \to \rho(x_e, x_n).$$

So the probability of finding an electron within a narrow window of width w centered on $x_e = 5$ and finding the neutron within a narrow window of width u centered on $x_n = 9$ is approximately $\rho(5, 9)wu$, and this approximation grows better and better as the two windows grow narrower and narrower.

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

$$\lim_{\Delta x \to 0} \frac{\psi_{i,j}}{\Delta x} = \psi(x_e, x_n). \tag{6.14}$$

This wavefunction has dimensions 1/[length].

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

First, the wavefunction (like amplitude in general) is a mathematical tool for calculating the results of experiments; it is not physically "real". I have mentioned this before, but it particularly stands out here. Even for a system as simple as two particles, the wavefunction does not exist in ordinary three-dimensional space, but in a six-dimensional space. (You might recall from a classical mechanics course that this space is called "configuration space".) I don't care how clever or talented an experimentalist you are: you cannot insert an instrument into six-dimensional space in order to measure wavefunction.[3]

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

Third, it might happen that the wavefunction factorizes:

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

In this case the electron has state $\psi_e(\vec{x}_e)$ and the neutron has state $\psi_n(\vec{x}_n)$. Such a peculiar case is called "non-entangled". But in all other cases the

---

[3] If you are familiar with the Coulomb gauge in electrodynamics, you might find it enlightening to compare wavefunction in quantum mechanics to scalar and vector potentials in electrodynamics. In the Coulomb gauge,the scalar and vector potentials at a "field point" change instantly when a charge is moved at a "source point", even if the two points are light years apart. But they change in such a way that the electromagnetic field at the field point does *not* change until some time interval later, during which interval the field effects propagate at finite speed c from source point to field point. The field is measurable, the potentials are not. (Pigeons have "magnetoreception" — the ability to detect magnetic field; bumblebees and some fishes have "electroreception" — the ability to detect electric field; but no organism has the ability to detect scalar or vector potential.) It is all right for potentials to change instantaneously, because potentials are abstract mathematical tools, not measurable or detectable quantities. Similarly wavefunction can change instantaneously, because it can't be measured or detected: it is an abstract mathematical tool.

state is called "entangled" and the individual particles making up the system *do not have states.* The system has a state, namely $\psi(\vec{x}_e, \vec{x}_n)$, but there is no state for the electron and no state for the neutron, in exactly the same sense that there is no position for a silver atom ambivating through an interferometer.

Fourth, quantum mechanics is intricate. To understand this point, contrast the description needed in classical versus quantum mechanics.

How does one describe the state of a single classical particle moving in one dimension? It requires two numbers: a position and a velocity. Two particles moving in one dimension require merely that we specify the state of each particle: four numbers. Similarly specifying the state of three particles require six numbers and N particles require $2N$ numbers. Exactly the same specification counts hold if the particle moves relativistically.

particles	real numbers needed to specify classical state
1	2
2	4
3	6
$\vdots$	$\vdots$
N	$2N$

How, in contrast, does one describe the state of a single quantal particle ambivating in one dimension? Here an issue arises at the very start, because the specification is given through a complex-valued wavefunction $\psi(x)$. Technically the specification requires an infinite number of numbers! Let's approximate the wavefunction through its value on a grid of, say, 100 points. This suggests that a specification requires 200 real numbers, a complex number at each grid point, but global phase freedom means that we can always set one of those numbers to zero through an overall phase factor, and one number is not independent through the normalization requirement. The specification actually requires 198 independent real numbers.

How does one describe the state of two quantal particles ambivating in one dimension? Now the wavefunction is a function of two variables, $\psi(x_e, x_n)$. The wavefunction of the system is a function of two-dimensional configuration space, so an approximation of the accuracy established previously requires a 100×100 grid of points. Each grid point carries one complex number, and again overall phase and normalization reduce the number of real numbers required by two. For two particles the specification requires $2 \times (100)^2 - 2 = 19\,998$ independent real numbers. To specify the two-particle states, we cannot get away with just specifying two one-particle states. Just as a particle might not have a position, so in a two-particle system an individual particle might not have a state.

Similarly, specifying the state of N quantal particles moving in one dimension requires a wavefunction in N-dimensional configuration space which (for a grid of the accuracy we've been using) is specified through $2 \times (100)^N - 2$ independent real numbers.

particles	real numbers needed to specify quantal state
1	$2(100) - 2 = 198$
2	$2(100)^2 - 2 = 19\,998$
3	$2(100)^3 - 2 = 1\,999\,998$
$\vdots$	$\vdots$
N	$2(100)^N - 2$

The specification of a quantal state not only requires more real numbers than the specification of the corresponding classical state, but that number increases exponentially rather than linearly with particle number N.

The fact that a quantal state holds more information than a classical state is the fundamental reason that a quantal computer can be (in principle) faster than a classical computer, and the basis for much of quantum information theory.

Relativity is different from classical physics, but no more complicated. Quantum mechanics, in contrast, is both *different* from and *richer* than classical physics. You may refer to this richness using terms like "splendor", or "abounding", or "intricate", or "ripe with possibilities". Or you may refer to it using terms like "complicated", or "messy", or "full of details likely to trip the innocent". It's your choice how to react to this richness, but you can't deny it.

Problem

6.1 **Properties of two-particle basis states**
Make a table like the one on page 176 concerning the continuous basis states for the system consisting of one electron and one neutron ambivating in one dimension.

6.3 What is wavefunction?

We have introduced the tool of wavefunction (or “amplitude density”). Wavefunction is sort of like magnetic field in that you can’t touch it or taste it or smell it, but in fact is even more abstract. For one thing wavefunction is complex-valued, not real-valued. For another it is determined, to some extent, by convention. For a third it exists in configuration space.

This abstractness has gnawed at people from the very beginnings of quantum mechanics: In the summer of 1926, Erich Hückel[4] composed the ditty, presented here in the free translation by Felix Bloch[5]

Erwin with his ψ can do
Calculations quite a few.
But one thing has not been seen:
Just what does ψ really mean?

Rather than worry about what wavefunction is, I recommend that you avoid traps of what wavefunction is not. It *can’t* be measured. It *doesn’t* exist in physical space. It *is* dependent on convention. It is a mathematical tool like the scalar and vector potentials of electromagnetism. The wavefunction ψ is a step in an algorithm: it has no more physical significance than the carries and borrows of integer arithmetic (see page 136).

6.4 How does wavefunction change with time?

In classical mechanics, the equation telling how position changes with time is $\vec{F} = m\vec{a}$. It is not possible to derive $\vec{F} = m\vec{a}$, but it *is* possible to motive it.

This section is uncovers the quantal equivalent of $\vec{F} = m\vec{a}$: the equation telling how position amplitude changes with time. As with $\vec{F} = m\vec{a}$, it is possible to motivate this equation but not to prove it. As such, the

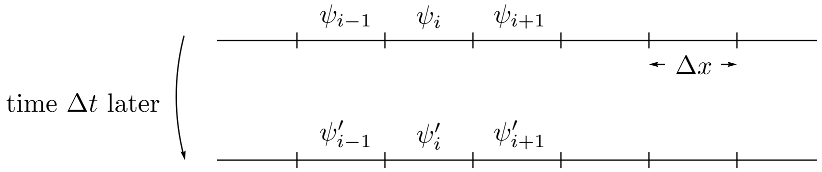
---

[4] Erich Hückel (1896–1980) was a German physicist whose work in molecular orbitals resulted in the first successful treatment of the carbon-carbon double bond.

[5] Felix Bloch (1905–1983) was a Jewish-Swiss-American physicist who made contribution to the quantum theory of solids and elsewhere. He won the Nobel Prize for his work in nuclear magnetic resonance. His memory of this poem comes from his “Reminiscences of Heisenberg and the early days of quantum mechanics” [*Physics Today* **29**(12) (December 1976) 23–27].

arguments in this section are suggestive, not definitive.[6] Indeed, in some circumstances (e.g. for a single charged particle in a magnetic field, or for a pair of entangled particles) the arguments are false.

The flow of amplitude



We begin with bin amplitudes evolving over a time step. By the end of the argument both the bin width Δx and the time step Δt will shrink to zero.

The amplitude for the particle to be within bin i is initially ψ_i, and after time Δt it changes to $\psi'_i = \psi_i + \Delta'\psi_i$. (In this section, change with time is denoted $\Delta'\psi$, while change with space is denoted $\Delta\psi$.)

Begin with the very reasonable surmise that

$$\psi'_i = A_i\psi_{i-1} + B_i\psi_i + C_i\psi_{i+1}. \tag{6.15}$$

This equation does nothing more than implement the rules for combining amplitude on page 60. It says than that the amplitude to be in bin i at the end of the time interval is the sum of

the amplitude to be in bin $i-1$ initially (ψ_{i-1}) times the amplitude to flow right (A_i)
plus
the amplitude to be in bin i initially (ψ_i) times the amplitude to stay in that bin (B_i)
plus
the amplitude to be in bin $i+1$ initially (ψ_{i+1}) times the amplitude to flow left (C_i).

---

[6] This section builds on R.P. Feynman, R.B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, volume 3: Quantum Mechanics (Addison-Wesley, Reading, Massachusetts, 1965) pages 16-1–16-4, and Gordon Baym, *Lectures on Quantum Mechanics* (Benjamin, Reading, Massachusetts, 1969) pages 46–53.

The key assumption we've made in writing down this surmise is that only adjacent bins are important: surely a reasonable assumption if the time interval Δt is short. (Some people like to call A_i and C_i "hopping amplitudes" rather than "flow amplitudes". And they call this bin picture the "Hubbard model".) From this "very reasonable surmise", plus a handful of ancillary assumptions, we will uncover the character of the amplitudes A_i, B_i, C_i, and motivate an equation (namely equation 6.26) governing the time evolution of wavefunction. The motivation arguments are long and technical, but please keep in mind that they do nothing more than elaborate these simple, familiar rules for combining amplitudes in series and in parallel.

The character of the change amplitudes

Note that the change amplitudes A_i, B_i, and C_i are *independent* of the position bin amplitudes ψ_{i-1}, ψ_i, and ψ_{i+1}. That is, A_i represents the amplitude to flow right regardless of what amplitude is originally in bin $i-1$. In other words, A_i, B_i, and C_i depend on the *situation* (e.g. the mass of the particle, the forces applied to the particle) but not on the *state*.

We surmise further that the flow amplitudes are independent of position and of direction, so all the A_i and C_i are independent of i, and equal to each other. This surmise seems at first to be silly: surely if the particle moves along a line containing a hill and a valley, the flow will be more likely downhill than uphill. However, this "surely" observation shows only that $A_i\psi_{i-1}$ will differ from $C_i\psi_{i+1}$, not that A_i will differ from C_i. We know that motion can happen even if there are no hills and valleys — that "a particle in motion remains in motion unless acted upon by an external force" — and the flow amplitudes concern this motion without external force. (The surmise that left flow amplitude equals right flow amplitude does, in fact, turn out to be false for a charged particle in a magnetic field.) On the other hand, the hill vs. valley argument means that B_i *will* depend on position.

Finally, realize that the amplitudes A and B_i *will* depend on Δx and Δt: we expect that the flow amplitude A will increase with increasing Δt (more time, more flow), and decrease with increasing Δx (with fat bins the flow at boundaries is less significant).

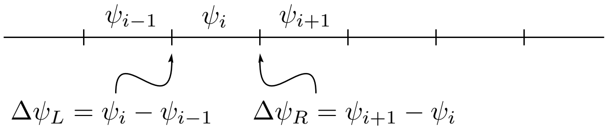
With these surmises in place, we have

$$\psi_i' = A\psi_{i-1} + B_i\psi_i + A\psi_{i+1}. \tag{6.16}$$

Now, I write B_i in a funny way as $B_i = -2A + 1 + D_i$. I do this so that the equation will turn into

$$\Delta'\psi_i = \psi'_i - \psi_i = A(\psi_{i-1} - \psi_i) + D_i\psi_i + A(\psi_{i+1} - \psi_i), \tag{6.17}$$

which emphasizes amplitude differences rather than amplitude totals. In terms of the differences sketched below



this equation is

$$\Delta'\psi_i = -A\Delta\psi_L + D_i\psi_i + A\Delta\psi_R. \tag{6.18}$$

Writing this way, in terms of differences, prepares for taking derivatives:

$$\Delta\psi_R - \Delta\psi_L = \Delta x\left(\frac{\Delta\psi_R}{\Delta x} - \frac{\Delta\psi_L}{\Delta x}\right).$$

The ratio $\Delta\psi_R/\Delta x$ clearly relates to a spatial derivative taken at the right boundary of bin i. Furthermore

$$\Delta\psi_R - \Delta\psi_L = (\Delta x)^2\left(\frac{\dfrac{\Delta\psi_R}{\Delta x} - \dfrac{\Delta\psi_L}{\Delta x}}{\Delta x}\right)$$

just as clearly relates to a second spatial derivative taken at the center of bin i.

At some point we need to switch over from talking about bin amplitude to talking about wavefunction, and this is a convenient point. Divide both sides of equation (6.18) by $\sqrt{\Delta x}$ and remember (equation 6.8) that, if x_i is the point at the center of bin i, then

$$\psi(x_i) = \lim_{\Delta x\to 0}\frac{\psi_i}{\sqrt{\Delta x}}. \tag{6.19}$$

Then, in an approximation that grows increasingly accurate as $\Delta x \to 0$,

$$\Delta'\psi(x_i) \approx A(\Delta x)^2\left(\frac{\partial^2\psi(x)}{\partial x^2}\right)_{x=x_i} + D_i\psi(x_i)$$

While this equation applies to the point at the center of bin i, of course it holds for any point. Defining $D(x_i) = D_i$ gives

$$\Delta'\psi(x) \approx A(\Delta x)^2\frac{\partial^2\psi}{\partial x^2} + D(x)\psi(x). \tag{6.20}$$

Normalization requirement

A technical requirement concerning normalization becomes useful soon. Because the probability that the particle is in *some* bin is one, the bin amplitudes are normalized to

$$\sum_i |\psi_i|^2 = 1$$

and

$$\sum_i |\psi_i'|^2 = 1.$$

The second equation can be written

$$\begin{aligned} 1 &= \sum_i \psi_i'^* \psi_i' \\ &= \sum_i (\psi_i^* + \Delta'\psi_i^*)(\psi_i + \Delta'\psi_i) \\ &= \sum_i (\psi_i^*\psi_i + \psi_i^*\Delta'\psi_i + \Delta'\psi_i^*\psi_i + \Delta'\psi_i^*\Delta'\psi_i). \end{aligned}$$

The first term on the last right-hand side sums to exactly 1, due to initial normalization. The next two terms are of the form $z + z^* = 2\,\Re e\{z\}$, so

$$0 = \sum_i 2\,\Re e\{\psi_i^*\Delta'\psi_i\} + \Delta'\psi_i^*\Delta'\psi_i.$$

When we go to the limit of very small Δt, then $\Delta'\psi_i$ will be very small, so $\Delta'\psi_i^*\Delta'\psi_i$, the product of two very small quantities, will be *ultra* small. Thus we neglect it and conclude that, due to normalization,

$$\Re e\left\{\sum_i \psi_i^*\Delta'\psi_i\right\} = 0. \tag{6.21}$$

We change over from bin amplitudes to wavefunction by observing that, for very small bins, this equation becomes

$$\Re e\left\{\sum_i \psi^*(x_i)\sqrt{\Delta x}\Delta'\psi(x_i)\sqrt{\Delta x}\right\} = 0$$

or

$$\Re e\left\{\int_{-\infty}^{+\infty} \psi^*(x)\Delta'\psi(x)\,dx\right\} = 0. \tag{6.22}$$

This is the desired “technical requirement concerning normalization”.

Applying (6.20) in (6.22) shows that

$$\begin{aligned}&\int_{-\infty}^{+\infty} \psi^*(x)\Delta'\psi(x)\,dx \qquad (6.23)\\ &= A(\Delta x)^2 \int_{-\infty}^{+\infty} \psi^*(x)\frac{\partial^2\psi}{\partial x^2}\,dx + \int_{-\infty}^{+\infty} \psi^*(x)D(x)\psi(x)\,dx\end{aligned}$$

must be pure imaginary. This requirement holds for *all wavefunctions* $\psi(x)$, and for *all situations* regardless of $D(x)$, so each of the two terms on the right must be pure imaginary. (We cannot count on a real part in first term on the right to cancel a real part in the second term on the right, because if they happened to cancel for one function $D(x)$, they *wouldn't* cancel for a different function $D(x)$. But the normalization condition has to hold for *all possible* functions $D(x)$.)

The first integral on the right-hand side of (6.23) can be performed by parts:

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

The part in square brackets vanishes... otherwise $\psi(x)$ is not normalized. The remaining integral is of the form

$$\int f^*(x)f(x)\,dx$$

which is pure real. Thus the constant A must be pure imaginary.

The second integral on the right-hand side of (6.23) is

$$\int_{-\infty}^{+\infty} \psi^*(x)D(x)\psi(x)\,dx$$

which must be imaginary. But $\psi^*(x)\psi(x)$ is pure real, so $D(x)$ must be pure imaginary.

Having discovered that the amplitudes A and $D(x)$ must be pure imaginary, we define the pure real quantities a and $d(x)$ through

$$A = ia \quad \text{and} \quad D(x) = id(x),$$

and the discrete-time amplitude equation (6.20) becomes

$$\Delta'\psi(x) \approx i\left[a(\Delta x)^2\frac{\partial^2\psi}{\partial x^2} + d(x)\psi(x)\right]. \qquad (6.24)$$

Dimensional analysis

Let's uncover more about the dimensionless quantity a. It's not plausible for the quantity a to depend on the phase of the moon, or the national debt. It can only depend on Δx, Δt, the particle mass m, and Planck's constant $\hbar$, from equation (1.2). (We've already pointed out that a involves flow, so it makes sense that a depends on the inertia of the particle m.)

quantity	dimensions
Δx	$[\ell]$
Δt	$[t]$
m	$[m]$
$\hbar$	$[m][\ell]^2/[t]$

The quantity $a(\Delta x)^2$ must be finite in the limit $\Delta x \to 0$, so a must depend on Δx through the proportionality

$$a \propto \frac{1}{(\Delta x)^2} \qquad \text{dimensions of right-hand side: } \frac{1}{[\ell]^2}.$$

To make a dimensionless we'll need to cancel the dimensions of length. The only way to do this is through $\hbar$:

$$a \propto \frac{\hbar}{(\Delta x)^2} \qquad \text{dimensions of right-hand side: } \frac{[m]}{[t]}.$$

Now we need to cancel out the dimensions of mass and time. Again there is only one way to do this:

$$a \propto \frac{\hbar}{(\Delta x)^2}\frac{\Delta t}{m} \qquad \text{dimensions of right-hand side: none.}$$

In short

$$a = \frac{\Delta t}{(\Delta x)^2}\frac{\hbar}{m}n_d$$

where n_d is a dimensionless real number. Note that, as anticipated immediately before equation (6.16), the quantity a increases with Δt and decreases with Δx.

With our new understanding we write equation (6.24) as

$$\Delta'\psi(x) \approx i\left[\frac{\hbar n_d}{m}\Delta t\frac{\partial^2\psi}{\partial x^2} + d(x)\psi(x)\right]$$

or

$$\frac{\Delta'\psi(x)}{\Delta t} \approx i\left[\frac{\hbar n_d}{m}\frac{\partial^2\psi}{\partial x^2} + \frac{d(x)}{\Delta t}\psi(x)\right]$$

which is conventionally written

$$\frac{\Delta'\psi(x)}{\Delta t} \approx -\frac{i}{\hbar}\left[-\frac{\hbar^2 n_d}{m}\frac{\partial^2\psi}{\partial x^2} - \frac{\hbar d(x)}{\Delta t}\psi(x)\right].$$

This conventional form has the advantage that the part in square brackets has the dimensions of energy times the dimensions of ψ.

The function $\hbar d(x)/\Delta t$ has the dimensions of energy, and we call it $v(x)$. Now taking the limit $\Delta t \to 0$ we find

$$\frac{\partial\psi(x,t)}{\partial t} = -\frac{i}{\hbar}\left[-\frac{\hbar^2 n_d}{m}\frac{\partial^2\psi(x,t)}{\partial x^2} - v(x)\psi(x,t)\right]. \qquad (6.25)$$

Exercise 6.D. Does it make physical sense that the "stay at home bin amplitude" D_i (see equation 6.17) should increase with increasing Δt?

Classical limit

To complete the specification of this equation, we must find values for n_d and $v(x)$. This can be done by applying the equation to a massive particle starting with a pretty-well defined position and seeing how that pretty-well defined position changes with time. In this so-called classical limit, the results of quantum mechanics must go over to match the results of classical mechanics. We are not yet equipped to do this, but we will find in section 6.9.4 (see problem 6.16) that enforcing the classical limit gives the result that $n_d = 1/2$ and $v(x)$ is the negative of the classical potential energy function $V(x)$.

This latter result astounds me. The classical potential energy function derives from considering a particle with a definite location. Why should it have anything to do with quantum mechanics? I don't know, but it surely does.

We will also see that the term

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

concerns kinetic energy, and sure enough we've been relating it to "flow" or "hopping". Again, I am astounded that the quantal expression corresponding to kinetic energy is so different from the classical expression, just as I am astounded that the quantal expression corresponding to potential energy is so similar to the classical expression. Again, it's true whether I find it astounding or not.

Conclusion

The wavefunction $\psi(x,t)$ evolves in time according to

$$\boxed{\frac{\partial \psi(x,t)}{\partial t} = -\frac{i}{\hbar}\left[-\frac{\hbar^2}{2m}\frac{\partial^2 \psi(x,t)}{\partial x^2} + V(x)\psi(x,t)\right]}, \qquad (6.26)$$

where $V(x)$ is the classical potential energy function. This equation was discovered in a completely different way by the 38-year-old Erwin Schrödinger during the Christmas season of 1925, at the alpine resort of Arosa, Switzerland, in the company of "an old girlfriend [from] Vienna", while his wife stayed at home in Zürich.[7] It is called the Schrödinger equation, and it plays the same central role in quantum mechanics that $\vec{F} = m\vec{a}$ plays in classical mechanics.

Do not think that we have derived the Schrödinger equation. . . instead we have taken it to pieces to see how it works. While the equation looks complicated and technical (two partial derivatives!), at heart it simply expresses the rules for combining amplitudes in series and in parallel (see equation 6.15), buttressed with some reasonable ancillary assumptions.

> ⟦What was the "completely different way" that Schrödinger used to come up with his equation? You know that there are many formulations of classical mechanics: Newtonian, Lagrangian, Hamiltonian, method of least action, etc. One of these is the Hamilton-Jacobi formulation, in which the time evolution of a classical system is analogous to the motion of light in ray optics. Just as ray optics is the short-wavelength limit of wave optics, so the Hamilton-Jacobi formulation of classical mechanics is the short-wavelength limit of quantum mechanics. Schrödinger started with this limit and, with the guidance of several experimental results, generalized the Hamilton-Jacobi formulation to work at all wavelengths. The result was the Schrödinger equation.[8]⟧

---

[7] Walter Moore, *Schrödinger: Life and Thought* (Cambridge University Press, 1989) page 194.

[8] See second paragraph of Erwin Schrödinger, "Quantisierung als Eigenwertproblem

Problem

6.2 **Schrödinger equation for wavefunction in polar form**
Write the wavefunction in polar form as

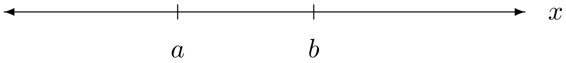
$$\psi(x,t) = R(x,t)e^{i\phi(x,t)}, \tag{6.27}$$

where the magnitude $R(x,t)$ and the phase $\phi(x,t)$ are pure real. Show that the Schrödinger equation is equivalent to the two real equations

$$\frac{\partial R}{\partial t} = -\frac{\hbar}{2m}\left[R\frac{\partial^2 \phi}{\partial x^2} + 2\frac{\partial R}{\partial x}\frac{\partial \phi}{\partial x}\right] \tag{6.28}$$

$$\frac{\partial \phi}{\partial t} = -\frac{1}{\hbar}\left\{-\frac{\hbar^2}{2m}\left[\frac{1}{R}\frac{\partial^2 R}{\partial x^2} - \left(\frac{\partial \phi}{\partial x}\right)^2\right] + V(x)\right\}. \tag{6.29}$$

6.5 How does probability change with time?



Before tackling this question, we introduce a parallel but more familiar situation. Water moves around in a long, narrow trough. It's not raining and the trough doesn't leak, so the amount of water is fixed ("conserved"). Think of a portion of the trough between positions a and b. The amount of water in this portion *does* change with time, because water can flow in or out at a and at b. In fact, if the current flowing toward the right at point x is called $j_w(x)$, then

$$\frac{d(\text{amount of water between } a \text{ and } b)}{dt} = j_w(a) - j_w(b). \tag{6.30}$$

Exercise 6.E. If the amount of water is measured in kilograms, what are the units of $j_w(x)$?

Now turn to the question of interest: The state of a particle ambivating in one dimension is represented by wavefunction $\psi(x,t)$. What is the probability that, when the particle's position is measured at time t, it is found between points a and b?

$$P_{a,b}(t) = \int_a^b |\psi|^2\,dx = \int_a^b \psi^*\psi\,dx \tag{6.31}$$

(Erste Mitteilung)", *Annalen der Physik* **79**, 361–376 (1926). Translated as "Quantization as a problem of proper values (part I)" in *Collected Papers on Wave Mechanics* (Chelsea Publishing Company, New York, 1978; reprint of the 1928 edition published by Blackie, London).

How does that probability change with time?

$$
\begin{aligned}
\frac{dP_{a,b}(t)}{dt} &= \int_a^b \left\{ \frac{\partial \psi^*}{\partial t}\psi + \psi^* \frac{\partial \psi}{\partial t} \right\} dx \\
&= \int_a^b \left\{ \left(\frac{i}{\hbar}\right) \left(-\frac{\hbar^2}{2m}\frac{\partial^2 \psi^*}{\partial x^2} + V(x)\psi^* \right) \psi \right. \\
&\qquad \left. + \psi^* \left(-\frac{i}{\hbar}\right) \left(-\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial x^2} + V(x)\psi \right) \right\} dx \\
&= -\frac{i}{\hbar}\left(-\frac{\hbar^2}{2m}\right) \int_a^b \left\{ -\frac{\partial^2 \psi^*}{\partial x^2}\psi + \psi^* \frac{\partial^2 \psi}{\partial x^2} \right\} dx \\
&= -\frac{i\hbar}{2m} \int_a^b \frac{\partial}{\partial x} \left\{ \frac{\partial \psi^*}{\partial x}\psi - \psi^* \frac{\partial \psi}{\partial x} \right\} dx \\
&= -\frac{i\hbar}{2m} \left\{ \left[\frac{\partial \psi^*}{\partial x}\psi - \psi^* \frac{\partial \psi}{\partial x} \right]_{x=b} - \left[\frac{\partial \psi^*}{\partial x}\psi - \psi^* \frac{\partial \psi}{\partial x} \right]_{x=a} \right\} \\
&= j(a,t) - j(b,t) \qquad (6.32)
\end{aligned}
$$

where we have defined the "probability current"

$$
j(x,t) = \frac{i\hbar}{2m} \left[\frac{\partial \psi^*}{\partial x}\psi - \psi^* \frac{\partial \psi}{\partial x} \right]. \qquad (6.33)
$$

Exercise 6.F. What are the dimensions of $P_{a,b}(t)$ and of $j(x,t)$?

Problem

6.3 **Other expressions for probability current**

Show that

$$
j(x,t) = \frac{\hbar}{m} \Im m \left\{ \psi^* \frac{\partial \psi}{\partial x} \right\} \qquad (6.34)
$$

and, for the polar form $\psi(x,t) = R(x,t)e^{i\phi(x,t)}$,

$$
j(x,t) = \frac{\hbar}{m} R^2 \frac{\partial \phi}{\partial x}. \qquad (6.35)
$$

6.4 **Equation of continuity**
Apply equation (6.32) in the limit that b moves very close to a to show that

$$\frac{\partial|\psi|^2}{\partial t} = -\frac{\partial j}{\partial x}. \tag{6.36}$$

This is called the "equation of continuity".

6.6 Operators and their representations

In abstract Hilbert space formulation, the Schrödinger equation for the time evolution of $|\psi(t)\rangle$ reads

$$\frac{d|\psi(t)\rangle}{dt} = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle. \tag{6.37}$$

In terms of wavefunction, the Schrödinger equation for the time evolution of $\psi(x,t) = \langle x|\psi(t)\rangle$ reads

$$\frac{\partial\psi(x,t)}{\partial t} = -\frac{i}{\hbar}\left[-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V(x)\psi(x,t)\right]. \tag{6.38}$$

How are these two equations related?

The position operator and functions of the position operator

The position operator is called $\hat{x}$. If we know the action of $\hat{x}$ on every member of the $\{|x\rangle\}$ basis (or any other basis!), then we know everything about the operator. But we do know that! If x_0 is some particular position,

$$\hat{x}|x_0\rangle = x_0|x_0\rangle.$$

Furthermore, we can find the action of $\hat{x}^2$ on every member of the $\{|x\rangle\}$ basis as follows:

$$\hat{x}^2|x_0\rangle = \hat{x}\left[\hat{x}|x_0\rangle\right] = \hat{x}\left[x_0|x_0\rangle\right] = x_0\left[\hat{x}|x_0\rangle\right] = x_0\left[x_0|x_0\rangle\right] = (x_0)^2|x_0\rangle.$$

Similarly, for any integer power n,

$$\hat{x}^n|x_0\rangle = (x_0)^n|x_0\rangle.$$

Exercise 6.G. Prove this using mathematical induction.

If $f(x)$ is a scalar function with Taylor series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n, \tag{6.39}$$

then we define the *operator* $f(\hat{x})$ through

$$f(\hat{x}) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} \hat{x}^n. \tag{6.40}$$

This enables us to find operators like $e^{c\hat{x}}$ corresponding to quantities like e^{cx}. The upshot is that for such operators, the position basis states are eigenstates:

$$f(\hat{x})|x_0\rangle = f(x_0)|x_0\rangle.$$

We've been examining the action of operators like $f(\hat{x})$ on position basis states. What if they act upon some other state? We find out by expanding the general state $|\psi\rangle$ into position states:

$$\begin{aligned}
f(\hat{x})|\psi\rangle &= f(\hat{x})\hat{1}|\psi\rangle \\
&= f(\hat{x}) \left[\int_{-\infty}^{+\infty} |x'\rangle\langle x'|\, dx' \right] |\psi\rangle \\
&= \int_{-\infty}^{+\infty} f(\hat{x})|x'\rangle\langle x'|\psi\rangle\, dx' \\
&= \int_{-\infty}^{+\infty} |x'\rangle f(x')\langle x'|\psi\rangle\, dx'.
\end{aligned}$$

To get a feel for this result, we look for the representation of the state $f(\hat{x})|\psi\rangle$ in the $\{|x\rangle\}$ basis:

$$\begin{aligned}
\langle x|f(\hat{x})|\psi\rangle &= \int_{-\infty}^{+\infty} \langle x|x'\rangle f(x')\langle x'|\psi\rangle\, dx' \\
&= \int_{-\infty}^{+\infty} \delta(x - x') f(x')\psi(x')\, dx' \\
&= f(x)\psi(x).
\end{aligned}$$

The representation of an operator $f(\hat{x})$ in the position basis is

$$\boxed{\langle x|f(\hat{x})|\psi\rangle = f(x)\langle x|\psi\rangle}. \tag{6.41}$$

And, as we've seen, if we know $\langle x|\hat{A}|\psi\rangle$ for general $|\psi\rangle$ and for general x, then we know everything there is to know about the operator.

So the relation between a function-of-position operator and its position basis representation is simple: erase the hats!

$$|\phi\rangle = f(\hat{x})|\psi\rangle \quad\Longleftrightarrow\quad \phi(x) = f(x)\psi(x). \tag{6.42}$$

Another application:

$$\begin{aligned}\langle\phi|f(\hat{x})|\psi\rangle &= \langle\phi|\hat{1}f(\hat{x})|\psi\rangle \\ &= \int_{-\infty}^{+\infty} dx\,\langle\phi|x\rangle\langle x|f(\hat{x})|\psi\rangle \\ &= \int_{-\infty}^{+\infty} \phi^*(x)f(x)\psi(x)\,dx. \end{aligned}\tag{6.43}$$

So you might think we know all we need to know. But no, because…

There are other operators

Momentum is a measurable so, according to our statement 2 on page 127, there must be a Hermitian operator associated with momentum. What is a sensible[9] definition of that operator?

As always, we know everything about an operator $\hat{A}$ if we know $\langle x|\hat{A}|\psi\rangle$ for all $|\psi\rangle$ and for every $|x\rangle$. Equations (6.37) and (6.38), put together, show that the Hamiltonian operator defined in this way is

$$\langle x|\hat{H}|\psi\rangle = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\langle x|\psi\rangle. \tag{6.44}$$

The sensible definition of the momentum operator is through

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}), \tag{6.45}$$

so

$$\langle x|\hat{p}^2|\psi\rangle = -\hbar^2\frac{\partial^2}{\partial x^2}\langle x|\psi\rangle. \tag{6.46}$$

---

[9] "Here and elsewhere in science, as stressed not least by Henri Poincaré, that view is out of date which used to say, 'Define your terms before you proceed.' All the laws and theories of physics, including the Lorentz force law $[\vec{F} = q\vec{E} + q\vec{v}\times\vec{B}]$, have this deep and subtle character, that they both define the concepts they use (here $\vec{E}$ and $\vec{B}$) and make statements about these concepts. Contrariwise, the absence of some body of theory, law, and principle deprives one of the means properly to define or even to use concepts. Any forward step in human knowledge is truly creative in this sense: that theory, concept, law, and method of measurement — forever inseparable — are born into the world in union." C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (W.H. Freeman and Company, San Francisco, 1973) page 71.

Exercise 3.P on page 110, "An operator squared", inspires us to define the momentum operator $\hat{p}$ similarly as

$$\boxed{\langle x|\hat{p}|\psi\rangle = -i\hbar\frac{\partial}{\partial x}\langle x|\psi\rangle}. \tag{6.47}$$

The operator with "$+i$" rather than "$-i$" out in front would have the same square, but would not have the correct classical limit. (See problems 6.5 and 6.15, and the sample problem below.)

Exercise 6.H. Would the phase-shifted convention

$$\langle x|\hat{p}|\psi\rangle = -i\hbar e^{i\delta}\frac{\partial}{\partial x}\langle x|\psi\rangle,$$

where δ is pure real, be acceptable?

6.6.1 *Sample Problem: Sign of the momentum operator.*

The function $\psi_R(x,t) = Ae^{i(+kx-\omega t)}$ represents a wave moving to the right, while $\psi_L(x,t) = Ae^{i(-kx-\omega t)}$ represents a wave moving to the left. (Take k and ω to be positive.) Apply each of our two candidate momentum operators

$$\hat{p}_1 \doteq -i\hbar\frac{\partial}{\partial x} \quad \text{and} \quad \hat{p}_2 \doteq +i\hbar\frac{\partial}{\partial x}$$

to both of these functions, and show that the first candidate makes more sense.

Solution:

$$\langle x|\hat{p}_1|\psi_R\rangle = -i\hbar\frac{\partial}{\partial x}Ae^{i(+kx-\omega t)} = -i\hbar(+ik)Ae^{i(+kx-\omega t)} = (+\hbar k)\psi_R(x,t)$$
$$\langle x|\hat{p}_1|\psi_L\rangle = -i\hbar\frac{\partial}{\partial x}Ae^{i(-kx-\omega t)} = -i\hbar(-ik)Ae^{i(-kx-\omega t)} = (-\hbar k)\psi_L(x,t)$$
$$\langle x|\hat{p}_2|\psi_R\rangle = +i\hbar\frac{\partial}{\partial x}Ae^{i(+kx-\omega t)} = +i\hbar(+ik)Ae^{i(+kx-\omega t)} = (-\hbar k)\psi_R(x,t)$$
$$\langle x|\hat{p}_2|\psi_L\rangle = +i\hbar\frac{\partial}{\partial x}Ae^{i(-kx-\omega t)} = +i\hbar(-ik)Ae^{i(-kx-\omega t)} = (+\hbar k)\psi_L(x,t)$$

Thus the eigenvalues for these four situations are:

candidate	wave	eigenvalue
$\hat{p}_1$	rightward moving	$+\hbar k$
$\hat{p}_1$	leftward moving	$-\hbar k$
$\hat{p}_2$	rightward moving	$-\hbar k$
$\hat{p}_2$	leftward moving	$+\hbar k$

Candidate 1 associates the rightward moving wave with a positive momentum eigenvalue and the leftward moving wave with a negative momentum eigenvalue. Candidate 2 does the opposite. Since we sensibly associate rightward motion with positive momentum, candidate 1 is superior.

Check on $\hat{p}^2$:

$$\begin{aligned}\langle x|\hat{p}^2|\psi\rangle &= \langle x|\hat{p}\hat{p}|\psi\rangle \qquad [\![\text{define } |\phi\rangle = \hat{p}|\psi\rangle]\!] \\ &= \langle x|\hat{p}|\phi\rangle \\ &= -i\hbar\frac{\partial}{\partial x}\langle x|\phi\rangle \\ &= -i\hbar\frac{\partial}{\partial x}\Big(\langle x|\hat{p}|\psi\rangle\Big) \\ &= -i\hbar\frac{\partial}{\partial x}\left(-i\hbar\frac{\partial}{\partial x}\langle x|\psi\rangle\right) \\ &= -\hbar^2\frac{\partial^2}{\partial x^2}\langle x|\psi\rangle\end{aligned}$$

Now that we know everything there is to know about the momentum operator, we of course want to find its eigenstates $|p\rangle$!

Problems

6.5 Sign of the momentum operator

Given the “very reasonable surmise” 6.15, show that the net amplitude to flow right across the boundary between bin i and bin $i+1$ is

$$A_{i+1}\psi_i - C_i\psi_{i+1}.$$

Then use the three ancillary results $A_{i+1} = C_i = A$,

$$A = i\frac{\Delta t}{(\Delta x)^2}\frac{\hbar}{m}n_d,$$

and $n_d = 1/2$ to show that this net amplitude of rightward flow is

$$-i\frac{\Delta t}{2(\Delta x)^{1/2}}\frac{\hbar}{m}\frac{\partial \psi}{\partial x},$$

hence justifying the $-i$ choice in the definition of momentum operator.

6.6 Probability current and mean momentum

a. Show that for a quantal particle with wavefunction $\psi(x,t)$, the mean momentum is

$$-i\hbar\int_{-\infty}^{+\infty}\psi^*(x,t)\frac{\partial \psi(x,t)}{\partial x}\,dx. \tag{6.48}$$

b. If the "amount of water" in equation (6.30) is taken to mean the "mass of water", show that the total momentum of the water in the trough is

$$\int_{-\infty}^{+\infty} j_w(x)\,dx. \tag{6.49}$$

c. From this we might guess that the mean momentum for a particle with wavefunction $\psi(x,t)$, in terms of the probability current (6.33), is

$$m\int_{-\infty}^{+\infty} j(x,t)\,dx. \tag{6.50}$$

Show that this guess is correct, provided that $\psi(x,t)$ vanishes as $x \to \pm\infty$.

⟦This result suggests again that we made the correct sign choice back at equation (6.47).⟧

6.7 Mean momentum using wavefunction in polar form

Writing the wavefunction in polar form as $\psi = Re^{i\phi}$ (see equation 6.27), show that the mean momentum is

$$\begin{aligned}\langle \hat{p}\rangle_t &= \int_{-\infty}^{+\infty}\psi^*(x,t)\left[-i\hbar\frac{\partial \psi(x,t)}{\partial x}\right]dx\\ &= \hbar\int_{-\infty}^{+\infty}R^2(x,t)\frac{\partial \phi}{\partial x}\,dx.\end{aligned} \tag{6.51}$$

6.7 The momentum basis

Position representation of momentum eigenstates

The operator $\hat{p}$ represents a physical measurement, so it is Hermitian, so it posses a basis of eigenstates $|p_0\rangle$ (technically a rigged, continuous eigenbasis). What are these states like? In particularly, what is the position representation $\pi_0(x) = \langle x|p_0\rangle$?

$$\begin{aligned}
\hat{p}|p_0\rangle &= p_0|p_0\rangle \\
\langle x|\hat{p}|p_0\rangle &= p_0\langle x|p_0\rangle \\
-i\hbar\frac{\partial}{\partial x}\langle x|p_0\rangle &= p_0\langle x|p_0\rangle \\
-i\hbar\frac{\partial\pi_0(x)}{\partial x} &= p_0\pi_0(x) \\
\frac{\partial\pi_0(x)}{\partial x} &= i\frac{p_0}{\hbar}\pi_0(x) \\
\pi_0(x) &= Ce^{i(p_0/\hbar)x} \qquad (6.52)
\end{aligned}$$

That's funny. When we solve an eigenproblem, we expect that only a few eigenvalues will result. That's what happened with ammonia. But there we had 2 × 2 matrices, and got two eigenvalues, whereas here we have $\infty \times \infty$ matrices, so we get an infinite number of eigenvalues! The eigenvalue p_0 can be any real number... positive, negative, even zero! (It cannot be complex valued, because a Hermitian operator must have only real eigenvalues.)

The constant C is just an overall normalization constant. The best convention is (see problem 6.8)

$$C = \frac{1}{\sqrt{2\pi\hbar}}. \qquad (6.53)$$

In summary, the operator $\hat{p}$ has as eigenvalues any real number p_0, with eigenvectors $|p_0\rangle$ (technically, rigged vectors) satisfying

$$\hat{p}|p_0\rangle = p_0|p_0\rangle \qquad (6.54)$$

$$\langle x|p_0\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{i(p_0/\hbar)x}. \qquad (6.55)$$

Exercise 6.I. Show that $|p\rangle$ has the dimensions of $1/\sqrt{\text{momentum}}$. What are the dimensions of $\langle x|p\rangle$?

Problem 6.8 will show that the momentum states are orthonormal

$$\langle p|p'\rangle = \delta(p-p') \tag{6.56}$$

and complete

$$\hat{1} = \int_{-\infty}^{+\infty} |p\rangle\langle p|\,dp, \tag{6.57}$$

and hence the set $\{|p\rangle\}$ constitutes a continuous ("rigged") basis.

Representing states in the momentum basis

We have been dealing with a state $|\psi\rangle$ through its representation in the position basis, that is, through its wavefunction (or position representation)

$$\psi(x) = \langle x|\psi\rangle. \tag{6.58}$$

It is equally legitimate to deal with that state through its representation in the momentum basis, that is, through its so-called momentum wavefunction (or momentum representation)

$$\tilde{\psi}(p) = \langle p|\psi\rangle. \tag{6.59}$$

Either representation carries complete information about the state $|\psi\rangle$, so you can obtain one from the other

$$\begin{aligned}
\tilde{\psi}(p) = \langle p|\psi\rangle &= \langle p|\hat{1}|\psi\rangle \\
&= \int_{-\infty}^{+\infty} \langle p|x\rangle\langle x|\psi\rangle\,dx \\
&= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty} e^{-i(p/\hbar)x}\psi(x)\,dx \qquad (6.60)\\
\psi(x) = \langle x|\psi\rangle &= \langle x|\hat{1}|\psi\rangle \\
&= \int_{-\infty}^{+\infty} \langle x|p\rangle\langle p|\psi\rangle\,dp \\
&= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{+\infty} e^{+i(p/\hbar)x}\tilde{\psi}(p)\,dp. \qquad (6.61)
\end{aligned}$$

Perhaps you have seen pairs of functions like this before in a math course. The position and momentum wavefunctions are related to each other through what mathematicians call a "Fourier transform".

Exercise 6.J. The concept of momentum wavefunction. I wrote back on page 178 that "I don't care how clever or talented an experimentalist you are: you cannot insert an instrument into six-dimensional [configuration] space in order to measure wavefunction." Draw a similar conclusion concerning momentum wavefunction.

Representing operators in the momentum basis

It is easy to represent momentum-related operators in the momentum basis. For example, using the fact the $\hat{p}$ is Hermitian,

$$\langle p|\hat{p}|\psi\rangle = [\langle\psi|\hat{p}|p\rangle]^* = [p\langle\psi|p\rangle]^* = p\langle p|\psi\rangle. \tag{6.62}$$

More generally, for any function of the momentum operator,

$$\langle p|f(\hat{p})|\psi\rangle = f(p)\langle p|\psi\rangle. \tag{6.63}$$

It's a bit more difficult to find the momentum representation of the position operator, that is, to find $\langle p|\hat{x}|\psi\rangle$. But we can do it, using a slick trick called "parametric differentiation".

First, I'll introduce parametric differentiation in a purely mathematical context. Suppose you need to evaluate the integral

$$\int_0^\infty xe^{-kx}\cos x\,dx$$

but you can only remember that

$$\int_0^\infty e^{-kx}\cos x\,dx = \frac{k}{k^2+1}.$$

You can differentiate both sides with respect to the parameter k finding

$$\begin{aligned}
\frac{\partial}{\partial k}\int_0^\infty e^{-kx}\cos x\,dx &= \frac{\partial}{\partial k}\frac{k}{k^2+1}\\
\int_0^\infty \frac{\partial e^{-kx}}{\partial k}\cos x\,dx &= \frac{(k^2+1)-k(2k)}{(k^2+1)^2}\\
\int_0^\infty (-xe^{-kx})\cos x\,dx &= \frac{-k^2+1}{(k^2+1)^2}\\
\int_0^\infty xe^{-kx}\cos x\,dx &= \frac{k^2-1}{(k^2+1)^2}
\end{aligned}$$

This is a lot easier than any other method I can think of to evaluate this integral.

Go back to the problem of finding $\langle p|\hat{x}|\psi\rangle$:

$$\begin{aligned}
\langle p|\hat{x}|\psi\rangle &= \langle p|\hat{x}\hat{1}|\psi\rangle \\
&= \int_{-\infty}^{+\infty} \langle p|\hat{x}|x\rangle\langle x|\psi\rangle\, dx \\
&= \int_{-\infty}^{+\infty} \langle p|x\rangle x\langle x|\psi\rangle\, dx \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-i(p/\hbar)x} x\langle x|\psi\rangle\, dx \\
&\qquad [\![\text{Now use parametric differentiation!}]\!] \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \frac{\hbar}{-i}\frac{\partial}{\partial p}\left[e^{-i(p/\hbar)x}\right] \langle x|\psi\rangle\, dx \\
&= +i\hbar \frac{1}{\sqrt{2\pi\hbar}}\frac{\partial}{\partial p}\left[\int_{-\infty}^{+\infty} e^{-i(p/\hbar)x}\langle x|\psi\rangle\, dx\right] \\
&= +i\hbar\frac{\partial}{\partial p}\left[\int_{-\infty}^{+\infty} \langle p|x\rangle\langle x|\psi\rangle\, dx\right] \\
&= +i\hbar\frac{\partial}{\partial p}\langle p|\psi\rangle \qquad (6.64)
\end{aligned}$$

There's a nice symmetry to this result, making it easy to remember: The momentum operator, represented in the position basis, is

$$\langle x|\hat{p}|\psi\rangle = -i\hbar\frac{\partial}{\partial x}\psi(x). \qquad (6.65)$$

while the position operator, represented in the momentum basis, is

$$\langle p|\hat{x}|\psi\rangle = +i\hbar\frac{\partial}{\partial p}\tilde{\psi}(p). \qquad (6.66)$$

Exercise 6.K. Show that

$$|\psi\rangle = \int_{-\infty}^{+\infty} \psi(x)\,|x\rangle\, dx = \int_{-\infty}^{+\infty} \tilde{\psi}(p)\,|p\rangle\, dp. \qquad (6.67)$$

Verify that both of these relations have the correct dimensions.

Other bases

For continuous systems, we have the position basis and the momentum basis. But there are other useful bases as well. Much of the rest of this book is devoted to the energy basis. Another basis of interest is the "gaussian orthogonal basis", consisting of elements that are "nearly classical".

Problems

6.8 **The states $\{|p\rangle\}$ constitute a continuous basis**

At equation (6.52) we showed that the inner product $\langle x|p\rangle$ must have the form

$$\langle x|p\rangle = Ce^{i(p/\hbar)x} \tag{6.68}$$

where C may be chosen for convenience.

a. Show that the operator

$$\hat{A} = \int_{-\infty}^{+\infty} |p\rangle\langle p|\, dp \tag{6.69}$$

is equal to

$$2\pi\hbar|C|^2\hat{1} \tag{6.70}$$

by evaluating

$$\langle\phi|\hat{A}|\psi\rangle = \langle\phi|\hat{1}\hat{A}\hat{1}|\psi\rangle \tag{6.71}$$

for arbitrary states $|\psi\rangle$ and $|\phi\rangle$. *Clues:* Set the first $\hat{1}$ equal to $\int_{-\infty}^{+\infty} |x\rangle\langle x|\, dx$, the second $\hat{1}$ equal to $\int_{-\infty}^{+\infty} |x'\rangle\langle x'|\, dx'$. The identity (G.1) for the Dirac delta function is useful here. Indeed, this is one of the most useful equations to be found anywhere!

b. Using the conventional choice $C = 1/\sqrt{2\pi\hbar}$, show that

$$\langle p|p'\rangle = \delta(p - p'). \tag{6.72}$$

The expression (G.1) is again helpful.

6.9 **Peculiarities of continuous basis states**

Recall that the members of a continuous basis set are peculiar in that they possess dimensions. That is not their only peculiarity. For any ordinary state $|\psi\rangle$, the wavefunction $\psi(x) = \langle x|\psi\rangle$ satisfies

$$\int_{-\infty}^{+\infty} \psi^*(x)\psi(x)\, dx = 1. \tag{6.73}$$

Show that the states $|x_0\rangle$ and $|p_0\rangle$ cannot obey this normalization.

6.10 **Hermiticity of the momentum operator**

Show that the momentum operator is Hermitian over the space of states $|\psi\rangle$ that have wavefunctions $\psi(x)$ which vanish at $x = \pm\infty$. *Clue:*

$$\langle\phi|\hat{p}|\psi\rangle = \int_{-\infty}^{+\infty} \phi^*(x)\left(-i\hbar\frac{d\psi(x)}{dx}\right) dx. \tag{6.74}$$

Integrate by parts.

6.11 **Commutator of $\hat{x}$ and $\hat{p}$**

Show that

$$[\hat{x}, \hat{p}] = i\hbar \tag{6.75}$$

by showing that $\langle\phi|[\hat{x},\hat{p}]|\psi\rangle = i\hbar\langle\phi|\psi\rangle$ for arbitrary $|\phi\rangle$ and $|\psi\rangle$. (*Clues:* First evaluate $\langle x|\hat{p}\hat{x}|\psi\rangle$ and $\langle x|\hat{x}\hat{p}|\psi\rangle$. It helps to define $|\chi\rangle = \hat{x}|\psi\rangle$.)

6.12 **Summary of momentum basis states**

Make a table like the one on page 176 summarizing the properties of momentum basis states.

6.13 **Momentum representation of the Schrödinger equation**

You know that the Schrödinger equation

$$\frac{d|\psi(t)\rangle}{dt} = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle \tag{6.76}$$

has the position representation

$$\frac{\partial\langle x|\psi(t)\rangle}{\partial t} = -\frac{i}{\hbar}\langle x|\hat{H}|\psi(t)\rangle \tag{6.77}$$

or

$$\frac{\partial\psi(x,t)}{\partial t} = -\frac{i}{\hbar}\left[-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V(x)\psi(x,t)\right]. \tag{6.78}$$

In this problem you will uncover the corresponding equation that governs the time evolution of

$$\tilde{\psi}(p,t) = \langle p|\psi(t)\rangle. \tag{6.79}$$

The left hand side of equation (6.76) is straightforward because

$$\langle p|\frac{d}{dt}|\psi(t)\rangle = \frac{\partial\tilde{\psi}(p,t)}{\partial t}. \tag{6.80}$$

To investigate the right hand side of equation (6.76) write

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + V(\hat{x}) \tag{6.81}$$

where $\hat{p}$ is the momentum operator and $V(\hat{x})$ the potential energy operator.

a. Use the Hermiticity of $\hat{p}$ to show that

$$\langle p|\hat{H}|\psi(t)\rangle = \frac{p^2}{2m}\tilde{\psi}(p,t) + \langle p|V(\hat{x})|\psi(t)\rangle. \tag{6.82}$$

Now we must investigate $\langle p|V(\hat{x})|\psi(t)\rangle$.

b. Show that

$$\langle p|\hat{V}|\psi(t)\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-i(p/\hbar)x}\, V(x)\psi(x,t)\, dx \qquad (6.83)$$

by inserting the proper form of $\hat{1}$ at the proper location.

c. Define the (modified) Fourier transform $\tilde{V}(p)$ of $V(x)$ through

$$\tilde{V}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-i(p/\hbar)x}\, V(x)\, dx \qquad (6.84)$$

$$= \int_{-\infty}^{+\infty} \langle p|x\rangle V(x)\, dx. \qquad (6.85)$$

Does $\tilde{V}(p)$ have the dimensions of energy? Show that

$$V(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{i(p/\hbar)x}\, \tilde{V}(p)\, dp \qquad (6.86)$$

$$= \int_{-\infty}^{+\infty} \langle x|p\rangle \tilde{V}(p)\, dp. \qquad (6.87)$$

You may use either forms (6.84) and (6.86), in which case the proof employs equation (G.1), or forms (6.85) and (6.87), in which case the proof involves completeness and orthogonality of basis states.

d. Hence show that

$$\langle p|\hat{V}|\psi(t)\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \tilde{V}(p-p')\tilde{\psi}(p',t)\, dp'. \qquad (6.88)$$

(Caution! Your intermediate expressions will probably involve three distinct variables that you'll want to call "p". Put primes on two of them!)

e. Put everything together to see that $\tilde{\psi}(p;t)$ obeys the integro-differential equation

$$\frac{\partial\tilde{\psi}(p,t)}{\partial t} = -\frac{i}{\hbar}\left[\frac{p^2}{2m}\tilde{\psi}(p,t) + \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \tilde{V}(p-p')\tilde{\psi}(p',t)\, dp'\right]. \qquad (6.89)$$

The time evolution equation is local in position space — that is the change $\partial\psi(x)/\partial t$ is affected only by the the values of ψ in the immediate vicinity of x. But it is *not* local in momentum space — the change $\partial\tilde{\psi}(p)/\partial t$ is affected by the values of $\tilde{\psi}$ all up and down the momentum axis.

This momentum representation of the Schrödinger equation is particularly useful in the study of superconductivity.

6.8 Position representation of time evolution solution

In our general treatment of time evolution we found (equation 5.44) that state $|\psi(t)\rangle$ evolved in time according to

$$|\psi(t)\rangle = \sum_n \psi_n(0) e^{-(i/\hbar)e_n t} \,|e_n\rangle, \tag{6.90}$$

where $|e_n\rangle$ is the energy eigenstate with energy eigenvalue e_n

$$\hat{H}|e_n\rangle = e_n|e_n\rangle \tag{6.91}$$

and where

$$\psi_n(0) = \langle e_n|\psi(0)\rangle. \tag{6.92}$$

How does this formal time evolution solution translate into the position representation for a single spinless particle ambivating in one dimension subject to the potential energy function $V(x)$?

The energy eigenfunctions (that is, the wavefunctions of the energy eigenvectors) are usually called

$$\eta_n(x) = \langle x|e_n\rangle \tag{6.93}$$

where the Greek letter η, pronounced "eta", suggests "energy" through alliteration. They satisfy the energy eigenequation

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\eta_n(x) = E_n\eta_n(x). \tag{6.94}$$

The initial wavefunction

$$\psi(x,0) = \langle x|\psi(0)\rangle \tag{6.95}$$

is represented in terms of energy eigenfunctions as

$$\psi(x,0) = \sum_{n=1}^{\infty} C_n\eta_n(x), \tag{6.96}$$

where

$$C_n = \langle e_n|\psi(0)\rangle = \int_{-\infty}^{+\infty} \eta_n^*(x)\psi(x,0)\,dx. \tag{6.97}$$

The wavefunction evolves in time as

$$\psi(x,t) = \sum_{n=1}^{\infty} C_n e^{-(i/\hbar)E_n t}\,\eta_n(x). \tag{6.98}$$

Exercise 6.L. Prove these statements to your own satisfaction.

6.9 The classical limit of quantum mechanics

I told you way back on page 2 that when quantum mechanics is applied to big things, it gives the results of classical mechanics. It's hard to see how my claim could possibly be correct: the whole structure of quantum mechanics differs so dramatically from the structure of classical mechanics — the character of a "state", the focus on potential energy function rather than on force, the fact that the quantal time evolution equation involves a first derivative with respect to time while the classical time evolution equation involves a second derivative with respect to time.

6.9.1 *How does mean position change with time?*

This nut is cracked by focusing, not on the full quantal state $\psi(x,t)$, but on the mean position

$$\langle x \rangle = \int_{-\infty}^{+\infty} \psi^*(x,t) x \psi(x,t)\, dx, \tag{6.99}$$

How does this mean position change with time?

The answer depends on the classical force function $F(x)$ — i.e., the classical force that would be exerted on a classical particle if it were at position x. (I'm not saying that the particle *is* at x, I'm not even saying that the particle has a position; I'm saying that's what the force *would be* if the particle *were* classical and at position x.)

The answer is that

$$\langle F(x) \rangle = m \frac{d^2 \langle x \rangle}{dt^2}, \tag{6.100}$$

a formula that certainly plucks our classical heartstrings! This result is called the **Ehrenfest**[10] **theorem**. We will prove this theorem later (at equations 6.109 and 6.110), but first discuss its significance.

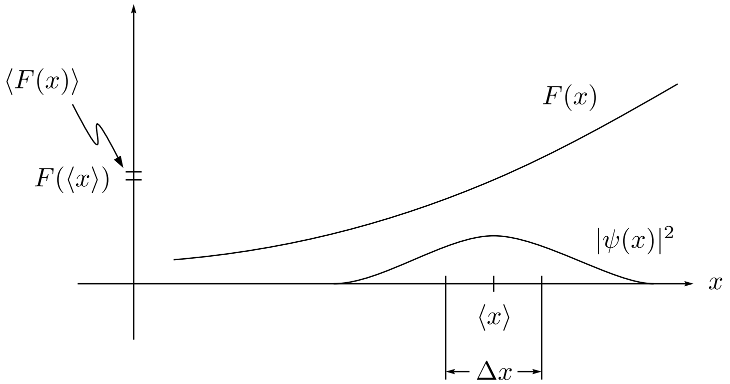
Although the theorem is true in all cases, it is most useful when the spread in position Δx is in some sense small, so the wavefunction is relativity compact. Such wavefunctions are called "wavepackets". In this

---

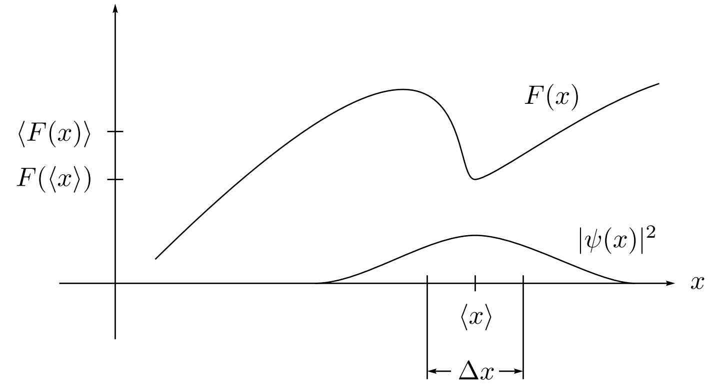
[10] Paul Ehrenfest (1880–1933), Austrian-Dutch theoretical physicist, known particularly for asking probing questions that clarified the essence and delineated the unsolved problems of any issue at hand. As a result, several telling arguments have names like "Ehrenfest's paradox" or "Ehrenfest's urn" or "the Ehrenfest dog-flea model". Particularly in this mode of questioner, he played a central role in the development of relativity, of quantum mechanics, and of statistical mechanics. He died tragically by his own hand.

situation we might hope for a useful approximation — the classical limit — by ignoring the quantal indeterminacy of position and focusing solely on mean position.

If the force function $F(x)$ varies slowly on the scale of Δx, then our hopes are confirmed: the spread in position is small, the spread in force is small, and to a good approximation the mean force $\langle F(x)\rangle$ is equal to the force at the mean position $F(\langle x\rangle)$.



But if the force function varies rapidly on the scale of Δx, then our hopes are dashed: the spread in position is small, but the spread in force is not, and the classical approximation is not appropriate.



To head off a misconception, I emphasize that Ehrenfest's theorem is *not* that

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

If this were true, then the mean position of a quantal particle would in all cases move exactly as a classical particle does. But (see problem 6.**??**, "Mean of function vs. function of mean", on page **??**) it's *not* true.

6.9.2 *Is the classical approximation good enough?*

If the quantal position indeterminacy Δx is small compared to the experimental uncertainty of your position-locating experimental apparatus, for the entire duration of your experiment, then the classical approximation is usually appropriate. So the central question is: How big is the quantal Δx in my situation? This will of course vary from case to case and from time to time within a given case. But there's an important theorem that connects the indeterminacy of position Δx with the indeterminacy of momentum Δp: in *all* situations

$$\Delta x \Delta p \geq \tfrac{1}{2}\hbar. \tag{6.101}$$

This theorem is the original *Heisenberg indeterminacy principle.* (The more general indeterminacy principle presented on page 133 was discovered one month later by Earle Hesse Kennard[11].) It is proven simply by applying the indeterminacy principle (4.13) to the commutator (6.75). It is important for two reasons: First, because it's important for determining whether the classical approximation is adequate in a given case. Second, because it was important in the historical development of quantum mechanics.

Quantum mechanics has a long and intricate (and continuing!) history, but one of the keystone events occurred in the spring of 1925. Werner Heisenberg,[12] a freshly minted Ph.D., had obtained a position as assistant

---

[11] "Zur Quantenmechanik einfacher Bewegungstypen" *Zeitschrift für Physik* **44** (April 1927) 326–352.

[12] German theoretical physicist (1901–1976) who nearly failed his Ph.D. oral exam due to his fumbling in experimental physics. He went on to discover quantum mechanics as we know it today. Although attacked by Nazis as a "white Jew", he became a principal scientist in the German nuclear program during World War II, where he focused on building nuclear reactors rather than nuclear bombs. After the war he worked to rebuild German science, and to extend quantum theory into relativistic and field theoretic domains. He enjoyed hiking, particularly in the Bavarian Alps, and playing the piano. After a three-month whirlwind romance, Heisenberg married Elisabeth Schumacher, sister of the *Small Is Beautiful* economist E.F. Schumacher, and they went on to parent seven children.

to Max Born at the University of Göttingen. There he realized that the key to formulating quantum mechanics was to develop a theory that fit atomic experiments, and that also had the correct classical limit. He was searching for such a theory when he came down with a bad case of allergies to spring pollen from the "mass of blooming shrubs, rose gardens and flower beds"[13] of Göttingen. He decided to travel to Helgoland, a rocky island and fishing center in the North Sea, far from pollen sources, arriving there by ferry on 8 June 1925.

Once his health returned, Heisenberg reproduced his earlier work, cleaning up the mathematics and simplifying the formulation. He worried that the mathematical scheme he invented might prove to be inconsistent, and in particular that it might violate the principle of energy conservation. In Heisenberg's own words:[14]

> One evening I reached the point where I was ready to determine the individual terms in the energy table, or, as we put it today, in the energy matrix, by what would now be considered an extremely clumsy series of calculations. When the first terms seemed to accord with the energy principle, I became rather excited, and I began to make countless arithmetical errors. As a result, it was almost three o'clock in the morning before the final result of my computations lay before me. The energy principle had held for all the terms, and I could no longer doubt the mathematical consistency and coherence of the kind of quantum mechanics to which my calculations pointed. At first, I was deeply alarmed. I had the feeling that, through the surface of atomic phenomena, I was looking at a strangely beautiful interior, and felt almost giddy at the thought that I now had to probe this wealth of mathematical structures nature had so generously spread out before me. I was far too excited to sleep, and so, as a new day dawned, I made for the southern tip of the island, where I had been longing to climb a rock jutting out into the sea. I now did so without too much trouble, and waited for the sun to rise.

Because the correct classical limit was essential in producing this theory, it was easy to fall into the misconception that an electron really did behave classically, with a single position, but that this single position is disturbed

---

[13] Werner Heisenberg, *Physics and Beyond* (Harper and Row, New York, 1971) page 37.

[14] *Physics and Beyond*, page 61.

by the measuring apparatus used to determine position. Indeed, Heisenberg wrote as much:[15]

> observation of the position will alter the momentum by an unknown and undeterminable amount.

But Neils Bohr repeatedly objected to this "disturbance" interpretation. For example, at a 1938 conference in Warsaw,[16] he

> warned specifically against phrases, often found in the physical literature, such as "disturbing of phenomena by observation."

Today, interference and entanglement experiments make clear that Bohr was right and that "measurement disturbs the system" is not a tenable position.[17] In an interferometer, there is *no local way* that a photon at path `a` can physically disturb an atom taking path `b`. For an entangled pair of atoms, there is *no local way* that an analyzer measuring the magnetic moment of the left atom can physically disturb the right atom. It is no defect in our measuring apparatus that it cannot determine what does not exist.

And this brings us to one last terminology note. What we have called the "Heisenberg indeterminacy principle" is called by some the "Heisenberg uncertainty principle".[18] The second name is less accurate because it gives the mistaken impression that an electron *really does* have a position and we are just uncertain as to what that position is. It also gives the mistaken impression that an electron *really does* have a momentum and we are just uncertain as to what that momentum is.

---

[15] Werner Heisenberg, *The Physical Principles of the Quantum Theory*, translated by Carl Eckart and F.C. Hoyt (University of Chicago Press, Chicago, 1930) page 20.

[16] Niels Bohr, "Discussion with Einstein on epistemological problems in atomic physics," in *Albert Einstein, Philosopher–Scientist*, edited by Paul A. Schilpp (Library of Living Philosophers, Evanston, Illinois, 1949) page 237.

[17] To be completely precise, "measurement disturbs the system locally" is not a tenable position. The "de Broglie–Bohm pilot wave" formulation of quantum mechanics can be interpreted as saying that "measurement disturbs the system", but the measurement at one point in space is felt instantly at points arbitrarily far away. When this formulation is applied to a two-particle system, a "pilot wave" situated in six-dimensional configuration space somehow physically guides the two particles situated in ordinary three-dimensional space.

[18] Heisenberg himself, writing in German, called it the "Genauigkeit Beziehung" — accuracy relationship. See "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik" *Zeitschrift für Physik* **43** (March 1927) 172–198.

6.9.3 *Sample Problem*

For the "Underground Guide to Quantum Mechanics" (described on page **??**), you decide to write a passionate persuasive paragraph or two concerning the misconception that "measurement disturbs the system". What do you write?

Possible Solution: For those of us who know and love classical mechanics, there's a band-aid, the idea that "measurement disturbs the system". This idea is that fundamentally classical mechanics actually holds, but that quantum mechanics is a mask layered over top of, and obscuring the view of, the classical mechanics because our measuring devices disturb the underlying classical system. That's not possible. It is no defect of our measuring instruments that they cannot determine what does not exist, just as it is no defect of a colorimeter that it cannot determine the color of love.

This idea that "measurement disturbs the system" is a psychological trick to comfort us, and at the same time to keep us from exploring, fully and openly, the strange world of quantum mechanics. I urge you, I implore you, to discard this security blanket, to go forth and discover the new world as it really is rather than cling to the familiar classical world. Like Miranda in Shakespeare's *Tempest*, take delight in this "brave new world, that has such people in't".

Unlike most band-aids, this band-aid does not protect or cover up. Instead it exposes a lack of imagination.

6.9.4 *Time evolution of mean quantities*

Our general treatment of time evolution found (equation 5.45) that for the measurable with associated operator $\hat{A}$, the mean value $\langle\hat{A}\rangle_t$ changes with time according to

$$\frac{d\langle\hat{A}\rangle_t}{dt} = -\frac{i}{\hbar}\langle[\hat{A},\hat{H}]\rangle_t. \tag{6.102}$$

For one particle ambivating in one dimension,

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + V(\hat{x}), \tag{6.103}$$

where $\hat{x}$ and $\hat{p}$ satisfy the commutation relation (see problem 6.11)

$$[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar. \tag{6.104}$$

Knowing this, let's see how the mean position $\langle \hat{x} \rangle_t$ changes with time. We must find

$$[\hat{x}, \hat{H}] = \frac{1}{2m}[\hat{x}, \hat{p}^2] + [\hat{x}, V(\hat{x})].$$

The commutator $[\hat{x}, V(\hat{x})]$ is easy:

$$[\hat{x}, V(\hat{x})] = \hat{x}V(\hat{x}) - V(\hat{x})\hat{x} = 0.$$

And the commutator $[\hat{x}, \hat{p}^2]$ is not much harder. We use the know commutator for $[\hat{x}, \hat{p}]$ to write

$$\hat{x}\hat{p}^2 = (\hat{x}\hat{p})\hat{p} = (\hat{p}\hat{x} + i\hbar)\hat{p} = \hat{p}\hat{x}\hat{p} + i\hbar\hat{p},$$

and then use it again to write

$$\hat{p}\hat{x}\hat{p} = \hat{p}(\hat{x}\hat{p}) = \hat{p}(\hat{p}\hat{x} + i\hbar) = \hat{p}^2\hat{x} + i\hbar\hat{p}.$$

Together we have

$$\hat{x}\hat{p}^2 = \hat{p}^2\hat{x} + 2i\hbar\hat{p}$$

or

$$[\hat{x}, \hat{p}^2] = 2i\hbar\hat{p}.$$

Plugging these commutators into the time-evolution result, we get

$$\frac{d\langle \hat{x} \rangle_t}{dt} = -\frac{i}{\hbar}\frac{1}{2m}2i\hbar\langle \hat{p} \rangle_t.$$

or

$$\frac{d\langle \hat{x} \rangle_t}{dt} = \frac{\langle \hat{p} \rangle_t}{m}, \tag{6.105}$$

a result that stirs our memories of classical mechanics!

Meanwhile, what happens for mean momentum $\langle \hat{p} \rangle_t$?

$$[\hat{p}, \hat{H}] = \frac{1}{2m}[\hat{p}, \hat{p}^2] + [\hat{p}, V(\hat{x})] = [\hat{p}, V(\hat{x})].$$

To evaluate $[\hat{p}, V(\hat{x})]$ we use the familiar idea that if we know $\langle x|\hat{A}|\psi\rangle$ for arbitrary $|x\rangle$ and $|\psi\rangle$, then we know everything there is to know about the operator $\hat{A}$. In this way, examine

$$\begin{aligned}
\langle x|[\hat{p}, V(\hat{x})]|\psi\rangle &= \langle x|\hat{p}V(\hat{x})|\psi\rangle - \langle x|V(\hat{x})\hat{p}|\psi\rangle \\
&= -i\hbar\frac{\partial}{\partial x}\langle x|V(\hat{x})|\psi\rangle - V(x)\langle x|\hat{p}|\psi\rangle \\
&= -i\hbar\frac{\partial}{\partial x}\left[V(x)\psi(x)\right] - V(x)\left[-i\hbar\frac{\partial}{\partial x}\psi(x)\right] \\
&= -i\hbar\left[\frac{\partial V(x)}{\partial x}\psi(x) + V(x)\frac{\partial\psi(x)}{\partial x} - V(x)\frac{\partial\psi(x)}{\partial x}\right] \\
&= -i\hbar\left[\frac{\partial V(x)}{\partial x}\psi(x)\right].
\end{aligned}$$

Now, the derivative of the classical potential energy function has a name. It's just the negative of the classical force function!

$$F(x) = -\frac{\partial V(x)}{\partial x}. \tag{6.106}$$

Continuing the evaluation begun above,

$$\begin{aligned}
\langle x|[\hat{p}, V(\hat{x})]|\psi\rangle &= i\hbar\left[F(x)\psi(x)\right] \\
&= i\hbar\langle x|F(\hat{x})|\psi\rangle.
\end{aligned}$$

Because this relation holds for any $|x\rangle$ and for any $|\psi\rangle$, we know that the operators are related as

$$[\hat{p}, V(\hat{x})] = i\hbar F(\hat{x}). \tag{6.107}$$

Going back to the time evolution of mean momentum,

$$\frac{d\langle\hat{p}\rangle_t}{dt} = -\frac{i}{\hbar}\langle[\hat{p}, \hat{H}]\rangle_t = -\frac{i}{\hbar}i\hbar\langle F(\hat{x})\rangle_t$$

or

$$\frac{d\langle\hat{p}\rangle_t}{dt} = \langle F(\hat{x})\rangle_t, \tag{6.108}$$

which is suspiciously close to Newton's second law!

These two results together,

$$\frac{d\langle\hat{x}\rangle_t}{dt} = \frac{\langle\hat{p}\rangle_t}{m} \tag{6.109}$$

$$\frac{d\langle\hat{p}\rangle_t}{dt} = \langle F(\hat{x})\rangle_t, \tag{6.110}$$

which tug so strongly on our classical heartstrings, are called the **Ehrenfest theorem**. You should remember two things about them: First, they are exact (within the assumptions of our derivation: non-relativistic, one-dimensional, no frictional or magnetic forces, etc.). Because they do tug our classical heartstrings, some people get the misimpression that they apply only in the classical limit. That's wrong — if you go back over the derivation you'll see that we never made any such assumption. Second, they are incomplete. This is because (1) knowing $\langle \hat{x} \rangle_t$ doesn't let you calculate $\langle F(\hat{x}) \rangle_t$, because in general $\langle F(\hat{x}) \rangle_t \neq F(\langle \hat{x} \rangle_t)$, and because (2) even if you did know both $\langle \hat{x} \rangle_t$ and $\langle \hat{p} \rangle_t$, that would not give you complete knowledge of the state.

Problems

6.14 **Alternative derivation.**
Derive result 6.107 by expanding $V(x)$ in a Taylor series.

6.15 **Sign of momentum operator**
If we had taken the opposite sign choice for the momentum operator at equation (6.47) (call this choice $\hat{p}_2$), then what would have been the commutator $[\hat{x}, \hat{p}_2]$? What would have been the result 6.105?

6.16 **Quantities in the Hamiltonian**
When we derived equation (6.25) we were left with an undetermined number n_d and an undetermined function $v(x)$. Repeat the derivation of the Ehrenfest equations with this form of the Schrödinger equation to determine that number and function by demanding the correct classical limit.

6.17 **Questions** (recommended problem)
Update your list of quantum mechanics questions that you started at problem 1.13 on page 56. Write down new questions and, if you have uncovered answers to any of your old questions, write them down briefly.

Chapter 7

Particle in an Infinite Square Well

7.1 Setup

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

$$x(t); \quad v(t).$$

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

$$\frac{d^2x(t)}{dt^2} = \frac{1}{m}F(x(t)) \tag{7.1}$$

subject to the given initial conditions

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

In quantum mechanics, the state of the particle is given through the wavefunction: that is, we want to know the two-variable function

$$\psi(x, t).$$

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

$$\frac{\partial\psi(x,t)}{\partial t} = -\frac{i}{\hbar}\left[-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V(x)\psi(x,t)\right], \tag{7.2}$$

subject to the given initial condition

$$\psi(x, 0) = \psi_0(x).$$

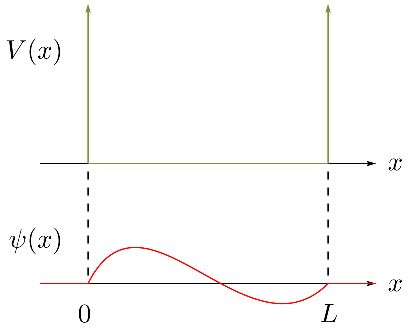
⟦The classical time evolution equation (7.1) is *second* order in time, so there are two initial conditions: initial position and initial velocity. The quantal time evolution equation (7.2) is *first* order in time, so there is only one initial condition: initial wavefunction.⟧

Infinite square well. For this, our first concrete problem involving position, let's choose the easiest potential energy function: the so-called infinite square well[1] or "particle in a box":

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

This is an approximate potential energy function for an electron added to a hydrocarbon chain molecule (a "conjugated polymer"), or for an atom trapped in a capped carbon nanotube.

The infinite square well is like the "perfectly rigid cylinder that rolls without slipping" in classical mechanics. It does not exactly exist in reality: any cylinder will be dented or cracked if hit hard enough. But it is a good model for some real situations. And it's *certainly* better to work with this model than it is to shrug your shoulders and say "I have no idea."



The infinite square well potential energy function $V(x)$ in olive green, and a possible wavefunction $\psi(x)$ in red.

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

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

and we adopt these conditions.

---

[1] Any potential energy function with a minimum is called a "well".

7.2 Solving the energy eigenproblem

We start by solving the energy eigenproblem (6.94)

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V(x)\right]\eta_n(x)=E_n\eta_n(x). \tag{7.3}$$

This will give us the allowed energies, which is often all we need. And if we wish to investigate time evolution, this is an important intermediate step.

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

$$-\frac{\hbar^2}{2m}\frac{d^2\eta_n(x)}{dx^2}=E_n\eta_n(x) \quad \text{with } \eta_n(0)=0; \eta_n(L)=0. \tag{7.4}$$

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

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

$$M\frac{d^2f(t)}{dt^2}=-kf(t),$$

you'd respond: "Of course, this is the ODE for a classical mass on a spring! The solution is

$$f(t)=A\cos(\omega t)+B\sin(\omega t) \quad \text{where } \omega=\sqrt{k/M}.\text{"}$$

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

$$\eta_n(x)=A_n\cos(\omega x)+B_n\sin(\omega x) \quad \text{where } \omega=\sqrt{2mE_n/\hbar^2}.$$

Writing this out neatly,

$$\eta_n(x)=A_n\cos((\sqrt{2mE_n}/\hbar)x)+B_n\sin((\sqrt{2mE_n}/\hbar)x). \tag{7.5}$$

When you solved the classical problem of a mass on a spring, you had to supplement the ODE solution with the initial values $f(0)=x_0$, $f'(0)=v_0$, to find the constants A and B. This is called an "initial value problem". For the problem of a particle in a box, we don't have an initial value problem; instead we are given $\eta_n(0)=0$ and $\eta_n(L)=0$, which is called a "boundary value problem".

Plugging $x = 0$ into equation (7.5) will be easier than plugging in $x = L$, so I'll do that first.[2] The result gives

$$\eta_n(0) = A_n \cos(0) + B_n \sin(0) = A_n,$$

so the boundary value $\eta_n(0) = 0$ means that $A_n = 0$ — for all values of n! Thus

$$\eta_n(x) = B_n \sin((\sqrt{2mE_n}/\hbar)x). \tag{7.6}$$

Now plug $x = L$ into equation (7.6), giving

$$\eta_n(L) = B_n \sin((\sqrt{2mE_n}/\hbar)L),$$

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

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

and it follows that

$$\eta_n(x) = B_n \sin((n\pi/L)x).$$

If you think about it for a minute, you'll realize that $n = 0$ gives rise to $\eta_0(x) = 0$. True, this is a solution to the differential equation, but it's not an interesting one. Similarly, the solution for $n = -3$ is just the negative of the solution for $n = +3$, so we get the same effect by changing the sign of B_3. We don't have to worry about negative or zero values for n.

The final (and often unnecessary) step in solving the energy eigenprobem is fixing the coefficient B_n by normalizing the energy eigenfunction.

Exercise 7.A. Show that the energy eigenfunction is normalized when $B_n = \sqrt{2/L}$ which, suprisingly, is independent of n. Does it have the correct dimensions?

In short, the solutions for the energy eigenproblem are

$$\eta_n(x) = \sqrt{\frac{2}{L}} \sin\left(n\pi \frac{x}{L}\right) \qquad \text{where } n = 1, 2, 3, \ldots \tag{7.7}$$

and with

$$E_n = n^2 \frac{\pi^2 \hbar^2}{2mL^2}. \tag{7.8}$$

We have accomplished the "unfair": we have not only solved the differential equation, we have also determined the permissible values of E_n.

---

[2] When faced with two conditions, I always invoke the easy one first. That way, if the result is zero, I won't have to bother invoking the second condition.

7.3 Solution to the time evolution problem

With the solution of the energy eigenproblem in hand, it is easy to solve the time development problem. The initial wavefunction $\psi_0(x)$ evolves in time to

$$\psi(x,t) = \sum_{n=1}^{\infty} C_n e^{-(i/\hbar)E_n t}\eta_n(x) \qquad \text{where} \qquad C_n = \int_{-\infty}^{+\infty} \eta_n^*(x)\psi_0(x)\,dx. \tag{7.9}$$

In one sense, this result is nothing but a very special case of the general time evolution theorem 5.44, but seeing the result come in this particular problem through separation of variables renders the more general theorem more tangible and less abstract.

7.4 What have we learned?

It's always satisfying to successfully conclude an intricate piece of mathematics. But we can't just stop and take a nap. We tackled the mathematical problem in the first place because we were interested in what the math can tell us about nature. So what is the math saying?

7.4.1 *Quantal recurrence*

The wavefunction evolves in time according to

$$\psi(x,t) = \sum_{n=1}^{\infty} C_n e^{-(i/\hbar)E_n t}\,\eta_n(x). \tag{7.10}$$

Suppose there were a time T_{rec} such that

$$e^{-(i/\hbar)E_n T_{\text{rec}}} = 1 \qquad \text{for } n = 1, 2, 3, \ldots. \tag{7.11}$$

What would the wavefunction $\psi(x,t)$ look like at time $t = T_{\text{rec}}$? It would be exactly equal to the initial wavefunction $\psi_0(x)$! If there is such a time, it's called the "recurrence time".

But it's not clear that such a recurrence time exists. After all, equation (7.11) lists an infinite number of conditions to be satisfied for recurrence to occur. Let's investigate. Because $e^{-i\,2\pi\,\text{integer}} = 1$ for any integer, the recurrence conditions (7.11) are equivalent to

$$(1/\hbar)E_n T_{\text{rec}} = 2\pi(\text{an integer}) \qquad \text{for } n = 1, 2, 3, \ldots.$$

Combined with the energy eigenvalues (7.8), these conditions are

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

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

$$T_{\text{rec}} = \frac{4mL^2}{\pi\hbar}. \tag{7.12}$$

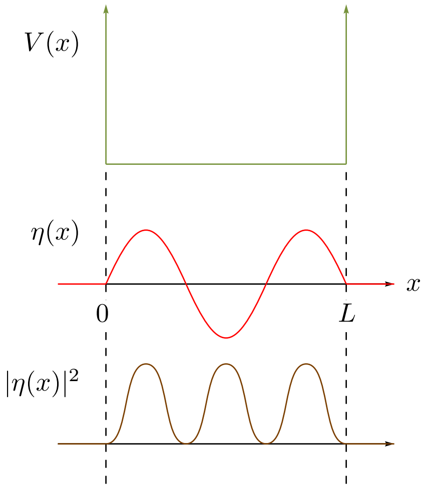
Cute and unexpected! This behavior is buried within equations (7.10) and (7.8), but no one would have uncovered it from a glance.

7.4.2 *Moving across a node*

Think about the wavefunction

$$\sqrt{2/L}\,\sin(3\pi x/L).$$

This wavefunction and corresponding probability density are graphed below the infinite square well potential energy function.



This particular wavefunction has two interior zeros, also called nodes. A common question is "There is zero probability of finding the particle at the node, so how can it move from one side of the node to the other?" People who ask this question suffer from the misconception that the particle is an infinitely small, infinitely hard version of a classical marble, which hence has a definite position. They think that the definite position of this infinitely small marble is changing rapidly, or changing erratically, or changing unpredictably, or changing subject to the slings and arrows of outrageous fortune. In truth, the quantal particle in this state doesn't have a definite position: it doesn't have a position at all! The quantal particle in the state above doesn't, *can't*, change its position from one side of the node to the other, because the particle *doesn't have a position.*

The "passing through nodes" question doesn't have an answer because the question assumes an erroneous picture for the character of a particle. It is as silly and as unanswerable as the question "If love is blue and passion is red-hot, how can passionate love exist?"

7.4.3 *Stationary states*

We see that, as far as time evolution is concerned, wavefunctions like $\sin(n\pi x/L)$ play a special role. What if the initial wavefunction $\psi_0(x)$ happens to have this form? We investigate $n = 3$. Once you see how things work in this case, you can readily generalize to any positive integer n.

So the initial wavefunction is

$$\psi_0(x) = \sqrt{2/L}\,\sin(3\pi x/L),$$

which evolves in time to

$$\psi(x,t) = e^{-(i/\hbar)E_3 t}\sqrt{2/L}\,\sin(3\pi x/L). \tag{7.13}$$

That's it! For this particular initial wavefunction, the system remains always in that same wavefunction, except multiplied by an time-dependent phase factor of $e^{-(i/\hbar)E_3 t}$. This uniform phase factor has *no effect whatsoever* on the probability density! Such states are called "stationary states".

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

$$\psi_0(x) = \frac{4}{5}\sqrt{\frac{2}{L}}\,\sin(3\pi x/L) + \frac{3}{5}\sqrt{\frac{2}{L}}\,\sin(7\pi x/L). \tag{7.14}$$

How does this state change with time?

Exercise 7.B. Show that the wavefunction given above (equation 7.14) is normalized.

Exercise 7.C. Show that the wavefunction given above (equation 7.14) evolves in time to

$$\psi(x,t) = \frac{4}{5}\sqrt{\frac{2}{L}}\, e^{-(i/\hbar)E_3 t} \sin(3\pi x/L) + \frac{3}{5}\sqrt{\frac{2}{L}}\, e^{-(i/\hbar)E_7 t} \sin(7\pi x/L). \tag{7.15}$$

Exercise 7.D. Show that the probability density of state (7.15) is

$$\begin{aligned} &\frac{16}{25}\frac{2}{L}\sin^2(3\pi x/L) + \frac{9}{25}\frac{2}{L}\sin^2(7\pi x/L) \\ &\quad + \frac{24}{25}\frac{2}{L}\cos((E_7 - E_3)t/\hbar)\sin(3\pi x/L)\sin(7\pi x/L), \end{aligned}$$

which *does* change with time, so this is *not* a stationary state.

Recall from page 99 that the word *eigen* means "characteristic of" or "peculiar to" or "belonging to". The state (7.13) "belongs to" the energy E_3. In contrast, the state (7.15) does *not* "belong to" any particular energy, because it involves both E_3 and E_7. Instead, this state has amplitude $\frac{4}{5}$ to have energy E_3 and amplitude $\frac{3}{5}$ to have energy E_7. We say that this state is a "superposition" of the energy states $\eta_3(x)$ and $\eta_7(x)$.

A particle trapped in a one-dimensional infinite square well cannot have any old energy: the only energies possible are the energy eigenvalues E_1, E_2, E_3, ... given in equation (7.8).

From the very first page of the very first chapter of this book we have been talking about quantization. But when we started it came from an experiment. Here quantization comes out of the theory, a theory predicting that the only possible energies are those listed in equation (7.8). We have reached a milestone in our development of quantum mechanics

Because the only possible energies are the energy eigenvalues E_1, E_2, E_3, ..., some people get the misimpression that the only possible *states* are the energy eigenstates $\eta_1(x)$, $\eta_2(x)$, $\eta_3(x)$, ... That's false. The state (7.15), for example, is a superposition of two energy states with different energies.

Analogy. A silver atom in magnetic moment state $|z+\rangle$ enters a vertical interferometer. It passes through the upper path. While traversing the interferometer, this atom has a position.

A different silver atom in magnetic moment state $|x-\rangle$ enters that same vertical interferometer. It ambivates through both paths. In more detail (see equation 2.18), it has amplitude $\langle z+|x-\rangle = -\frac{1}{\sqrt{2}}$ to take the upper path and amplitude $\langle z-|x-\rangle = \frac{1}{\sqrt{2}}$ to take the lower path, but it doesn't take a path. While traversing the interferometer, this atom has no position in the same way that love has no color.

A particle trapped in an infinite square well has state $\eta_6(x)$. This particle has energy E_6.

A different particle trapped in that same infinite square well has state

$$\frac{1}{\sqrt{2}}\eta_3(x) - \frac{1}{\sqrt{2}}\eta_4(x).$$

This particle does not have an energy. In more detail, it has amplitude $\frac{1}{\sqrt{2}}$ to have energy E_3 and amplitude $-\frac{1}{\sqrt{2}}$ to have energy E_4, but it doesn't have an energy in the same way that love doesn't have a color.

Problems

7.1 Quantal recurrence

In section 7.4.1 we found the quantal recurrence time for any initial state in the infinite square well. (Remarkably, we found it knowing only the energy eigenvalues. . . we did not exploit our knowledge of the energy eigenfunctions.) What happens after one-half of a recurrence time has passed? (This part requires some knowledge of the energy eigenfunctions.)

7.2 Time evolution of an average

Recall that for n odd, the energy eigenfunction is even under reflection about the center of the well, whereas for n even, the energy eigenfunction is odd under such reflection. Prove the following: If the initial wavefunction is a superpositon of only odd-numbered energy eigenfunctions, then as time goes on the probability density dances merrily, but the mean position is always in the exact center of the well. What if the initial wavefunction is a superpostion of only even-numbered energy eigenfunctions?

7.3 Explore time evolution

Equation (7.9) contains all there is to know about time evolution in the infinite square well. But that knowledge is hidden and hard to unpack. Find on the Internet a computer simulation that displays this time evolution. (I recommend the simulation "Infinite Square Well: Wave Packet Dynamics", with the initial condition "Start p0 = 40pi", part of *Physlet Quantum Physics* by Mario Belloni, Wolfgang Christian, and Anne J. Cox. However — depending upon the unpredictable trajectory of computer technology — by the time you read this book, that simulation might be unavailable.) Run the simulation through at least one recurrence time (7.12) and write a few sentences recording your impressions. Here is a scattering of my impressions, which might help you get started:

> This behavior is very rich, in stark contrast to the classical behavior for the same system.
>
> All that richness is packed into one tiny equation!
>
> The system does *not* time evolve into the ground state.
>
> All that richness derives from simple ideas about hopping from bin to adjacent bin (6.15), through the rules for combining amplitudes in series and in parallel (page 60). It's like the game of chess, where simple rules are applied over and over again to produce a complex and subtle game.
>
> "We should take comfort in two conjoined features of nature: first, that our world is incredibly strange and therefore supremely fascinating. . . second, that however bizarre and arcane our world might be, nature remains potentially comprehensible to the human mind."
>
> — Stephen Jay Gould (*Dinosaur in a Haystack*, page 386)

7.4 **Characteristics of the ground energy level**

The ground state energy for the infinite square well is

$$\frac{\pi^2\hbar^2}{2mL^2}.$$

Does it makes sense that. . .

a. . . . this energy vanishes as $\hbar \to 0$? (*Clue:* Consider the classical limit.)

b. . . . this energy vanishes as $L \to \infty$? (*Clue:* Think about the Heisenberg indeterminacy principle. Compare problem **??**.**??**.)

c. . . . this energy varies as $1/m$?

7.5 **Questions** (recommended problem)

Update your list of quantum mechanics questions that you started at problem 1.13 on page 56. Write down new questions and, if you have uncovered answers to any of your old questions, write them down briefly.

Chapter 8

The Free Particle

The free particle (that is, a particle subject to no force) is an imperfect model, just like the infinite square well (see page 218): any real particle in your laboratory will eventually run into the laboratory walls. But atoms are so small relative to laboratory walls that the free particle is a *valuable* imperfect model.

So, how does a free particle behave?

8.1 Strategy

This is the first problem you encountered in introductory classical mechanics, and you know the answer: "A particle in motion remains in uniform motion unless acted upon by a force." Expressed mathematically for a particle in one dimension, the answer is that the position evolves in time as

$$x(t) = x_0 + \frac{p_0}{m}t, \tag{8.1}$$

whereas the momentum is constant

$$p(t) = p_0. \tag{8.2}$$

It's often helpful to shift the coordinate origin so that $x_0 = 0$, and to rotate the coordinate axis so that p_0 is non-negative.

And that's all there is to force-free motion in classical mechanics.

In quantum mechanics the time-development problem is more intricate. The usual approach is the one we've seen many times before: First, solve the energy eigenproblem — we already know how energy eigenstates change with time. Second, use superposition to find out how your particular initial state changes with time.

In mathematical terms, using the position. representation. First find the energy eigenvalues E_n and eigenfunctions $\eta_n(x)$. You know that the eigenfunction evolves in time as

$$e^{-(i/\hbar)E_n t}\eta_n(x). \tag{8.3}$$

Second, using superposition, express the initial wavefunction as

$$\psi_0(x) = \sum_n c_n \eta_n(x), \tag{8.4}$$

where

$$c_n = \int_{-\infty}^{+\infty} \eta_n^*(x)\psi_0(x)\, dx. \tag{8.5}$$

This wavefunction evolves in time to

$$\psi(x,t) = \sum_n c_n e^{-(i/\hbar)E_n t}\eta_n(t). \tag{8.6}$$

With this expression for the state $\psi(x,t)$ in hand, we can uncover anything we desire: mean position, indeterminacy in position, mean momentum, indeterminacy in momentum, mean energy, indeterminacy in energy, anything. (It might be difficult to do the uncovering, but it is always possible.)

8.2 Apply the strategy to a free particle

When this general strategy is applied to the free particle, there's one lucky break and one unlucky break.

The lucky break is that we've already solved the energy eigenproblem. There is no potential energy function, so the Hamiltonian is nothing but

$$\hat{H} = \frac{\hat{p}^2}{2m}. \tag{8.7}$$

The momentum states $|p_0\rangle$, introduced in section 6.7 ("Position representation of momentum eigenstates"), are also energy eigenstates, with

$$\hat{H}|p_0\rangle = \frac{{p_0}^2}{2m}|p_0\rangle. \tag{8.8}$$

It's worth noting that there's always a degeneracy: The states $|+p_0\rangle$ and $|-p_0\rangle$ share the energy eigenvalue

$$E(p_0) = \frac{{p_0}^2}{2m}. \tag{8.9}$$

The unlucky break is that the eigenvalues p_0 are continuous, not discrete, so whereas equation (8.6) contemplates an infinite sum, for the free particle we will have to execute an infinite integral. In light of equations (6.60) and (6.61), the general strategy above must be modified to

$$\tilde{\psi}_0(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-i(p/\hbar)x} \psi_0(x)\, dx \tag{8.10}$$

$$\psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{+i(p/\hbar)x} e^{-(i/\hbar)E(p)t} \tilde{\psi}_0(p)\, dp. \tag{8.11}$$

8.3 Time evolution of the energy eigenfunction

Before jumping into time evolution for an arbitrary initial wavefunction, we investigate the time evolution of an energy state with with momentum p_0. (We'll rotate the coordinate axis to make $p_0 \geq 0$.) Following equation (6.52), we write the wavefunction for this state as

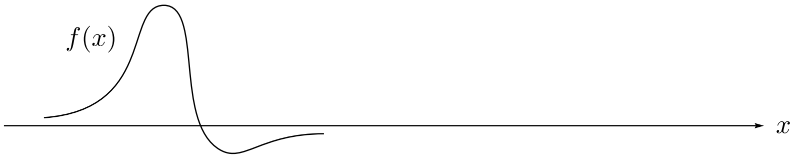
$$\pi_0(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{i(p_0/\hbar)x}, \tag{8.12}$$

and it evolves in time to

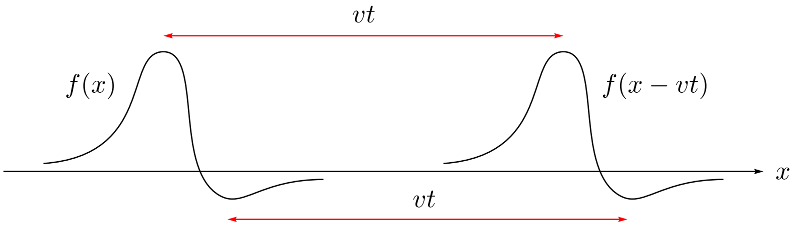
$$\begin{aligned} & \frac{1}{\sqrt{2\pi\hbar}} e^{-(i/\hbar)E(p_0)t} e^{i(p_0/\hbar)x} \\ = & \frac{1}{\sqrt{2\pi\hbar}} e^{i[p_0 x - E(p_0)t]/\hbar}. \end{aligned} \tag{8.13}$$

This equation has the form $f(x - vt)$ where $v = E(p_0)/p_0$.

Any function of the form $f(x - vt)$ represents a wave traveling right with wave velocity v. Start with the initial function $f(x)$ at time $t = 0$:



Now, to find the value of $f(x-vt)$ at some point x at a later time t, start at the x, then subtract vt, then find out what the initial function was at the point $x - vt$. The result is the initial function shifted right by a distance vt.



Now of course the function in equation (8.13) is more difficult to plot, because it is complex valued, but it still represents a wave moving right at speed v.

Repeating, equation (8.13) represents a wave moving at

$$\text{wave velocity} = \frac{E(p_0)}{p_0} = \frac{p_0}{2m} \tag{8.14}$$

which is puzzling, because the classical momentum is mV, so this wave velocity is $V/2$. We will resolve this apparent paradox at equation (8.22).

Finally, we connect with the result for classical sinusoidal waves of the form

$$A \sin[kx - \omega t].$$

The imaginary part of wave (8.13) is

$$\frac{1}{\sqrt{2\pi\hbar}} \sin[(p_0/\hbar)x - (E/\hbar)t],$$

and making the correspondence results in

$$k = \frac{p_0}{\hbar} \quad \text{or} \quad \lambda = \frac{h}{p_0}, \tag{8.15}$$

where λ is the so-called de Broglie[1] wavelength, and

$$\omega = \frac{E}{\hbar} \quad \text{or} \quad E = hf \tag{8.16}$$

which is the Einstein formula for quantized energy in terms of frequency f. These two formulas were extraordinary important in the historical development of quantum mechanics, but we use them rarely in this book.

---

[1] Louis de Broglie (1892–1987) was born into the French nobility, and is sometimes called "Prince de Broglie", although I am told that he was actually a duke. He earned an undergraduate degree in history, but then switched into physics and introduced the concept of particle waves in his 1924 Ph.D. thesis.

Problem

8.1 **Energy eigenstates**

This section examined the behavior of a free particle in a state of definite momentum. Such states have a definite energy, but they are not the only possible states of definite energy.

a. Show that the state

$$|\psi(0)\rangle = A|+p_0\rangle + B|-p_0\rangle, \tag{8.17}$$

where $|A|^2 + |B|^2 = 1$, has definite energy $E(p_0) = p_0^2/2m$. (That is, $|\psi(0)\rangle$ is an energy eigenstate with eigenvalue $p_0^2/2m$).

b. Show that the "wavefunction" corresponding to $|\psi(t)\rangle$ evolves in time as

$$\psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\left[Ae^{i(+p_0x-E(p_0)t)/\hbar} + Be^{i(-p_0x-E(p_0)t)/\hbar}\right]. \tag{8.18}$$

I use the term wavefunction in quotes because $\psi(x,t)$ is not $\langle x|\text{normal state}\rangle$ but rather a sum of two terms like $\langle x|\text{rigged basis state}\rangle$.

c. Show that the "probability density" $|\psi(x,t)|^2$ is independent of time and given by

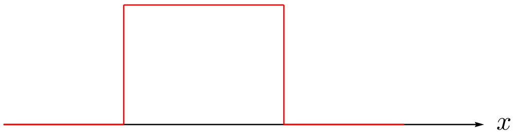
$$|\psi(x,t)|^2 = \frac{1}{2\pi\hbar}\left[1 + 2\,\Re e\{A^*B\}\cos\left(\frac{2p_0x}{\hbar}\right) + 2\,\Im m\{A^*B\}\sin\left(\frac{2p_0x}{\hbar}\right)\right]. \tag{8.19}$$

8.4 Which initial wavefunction should we use?

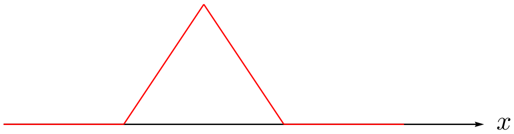
Now we are ready to implement the strategy of equations (8.10) and (8.11). But which initial wavefunction should we use?

This is a matter of choice. For our first problem, I'd like to use an initial wavefunction that is sort of like a classical particle, so that we can compare the classical and quantal results. A classical particle has an exact position and momentum at the same time, and no wavefunction can have that, but I'll seek an initial wavefunction that is pretty-well localized in both position and momentum.

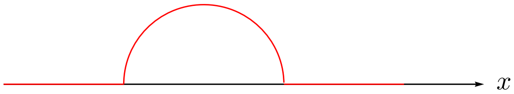
One possibility would be a "top-hat" wavefunction, with the shape



Another would be a "tent" wavefuction,



And a third would be a "semicircular" wavefunction,



All of these are legitimate initial wavefunctions. I happen to know, however, that while all of them have simple descriptions in position space, they have very complicated descriptions in momentum space, so that while integral (8.10) is straightforward, integral (8.11) is horrendous.

We are better off choosing an initial wavefunction with no jumps or kinks. At first we might think to use

$$e^{-x^2},$$

which is pretty-well localized in space. But this formula can't be correct as written, because the exponent must be dimensionless, and as written the exponent has the dimensions of $[\text{length}]^2$. So our second thought is to use

$$e^{-x^2/\sigma^2},$$

where σ is some parameter with the dimensions of length. Wavefunctions with a big σ are relatively wide, those with a small σ are relatively narrow.

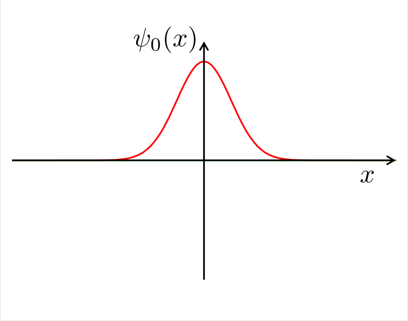
This choice can't be exactly correct, however, because it's dimensionless and unnormalized. Because wavefunction has the dimensions of $1/\sqrt{[\text{length}]}$, and the only length in the problem is σ, we write the normalized version as

$$\frac{A}{\sqrt{\sigma}} e^{-x^2/\sigma^2},$$

where A is a dimensionless number to be determined. Finally, recognizing that wavefunction might be complex, we choose the initial wavefunction

$$\psi_0(x) = \frac{A}{\sqrt{\sigma}} e^{-x^2/\sigma^2} e^{ip_0x/\hbar}, \tag{8.20}$$

where p_0 is some constant with the dimensions of momentum. This is called a "Gaussian[2] wavefunction" or "Gaussian wavepacket". (A "wavepacket" is a wavefunction that is pretty-well localized in space.) For the case $p_0 = 0$, it looks[3] like this



There are, of course, other possible initial wavefuctions. But this is the one I choose to investigate.

8.5 Character of the initial momentum wavefunction

I will give you the satisfaction of working out for yourself the details of the initial momentum wavefunction, and the time evolution (problems 8.2 through 8.5). Here I'll discuss what those details tell us about nature.

---

[2] Carl Friedrich Gauss (1777–1855) is best known as a prolific German mathematician, but also worked in electromagnetism ("the Gaussian system of units"), in astronomy, and in geodesy. He and his colleague Wilhelm Eduard Weber invented but never commercialized the telegraph.

[3] A mathematician looks at this wavefunction and says "As $x \to \pm\infty$, the function approaches but never reaches zero." A physicist looks at the same wavefunction and says "If it's smaller than my ability to measure it, I'll call it zero." When I drew this graph, I said "The line representing the x-axis has a finite width, and when the value of the function is less than that width, the black line representing the axis overlies the red line representing wavefunction."

In problem 8.4, “Static properties of a Gaussian wavepacket”, you will find these properties for the initial wavefunction:

$$\psi_0(x) = \frac{A}{\sqrt{\sigma}} e^{-x^2/\sigma^2} e^{ip_0x/\hbar} \qquad \tilde{\psi}_0(p) = A\sqrt{\frac{\sigma}{2\hbar}}\, e^{-[(p-p_0)\sigma/2\hbar]^2}$$

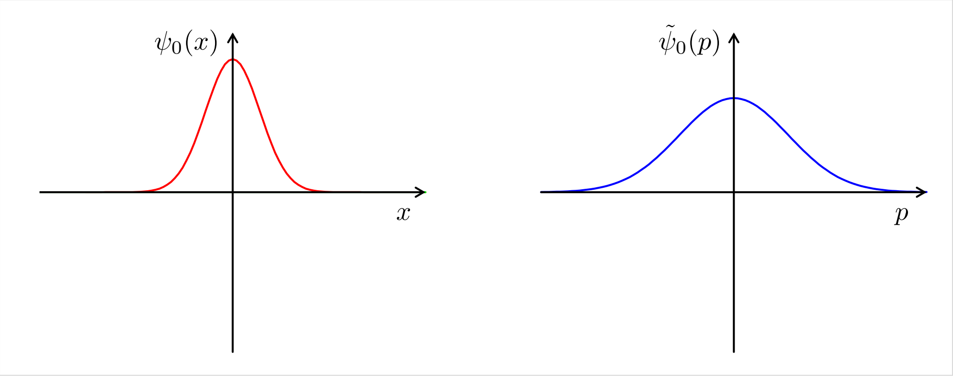
$$|\psi_0(x)|^2 = \frac{A^2}{\sigma} e^{-2x^2/\sigma^2} \qquad |\tilde{\psi}_0(p)|^2 = A^2 \frac{\sigma}{2\hbar}\, e^{-2[(p-p_0)\sigma/2\hbar]^2}$$

$$\langle \hat{x} \rangle = 0 \qquad \langle \hat{p} \rangle = p_0$$

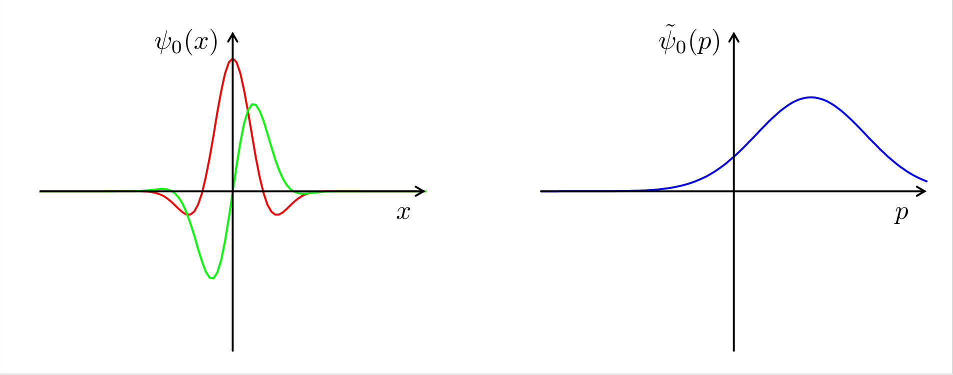
$$\Delta x = \sigma/2 \qquad \Delta p = \hbar/\sigma$$

Remarkably, the momentum wavefunction is pure real. Also remarkably, the Fourier transform of a position wavepacket of Gaussian form is a momentum wavepacket of Gaussian form.

Here are the position and momentum wavefunctions when $p_0 = 0$:



And for $p_0 > 0$:



In this second case the initial position wavefunction is complex-valued: We use red to represent the real part and green to represent the imaginary part.

Notice that the position probability density $|\psi_0(x)|^2$ is *independent* of $\langle \hat{p} \rangle = p_0$. The position probability density tells you everything you might want to know about position, but it tells you *absolutely nothing* about momentum. (Similarly, the momentum probability density $|\tilde{\psi}_0(p)|^2$ tells you everything you might want to know about momentum, but says absolutely nothing about position.)

Some people hold the misconception that the "probability cloud" $|\psi_0(x)|^2$ is the central entity of quantum mechanics. No. Because it is a real, positive probability, rather than a complex probability amplitude, it says nothing about interference. As we see here, it says nothing about momentum and hence cannot determine what will happen in the future. Overemphasis of the "probability cloud" clouds our vision of quantum mechanics.

8.6 Character of the time evolved wavefunction

When you work problem 8.5, "Force-free time evolution of a Gaussian wavepacket", you will find that the mean position (the "peak of the probability density wavepacket") moves with velocity p_0/m and that the probability density wavepacket "spreads out" with ever-increasing position indeterminacy.

We return to the wave velocity paradox presented at equation (8.14). Each Fourier component moves with a different phase velocity

$$\text{phase velocity} = \frac{E(p_0)}{p_0} = \frac{p_0}{2m}. \tag{8.21}$$

The hump moves at a different speed. It moves at the group velocity

$$\text{group velocity} = \frac{p_0}{m}. \tag{8.22}$$

It is the group velocity, and not any of the many different phase velocities of the many different components, that corresponds to the classical velocity.

If you have studied the classical wave equation

$$\frac{\partial^2 \phi(x,t)}{\partial x^2} - \frac{1}{v_w^2}\frac{\partial^2 \phi(x,t)}{\partial t^2},$$

where v_w is the wave velocity, you'll notice a big difference. For the classical wave equation, waves of every shape move at the same speed, namely v_w. Hence every Fourier component moves at the same phase velocity. Hence the group velocity is the same as the phase velocity. Classical wave packets don't "spread out".

Notice again that there is "more to life than probability density". Every initial wavepacket $\psi_0(t)$ has the same probability density, regardless of the value of p_0, yet they will result in vastly different outcomes.

So far, we've been discussing time evolution of the position probability density. What about the momentum probability density? As the position probability density spreads out, does the momentum probability density narrow in? No. The probability For any given momentum p, the phase of $\tilde{\psi}(p,t)$ changes with time, but the magnitude remains vigorously constant. You can see this as a consequence of the Fourier transform computation, but you can see it without computation as a consequence of our third theorem on time evolution, "Time evolution of projection probabilities" (page 164): $|\tilde{\psi}(p,t)|^2 = |\tilde{\psi}(p,0)|^2$ because $[\hat{p}, \hat{H}] = 0$.

8.7 More to do

We have examined the time evolution of a free particle that's initially in a Gaussian wavepacket. Many more questions could be asked. What is the time evolution of the top-hat, tent, and semicircular wavefunctions introduced on page 232? How about the Lorentzian wavepacket of equation (8.37)? How about a Gaussian wavepacket with a more elaborate structure in the exponent, such as

$$\psi_0(x) = \frac{A}{\sqrt{\sigma}} e^{-x^2/\sigma^2} e^{i[p_1 x + p_2 x^2/\sigma]/\hbar} \ ? \tag{8.23}$$

How about an initial wavepacket with two humps rather than one (such as the "double Gaussian")? (When the two humps are far enough apart, this is called a "Schrödinger cat state".)

You can see that we could spend many years investigating such questions. Instead, we change our focus to particles that aren't force-free.

8.8 Problems

8.2 A useful integral

Starting with $\int_{-\infty}^{+\infty} e^{-u^2}\, du = \sqrt{\pi}$, show that

$$\int_{-\infty}^{+\infty} e^{-ax^2+bx}\, dx = \sqrt{\frac{\pi}{a}}\, e^{b^2/4a} \tag{8.24}$$

where a and b are complex numbers with $\Re e\{a\} \geq 0$ and $a \neq 0$. This result is called "the Gaussian integral".

Clue: Complete the square by writing

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

8.3 A somewhat less useful integral

Using $\int_{-\infty}^{+\infty} e^{-u^2}\, du = \sqrt{\pi}$, show that

$$\int_{-\infty}^{+\infty} x^2 e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2}. \tag{8.25}$$

(*Clue:* Integrate by parts.)

8.4 Static properties of a Gaussian wavepacket

Consider the state represented by wavefunction

$$\psi_0(x) = \frac{A}{\sqrt{\sigma}} e^{-x^2/\sigma^2} e^{ip_0 x/\hbar}. \tag{8.26}$$

a. Show that the wavefunction is properly normalized when $A = \sqrt[4]{2/\pi}$.

b. Show that in this state $\langle \hat{x} \rangle = 0$ (trivial), and $\Delta x = \sqrt{\langle (\hat{x} - \langle \hat{x} \rangle)^2 \rangle} = \sigma/2$ (easy).

c. Use the Gaussian integral (8.24) to show that

$$\tilde{\psi}_0(p) = A\sqrt{\frac{\sigma}{2\hbar}}\, e^{-[(p-p_0)\sigma/2\hbar]^2}. \tag{8.27}$$

Remarkably, this momentum-space wavefunction is pure real.

d. Hence show that $\langle \hat{p} \rangle = p_0$ and $\Delta p = \hbar/\sigma$.

e. You know from the Heisenberg indeterminacy principle (6.101) that for any wavefunction $\Delta x \Delta p \geq \frac{1}{2}\hbar$. What is $\Delta x \Delta p$ for this particular Gaussian wavepacket? (Unsurprisingly, it is called a "minimum indeterminacy wavepacket").

8.5 **Force-free time evolution of a Gaussian wavepacket**

A particle with the initial momentum wavefunction $\tilde{\psi}(p,0)$ evolves as

$$\tilde{\psi}(p,t) = e^{-(i/\hbar)E(p)t}\,\tilde{\psi}_0(p), \tag{8.28}$$

where $E(p) = p^2/2m$, so that

$$\psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{i(px-E(p)t)/\hbar}\,\tilde{\psi}_0(p)\,dp. \tag{8.29}$$

a. Plug in initial momentum wavefunction $\tilde{\psi}_0(p)$ given in equation (8.27), change the integration variable to k where $\hbar k = p - p_0$, and show that

$$\psi(x,t) = A\sqrt{\frac{\sigma}{4\pi}}\, e^{i(p_0 x - E(p_0)t)/\hbar} \times \int_{-\infty}^{+\infty} e^{-k^2(\sigma^2/4 + i\hbar t/2m)}\, e^{ik(x-(p_0/m)t)}\,dk. \tag{8.30}$$

(*Clue:* Change variable first to $p' = p - p_0$, then to $k = p'/\hbar$.)

b. Define the complex dimensionless quantity

$$\beta = 1 + i\frac{2\hbar}{\sigma^2 m}t \tag{8.31}$$

and evaluate using the Gaussian integral (8.24), giving

$$\psi(x,t) = A\frac{1}{\sqrt{\sigma\beta}}\, e^{i(p_0 x - E(p_0)t)/\hbar}\, e^{-(x-(p_0/m)t)^2/\sigma^2\beta}. \tag{8.32}$$

c. Hence show that

$$|\psi(x,t)|^2 = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma|\beta|}e^{-2(x-(p_0/m)t)^2/\sigma^2|\beta|^2}. \tag{8.33}$$

By comparing $|\psi(x,t)|^2$ with $|\psi_0(x)|^2$, read off the results

$$\langle \hat{x} \rangle = \frac{p_0}{m}t, \qquad \Delta x = \frac{\sigma|\beta|}{2} = \frac{\sigma}{2}\sqrt{1 + \left(\frac{2\hbar}{\sigma^2 m}t\right)^2}. \tag{8.34}$$

(No computation is required!)

8.6 **Pauli data**

In his famous *Handbuch* article on quantum mechanics,[4] Wolfgang Pauli poses the following question: If for some quantal state you know the position probability density function, you know everything there is to know about position: you know the mean position, the indeterminacy in position, the mean value of x^5, everything. But you know very little about momentum. Meanwhile, if you know the momentum probability density function, you know everything there is to know about momentum: the mean momentum, the indeterminacy in momentum, the mean value of p^5, everything. But you know very little about position.

Now, the position probability density *plus* the momentum probability density are called the "Pauli data". If you know the Pauli data, you know everything about position and everything about momentum. Does that mean you know everything, and hence can determine the wavefunction (up to an overall phase factor)? The answer is no, because you don't know the correlations between position and momentum. This problem presents two different wavefunctions corresponding to two different states that have the same Pauli data but different position-momentum correlation properties.

I claim that for the Gaussian wavefunction

$$\psi(x) = \frac{A}{\sqrt{\sigma}} e^{-(1+i\alpha)x^2/\sigma^2}, \tag{8.35}$$

with α and σ pure real, and $\sigma > 0$, the position probability density $|\psi(x)|^2$ and the momentum probability density $|\tilde{\psi}(p)|^2$ are independent of the sign of α: for example, the wavefunctions with $\alpha = +5$ and with $\alpha = -5$ have the same position and momentum probability densities. Nevertheless these represent two distinct states. For example, they have different mean values for the measurable quantity corresponding to the the Hermitian operator

$$\hat{x}\hat{p} + \hat{p}\hat{x}. \tag{8.36}$$

This measurable quantity is called the position-momentum correlation.

Prove these claims by finding explicit formulas for $|\psi(x)|^2$, $|\tilde{\psi}(p)|^2$, and $\langle \hat{x}\hat{p} + \hat{p}\hat{x} \rangle$.

---

[4] Wolfgang Pauli, "Die allgemeinen Prinzipien der Wellenmechanik," in A. Smekal, editor, *Handbuch der Physik* (Julius Springer, Berlin, 1933), volume 24, part 1, *Quantentheorie*, footnote on page 98. This article was republished with small changes in Siegfried Flügge, editor, *Handbuch der Physik* (Springer-Verlag, Berlin, 1958), volume 5, part 1, footnote on page 17. English translation by P. Achuthan and K. Venkatesan, *General Principles of Quantum Mechanics* (Springer-Verlag, Berlin, 1980), footnote on page 17.

8.7 Non-Gaussian wavepackets

A “Lorentzian” wavepacket has

$$\psi(x) = \frac{A}{x^2 + \gamma^2} e^{ikx}, \tag{8.37}$$

where γ and k are fixed parameters.

a. What is the normalization constant A?
b. What is the mean kinetic energy?

Chapter 9

Energy Eigenproblems

Energy eigenproblems are important: they determine the "allowed" energy eigenvalues, and such energy quantization is the most experimentally accessible facet of quantum mechanics. Also, the most direct way to solve the time evolution problem is to first solve the energy eigenproblem.

In fact, Erwin Schrödinger discovered the energy eigenproblem first (in December 1925) and five months later discovered the time evolution equation, which he called "the true wave equation". Today, both equations carry the name "Schrödinger equation", which can result in confusion.

Now that we've looked at the energy eigenproblem for both the infinite square well and for the free particle, it is time to look at the problem more generally. There are large numbers of analytic and numerical techniques for solving eigenproblems. Most of these are effective but merely technical: they find the answer, but don't give any insight into the character of the resulting energy eigenfunctions. For example, the energy eigenfunctions of the simple harmonic oscillator, are given at equation (10.29). These results are correct, but provide little insight.

This chapter presents two of the many solution techniques available. First we investigate an informal, rough-and-ready technique for sketching energy eigenfunctions that doesn't give rigorous solutions, but that does provide a lot of insight. Second comes a numerical technique of wide applicability.

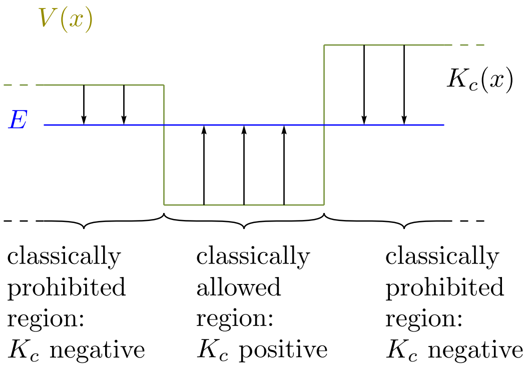
Put both of these techniques into your problem-solving toolkit. You'll find them valuable not only in quantum mechanics, but whenever you need to solve a second-order ordinary differential equation.

9.1 Sketching energy eigenfunctions

Since this chapter is more mathematical than physical in character, I start off by writing the energy eigenequation (6.94) in the mathematically suggestive form

$$\frac{d^2\eta(x)}{dx^2} = -\frac{2m}{\hbar^2}[E - V(x)]\eta(x) = -\frac{2m}{\hbar^2}K_c(x)\eta(x) \tag{9.1}$$

which defines the "classical kinetic energy function" $K_c(x)$. This parallels the potential energy function: $V(x)$ is the potential energy that the classical system would have if the particle were located at x. I'm not saying that the particle *is* classical nor that it *does have* a location; indeed a quantal particle might not have a location. But $V(x)$ is the potential energy that the system would have if it were classical with the particle located at point x. In the same way $K_c(x)$ is the kinetic energy that a classical particle would have if the particle were located at x and total energy were E. Whereas no classical particle can ever have a negative kinetic energy, it is perfectly permissible for the classical kinetic energy function to be negative: in the graph that follows, $K_c(x)$ is negative on the left, positive in the center, and strongly negative on the right.



A region were $K_c(x)$ is positive or zero is called a "classically allowed region"; otherwise it is a "classically prohibited region".

Remember that

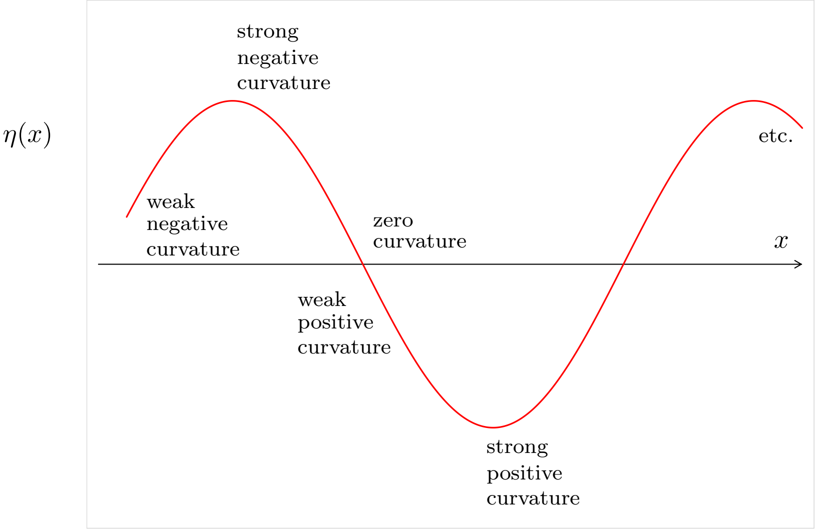
$$\frac{d\eta}{dx} \text{ represents slope;} \qquad \frac{d^2\eta}{dx^2} \text{ represents curvature.}$$

When curvature is positive, the slope increases as x increases (e.g. from negative to positive, or from positive small to positive large). When curvature is negative, the slope decreases as x increases.

Start off by thinking of a **classically allowed region** where $K_c(x)$ is constant and positive. Equation (9.1) says that if $\eta(x)$ is positive, then the curvature is negative, whereas if $\eta(x)$ is negative, then the curvature is positive. Furthermore, the size of the curvature depends on the size of $\eta(x)$:

when $\eta(x)$ is...	curvature is...
strongly positive	strongly negative
weakly positive	weakly negative
zero	zero
weakly negative	weakly positive
strongly negative	strongly positive

These observations allow us to find the character of $\eta(x)$ without finding a formal solution. If at one point $\eta(x)$ is positive with positive slope, then moving to the right $\eta(x)$ will grow because of the positive slope, but that growth rate will decline because of the negative curvature. Eventually the slope becomes zero and then negative, but the curvature continues negative. Because of the negative slope, $\eta(x)$ eventually plunges through $\eta(x) = 0$ (where its curvature is zero) and into regions where $\eta(x)$ is negative and hence the curvature is positive. The process repeats to produce the following graph:



⟦You can solve differential equation (9.1) formally to obtain

$$\eta(x) = A\sin((\sqrt{2mK_c}/\hbar)x + \phi) \tag{9.2}$$

where A and ϕ are adjusted to fit the initial or boundary conditions. In fact, this is exactly the equation that we already solved at (7.4). The formal approach has the advantage of finding an exact expression for the wavelength. The informal approach has the advantage of building your intuition.⟧

The direct way of keeping track of curvature in this classically allowed region is

negative curvature when $\eta(x)$ is positive;
positive curvature when $\eta(x)$ is negative.

But this is sort of clunky: to keep track of curvature, you have to keep track of height. A compact way of keeping track of the signs is that

$$\begin{gathered}\text{in a classically allowed region,}\\ \text{the eigenfunction curves toward the axis.}\end{gathered} \tag{9.3}$$

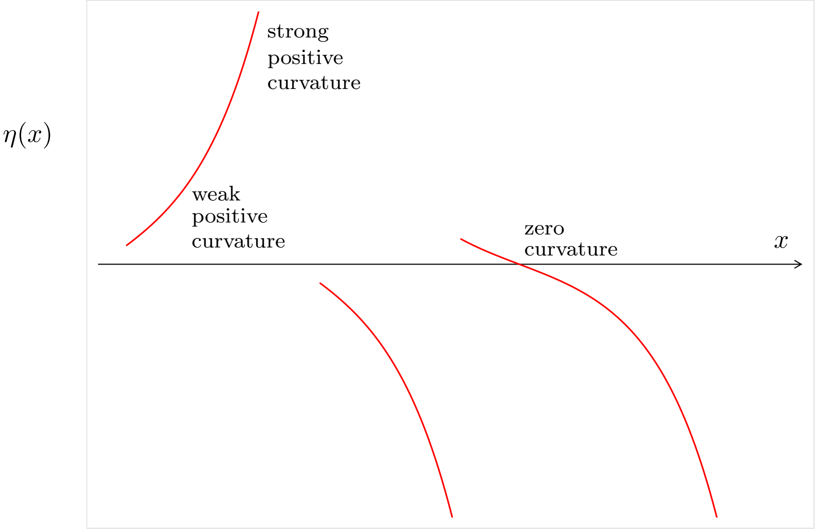
It doesn't *slope* toward the axis, as you can see from the graph, it *curves* toward the axis. Draw a tangent to the energy eigenfunction: in a classically allowed region, the eigenfunction will fall between that tangent line and the axis.

In fact, the informal approach uncovers more than just the oscillatory character of $\eta(x)$. Equation (9.1) shows that when K_c is large and positive, the "curving toward" impetus is strong; when K_c is small and positive, that impetus is weak. Thus when K_c is large, the wavefunction takes tight turns and snaps back toward the axis; when K_c is small, it lethargically bends back toward the axis. And sure enough the formal approach at equation (9.2) shows that the wavelength λ depends on K_c through

$$\lambda = \frac{2\pi\hbar}{\sqrt{2mK_c}}, \tag{9.4}$$

so a large K_c results in a short wavelength — a "tight turn" toward the axis.

Now turn your attention to a **classically prohibited region** where $K_c(x)$ is constant and negative. Equation (9.1) says that if $\eta(x)$ is positive, then the curvature is positive. Once again we can uncover the character of $\eta(x)$ without finding a formal solution. If at one point $\eta(x)$ is positive with positive slope, then moving to the right $\eta(x)$ will grow because of the positive slope, and that growth rate increases because of the positive curvature. The slope becomes larger and larger and $\eta(x)$ rockets to infinity. Or, if $\eta(x)$ starts out negative with negative slope, then it rockets down to negative infinity. Or, if $\eta(x)$ starts out positive with negative slope, it might cross the axis before rocketing down to negative infinity, or it might dip down toward the axis without crossing it, before rocketing up to positive infinity.



⟦You can solve differential equation (9.1) formally to obtain

$$\eta(x) = Ae^{+(\sqrt{2m|K_c|}/\hbar)x} + Be^{-(\sqrt{2m|K_c|}/\hbar)x}$$

where A and B are adjusted to fit the initial or boundary conditions.⟧

The direct way of keeping track of curvature in this classically prohibited region is

positive curvature when $\eta(x)$ is positive;
negative curvature when $\eta(x)$ is negative.

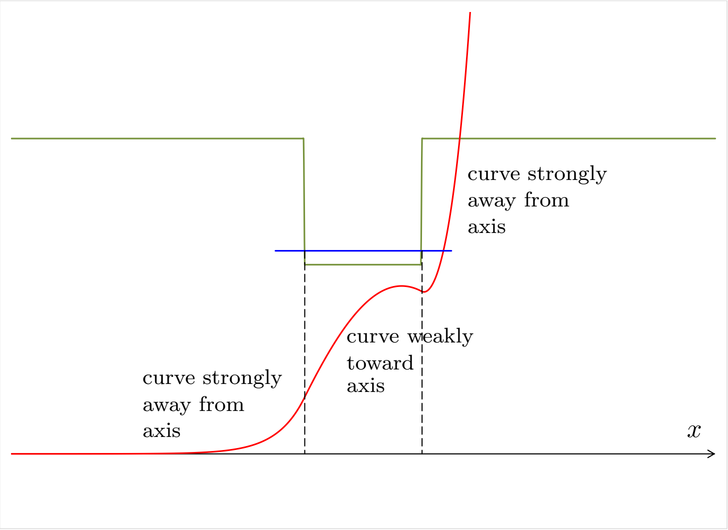
But a compact way is remembering that

in a classically prohibited region, (9.5)
the eigenfunction curves away from the axis.

Draw a tangent to the energy eigenfunction: in a classically prohibited region, that tangent line will fall between the eigenfunction and the axis.

Let's apply all these ideas to finding the character of energy eigenfunctions in a **finite square well**. Solve differential equation (9.1) for an energy E just above the bottom of the well. (I will draw the potential energy function in olive green, the energy E in blue, and the solution $\eta(x)$ in red.)

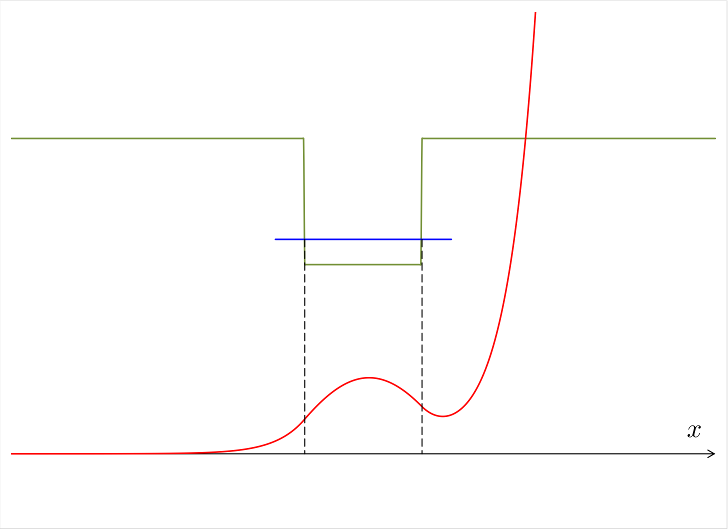
Suppose the wavefunction starts out on the left small and just above the axis. The region is strongly prohibited, that is $K_c(x)$ is strongly negative, so $\eta(x)$ curves strongly away from the axis. Then (at the dashed vertical line) the solution moves into a classically allowed region. But $K_c(x)$ is only weakly positive, so $\eta(x)$ curves only weakly toward the axis. By the time the solution gets to the right-hand classically prohibited region at the next dashed vertical line, $\eta(x)$ has only a weakly negative slope. In the prohibited region the slope increases as $\eta(x)$ curves strongly away from the axis and rockets off to infinity.



You should check that the curvatures and tangents of this energy eigenfunction strictly obey the rules set down at (9.3) and (9.5). What happens when $\eta(x)$ crosses a dashed vertical line, the boundary between a classically prohibited and a classically allowed region?

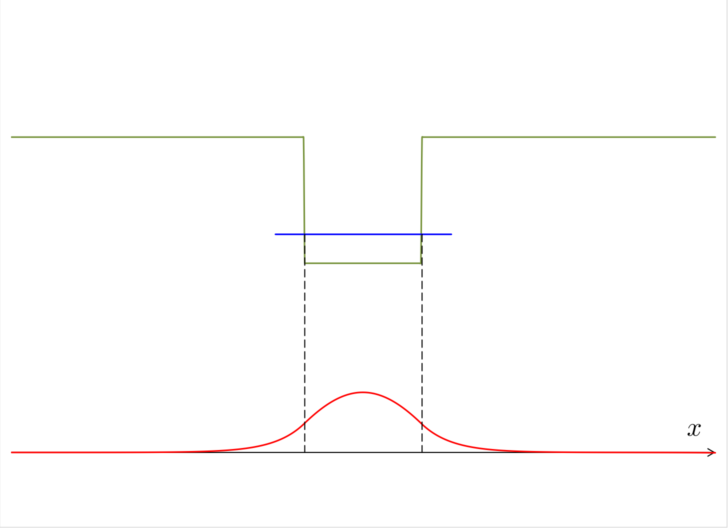
If you have studied differential equations you know that for any value of E, equation (9.1) has two linearly independent solutions. We've just sketched one of them. The other is the mirror image of it: small to the right and rocketing to infinity toward the left. Because of the "rocketing off to infinity" neither solution is normalizable. So these two solutions don't correspond to any physical energy eigenstate. To find such a solution we have to try a different energy.

So we try an energy slightly higher. Now the region on the left is not so strongly prohibited as it was before, so $\eta(x)$ curves away from the axis less dramatically. Then when it reaches the classically allowed region it curves more sharply toward the axis, so that it's strongly sloping downward when it reaches the right-hand prohibited region. But not strongly enough: it curves away from the axis and again rockets off to infinity — although this time not so dramatically.

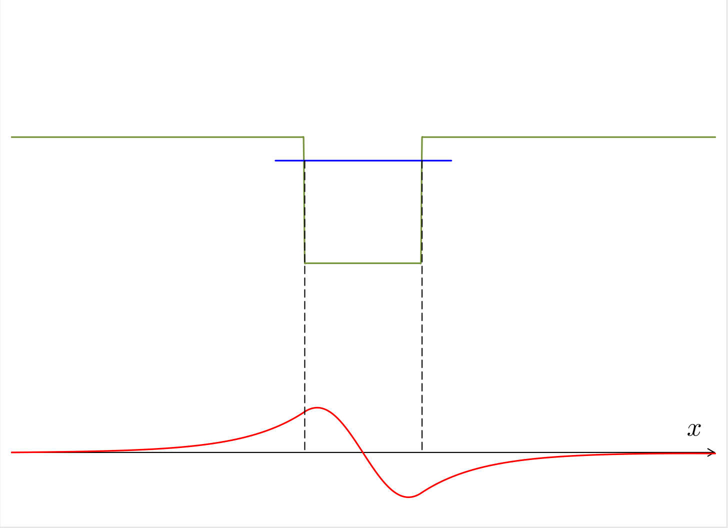


Once again we find a solution (and its mirror image is also a solution), but it's a non-physical, unnormalizable solution.

As we try energies higher and higher, the "rocketing to infinity" happens further and further to the right, until at one special energy it doesn't happen at all. Now the wavefunction *is* normalizable, and now we have found an energy eigenfunction.



What happens when we try an energy slightly higher still? At the right-hand side the wavefunction now rockets off to *negative* infinity! With increased energies, the wavefunction rockets down to negative infinity with increased drama. But then at some point, the drama decreases: as the energy rises the wavefunction continues to go to negative infinity, but it does so more and more slowly. Finally at one special energy the wavefunction settles down exactly to zero as $x \to \infty$, and we've found a second energy eigenfunction.

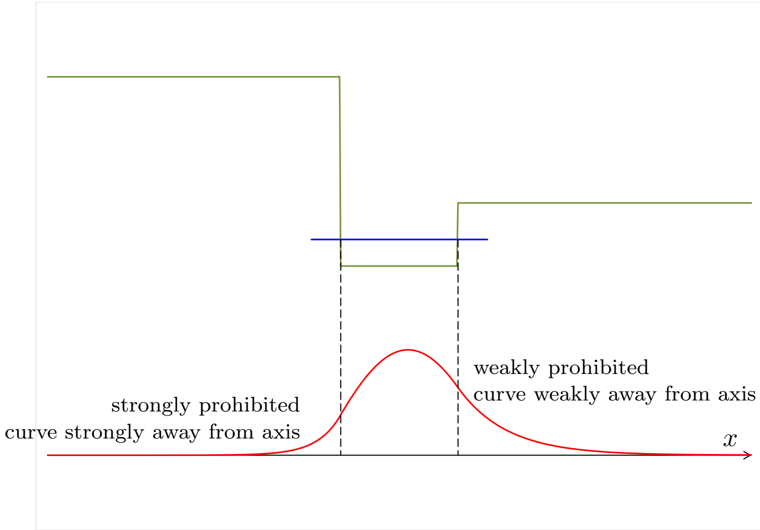


(The misconception concerning “pointlike particles moving across a node”, discussed on page 222, applies to this state as well.)

The process continues: with still higher values of E, the wavefunction $\eta(x)$ diverges to positive infinity as $x \to \infty$ until we reach a third special energy eigenvalue, then to negative infinity until we reach a fourth. Higher and higher energies result in higher and higher values of K_c and hence stronger and stronger snaps back toward the axis. The first (lowest) eigenfunction has no nodes, the second has one node, the third will have two nodes, and in general the nth energy eigenfunction will have $n - 1$ nodes. (See also the discussion on page 254.)

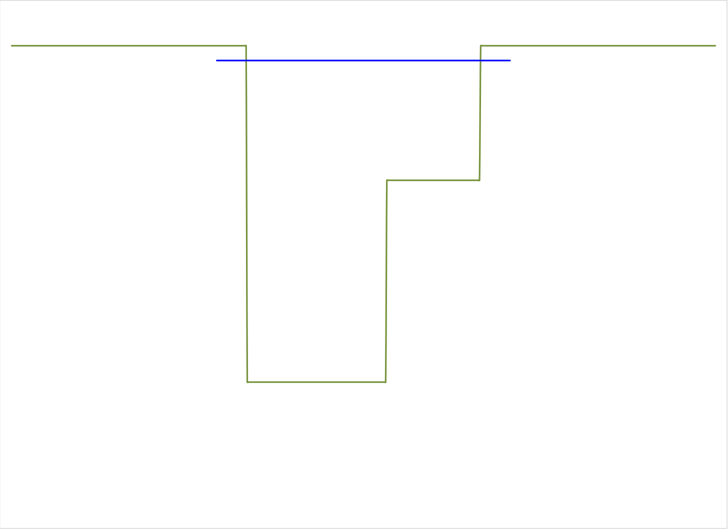
Notice that for a potential energy function symmetric about a point, the energy eigenfunction is *either* symmetric or antisymmetric about that point. The energy eigenfunction does *not* need to possess the same symmetry as the potential energy function. (See also problem 9.6, “Parity”.)

What about a **“lopsided” square well** that lacks symmetry? In the case sketched below the energy is strongly prohibited to the left, weakly prohibited to the right. Hence the wavefunction curves away sharply to the left, mildly to the right. The consequence is that the tail is short on the left, long on the right.

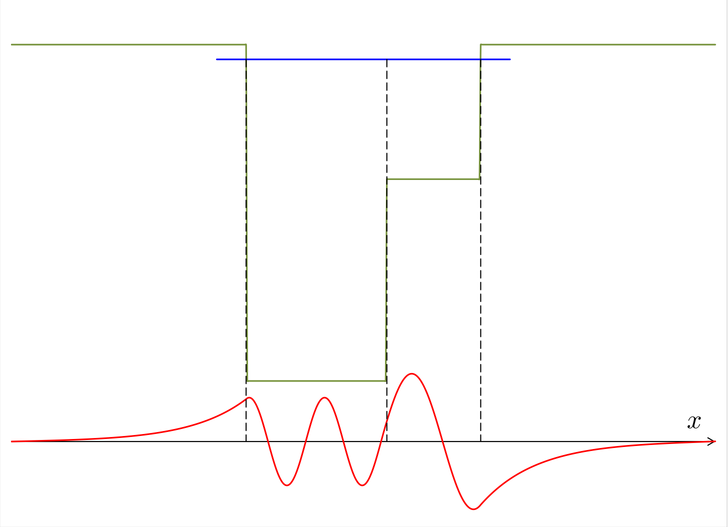


In some way it makes sense that the wavefunction tail should be longer where the classical prohibition is milder.

Now try a **square well with two different floor levels**:

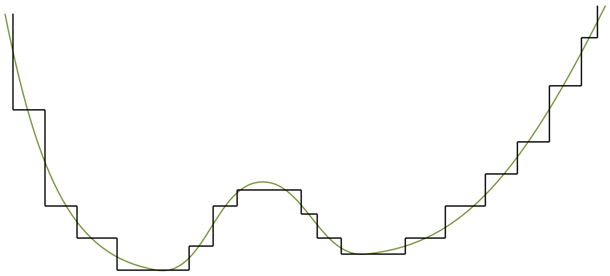


Within the deep left side of the well, K_c is relatively high, so the tendency for η to curve toward the axis is strong; within the shallow right side K_c is relatively low, so the tendency to curve toward the axis is weak. Thus within the deep side of the well, $\eta(x)$ snaps back toward the axis, taking the curves like an expertly driven sports car; within the shallow side $\eta(x)$ leisurely curves back toward the axis, curving like a student driver in a station wagon. Within the deep side, wavelength will be short and amplitude will be small; within the shallow side, wavelength will be longer and amplitude will be large (or at least the same size). One finds smaller amplitude at the deeper side of the well, and hence, all other things being equal, smaller probability for the particle to be in the deep side of the well.



This might seem counterintuitive: Shouldn't it be more probable for the particle to be in the deep side? After all, if you throw a classical marble into a bowl it comes to rest at the deepest point and spends most of its time there. The problem with this analogy is that it compares a classical marble rolling with friction to a quantal situation without friction. Imagine a classical marble rolling instead in a frictionless bowl: it never does come to rest at the deepest point of the bowl. In fact, at the deepest point it moves fastest: the marble spends little time at the deepest point and a lot of time near the edges, where it moves slowly. The classical and quantal pictures don't correspond exactly (there's no such thing as an energy eigenstate in classical mechanics, the classical marble always has a position, and its description never has a node), but the two pictures agree that the particle has high probability of appearing where the potential energy function is shallow, not deep.

Similar results hold for three-level square wells, for four-level square wells, and so forth. And because any potential energy function can be approximated by a series of steps, similar results hold for any potential energy function.



Number of nodes. For the infinite square well, the energy eigenfunction $\eta_n(x)$ has $n-1$ interior nodes. The following argument[1] shows that same holds for any one-dimensional potential energy function $V(x)$. Imagine a modified potential

$$V_a(x) = \begin{cases} \infty & x \leq -a \\ V(x) & -a < x < +a \\ \infty & +a \leq x \end{cases} .$$

When a is very small this is virtually an infinite square well, whose energy eigenfunctions we know. As a grows larger and larger, this potential becomes more and more like the potential of interest $V(x)$. During this expansion, can an extra node pop into an energy eigenfunction? If it does, then at the point x_p where it pops in the wavefunction vanishes, $\eta(x_p) = 0$, and its slope vanishes, $\eta'(x_p) = 0$. But the energy eigenproblem is a second-order ordinary differential equation: the only solution with $\eta(x_p) = 0$ and $\eta'(x_p) = 0$ is $\eta(x) = 0$ everywhere. This is not an eigenfunction. This can never happen.

---

[1] M. Moriconi, "Nodes of wavefunctions" *American Journal of Physics* **75** (March 2007) 284–285.

Summary

In classically prohibited regions, the eigenfunction magnitude declines while stepping away from the well: the stronger the prohibition, the more rapid the decline.

In classically allowed regions, the eigenfunction oscillates: in regions that are classically fast, the oscillation has small amplitude and short wavelength; in regions that are classically slow, the oscillation has large amplitude and long wavelength.

If the potential energy function is symmetric under reflection about a point, the eigenfunction will be either symmetric or antisymmetric under the same reflection.

The nth energy eigenfunction has $n - 1$ nodes.

Quantum mechanics involves situations (very small) and phenomena (interference, entanglement) remote from daily experience. And the energy eigenproblem, so central to quantum mechanics, does not arise in classical mechanics at all. Some people conclude from these facts that one cannot develop intuition about quantum mechanics, but that is false: the techniques of this section *do* allow you to develop a feel for the character of energy eigenstates. Just as chess playing or figure skating must be studied and practiced to develop proficiency, so quantum mechanics must be studied and practiced to develop intuition. If people don't develop intuition regarding quantum mechanics, it's not because quantum mechanics is intrinsically fantastic; it's because these people never try.

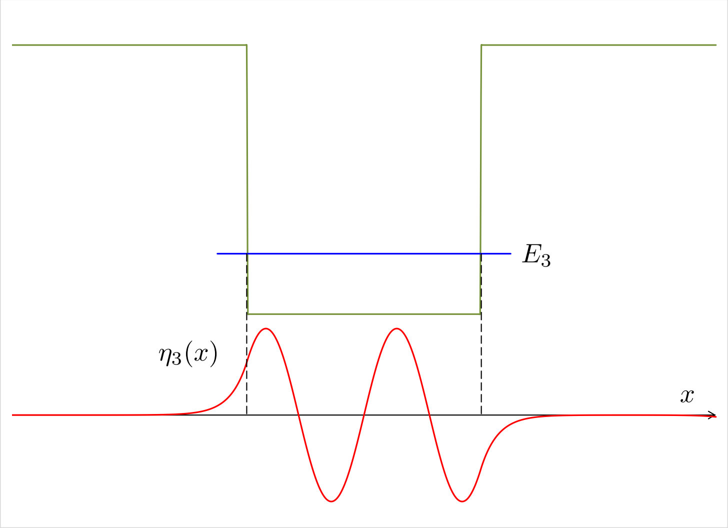
Problems

9.1 **Would you buy a used eigenfunction from this man?**
(recommended problem)
The four drawings below and on the next pages show four one-dimensional potential energy functions $V(x)$ (in olive green) along with candidate energy eigenfunctions $\eta(x)$ (in red) that purport to associate with those potential energy functions. There is something wrong with every candidate. Using the letter codes below, identify all eigenfunction errors, *and* sketch a qualitatively correct eigenfunction for each potential.

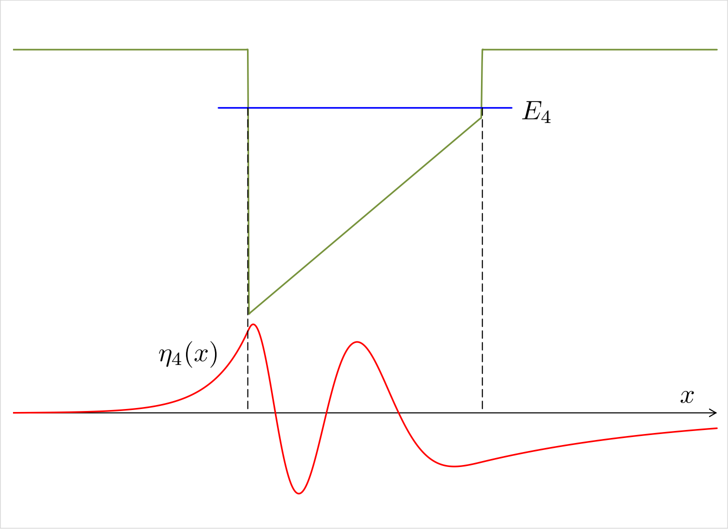
The energy eigenfunction is drawn incorrectly because:

A. Wrong curvature. (It curves toward the axis in a classically prohibited region or away from the axis in a classically allowed region.)
B. Its wavy part has the wrong number of nodes.
C. The amplitude of the wavy part varies incorrectly.
D. The wavelength of the wavy part varies incorrectly.
E. One or more of the declining tails has the wrong length.

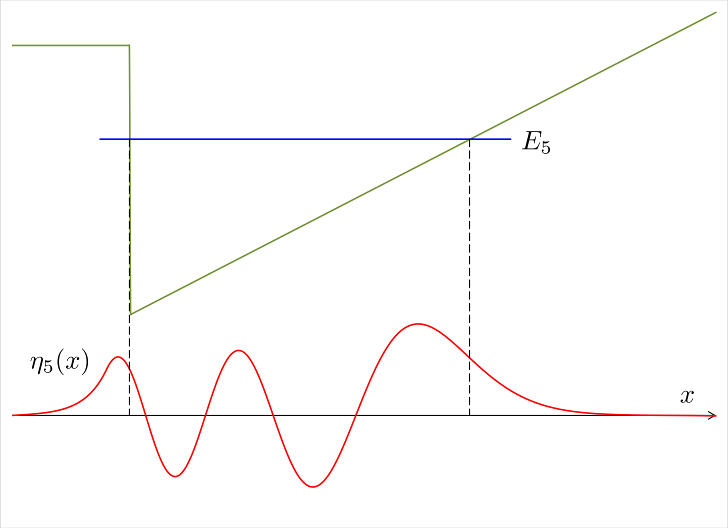
a.



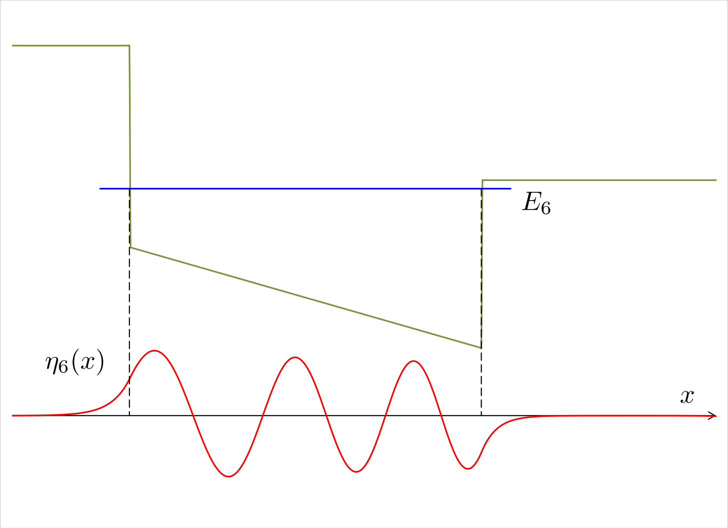
b.



c.

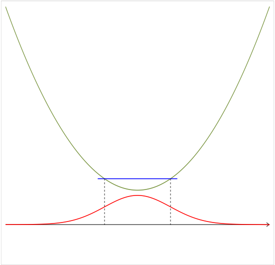


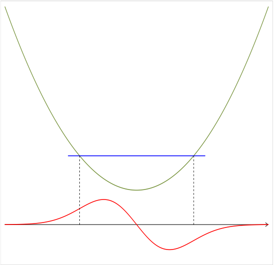
d.

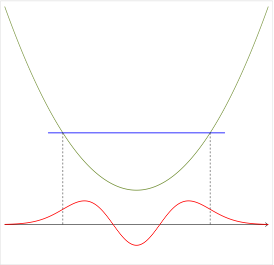


9.2 Simple harmonic oscillator energy eigenfunctions

Here are sketches of the three lowest-energy eigenfunctions for the potential energy function $V(x) = \frac{1}{2}kx^2$ (called the "simple harmonic oscillator"). In eight sentences or fewer, describe how these energy eigenfunctions do (or don't!) display the characteristics discussed in the summary on page 255.







9.3 **Wavelength as a function of K_c**

Before equation (9.4) we provided an informal argument that the wavelength λ would decrease with increasing K_c. This argument didn't say whether λ would vary as $1/K_c$, or as $1/\sqrt{K_c}$, or even as $e^{-K_c/(\text{constant})}$. Produce a dimensional argument showing that if λ depends only on $\hbar$, m, and K_c, then it must vary as $\hbar/\sqrt{mK_c}$.

9.4 **"At least the same size amplitude"**

Page 252 claims that in the two-level square well, the amplitude of $\eta(x)$ on the right would be larger "or at least the same size" as the amplitude on the left. Under what conditions will the amplitude be the same size?

9.5 **Placement of nodes**

Let $\eta_n(x)$ and $\eta_m(x)$ be solutions to

$$-\frac{\hbar^2}{2M}\eta_m''(x) + V(x)\eta_m(x) = E_m\eta_m(x) \tag{9.6}$$

$$-\frac{\hbar^2}{2M}\eta_n''(x) + V(x)\eta_n(x) = E_n\eta_n(x) \tag{9.7}$$

with $E_m > E_n$. The Sturm comparison theorem states that between any two nodes of $\eta_n(x)$ there exists at least one node of $\eta_m(x)$. Prove the theorem through contradiction by following these steps:

a. Multiply (9.6) by η_n, multiply (9.7) by η_m, and subtract to show that

$$-\frac{\hbar^2}{2M}[\eta_m'(x)\eta_n(x) - \eta_m(x)\eta_n'(x)]' = (E_m - E_n)\eta_m(x)\eta_n(x). \tag{9.8}$$

b. Call two adjacent nodes of $\eta_n(x)$ by the names x_1 and x_2. Argue that we can select $\eta_n(x)$ to be always positive for $x_1 < x < x_2$, and show that with this selection $\eta_n'(x_1) > 0$ while $\eta_n'(x_2) < 0$.

c. Integrate equation (9.8) from x_1 to x_2, producing

$$\begin{aligned}-\frac{\hbar^2}{2M}&[-\eta_m(x_2)\eta_n'(x_2) + \eta_m(x_1)\eta_n'(x_1)]\\ &= (E_m - E_n)\int_{x_1}^{x_2}\eta_m(x)\eta_n(x)\,dx.\end{aligned} \tag{9.9}$$

d. If $\eta_m(x)$ does *not* have a zero within $x_1 < x < x_2$, then argue that we can select $\eta_m(x)$ always positive on the same interval, including the endpoints.

The assumption that "$\eta_m(x)$ does *not* have a zero" hence implies that the left-hand side of (9.9) is strictly negative, while the right-hand side is strictly positive. This assumption, therefore, must be false.

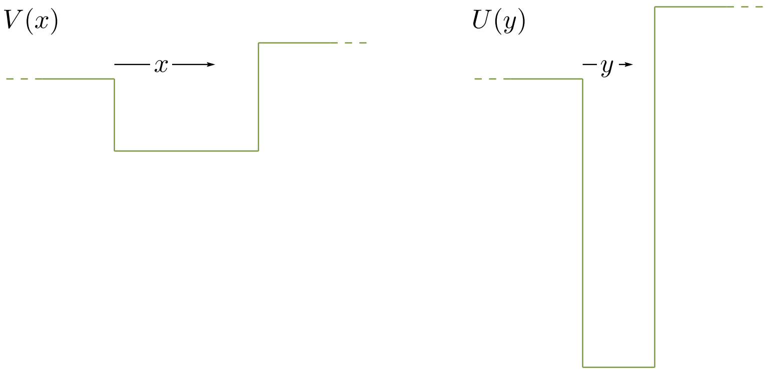
9.6 **Parity**

a. Think of an arbitrary potential energy function $V(x)$. Now think of its mirror image potential energy function $U(x) = V(-x)$ Show that if $\eta(x)$ is an eigenfunction of $V(x)$ with energy E, then $\sigma(x) = \eta(-x)$ is an eigenfunction of $U(x)$ with the same energy.
b. If $V(x)$ is symmetric under reflection about the origin, that is $U(x) = V(x)$, you might think that $\sigma(x) = \eta(x)$. But no! This identification ignores global phase freedom (pages 75 and **??**). Show that in fact $\sigma(x) = r\eta(x)$ where the "overall phase factor" r is a complex number with magnitude 1.
c. The overall phase factor r is a number, not a function of x: the same phase factor r applies at $x = 2$ ($\eta(-2) = r\eta(2)$), at $x = 7$ ($\eta(-7) = r\eta(7)$), and at $x = -2$ ($\eta(2) = r\eta(-2)$). Conclude that r can't be any old complex number with magnitude 1, it must be either $+1$ or -1.

Energy eigenfunctions symmetric under reflection, $\eta(x) = \eta(-x)$, are said to have "even parity" while those antisymmetric under reflection, $\eta(x) = -\eta(-x)$, are said to have "odd parity".

9.7 **Scaling**

Think of an arbitrary potential energy function $V(x)$, for example perhaps the one sketched on the left below. Now think of another potential energy function $U(y)$ that is half the width and four times the depth/height of $V(x)$, namely $U(y) = 4V(x)$ where $y = x/2$. Without solving the energy eigenproblem for either $V(x)$ or $U(y)$, I want to find how the energy eigenvalues of $U(y)$ relate to those of $V(x)$.



Show that if $\eta(x)$ is an eigenfunction of $V(x)$ with energy E, then $\sigma(y) = \eta(x)$ is an eigenfunction of $U(y)$. What is the corresponding energy? After working this problem for the scale factor 2, repeat for a general scale factor s so that $U(y) = s^2V(x)$ where $y = x/s$.

⟦This problem has a different cast from most: instead of giving you a problem and asking you to solve it, I'm asking you to find the relationship between the solutions of two different problems, neither of which you've solved. My thesis adviser, Michael Fisher, called this "Juicing an orange without breaking its peel."⟧

9.2 Scaled quantities

Look again at the quantal energy eigenproblem (9.1)

$$\frac{d^2\eta(x)}{dx^2} = -\frac{2m}{\hbar^2}[E - V(x)]\eta(x). \tag{9.10}$$

Suppose you want to write a computer program to solve this problem for the lopsided square well with potential energy function

$$V(x) = \begin{cases} V_1 & x < 0 \\ 0 & 0 < x < L \\ V_2 & L < x \end{cases} . \tag{9.11}$$

The program would have to take as input the particle mass m, the energy E, the potential well length L, and the potential energy values V_1 and V_2. Five parameters! Once the program is written, you'd have to spend a lot of time typing in these parameters and exploring the five-dimensional parameter space to find interesting values. Furthermore, these parameters have inconvenient magnitudes like the electron's mass 9.11×10^{-31} kg or the length of a typical carbon nanotube 1.41×10^{-10} m. Isn't there an easier way to set up this problem?

There is. The characteristic length for this problem is L. If you try to combine the parameters L, m, and $\hbar$ to form a quantity with the dimensions of energy (see sample problem 9.2.1 on page 264) you will find that there is only one way: this problem's characteristic energy is $E_c = \hbar^2/mL^2$. Define the dimensionless length variable $\tilde{x} = x/L$, the dimensionless energy parameter $\tilde{E} = E/E_c$, and the dimensionless potential energy function $\tilde{V}(\tilde{x}) = V(\tilde{x}L)/E_c = V(x)/E_c$.

In terms of these new so-called "scaled quantities" the quantal energy eigenproblem is

$$\frac{d^2\eta(\tilde{x})}{d\tilde{x}^2}\frac{1}{L^2} = -\frac{2m}{\hbar^2}\left[\frac{\hbar^2}{mL^2}\right][\tilde{E} - \tilde{V}(\tilde{x})]\eta(\tilde{x})$$

or

$$\frac{d^2\eta(\tilde{x})}{d\tilde{x}^2} = -2[\tilde{E} - \tilde{V}(\tilde{x})]\eta(\tilde{x}) \tag{9.12}$$

where

$$\tilde{V}(\tilde{x}) = \begin{cases} \tilde{V}_1 & \tilde{x} < 0 \\ 0 & 0 < \tilde{x} < 1 \\ \tilde{V}_2 & 1 < \tilde{x} \end{cases} . \tag{9.13}$$

The scaled problem has many advantages. Instead of five there are only three parameters: $\tilde{E}$, $\tilde{V_1}$, and $\tilde{V_2}$. And those parameters have nicely sized values like 1 or 0.5 or 6. But it has the disadvantage that you have to write down all those tildes. Because no one likes to write down tildes, we just drop them, writing the problem as

$$\frac{d^2\eta(x)}{dx^2} = -2[E - V(x)]\eta(x) \tag{9.14}$$

where

$$V(x) = \begin{cases} V_1 & x < 0 \\ 0 & 0 < x < 1 \\ V_2 & 1 < x \end{cases} \tag{9.15}$$

and saying that these equations are written down "using scaled quantities".

When you compare these equations with equations (9.10) and (9.11), you see that we would get the same result if we had simply said "let $\hbar = m = L = 1$". This phrase as stated is of course absurd: $\hbar$ is *not* equal to 1; $\hbar$, m, and L *do* have dimensions. But some people don't like to explain what they're doing so they do say this as shorthand. Whenever you hear this phrase, remember that it covers up a more elaborate — and more interesting — truth.

9.2.1 *Sample Problem: Characteristic energy*

Show that there is only one way to combine the quantities L, m, and $\hbar$ to form a quantity with the dimensions of energy, and find an expression for this so-called characteristic energy E_c.

Solution:

quantity	dimensions
L	[length]
m	[mass]
$\hbar$	[mass] × $[\text{length}]^2/[\text{time}]$
E_c	[mass] × $[\text{length}]^2/[\text{time}]^2$

If we are to build E_c out of L, m, and $\hbar$, we must start with $\hbar$, because that's the only source of the dimension [time]. And in fact we must start with $\hbar^2$, because that's the only way to make a $[\text{time}]^2$.

quantity	dimensions
L	[length]
m	[mass]
$\hbar^2$	$[\text{mass}]^2 \times [\text{length}]^4/[\text{time}]^2$
E_c	[mass] × $[\text{length}]^2/[\text{time}]^2$

But $\hbar^2$ has too many factors of [mass] and [length] to make an energy. There is only one way to get rid of them: to divide by m once and by L twice.

quantity	dimensions
$\hbar^2/mL^2$	[mass] × $[\text{length}]^2/[\text{time}]^2$
E_c	[mass] × $[\text{length}]^2/[\text{time}]^2$

There is only one possible characteristic energy, and it is $E_c = \hbar^2/mL^2$.

Problems

9.8 **Characteristic time**

Find the characteristic time for the square well problem by combining the parameters L, m, and $\hbar$ to form a quantity with the dimensions of time. Compare this characteristic time to the infinite square well recurrence time found at equation (7.12).

9.9 **Scaling for the simple harmonic oscillator**

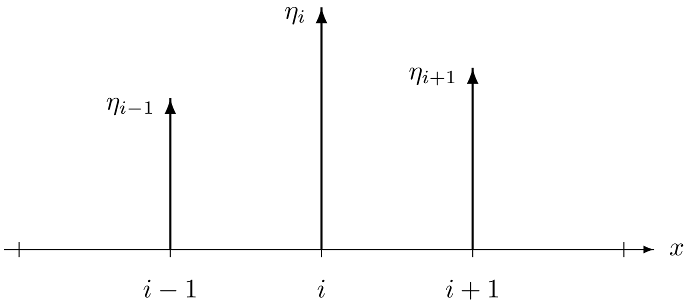
(recommended problem)

Execute the scaling strategy for the simple harmonic oscillator potential energy function $V(x) = \frac{1}{2}kx^2$. What is the characteristic length in terms of k, $\hbar$, and m? What is the resulting scaled energy eigenproblem? If you didn't like to explain what you were doing, how would you use shorthand to describe the result of this scaling strategy?

9.3 Numerical solution of the energy eigenproblem

Now that the quantities are scaled, we return to our task of writing a computer program to solve, numerically, the energy eigenproblem. In order to fit the potential energy function $V(x)$ and the energy eigenfunction $\eta(x)$ into a finite computer, we must of course approximate those continuous functions through their values on a finite grid. The grid points are separated by a small quantity Δ. It is straightforward to replace the function $V(x)$ with grid values V_i and the function $\eta(x)$ with grid values η_i. But what should we do with the second derivative $d^2\eta/dx^2$?

Start with a representation of the grid function η_i:

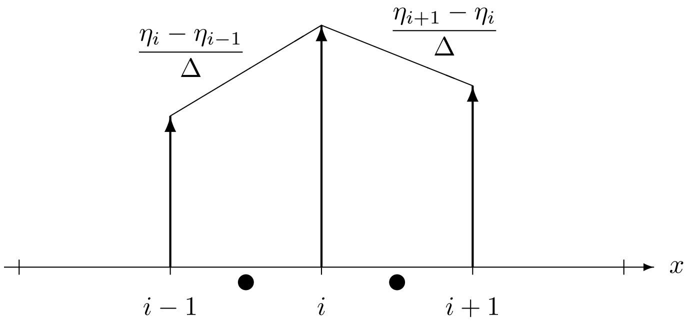


The slope at a point halfway between points $i-1$ and i (represented by the left dot in the figure below) is approximately

$$\frac{\eta_i - \eta_{i-1}}{\Delta},$$

while the slope half way between the points i and $i+1$ (represented by the right dot) is approximately

$$\frac{\eta_{i+1} - \eta_i}{\Delta}.$$



An approximation for the second derivative at point i is the change in slope divided by the change in distance

$$\frac{\dfrac{\eta_{i+1} - \eta_i}{\Delta} - \dfrac{\eta_i - \eta_{i-1}}{\Delta}}{\Delta}$$

so at point i we approximate

$$\frac{d^2\eta}{dx^2} \approx \frac{\eta_{i+1} - 2\eta_i + \eta_{i-1}}{\Delta^2}. \tag{9.16}$$

The discretized version of the energy eigenproblem (9.14) is thus

$$\frac{\eta_{i+1} - 2\eta_i + \eta_{i-1}}{\Delta^2} = -2[E - V_i]\eta_i \tag{9.17}$$

which rearranges to

$$\eta_{i+1} = 2[1 + \Delta^2(V_i - E)]\eta_i - \eta_{i-1}. \tag{9.18}$$

The algorithm then proceeds from left to right. Start in a classically prohibited region and select $\eta_1 = 0$, $\eta_2 = 0.001$. Then find

$$\eta_3 = 2[1 + \Delta^2(V_2 - E)]\eta_2 - \eta_1.$$

Now that you know η_3, find

$$\eta_4 = 2[1 + \Delta^2(V_3 - E)]\eta_3 - \eta_2.$$

Continue until you know η_i at every grid point.

For most values of E, this algorithm will result in a solution that rockets to $\pm\infty$ at the far right. When you pick a value of E where the solution approaches zero at the far right, you've found an energy eigenvalue. The algorithm is called "shooting", because it resembles shooting an arrow at a fixed target: your first shot might be too high, your second too low, so you try something between until you home in on your target.

Problems

9.10 Program

a. Implement the shooting algorithm using a computer spreadsheet, your favorite programming language, or in any other way. You will have to select reasonable values for Δ and η_2.

b. Check your implementation by solving the energy eigenproblem for a free particle and for an infinite square well.

c. Find the three lowest-energy eigenvalues for a square well with $V_1 = V_2 = 30$. Do the corresponding eigenfunctions have the qualitative character you expect?

d. Repeat for a square well with $V_1 = 50$ and $V_2 = 30$.

9.11 **Algorithm parameter**
Below equation (9.18) I suggested that you start the stepping algorithm with $\eta_1 = 0$, $\eta_2 = 0.001$. What would have happened had you selected $\eta_1 = 0$, $\eta_2 = 0.003$ instead?

9.12 **Simple harmonic oscillator**
(Work problem 9.9 on page 265 before working this one.)
Implement the algorithm for a simple harmonic oscillator using scaled quantities. Find the five lowest-energy eigenvalues, and compare them to the analytic results 0.5, 1.5, 2.5, 3.5, and 4.5.

9.13 **Questions** (recommended problem)
Update your list of quantum mechanics questions that you started at problem 1.13 on page 56. Write down new questions and, if you have uncovered answers to any of your old questions, write them down briefly.

⟦For example, one of my questions would be: "For any value of E — energy eigenvalue or no — equation (9.1) has two linearly independent solutions. We saw on page 247 that often the two linearly independent solutions are mirror images, one rocketing off to infinity as $x \to +\infty$ and the other rocketing off to infinity as $x \to -\infty$. But what about the energy eigenfunctions, which go to zero as $x \to \pm\infty$? What does the other linearly independent solution look like then?"⟧

Chapter 10

The Simple Harmonic Oscillator

The simple harmonic oscillator is a mainstay for both classical and quantum mechanics. In classical mechanics we often speak of a “mass on a spring” or of a “pendulum undergoing small oscillations”. In quantum mechanics we don’t typically attach electrons to springs! But the simple harmonic oscillator remains important, for example in treating small oscillations of diatomic molecules. And, remarkably, the electromagnetic field turns out to be equivalent to a large number of independent simple harmonic oscillators.

10.1 The classical simple harmonic oscillator

Recall that in the classical simple harmonic oscillator, the particle’s equilibrium position is conventionally taken as the origin, and the “restoring force” that pushes a displaced particle back toward the origin is

$$F(x) = -kx. \tag{10.1}$$

The potential energy function is thus

$$V(x) = \tfrac{1}{2}kx^2, \tag{10.2}$$

and the particle’s total energy

$$E = \frac{p^2}{2m} + \frac{kx^2}{2} \tag{10.3}$$

can range anywhere from 0 to $+\infty$.

If the initial position is x_0 and the initial momentum is p_0, then the motion is

$$\begin{aligned} x(t) &= x_0\cos(\omega t) + (p_0/m\omega)\sin(\omega t) \\ p(t) &= p_0\cos(\omega t) - (x_0 m\omega)\sin(\omega t), \end{aligned} \tag{10.4}$$

where the "angular frequency" is $\omega = \sqrt{k/m}$. Just as we found it convenient to shift the position origin so that the particle's equilibrium position is $x = 0$, so we may shift the time origin so that

$$\begin{aligned} x(t) &= A\cos(\omega t), \\ p(t) &= -(Am\omega)\sin(\omega t). \end{aligned} \qquad (10.5)$$

You may generalize this discussion enormously, for example by talking about the damped, driven harmonic oscillator, but that's all there is to say about he *simple* harmonic oscillator in classical mechanics.

10.2 Setting up the quantal problem

For the simple harmonic oscillator, $V(x) = \frac{1}{2}kx^2$, so quantal time evolution is governed by

$$\frac{\partial\psi(x,t)}{\partial t} = -\frac{i}{\hbar}\left[-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + \frac{1}{2}kx^2\psi(x,t)\right]. \qquad (10.6)$$

(In quantum mechanics, the letter k can denote either the spring constant, as above, or the wave number, as in sample problem 6.6.1. Make sure from context what meaning is intended!) The solutions can depend upon only these three parameters:

parameter	dimensions
m	$[M]$
k	$[(\text{force})/L] = [(ML/T^2)/L = [M/T^2]$
$\hbar$	$[(\text{momentum})L] = [(ML/T)L] = [ML^2/T]$

What, then, is the characteristic time t_c for this problem? Any formula for time has got to have contributions from k or $\hbar$, because these are the only parameters that include the dimensions of time. But if the formula contained $\hbar$, there would be dimensions of length that could not be canceled through any other parameter, so it must be independent of $\hbar$. To build a quantity with dimensions of time from k, you have to get rid of those mass dimensions, and the only way to do that is through division by m. In conclusion there is only one way to build up a quantity with the dimensions of time from the three parameters m, k, and $\hbar$, and that is

$$t_c = \sqrt{\frac{m}{k}}. \qquad (10.7)$$

Similar but slightly more elaborate reasoning shows that there is only one way to build a characteristic length x_c from these three parameters, and that is

$$x_c = \sqrt[4]{\frac{\hbar^2}{mk}}. \tag{10.8}$$

Finally, the characteristic energy is

$$e_c = \hbar\sqrt{k/m} = \hbar\omega \tag{10.9}$$

where $\omega = \sqrt{k/m}$ is the classical angular frequency of oscillation.

Exercise 10.A. Execute the "similar but slightly more elaborate reasoning" required to uncover that characteristic length and energy.

10.3 Resume of energy eigenproblem

The energy eigenproblem for the simple harmonic oscillator is

$$-\frac{\hbar^2}{2m}\frac{d^2\eta_n(x)}{dx^2} + \frac{k}{2}x^2\eta_n(x) = E_n\eta_n(x). \tag{10.10}$$

This is a second-order linear ordinary differential equation, and the theory of differential equations assures us that for every value of E_n, there are two linearly independent solutions to this equation.

This does *not*, however, mean that every E_n is an energy eigenvalue with two energy eigenfunctions. Nearly all of these solutions turn out to be unnormalizable,

$$\int_{-\infty}^{+\infty} \eta^*(x)\eta(x)\,dx = \infty,$$

so they do not represent physical states. The problem of solving the energy eigenproblem is simply the problem of plowing through the vast haystack of solutions of (10.10) to find those few needles with finite norm.

10.4 Solution of the energy eigenproblem: Differential equation approach

Problem: Given m and k, find values E_n such that the corresponding solutions $\eta_n(x)$ of

$$-\frac{\hbar^2}{2m}\frac{d^2\eta_n(x)}{dx^2}+\frac{k}{2}x^2\eta_n(x)=E_n\eta_n(x) \tag{10.11}$$

are normalizable wavefunctions. Such E_n are the energy eigenvalues, and the corresponding solutions $\eta_n(x)$ are energy eigenfunctions.

Strategy: The following four-part strategy is effective for most differential equation eigenproblems:

(1) Convert to dimensionless variables.
(2) Remove asymptotic behavior of solutions.
(3) Find non-asymptotic behavior using the series method.
(4) Invoke normalization to terminate the series as a polynomial.

In this treatment, I'll play fast and loose with asymptotic analysis. But everything I'll do is reasonable and, if you push hard enough, rigorously justifiable.[1]

1. Convert to dimensionless variables: Using the characteristic length x_c and the characteristic energy e_c, define the dimensionless scaled lengths and energies

$$\tilde{x}=x/x_c \qquad \text{and} \qquad \tilde{E}_n=E_n/e_c. \tag{10.12}$$

Exercise 10.B. Show that, in terms of these variables, the ordinary differential equation (10.11) is

$$\frac{d^2\eta_n(\tilde{x})}{d\tilde{x}^2}+\left(2\tilde{E}_n-\tilde{x}^2\right)\eta_n(\tilde{x})=0. \tag{10.13}$$

Exercise 10.C. We're using this equation merely as a stepping-stone to reach the full answer, but in fact it contains a lot of information already. For example, suppose we had two electrons in two far-apart simple harmonic oscillators, the second one with three times the "stiffness" of the first (that is, the spring constants are related through $k^{(2)}=3k^{(1)}$). We don't yet know the energy of the fourth excited state for either oscillator, yet we can easily find their ratio. What is it?

---

[1] See for example C.M. Bender and S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978).

2. Remove asymptotic behavior of solutions: Consider the limit as $\tilde{x}^2 \to \infty$. In this limit, the ODE (10.13) becomes approximately

$$\frac{d^2\eta_n(\tilde{x})}{d\tilde{x}^2} - \tilde{x}^2\eta_n(\tilde{x}) = 0, \tag{10.14}$$

but it is hard to solve even this simplified equation! Fortunately, it's not necessary to find an exact solution, only to find the asymptotic character of the solutions.

Pick the trial solution

$$f(\tilde{x}) = e^{-\tilde{x}^2/2}. \tag{10.15}$$

When we test to see whether this is a solution, we find

$$\begin{aligned} &\frac{d^2 f(\tilde{x})}{d\tilde{x}^2} - \tilde{x}^2 f(\tilde{x}) \\ &= \left(\tilde{x}^2 e^{-\tilde{x}^2/2} - e^{-\tilde{x}^2/2}\right) - \tilde{x}^2 e^{-\tilde{x}^2/2} = -e^{-\tilde{x}^2/2} \end{aligned}$$

So the function (10.15) does *not* solve the ODE (10.14). On the other hand, the amount by which it "misses" solving (10.14) is small in the sense that

$$\lim_{\tilde{x}^2\to\infty} \frac{d^2 f/d\tilde{x}^2 - \tilde{x}^2 f}{\tilde{x}^2 f} = \lim_{\tilde{x}^2\to\infty} \frac{-e^{-\tilde{x}^2/2}}{\tilde{x}^2 e^{-\tilde{x}^2/2}} = \lim_{\tilde{x}^2\to\infty} \frac{-1}{\tilde{x}^2} = 0.$$

A similar result holds for $g(x) = e^{+\tilde{x}^2/2}$.

Our conclusion is that, in the limit $\tilde{x}^2 \to \infty$, the solution $\eta_n(\tilde{x})$ behaves like

$$\eta_n(\tilde{x}) \approx Ae^{-\tilde{x}^2/2} + Be^{+\tilde{x}^2/2}.$$

If $B \neq 0$, then $\eta_n(\tilde{x})$ will not be normalizable because the probability density would become infinite as $\tilde{x}^2 \to \infty$. Thus the solutions we want — the normalizable solutions — behave like

$$\eta_n(\tilde{x}) \approx Ae^{-\tilde{x}^2/2}$$

in the limit that $\tilde{x}^2$ becomes very large.

The paragraph above motivates us to define a new function $v_n(\tilde{x})$ through

$$\eta_n(\tilde{x}) = e^{-\tilde{x}^2/2} v_n(\tilde{x}). \tag{10.16}$$

(I *could* have just produced this definition by fiat, without motivation. But then you wouldn't know how to come up with the proper motivation yourself when you're faced with a new and unfamiliar differential equation.) In terms of this new function, the exact ODE (10.13) becomes

$$\frac{d^2v_n(\tilde{x})}{d\tilde{x}^2} - 2\tilde{x}\frac{dv_n(\tilde{x})}{d\tilde{x}} + \left(2\tilde{E}_n - 1\right)v_n(\tilde{x}) = 0. \tag{10.17}$$

For brevity we introduce the shorthand notation

$$e_n = 2\tilde{E}_n - 1. \tag{10.18}$$

3. Find non-asymptotic behavior using the series method: Okay, but how are we going to solve equation (10.17) for $v_n(\tilde{x})$? Through the power series method!

Try a solution of the form

$$\begin{aligned}
v(\tilde{x}) &= \sum_{k=0}^{\infty} a_k\tilde{x}^k \\
v'(\tilde{x}) &= \sum_{k=0}^{\infty} ka_k\tilde{x}^{k-1} \qquad\qquad \tilde{x}v'(\tilde{x}) = \sum_{k=0}^{\infty} ka_k\tilde{x}^k \\
v''(\tilde{x}) &= \sum_{k=0}^{\infty} k(k-1)a_k\tilde{x}^{k-2} \qquad [\![\text{ note that first two terms vanish} \ldots]\!] \\
&= \sum_{k=2}^{\infty} k(k-1)a_k\tilde{x}^{k-2} \qquad [\![\text{ change summation index to } k' = k-2 \ldots]\!] \\
&= \sum_{k'+2=2}^{\infty} (k'+2)(k'+1)a_{k'+2}\tilde{x}^{k'} \qquad [\![\text{ rename dummy index } k' \text{ to } k \ldots]\!] \\
&= \sum_{k=0}^{\infty} (k+2)(k+1)a_{k+2}\tilde{x}^k.
\end{aligned}$$

Then equation (10.17) becomes

$$\sum_{k=0}^{\infty}[(k+2)(k+1)a_{k+2} - 2ka_k + e_na_k]\tilde{x}^k = 0. \tag{10.19}$$

Each term in square brackets must vanish, whence the recursion relation

$$a_{k+2} = \frac{2k - e_n}{(k+2)(k+1)}a_k \qquad k = 0, 1, 2, \ldots. \tag{10.20}$$

Like any second order linear ODE, equation (10.17) has two linearly independent solutions:

- An *even solution* of equation (10.17) comes by taking $a_0 = 1$, $a_1 = 0$. It is
$$v^{(e)}(\tilde{x}) = 1 - \frac{e_n}{2!}\tilde{x}^2 + \frac{(e_n-4)e_n}{4!}\tilde{x}^4 - \frac{(e_n-8)(e_n-4)e_n}{6!}\tilde{x}^6 + \cdots . \tag{10.21}$$
- An *odd solution* of equation (10.17) comes by taking $a_0 = 0$, $a_1 = 1$. It is
$$\begin{aligned} v^{(o)}(\tilde{x}) = \tilde{x} - \frac{e_n-2}{3!}\tilde{x}^3 + \frac{(e_n-6)(e_n-2)}{5!}\tilde{x}^5 \\ - \frac{(e_n-10)(e_n-6)(e_n-2)}{7!}\tilde{x}^7 + \cdots . \end{aligned} \tag{10.22}$$

What is the asymptotic behavior of such solutions $v_n(\tilde{x})$ as $\tilde{x}^2 \to \infty$? Well, the large $\tilde{x}$ behavior will be dominated by the high-order terms of the series. Generally, as $k \to \infty$,
$$\frac{a_{k+2}}{a_k} = \frac{2k - e_n}{(k+2)(k+1)} \to \frac{2}{k}. \tag{10.23}$$

Compare this behavior to the expansion
$$e^{\tilde{x}^2} = b_0 + b_2\tilde{x}^2 + b_4\tilde{x}^4 + \cdots \tag{10.24}$$
which has
$$\frac{b_{k+2}}{b_k} = \frac{1}{(k/2)+1} \to \frac{2}{k}. \tag{10.25}$$
So whenever this happens,
$$v_n(\tilde{x}) \approx e^{\tilde{x}^2} \qquad \text{and} \qquad \eta_n(\tilde{x}) = e^{-\tilde{x}^2/2} v_n(\tilde{x}) \approx e^{\tilde{x}^2/2},$$
Thus giving us the very same unnormalizable behavior we've been trying so hard to avoid!

Is there no way to salvage the situation?

4. Invoke normalization to terminate the series as a polynomial: The candidate wavefunction $\eta_n(\tilde{x})$ is not normalizable when $a_{k+2}/a_k \to 2/k$ (see equation 10.23). There is only one way to avoid this limit: when the series for $v_n(\tilde{x})$ terminates as a polynomial.[2] This termination occurs when, for some non-negative integer n, we have $2n = e_n$ whence (by recursion relation 10.20), $a_k = 0$ for all $k > n$, and the solution is a polynomial of order n. Hence the only physical states correspond to energies with
$$2n = e_n = 2\tilde{E}_n - 1.$$
Rephrasing, and converting back from scaled to conventional units,

---

[2] This is why we removed the asymptotic behavior and concentrated on $v_n(\tilde{x})$ rather than on $\eta_n(\tilde{x}) = e^{-\tilde{x}^2/2}v_n(\tilde{x})$. If we had solved differential equation (10.14) directly using the power series method, the expansion would not terminate for any value of $\tilde{E}_n$.

Energy (eigen)states can exist only if they correspond to the energy (eigen)values

$$E_n = \hbar\omega(n + \tfrac{1}{2}) \qquad n = 0, 1, 2, 3, \ldots \tag{10.26}$$

What are the wavefunctions of the energy eigenstates?

For n even, $v_n^{(e)}(\tilde{x})$ terminates and $v_n^{(o)}(\tilde{x})$ doesn't.
For n odd, $v_n^{(o)}(\tilde{x})$ terminates and $v_n^{(e)}(\tilde{x})$ doesn't.

In both cases, By tradition one defines the *Hermite*[3] *polynomial* of nth order $H_n(\tilde{x})$:

$$n \text{ even:} \quad H_n(\tilde{x}) = (-1)^{n/2}\frac{n!}{(n/2)!}v_n^{(e)}(\tilde{x}) \tag{10.27}$$

$$n \text{ odd:} \quad H_n(\tilde{x}) = (-1)^{(n-1)/2}\frac{2n!}{((n-1)/2)!}v_n^{(o)}(\tilde{x}) \tag{10.28}$$

so that

$$\eta_n(x) = A_n e^{-\tilde{x}^2/2}H_n(\tilde{x}) \qquad \tilde{x} = \sqrt[4]{\frac{mk}{\hbar^2}}\,x \tag{10.29}$$

where A_n is a normalization factor.

10.4.1 *Sample Problem: Gaussian wavefunctions in the simple harmonic oscillator*

We encountered the Gaussian wavefunction

$$\psi(x) = \frac{A}{\sqrt{\sigma}}e^{-(x/\sigma)^2} \tag{10.30}$$

at equation (8.20). It is generally *not* an energy eigenfunction of the simple harmonic oscillator, so it generally *doesn't have* an energy. Nevertheless it *does have* a mean energy $\langle \hat{H} \rangle$. Calculate the mean potential and kinetic energies. (Use unscaled variables. You may employ the results uncovered in problem 8.4, "Static properties of a Gaussian wavepacket".) The mean potential energy will approach infinity for very wide wavefunctions ($\sigma \to \infty$), the mean kinetic energy will approach infinity for very narrow wavefunctions ($\sigma \to 0$). Discuss qualitatively why this is so. There will be one σ in

---

[3] Biographical information on Charles Hermite is given on page 112.

between that minimizes the mean total energy. Find it and compare to the ground state wavefunction $\eta_0(x)$

Solution: This is just the wavefuction of equation (8.20) with $p_0 = 0$. We could find the mean potential energy

$$\langle\widehat{\mathrm{PE}}\rangle = \tfrac{1}{2}k\langle\hat{x}^2\rangle,$$

but we've already done that in problem 8.4, part b, where we found $\Delta x = \sigma/2$, so $\langle\hat{x}^2\rangle = \sigma^2/4$, so

$$\langle\widehat{\mathrm{PE}}\rangle = \tfrac{1}{8}k\sigma^2.$$

The qualitative behavior is easy to explain: If the wavefunction is narrow (small σ) the particle is very likely to be found in the low potential energy region near the origin. If the wavefunction is wide (large σ) there is a good chance that it will be found far from the origin in a high potential energy situation.

We could find the mean kinetic energy

$$\langle\widehat{\mathrm{KE}}\rangle = \frac{\langle\hat{p}^2\rangle}{2m},$$

but we've already done that in problem 8.4, part d, where we found $\Delta p = \hbar/\sigma$, so $\langle\hat{p}^2\rangle = \hbar^2/\sigma^2$, so

$$\langle\widehat{\mathrm{KE}}\rangle = \frac{\hbar^2}{2m\sigma^2}.$$

The qualitative behavior is easy to explain: If the wavefunction is wide in position space (large σ), then it is narrow in momentum space so the particle is very likely to be found in the low kinetic energy region with small momentum magnitudes. The opposite holds if the wavefunction is narrow in position space.

To find the minimum energy

$$\langle\hat{H}\rangle = \frac{k\sigma^2}{8} + \frac{\hbar^2}{2m\sigma^2}$$

just take the derivative with respect to σ and set it equal to zero. The result is that the minimum falls when

$$\sigma^2 = \frac{2\hbar}{\sqrt{mk}}.$$

At this value the mean potential energy equals the mean kinetic energy and the mean total energy is

$$\langle\hat{H}\rangle_{\min} = \frac{\hbar}{2}\sqrt{\frac{k}{m}} = \tfrac{1}{2}\hbar\omega.$$

By shear good fortune, we have stumbled upon the ground state!

Problems

10.1 **Explicit eigenfunctions**
Write out the unnormalized eigenfunctions $\eta_0(x)$ through $\eta_5(x)$. Use scaled variables. Do the eigenfunctions always display the symmetry of the potential energy function?

10.5 Character of the energy eigenfunctions

The energy of the classical simple harmonic oscillator is continuous, the energy of the quantal simple harmonic oscillator is discrete. That is to be expected given the name "quantum mechanics".

The minimum energy of the classical simple harmonic oscillator is zero, and the ground state consists of a particle stationary (momentum zero) at the very bottom of the well (position zero). The minimum energy of the quantal simple harmonic oscillator is $\frac{1}{2}\hbar\omega$, and the ground state is a Gaussian wavefunction that, by its quantal character, *has* an energy and hence *cannot have* a momentum (including momentum zero) and *cannot have* a position (including position zero).

The difference between the minimum classical energy and the minimum quantal energy is called the "zero-point energy" (or the "vacuum energy") and some people find it more disturbing than the quantization of energy. You can get rid of it by remembering that only changes in energy are physically significant — just as we shifted the origin of position so that the equilibrium position was zero (equation 10.1), and just as we shifted the origin of time so that the initial momentum was zero (equation 10.5), so we can shift the zero of energy up by $\frac{1}{2}\hbar\omega$ so that the ground state energy is zero.

Or you can gain insight into zero-point energy by considering mean potential and kinetic energies for a range of wavefunctions: this was the objective of sample problem refSP:GaussianWavefunctionsSHO, "Gaussian wavefunctions in the simple harmonic oscillator".

But what you *can't* do is exploit zero-point energy. If we could extract zero-point energy and use it to power cars and airplanes and computers with pollution-free energy, it would produce enormous societal gains. Indeed, on 27 May 2008, U.S. Patent 7379286 for "Quantum Vacuum Energy Extraction" was issued to the Jovion Corporation. The misconception that one can "extract" zero-point energy flows from the misconception that classical mechanics is correct, and that quantum mechanics is some sort of overlaid screen to obscure our vision and prevent us from getting to the correct, underlying classical mechanics. The truth is the other way around: quantum mechanics is correct and classical mechanics is an approximation accurate only when quantum mechanics is applied to big things. There is a reason that the Jovion Corporation has not produced a useful product since its patent was issued in 2008: that patent is based on a misconception.

10.6 Solution of the energy eigenproblem: Operator factorization approach

The differential equation approach works. It's hard. It's inefficient in that we find an infinite number of solutions and then throw most of them away. It depends on a particular representation, namely the position representation. Worst of all, it's hard to use. For example, suppose we wanted to find the mean value of the potential energy in the n-th energy eigenstate. It is

$$\begin{aligned}\langle \hat{U} \rangle_n &= \frac{k}{2}\langle \eta_n | \hat{x}^2 | \eta_n \rangle = \frac{k}{2}\int_{-\infty}^{+\infty} x^2 \eta_n^2(x)\, dx \\ &= \frac{k}{2}\frac{\hbar}{\sqrt{mk}} \frac{\displaystyle\int_{-\infty}^{+\infty} \tilde{x}^2 e^{-\tilde{x}^2} H_n^2(\tilde{x})\, d\tilde{x}}{\displaystyle\int_{-\infty}^{+\infty} e^{-\tilde{x}^2} H_n^2(\tilde{x})\, d\tilde{x}}. \end{aligned} \tag{10.31}$$

Unless you happen to relish integrating Hermite polynomials, these last two integrals are intimidating.

I'll show you a method, invented by Dirac, that avoids all these problems. On the other hand the method is hard to motivate. It required no special insight or talent to use the differential equation approach — while difficult, it was just a straightforward "follow your nose" application of standard differential equation solution techniques. In contrast the operator factorization method clearly springs from the creative mind of genus.

Start with the Hamiltonian

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{m\omega^2}{2}\hat{x}^2. \tag{10.32}$$

(I follow quantal tradition here by writing the spring constant k as $m\omega^2$, where $\omega = \sqrt{k/m}$ is the classical angular frequency of oscillation.) Since we're in a mathematical mode, it makes sense to define the dimensionless operators

$$\hat{X} = \sqrt{\frac{m\omega}{2\hbar}}\,\hat{x} \quad \text{and} \quad \hat{P} = \frac{1}{\sqrt{2m\hbar\omega}}\hat{p}, \tag{10.33}$$

that satisfy

$$[\hat{X}, \hat{P}] = \sqrt{\frac{m\omega}{2\hbar}}\frac{1}{\sqrt{2m\hbar\omega}}[\hat{x}, \hat{p}] = \frac{i}{2}\hat{1}, \tag{10.34}$$

and write

$$\hat{H} = \hbar\omega(\hat{X}^2 + \hat{P}^2). \tag{10.35}$$

Now, one of the oldest and most fundamental tools of problem solving is breaking something complex into its simpler pieces. ("All Gaul is divided into three parts." — Julius Caesar.) If you had an expression like

$$x^2 - p^2$$

you might well break it into simpler pieces as

$$(x - p)(x + p).$$

Slightly less intuitive would be to express

$$x^2 + p^2$$

as

$$(x - ip)(x + ip).$$

But in our case, we're factoring an *operator*, and we have to ask about the expression

$$\begin{aligned}(\hat{X} - i\hat{P})(\hat{X} + i\hat{P}) &= \hat{X}^2 + i\hat{X}\hat{P} - i\hat{P}\hat{X} + \hat{P}^2 \\ &= \hat{X}^2 + i[\hat{X}, \hat{P}] + \hat{P}^2 \\ &= \hat{X}^2 + \hat{P}^2 - \tfrac{1}{2}\hat{1}. \end{aligned} \tag{10.36}$$

So we haven't quite succeeded in factorizing our Hamiltonian — there's a bit left over due to non-commuting operators — but the result is

$$\hat{H} = \hbar\omega[(\hat{X} - i\hat{P})(\hat{X} + i\hat{P}) + \tfrac{1}{2}]. \tag{10.37}$$

From here, define

$$\hat{a} = \hat{X} + i\hat{P}. \tag{10.38}$$

The Hermitian adjoint of $\hat{a}$ is

$$\hat{a}^\dagger = \hat{X} - i\hat{P}. \tag{10.39}$$

Note that the operators $\hat{a}$ and $\hat{a}^\dagger$ are *not* Hermitian. There is no observable corresponding to $\hat{a}$. The commutator is

$$[\hat{a}, \hat{a}^\dagger] = \hat{1}. \tag{10.40}$$

Exercise 10.D. Verify the above commutator.

Exercise 10.E. Show that

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}\,(\hat{a} + \hat{a}^\dagger) \tag{10.41}$$

$$\hat{p} = -i\sqrt{\frac{m\hbar\omega}{2}}\,(\hat{a} - \hat{a}^\dagger). \tag{10.42}$$

And in terms of $\hat{a}$ and $\hat{a}^\dagger$, the Hamiltonian is

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \tfrac{1}{2}). \tag{10.43}$$

Our task: Using only the fact that $[\hat{a}, \hat{a}^\dagger] = \hat{1}$, where $\hat{a}^\dagger$ is the Hermitian adjoint of $\hat{a}$, solve the energy eigenproblem for $\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})$.

We are *not* going to use the facts that $\hat{a}$ and $\hat{a}^\dagger$ are related to $\hat{x}$ and $\hat{p}$. We are *not* going to use the definitions of $\hat{a}$ or $\hat{a}^\dagger$ at all. We are going to use only the commutator.

We will do this by solving the eigenproblem for the operator $\hat{N} = \hat{a}^\dagger\hat{a}$. Once these are known, we can immediately read off the solution for the eigenproblem for $\hat{H}$. So, we look for the eigenvectors $|n\rangle$ with eigenvalues n such that

$$\hat{N}|n\rangle = n|n\rangle. \tag{10.44}$$

Because $\hat{N}$ is Hermitian, its eigenvalues are real. Furthermore, they are non-negative because, where we define the vector $|\phi\rangle$ through $|\phi\rangle = \hat{a}|n\rangle$,

$$n = \langle n|\hat{N}|n\rangle = \langle n|\hat{a}^\dagger\hat{a}|n\rangle = \langle n|\hat{a}^\dagger|\phi\rangle = \langle\phi|\hat{a}|n\rangle^* = \langle\phi|\phi\rangle^* \geq 0. \tag{10.45}$$

Now I don't know much about energy state $|n\rangle$, but I do know that at least one exists. So for this particular one, I can ask "What is $\hat{a}|n\rangle$?". Well,

$$\begin{aligned}\hat{a}|n\rangle &= \hat{1}\hat{a}|n\rangle\\ &= (\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a})\hat{a}|n\rangle\\ &= \hat{a}\hat{N}|n\rangle - \hat{N}\hat{a}|n\rangle\\ &= n\hat{a}|n\rangle - \hat{N}\hat{a}|n\rangle.\end{aligned}$$

So if I define $|\phi\rangle = \hat{a}|n\rangle$ (an unnormalized vector), then

$$\begin{aligned}|\phi\rangle &= n|\phi\rangle - \hat{N}|\phi\rangle\\ \hat{N}|\phi\rangle &= n|\phi\rangle - |\phi\rangle = (n-1)|\phi\rangle.\end{aligned}$$

In other words, the vector $|\phi\rangle$ is an eigenvector of $\hat{N}$ with eigenvalue $n-1$. Wow!

$$|\phi\rangle = C|n-1\rangle.$$

We need to find the normalization constant C:

$$\begin{aligned}\langle\phi|\phi\rangle &= |C|^2\langle n-1|n-1\rangle = |C|^2\\ \langle\phi|\phi\rangle &= \langle n|\hat{a}^\dagger\hat{a}|n\rangle = \langle n|\hat{N}|n\rangle = n.\end{aligned}$$

So $C = \sqrt{n}$ and

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle \tag{10.46}$$

The operator $\hat{a}$ is called a "lowering operator".

So, we started off with one eigenstate $|n\rangle$. We applied $\hat{a}$ to get another eigenstate — with smaller eigenvalue. We can apply $\hat{a}$ to this new state to get yet another eigenstate with an even smaller eigenvalue. But this seems to raise a paradox. We saw at equation (10.45) that the eigenvalues were positive or zero. This seems to provide a mechanism for getting negative eigenvalues — in fact, eigenvalues as small as desired! For example if we started with a state of eigenvalue 2.3, we could lower it to produce a state of eigenvalue 1.3. We could lower this to produce a state of eigenvalue 0.3, and we could lower once more to produce a state of eigenvalue -0.7. But we know there are no states with negative eigenvalues! Thus there can't be a state of eigenvalue 2.3 to start off with.

However, if we start instead with a state of eigenvalue 2, we could lower that to get $|1\rangle$, lower that to get $|0\rangle$, and what happens when we try to lower $|0\rangle$? From equation (10.46), we find

$$\hat{a}|0\rangle = \sqrt{0}|-1\rangle = 0.$$

When we lower the state $|0\rangle$, we don't get the state $|-1\rangle$. Instead we get nothing!

In conclusion, there are no fractional eigenvalues. The only eigenvalues of $\hat{N}$ are the non-negative integers.

We've gotten a lot out of the use of $\hat{a}$. What happens when we use $\hat{a}^\dagger$?

$$\begin{aligned}\hat{a}^\dagger|n\rangle &= \hat{a}^\dagger\hat{1}|n\rangle\\ &= \hat{a}^\dagger(\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a})|n\rangle\\ &= \hat{N}\hat{a}^\dagger|n\rangle - \hat{a}^\dagger\hat{N}|n\rangle\\ &= \hat{N}\hat{a}^\dagger|n\rangle - n\hat{a}^\dagger|n\rangle.\end{aligned}$$

So if I define $|\chi\rangle = \hat{a}^\dagger|n\rangle$ (an unnormalized vector), then

$$\begin{aligned} |\chi\rangle &= \hat{N}|\chi\rangle - n|\chi\rangle \\ \hat{N}|\chi\rangle &= n|\chi\rangle + |\chi\rangle = (n+1)|\chi\rangle. \end{aligned}$$

In other words, the vector $|\chi\rangle$ is an eigenvector of $\hat{N}$ with eigenvalue $n+1$:

$$|\chi\rangle = C|n+1\rangle.$$

The operator $\hat{a}^\dagger$ is a "raising operator"!

Exercise 10.F. Find the normalization constant C and conclude that

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{10.47}$$

The eigenproblem is solved entirely. Given only $[\hat{a}, \hat{a}^\dagger] = \hat{1}$, where $\hat{a}^\dagger$ is the Hermitian adjoint of $\hat{a}$, the operator

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \tfrac{1}{2})$$

has

$$\begin{array}{rcccc} \text{eigenstates} & |0\rangle, & |1\rangle, & |2\rangle, & \ldots \\ \text{with eigenvalues} & \hbar\omega(\frac{1}{2}), & \hbar\omega(\frac{3}{2}), & \hbar\omega(\frac{5}{2}), & \ldots \end{array}$$

These eigenstates are related through

$$\begin{aligned} \hat{a}|n\rangle &= \sqrt{n}\,|n-1\rangle \qquad \text{"lowering operator"} \\ \hat{a}^\dagger|n\rangle &= \sqrt{n+1}\,|n+1\rangle \qquad \text{"raising operator"} \end{aligned}$$

The operators $\hat{a}$ and $\hat{a}^\dagger$ are collectively called "ladder operators".

Let's try this scheme on the problem of mean potential energy that we found so intimidating at equation (10.31). Using equation (10.41) for $\hat{x}$ in terms of ladder operators,

$$\begin{aligned} \langle\hat{U}\rangle_n &= \frac{m\omega^2}{2}\langle n|\hat{x}^2|n\rangle \\ &= \frac{m\omega^2}{2}\frac{\hbar}{2m\omega}\langle n|(\hat{a}+\hat{a}^\dagger)^2|n\rangle \\ &= \tfrac{1}{4}\hbar\omega\langle n|(\hat{a}\hat{a} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} + \hat{a}^\dagger\hat{a}^\dagger)|n\rangle. \end{aligned}$$

But

$$\begin{aligned} \langle n|\hat{a}\hat{a}|n\rangle &= \sqrt{n}\,\langle n|\hat{a}|n-1\rangle \\ &= \sqrt{n}\,\sqrt{n-1}\,\langle n|n-2\rangle \\ &= 0. \end{aligned}$$

Similarly, you can see without doing any calculation that $\langle n|\hat{a}^\dagger\hat{a}^\dagger|n\rangle = 0$. Now

$$\begin{aligned}\langle n|\hat{a}\hat{a}^\dagger|n\rangle &= \sqrt{n+1}\,\langle n|\hat{a}|n+1\rangle\\ &= \sqrt{n+1}\,\sqrt{n+1}\,\langle n|n\rangle\\ &= n+1\end{aligned}$$

while

$$\langle n|\hat{a}^\dagger\hat{a}|n\rangle = \langle n|\hat{N}|n\rangle = n,$$

so

$$\langle\hat{U}\rangle_n = \tfrac{1}{2}(n+\tfrac{1}{2})\hbar\omega. \tag{10.48}$$

We did it without Hermite polynomials, we did it without integrals. What seemed at first to be impossibly difficult was actually sort of fun.

Our excursion into raising and lowering operators seemed like a flight of pure fantasy, but it resulted in a powerful and practical tool.

10.7 Time evolution in the simple harmonic oscillator

This book, like any quantum mechanics book, devotes considerable space to solving the energy eigenproblem. There are two reasons for this: First, energy is the quantity easiest to measure in atomic systems, so energy quantization is the most direct way to see quantum mechanics at work. Second, the most straightforward way to solve the time evolution problem is to first solve the energy eigenproblem, then invoke the "Formal solution of the Schrödinger equation" given in equation (5.44).

But while the energy eigenproblem is important, it is not the whole story. It is true that the energy eigenvalues are the only allowed energy values. It is false that the energy eigenstates are the only allowed states. There are position states, momentum states, potential energy states, kinetic energy states, angular momentum states, and states (such as the Gaussian wavepacket) that are not eigenstates of *any* observable!

This section investigates how an quantal states evolve with time in the simple harmonic oscillator. This investigation is not so important as it was in classical mechanics, because it's hard to measure the position of an electron, but it's important conceptually, and it's important for understanding the classical limit of quantum mechanics.

There are two possible approaches to this problem. First, we could take some specific class of initial wavefunctions $\psi(x,0)$ and work out $\psi(x,t)$ exactly. We took this approach when we investigated the time evolution of free Gaussian wavepackets in problem 8.5, "Force-free time evolution of a Gaussian wavepacket", on page 238. (We never asked about the time evolution of, say, a Lorentzian wavepacket.) Second, we could consider an arbitrary initial wavefunction and then work out not the full wavefunction, but but just some values such as the mean position $\langle\hat{x}\rangle_t$, the mean momentum $\langle\hat{p}\rangle_t$, the indeterminacy in position $(\Delta x)_t$, etc. We take this second approach here.

10.7.1 *Time evolution of mean quantities*

The Ehrenfest theorem (page 214) says that

$$\begin{aligned}\frac{d\langle\hat{x}\rangle_t}{dt} &= \frac{\langle\hat{p}\rangle_t}{m},\\ \frac{d\langle\hat{p}\rangle_t}{dt} &= \langle F(\hat{x})\rangle_t.\end{aligned}$$

For the simple harmonic oscillator,

$$F(x) = -\frac{\partial V(x)}{\partial x} = -kx,$$

so

$$\frac{d\langle\hat{x}\rangle_t}{dt} = \frac{\langle\hat{p}\rangle_t}{m}, \tag{10.49}$$

$$\frac{d\langle\hat{p}\rangle_t}{dt} = -k\langle\hat{x}\rangle_t. \tag{10.50}$$

These equations for $\langle\hat{x}\rangle_t$ and $\langle\hat{p}\rangle_t$ are *exactly* the same as the classical equations for $x(t)$ and $p(t)$, so of course they have *exactly* the same solutions. The initial values $\langle\hat{x}\rangle_0$ and $\langle\hat{p}\rangle_0$ evolve with time into

$$\langle\hat{x}\rangle_t = \langle\hat{x}\rangle_0\cos(\omega t) + (\langle\hat{p}\rangle_0/m\omega)\sin(\omega t), \tag{10.51}$$

$$\langle\hat{p}\rangle_t = \langle\hat{p}\rangle_0\cos(\omega t) - (\langle\hat{x}\rangle_0 m\omega)\sin(\omega t). \tag{10.52}$$

where, as in classical mechanics, $\omega = \sqrt{k/m}$.

Exercise 10.G. Verify that these solutions satisfy the differential equations and initial conditions. (*Clue:* Physically, all the brackets and hats in $\langle\hat{x}\rangle_t$ help keep track of its meaning. Mathematically, they just get in the way. For this mathematical problem, you may write $\langle\hat{x}\rangle_t$ as just $x(t)$.)

The takeaway is that in a simple harmonic oscillator, the quantal mean position and momentum oscillate back and forth exactly as a classical particle would oscillate: with the same period, for example. This holds for *any* initial wavefunction, not just Gaussian wavepackets.

Exercise 10.H. My claim is that for any initial wavefunction, "the quantal mean position and momentum oscillate back and forth exactly as a classical particle would oscillate". But if the initial wavefunction is a stationary state, the mean values don't oscillate at all. Is this a violation of my claim?

10.7.2 *Time evolution of indeterminacy*

Does the wavefunction in a simple harmonic oscillator simply spread out with time, as it does for a free particle? (See problem 8.5, "Force-free time evolution of a Gaussian wavepacket", on page 238.)

We find out by tracing the time evolution of

$$(\Delta x)_t^2 = \langle \hat{x}^2 \rangle_t - \langle \hat{x} \rangle_t^2 .$$

We have just found $\langle \hat{x} \rangle_t$, so we only need $\langle \hat{x}^2 \rangle_t$, which we can find through

$$\frac{d\langle \hat{x}^2 \rangle_t}{dt} = -\frac{i}{\hbar}\langle [\hat{x}^2, \hat{H}] \rangle_t . \tag{10.53}$$

And in order to find *this*, we must evaluate the commutator $[\hat{x}^2, \hat{H}]$.

Our approach to this commutator uses two theorems from problem 3.13, "Commutator algebra", on page 123, namely

$$\begin{aligned} [\hat{A}, \hat{B}\hat{C}] &= \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C}, \\ [\hat{A}\hat{B}, \hat{C}] &= \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}. \end{aligned}$$

Recalling that

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{k}{2}\hat{x}^2 ,$$

and that $[\hat{x},\hat{p}] = i\hbar$, we find

$$\begin{aligned}
[\hat{x}^2,\hat{H}] &= \frac{1}{2m}[\hat{x}^2,\hat{p}^2] + \frac{k}{2}[\hat{x}^2,\hat{x}^2] \\
&= \frac{1}{2m}[\hat{x}^2,\hat{p}^2] \\
&= \frac{1}{2m}\left\{\hat{x}[\hat{x},\hat{p}^2] + [\hat{x},\hat{p}^2]\hat{x}\right\} \\
&= \frac{1}{2m}\left\{\hat{x}\hat{p}[\hat{x},\hat{p}] + \hat{x}[\hat{x},\hat{p}]\hat{p} + \hat{p}[\hat{x},\hat{p}]\hat{x} + [\hat{x},\hat{p}]\hat{p}\hat{x}\right\} \\
&= \frac{1}{2m}\left\{2i\hbar\left(\hat{x}\hat{p}+\hat{p}\hat{x}\right)\right\} \\
&= \frac{i\hbar}{m}\left(\hat{x}\hat{p}+\hat{p}\hat{x}\right),
\end{aligned} \tag{10.54}$$

so

$$\frac{d\langle\hat{x}^2\rangle_t}{dt} = -\frac{i}{\hbar}\frac{i\hbar}{m}\langle(\hat{x}\hat{p}+\hat{p}\hat{x})\rangle = \frac{1}{m}\langle\hat{x}\hat{p}+\hat{p}\hat{x}\rangle, \tag{10.55}$$

which seems to do us no good at all because we don't know the behavior of $\langle\hat{x}\hat{p}\rangle$ or $\langle\hat{p}\hat{x}\rangle$. Don't give up.

$$\frac{d\langle\hat{x}\hat{p}\rangle_t}{dt} = -\frac{i}{\hbar}\langle[\hat{x}\hat{p},\hat{H}]\rangle_t,$$

so we find

$$\begin{aligned}
[\hat{x}\hat{p},\hat{H}] &= \frac{1}{2m}[\hat{x}\hat{p},\hat{p}^2] + \frac{k}{2}[\hat{x}\hat{p},\hat{x}^2] \\
&= \frac{1}{2m}\left\{\hat{x}[\hat{p},\hat{p}^2] + [\hat{x},\hat{p}^2]\hat{p}\right\} + \frac{k}{2}\left\{\hat{x}[\hat{p},\hat{x}^2] + [\hat{x},\hat{x}^2]\hat{p}\right\} \\
&= \frac{1}{2m}\left\{[\hat{x},\hat{p}^2]\hat{p}\right\} + \frac{k}{2}\left\{\hat{x}[\hat{p},\hat{x}^2]\right\} \\
&= \frac{1}{2m}\left\{\hat{p}[\hat{x},\hat{p}]\hat{p} + [\hat{x},\hat{p}]\hat{p}^2\right\} + \frac{k}{2}\left\{\hat{x}^2[\hat{p},\hat{x}] + \hat{x}[\hat{p},\hat{x}]\hat{x}\right\} \\
&= 2i\hbar\left\{\frac{1}{2m}\hat{p}^2 - \frac{k}{2}\hat{x}^2\right\} \\
&= 2i\hbar\left\{\hat{H} - k\hat{x}^2\right\}
\end{aligned} \tag{10.56}$$

whence

$$\frac{d\langle\hat{x}\hat{p}\rangle_t}{dt} = 2\langle\hat{H} - k\hat{x}^2\rangle_t. \tag{10.57}$$

A parallel calculation shows that

$$\frac{d\langle\hat{p}\hat{x}\rangle_t}{dt} = 2\langle\hat{H} - k\hat{x}^2\rangle_t. \tag{10.58}$$

Exercise 10.I. Execute this parallel calculation.

Putting these equations together shows that

$$\frac{d^2\langle\hat{x}^2\rangle_t}{dt^2} = -4\omega^2\langle\hat{x}^2\rangle_t + \frac{4}{m}\langle\hat{H}\rangle. \tag{10.59}$$

The quantity $\langle\hat{H}\rangle$ is time-constant, because energy is conserved.

Can we solve this differential equation? You might remember that "the general solution of a linear inhomogeneous ordinary differential equation is the general solution of the homogeneous equation plus any particular solution of the inhomogeneous equation." A particular solution is

$$\langle\hat{x}^2\rangle_t = \frac{\langle\hat{H}\rangle}{m\omega^2}.$$

And the homogeneous equation

$$\frac{d^2\langle\hat{x}^2\rangle_t}{dt^2} = -4\omega^2\langle\hat{x}^2\rangle_t$$

is just the equation for oscillation at frequency 2ω, with general solution

$$\langle\hat{x}^2\rangle_t = C\cos(2\omega t) + D\sin(2\omega t),$$

where C and D are adjustable parameters. Thus the general solution of differential equation (10.59) is

$$\langle\hat{x}^2\rangle_t = C\cos(2\omega t) + D\sin(2\omega t) + \frac{\langle\hat{H}\rangle}{m\omega^2}. \tag{10.60}$$

The indeterminacy squared is thus (using equation (10.51) but replacing the constants with A and B)

$$\begin{aligned}(\Delta x)_t^2 &= \langle x^2\rangle_t - \langle x\rangle_t^2 \\ &= C\cos(2\omega t) + D\sin(2\omega t) + \frac{\langle\hat{H}\rangle}{m\omega^2} - (A\cos(\omega t) + B\sin(\omega t))^2 \\ &= a\cos(2\omega t) + b\sin(2\omega t) + c. \end{aligned} \tag{10.61}$$

In the simple harmonic oscillator, the indeterminacy does *not* simply increase with time, as it does for a free particle. Instead it rises and falls but remains bounded. The oscillation period for the indeterminacy is half the oscillation period for the mean location: During a half cycle of the location — say from left to right — the indeterminacy executes a full cycle — say from wide to narrow to wide, or from narrow to wide to narrow, or from middling to wide to narrow to middling.

Exercise 10.J. Back up the derivation of (10.61) by showing that

$$(A\cos(\omega t)+B\sin(\omega t))^2 = \tfrac{1}{2}(A^2-B^2)\cos(2\omega t)+AB\sin(2\omega t)+\tfrac{1}{2}(A^2+B^2).$$

10.8 Wavepackets with rigidly sliding probability density

Are there any wavepackets in the simple harmonic oscillator potential, where the probability density slides around, changing position but without expanding, contracting, or in any other way changing shape? There are. Any of the energy eigenstates, when displaced from their location as stationary states, will move in this fashion. This section proves this remarkable result.[4] It is not needed for anything following, but it does nicely illustrate this book's epigraph (page iii).

First of all, how would we recognize such a rigidly sliding probability density? In most cases, the probability density $\rho(x,t) = |\psi(x,t)|^2$ will change shape. But if ρ depends upon x and t only through the combination

$$\xi = x - f(t), \tag{10.62}$$

then $\rho(x,t)$ has always the same shape. Suppose, for example, that the function $h(x)$ has a sharp peak at $x = 1$. Then $h(x-5)$ is the same function displaced by 5 to the right, so it has a sharp peak at $x = 6$. More generally, $h(\xi) = h(x - f(t))$ is the same shape always but displaced by $f(t)$.

We already know, from equation (10.51), that for any wavepacket in a simple harmonic oscillator potential, the function $f(t)$ is simple harmonic motion with frequency $\omega = \sqrt{k/m}$. But we won't yet exploit that knowledge.

If $\rho(x,t) = h(x - f(t))$, then the time and space derivatives are related through

$$\frac{\partial \rho}{\partial t} = \frac{dh}{d\xi}\frac{\partial \xi}{\partial t} = -\frac{dh}{d\xi}\dot{f}(t) = -\frac{\partial}{\partial x}\left(\rho \dot{f}(t)\right). \tag{10.63}$$

We use this connection in the equation of continuity (6.36), employing the polar form of the probability current equation (6.35) to find

$$-\frac{\partial}{\partial x}\left(\rho \dot{f}(t)\right) = -\frac{\partial}{\partial x}\left(\frac{\hbar}{m}\rho\frac{\partial \phi}{\partial x}\right) \tag{10.64}$$

whence

$$\dot{f}(t) = \frac{\hbar}{m}\frac{\partial \phi}{\partial x} + \alpha(t), \tag{10.65}$$

---

[4] M.E. Marhic, "Oscillating Hermite-Gaussian wave functions of the harmonic oscillator" *Lettere al Nuovo Cimento* **22** (1978) 376–378. C.C. Yan, "Soliton like solutions of the Schrödinger equation for simple harmonic oscillator" *American Journal of Physics* **62** (1994) 147–151.

where $\alpha(t)$ is any function of time alone. Thus

$$\frac{\partial \phi}{\partial x} = \frac{m}{\hbar}[\dot{f}(t) - \alpha(t)], \tag{10.66}$$

but from equation (6.51) for mean momentum,

$$\begin{aligned}\langle \hat{p} \rangle_t &= \hbar \int_{-\infty}^{+\infty} R^2(x,t) \frac{m}{\hbar}[\dot{f}(t) - \alpha(t)]\, dx \\ &= m[\dot{f}(t) - \alpha(t)]. \end{aligned} \tag{10.67}$$

Now comparison with the Ehrenfest equation (10.49) demonstrates that $\alpha(t) = 0$ and we conclude that

$$\phi(x,t) = \frac{m}{\hbar}\dot{f}(t)x + g(t), \tag{10.68}$$

where $g(t)$ is the constant of integration over x.

To uncover how this imposes conditions on the wavefunction magnitude $R(x,t) = R(\xi)$, use equation (6.29):

$$\begin{aligned}\frac{\partial \phi}{\partial t} &= -\frac{1}{\hbar}\left\{ -\frac{\hbar^2}{2m}\left[\frac{1}{R}\frac{\partial^2 R}{\partial x^2} - \left(\frac{\partial \phi}{\partial x}\right)^2\right] + V(x)\right\} \\ \frac{m}{\hbar}\ddot{f}(t)x + \dot{g}(t) &= -\frac{1}{\hbar}\left\{ -\frac{\hbar^2}{2m}\left[\frac{1}{R}\frac{\partial^2 R}{\partial x^2} - \left(\frac{m}{\hbar}\dot{f}(t)\right)^2\right] + \tfrac{1}{2}kx^2\right\}.\end{aligned}$$

Next change from variables x and t with function $\xi(x,t) = x - f(t)$ to variables ξ and t with function $x(\xi,t) = \xi + f(t)$. The calculation is straightforward and results in

$$\begin{aligned}-\frac{\hbar^2}{2m}\frac{1}{R}\frac{d^2 R}{d\xi^2} &+ \tfrac{1}{2}k\xi^2 + [m\ddot{f}(t) + kf(t)]\xi \\ &= -mf(t)\ddot{f}(t) - \tfrac{1}{2}m\dot{f}^2(t) - \hbar\dot{g}(t) - \tfrac{1}{2}kf^2(t).\end{aligned} \tag{10.69}$$

Notice that when the variables were x and t, we use a partial derivative for $R(x,t)$, but when the variables are ξ and t, we use an ordinary derivative because for rigidly sliding wavepackets $R(\xi)$ is a function of ξ alone.

Now we invoke that fact, from equation (10.51), that $f(t)$ is a classical simple harmonic oscillation, so $m\ddot{f}(t) + kf(t)$ vanishes, and the right-hand side of equation (10.69) simplifies to

$$-\frac{\hbar^2}{2m}\frac{1}{R(\xi)}\frac{d^2 R}{d\xi^2} + \tfrac{1}{2}k\xi^2 = \tfrac{1}{2}kf^2(t) - \tfrac{1}{2}m\dot{f}^2(t) - \hbar\dot{g}(t). \tag{10.70}$$

This equation has the "separation of variables" form "function of ξ alone = function of t alone", so each side must equal the same constant, call it E. Equation (10.70) then becomes

$$-\frac{\hbar^2}{2m}\frac{d^2R}{d\xi^2}+\tfrac{1}{2}k\xi^2R(\xi)=ER(\xi). \tag{10.71}$$

This is exactly the same as the simple harmonic oscillator energy eigenequation (10.10), with $R(\xi)$ replacing $\eta(x)$, and hence has exactly the same solutions.

Problems

10.2 Ground state of the simple harmonic oscillator

You may have been surprised that the lowest possible energy for the simple harmonic oscillator was $E_0 = \frac{1}{2}\hbar\omega$ rather than $E_0 = 0$. This problem attempts to explain the non-zero ground state energy in seat-of-the-pants, semiclassical terms rather than in rigorous, formal, mathematical terms. It goes on to use these ideas plus the Heisenberg indeterminacy principle to guess at a value for the ground state energy. The arguments are not rigorous, but this style of argument allows you to make informed guesses in situations that are too complicated to yield to rigorous mathematics.

In classical mechanics the SHO ground state has zero potential energy (the particle is at the origin) and zero kinetic energy (it is motionless). However in quantum mechanics if a particle is localized precisely at the origin, and hence has zero potential energy, then it has a considerable spread of momentum values and hence a non-zero mean kinetic energy. That mean kinetic energy can be reduced by decreasing the spread of momentum values, but only by increasing the spread of position values and hence by increasing the mean potential energy. The ground state is the state in which this trade off between kinetic and potential energies results in a minimum total energy.

Assume that the spread in position extends over some distance d about the origin (i.e. the particle will very likely be found between $x = -d/2$ and $x = +d/2$). This will result in a potential energy somewhat less than

$$\frac{1}{2}m\omega^2\left(\frac{d}{2}\right)^2.$$

This argument is not intended to be rigorous, so let's forget the "somewhat less" part of the last sentence. Furthermore, a position spread of $\Delta x = d$ implies through the uncertainty principle a momentum spread of $\Delta p \geq \hbar/2d$. (The mean momentum is zero.) Continuing in our nonrigorous vein, let's set $\Delta p = \hbar/2d$ and kinetic energy equal to

$$\frac{1}{2m}\left(\frac{\Delta p}{2}\right)^2 .$$

Sketch potential energy, kinetic energy and total energy as a function of d. Find the minimum value of $E(d)$ and compare with the true ground state energy $E_0 = \frac{1}{2}\hbar\omega$. (Note that if $\hbar$ were zero, the energy minimum would fall at $E(d) = 0$!)

10.3 **Expressions for simple harmonic oscillator ladder operators**
Show that the lowering operator $\hat{a}$ has the outer product expression

$$\hat{a} = \sum_{n=0}^{\infty} \sqrt{n}\, |n-1\rangle\langle n|$$

and the matrix representation (in the energy basis)

$$\begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & 0 & \\ 0 & 0 & \sqrt{2} & 0 & 0 & \\ 0 & 0 & 0 & \sqrt{3} & 0 & \cdots \\ 0 & 0 & 0 & 0 & \sqrt{4} & \\ 0 & 0 & 0 & 0 & 0 & \\ & & \vdots & & & \ddots \end{pmatrix}.$$

Write down the outer product expression and matrix representation for the raising operator $\hat{a}^\dagger$.

10.4 **Ladder operators for the simple harmonic oscillator**

a. Calculate the following simple harmonic oscillator matrix elements:

$$\begin{array}{lll} \langle m|\hat{a}|n\rangle & \langle m|\hat{p}|n\rangle & \langle m|\hat{x}\hat{p}|n\rangle \\ \langle m|\hat{a}^\dagger|n\rangle & \langle m|\hat{x}^2|n\rangle & \langle m|\hat{p}\hat{x}|n\rangle \\ \langle m|\hat{x}|n\rangle & \langle m|\hat{p}^2|n\rangle & \langle m|\hat{H}|n\rangle \end{array}$$

b. Show that, in any SHO energy eigenstate, the mean of the potential energy equals the mean of the kinetic energy. (You might recall that for a classical simple harmonic oscillator, the time average potential energy equals the time average kinetic energy, but this problem investigates quantal averages, not classical time averages.)

c. Find Δx, Δp, and $\Delta x \Delta p$ for the energy eigenstate $|n\rangle$.

10.5 **Simple harmonic oscillator states**

Use scaled variables throughout this problem

a. Concerning the ground energy state: What is $\eta_0(x)$ at $x = 0.5$? What is the probability density $\rho_0(x)$ there?
b. Concerning the first excited energy state: What is $\eta_1(x)$ at $x = 0.5$? What is the probability density $\rho_1(x)$ there?
c. Concerning the "50–50 combination" $\psi_A(x) = (\eta_0(x)+\eta_1(x))/\sqrt{2}$: What is $\psi_A(x)$ at $x = 0.5$? What is the probability density $\rho_A(x)$ there?
d. Concerning another "50–50 combination" $\psi_B(x) = (\eta_0(x) - \eta_1(x))/\sqrt{2}$: What is $\psi_B(x)$ at $x = 0.5$? What is the probability density $\rho_B(x)$ there?
e. Veronica argues that "Probability is central to quantum mechanics, so the probability density of any 50–50 combination of $\eta_0(x)$ and $\eta_1(x)$ will be half-way between $\rho_0(x)$ and $\rho_1(x)$." Prove Veronica wrong. What phenomenon of quantum mechanics has she ignored?
f. (Optional, for the mathematically inclined.) Prove that for any 50–50 combination of $\eta_0(x)$ and $\eta_1(x)$, the probability density at x will range from $\rho_A(x)$ to $\rho_B(x)$. (*Clue:* Use the triangle inequality.)

10.6 **Coincidence?**

Is it just a coincidence that the right-hand-sides are the same in equations (10.57) and (10.58)? Use the commutator $[\hat{x}, \hat{p}] = i\hbar$ to show that (for any one-dimensional system, not just the simple harmonic oscillator)

$$\Re e\{\langle \hat{x}\hat{p}\rangle\} = \Re e\{\langle \hat{p}\hat{x}\rangle\}\,. \tag{10.72}$$

Use the Hermiticity of $\hat{x}$ and $\hat{p}$ to show that

$$\langle \hat{x}\hat{p}\rangle = \langle \hat{p}\hat{x}\rangle^*. \tag{10.73}$$

Conclude that

$$\langle \hat{x}\hat{p} + \hat{p}\hat{x}\rangle = 2\,\Re e\{\langle \hat{x}\hat{p}\rangle\}\,. \tag{10.74}$$

What is $\Im m\{\langle \hat{x}\hat{p}\rangle\}$?

10.7 **Time evolution project**

Generalize the treatment of time evolution in section 10.7 from the simple harmonic oscillator $V(x) = \frac{1}{2}kx^2$ to an arbitrary potential energy function $V(x)$ (and where $F(x) = -\partial V/\partial x$). This is a project, so the exact direction is up to you, but you might want to prove any of these results:

$$\frac{d\langle\hat{x}\hat{p}\rangle_t}{dt} = \frac{\langle\hat{p}^2\rangle_t}{m} + \langle\hat{x}F(\hat{x})\rangle_t \quad \text{(“quantal virial theorem”)} \tag{10.75}$$

$$\frac{d\langle\hat{p}^2\rangle_t}{dt} = \langle\hat{p}F(\hat{x}) + F(\hat{x})\hat{p}\rangle_t \tag{10.76}$$

$$\frac{d^2\langle\hat{x}^2\rangle_t}{dt^2} = \frac{2}{m^2}\langle\hat{p}^2\rangle_t + \frac{2}{m}\langle\hat{x}F(\hat{x})\rangle_t \tag{10.77}$$

$$\frac{d^2(\Delta x)_t^2}{dt^2} = \frac{2}{m^2}(\Delta p)_t^2 + \frac{2}{m}\left\{\langle\hat{x}F(\hat{x})\rangle_t - \langle\hat{x}\rangle_t\langle F(\hat{x})\rangle_t\right\} \tag{10.78}$$

You might then apply these equations to the case of a constant force, or to the case of zero force (in which case your results should agree with equation 8.34).

Chapter 11

Perturbation Theory

11.1 The $\mathcal{O}$ notation

Most problems can't be solved exactly. This is true not only in quantum mechanics, not only in physics, not only in science, but everywhere: For example, whenever a war breaks out, diplomats look for a similar war in the past and try to stop the current war by using a small change to the solution for the previous war.

Approximations are an important part of physics, and an important part of approximation is to ensure their reliability and consistency. The $\mathcal{O}$ notation (pronounced "the big-oh notation") is a practical tool for making approximations reliable and consistent.

The technique is best illustrated through an example. Suppose you desire an approximation for

$$f(x) = \frac{e^{-x}}{1-x} \tag{11.1}$$

valid for small values of x, that is, for $x \ll 1$. You know that

$$e^{-x} = 1 - x + \tfrac{1}{2}x^2 - \tfrac{1}{6}x^3 + \cdots \tag{11.2}$$

and that

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots , \tag{11.3}$$

so it seems that reasonable approximations are

$$e^{-x} \approx 1 - x \tag{11.4}$$

and

$$\frac{1}{1-x} \approx 1 + x, \tag{11.5}$$

whence

$$\frac{e^{-x}}{1-x} \approx (1-x)(1+x) = 1 - x^2. \tag{11.6}$$

Let's try out this approximation at $x_0 = 0.01$. A calculator shows that

$$\frac{e^{-x_0}}{1-x_0} = 1.0000503\ldots \tag{11.7}$$

while the value for the approximation is

$$1 - x_0^2 = 0.9999000. \tag{11.8}$$

This is a very poor approximation indeed. . . the deviation from $f(0) = 1$ is even of the wrong sign!

Let's do the problem over again, but this time keeping track of exactly how much we've thrown away while making each approximation. We write

$$e^{-x} = 1 - x + \tfrac{1}{2}x^2 - \tfrac{1}{6}x^3 + \cdots \tag{11.9}$$

as

$$e^{-x} = 1 - x + \tfrac{1}{2}x^2 + \mathcal{O}(x^3), \tag{11.10}$$

where the notation $\mathcal{O}(x^3)$ stands for the small terms that we haven't bothered to write out explicitly. The symbol $\mathcal{O}(x^3)$ means "terms that are about the magnitude of x^3, or smaller" and is pronounced "terms of order x^3". The $\mathcal{O}$ notation will allow us to make controlled approximations in which we keep track of exactly how good the approximation is.

Similarly, we write

$$\frac{1}{1-x} = 1 + x + x^2 + \mathcal{O}(x^3), \tag{11.11}$$

and find the product

$$\begin{aligned}
f(x) &= \left[1 - x + \tfrac{1}{2}x^2 + \mathcal{O}(x^3)\right] \times \left[1 + x + x^2 + \mathcal{O}(x^3)\right] && (11.12)\\
&= \quad \left[1 - x + \tfrac{1}{2}x^2 + \mathcal{O}(x^3)\right] && (11.13)\\
&\quad + \left[1 - x + \tfrac{1}{2}x^2 + \mathcal{O}(x^3)\right] x && (11.14)\\
&\quad + \left[1 - x + \tfrac{1}{2}x^2 + \mathcal{O}(x^3)\right] x^2 && (11.15)\\
&\quad + \left[1 - x + \tfrac{1}{2}x^2 + \mathcal{O}(x^3)\right] \mathcal{O}(x^3). && (11.16)
\end{aligned}$$

Note, however, that $x \times \frac{1}{2}x^2 = \mathcal{O}(x^3)$, and that $x^2 \times \mathcal{O}(x^3) = \mathcal{O}(x^3)$, and so forth, whence

$$\begin{aligned}
f(x) = &\quad \left[1 - x + \tfrac{1}{2}x^2 + \mathcal{O}(x^3)\right] && (11.17)\\
&+ \left[x - x^2 + \mathcal{O}(x^3)\right] && (11.18)\\
&+ \left[x^2 + \mathcal{O}(x^3)\right] && (11.19)\\
&+ \mathcal{O}(x^3) && (11.20)\\
&= 1 + \tfrac{1}{2}x^2 + \mathcal{O}(x^3). && (11.21)
\end{aligned}$$

Thus we have the approximation

$$f(x) \approx 1 + \tfrac{1}{2}x^2. \tag{11.22}$$

Furthermore, we know that this approximation is accurate to terms of order $\mathcal{O}(x^2)$ (i.e. that the first neglected terms are of order $\mathcal{O}(x^3)$). Evaluating this approximation at $x_0 = 0.01$ gives

$$1 + \tfrac{1}{2}x_0^2 = 1.0000500, \tag{11.23}$$

far superior to our old approximation.

What went wrong on our first try? The $-x^2$ in approximation (11.6) is the same as the $-x^2$ on line (11.18). However, lines (11.17) and (11.19) demonstrate that there were other terms of about the same size (i.e. other "terms of order x^2") that we neglected in our first attempt.

The $\mathcal{O}$ notation is superior to the "dot notation" (such as $\cdots$) in that dots stand for "a bunch of small terms", but the dots don't tell you just how small they are. The symbol $\mathcal{O}(x^3)$ also stands for "a bunch of small terms", but in addition it tells you precisely how small those terms are. The $\mathcal{O}$ notation allows us to approximate in a consistent manner, unlike the uncontrolled approximations where we ignore a "small term" without knowing whether we have already retained terms that are even smaller.

Problem

11.1 **Tunneling for small times — $\mathcal{O}$ notation version**

Problem 5.3, part e, raised the paradox that, according to an approximation produced using truncation rather than $\mathcal{O}$ notation, the total probability was greater than 1. This problem resolves the paradox using $\mathcal{O}$ notation.

a. Approximate time evolution through

$$|\psi(\Delta t)\rangle = \left[\hat{1} - \frac{i}{\hbar}\hat{H}\Delta t - \frac{1}{2\hbar^2}\hat{H}^2(\Delta t)^2 + \mathcal{O}(\Delta t^3)\right]|\psi(0)\rangle. \tag{11.24}$$

Find the representation of this equation in the $\{|1\rangle, |2\rangle\}$ basis.

b. Conclude that for initial condition $|\psi(0)\rangle = |1\rangle$,

$$\begin{pmatrix} \psi_1(\Delta t) \\ \psi_2(\Delta t) \end{pmatrix} = \begin{pmatrix} 1 - (i/\hbar)E\Delta t - (1/2\hbar^2)(E^2 + A^2)(\Delta t)^2 + \mathcal{O}(\Delta t^3) \\ -(i/\hbar)Ae^{-i\phi}\Delta t - (1/\hbar^2)EAe^{-i\phi}(\Delta t)^2 + \mathcal{O}(\Delta t^3) \end{pmatrix}. \tag{11.25}$$

c. Find the resulting probabilities for the system to be found in $|1\rangle$ and in $|2\rangle$, correct to second order in Δt, and show that these probabilities sum to 1, correct to second order in Δt.

11.2 Perturbation theory for cubic equations

Perturbation theory is any technique for approximately solving one problem, when an exact solution for a similar problem is available.

It's a general mathematical technique, applicable to many problems. (It was first developed in the context of classical mechanics: We have an exact solution for the problem two gravitating bodies, such as the ellipse of the Earth orbiting the Sun. But we don't have an exact solution for the problem of three gravitating bodies, such as the Earth plus the Sun plus Jupiter. Perturbation theory was developed to understand how the attraction by Jupiter "perturbed" the motion of the Earth away from the pure elliptical orbit that it would execute if Jupiter didn't exist.) Before we apply perturbation theory to quantum mechanics, we'll apply it in a simpler, and purely mathematical, context.

I wish to solve the cubic equation

$$x^3 - 4.001\,x + 0.002 = 0. \tag{11.26}$$

There is a formula for finding the three roots of a cubic equation, and we could use it to solve this problem. On the other hand, that formula is very complicated and awkward. And while there's no straightforward exact solution to the problem as stated, that problem is very close to the problem

$$x^3 - 4\,x = 0, \tag{11.27}$$

which *does* have straightforward exact solutions, namely

$$0, \pm 2. \tag{11.28}$$

Can I use the exact solution of this "nearby" problem to find an approximate solution for the problem of interest?

I'll write the cubic equation as the sum of a part we can solve plus a "small" perturbing part, namely

$$x^3 - 4x + (-0.001\,x + 0.002) = 0. \tag{11.29}$$

I place the word "small" in quotes because its meaning is not precisely clear. On one hand, for a typical value of x, say $x = 1$, the "big" part is -3 while the small part is only 0.001. On the other hand, for the value $x = 0$, the "big" part is zero and the "small" part is 0.002. So for some values of x the "small" part is bigger than the "big" part. Mathematicians spend a lot of time figuring out a precise meaning of "big" versus "small" in this context, but we don't need to follow their figurings. It's enough for us that the perturbing part is, in some general way, small compared to the remaining part of the problem, the part that we can solve exactly.

To save space, I'll introduce the constant T to mean "thousandths", and write our problem as

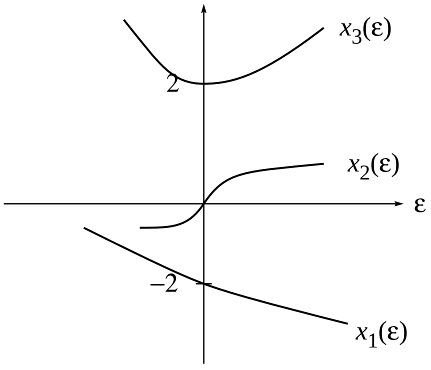
$$x^3 - 4x + T(-x + 2) = 0. \tag{11.30}$$

And now I'll generalize this problem by inserting a variable ϵ in front of the "small" part:

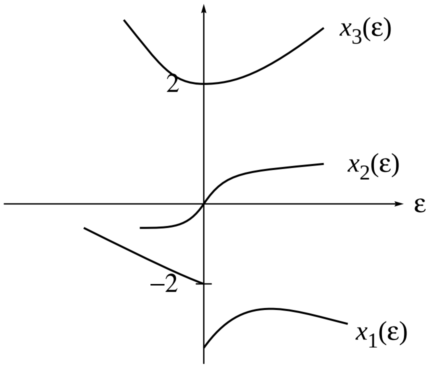
$$x^3 - 4x + \epsilon T(-x + 2) = 0. \tag{11.31}$$

The variable ϵ enables us to interpolate smoothly from the problem we're interested in, with $\epsilon = 1$, to the problem we know how to solve, with $\epsilon = 0$.

Instead of solving one cubic equation, the problem with $\epsilon = 1$, we're going to try to solve an infinite number of cubic equations, those with $0 \le \epsilon \le 1$. For example, I can call the smallest of these solutions $x_1(\epsilon)$. I don't know much about $x_1(\epsilon)$ — I know only that $x_1(0) = -2$ — but I have an expectation: I expect that $x_1(\epsilon)$ will behave smoothly as a function of ϵ, for example something like this



and I expect that it won't have jumps or kinks like this



Because of this expectation, I expect that I can write $x_1(\epsilon)$ as a Taylor series:

$$\begin{aligned} x_1(\epsilon) &= \sum_{k=0}^{\infty} a_k \epsilon^k && (11.32) \\ &= -2 + a_1\epsilon + a_2\epsilon^2 + \mathcal{O}(\epsilon^3) && (11.33) \end{aligned}$$

This function $x_1(\epsilon)$ has to satisfy

$$x_1^3(\epsilon) - (4+\epsilon T)x_1(\epsilon) + 2\epsilon T = 0. \qquad (11.34)$$

I can write the middle term above as an expansion in powers of ϵ using equation (11.33):

$$\begin{aligned} -4x_1(\epsilon) &= 8 - \epsilon(4a_1) && - \epsilon^2(4a_2) && + \mathcal{O}(\epsilon^3) \\ -\epsilon T x_1(\epsilon) &= + \epsilon(2T) && - \epsilon^2(Ta_1) && + \mathcal{O}(\epsilon^3) \\ -(4+\epsilon T)x_1(\epsilon) &= 8 + \epsilon(-4a_1 + 2T) && + \epsilon^2(-4a_2 - Ta_1) && + \mathcal{O}(\epsilon^3) \end{aligned}$$

With just a bit more effort, I can work out the left-most term in equation (11.34) as an expansion:

$$\begin{aligned} x_1^2(\epsilon) &= 4 \quad - \epsilon(4a_1) && + \epsilon^2(-4a_2 + a_1^2) && + \mathcal{O}(\epsilon^3) \\ x_1^3(\epsilon) &= -8 - \epsilon(-12a_1) && + \epsilon^2(12a_2 - 6a_1^2) && + \mathcal{O}(\epsilon^3) \end{aligned}$$

So finally, I have worked out the expansion of every term in equation (11.34):

$$\begin{aligned} x_1^3(\epsilon) &= -8 - \epsilon(-12a_1) && + \epsilon^2(12a_2 - 6a_1^2) && + \mathcal{O}(\epsilon^3) \\ -(4+\epsilon T)x_1(\epsilon) &= 8 \quad + \epsilon(-4a_1 + 2T) && + \epsilon^2(-4a_2 - Ta_1) && + \mathcal{O}(\epsilon^3) \\ 2\epsilon T &= \quad + \epsilon(2T) \end{aligned}$$

Summing the three equations above must, according to equation (11.34), produce zero:

$$\begin{aligned} 0 &= (-8+8) + \epsilon(12a_1 - 4a_1 + 4T) + \epsilon^2(12a_2 - 6a_1^2 - 4a_2 - Ta_1) + \mathcal{O}(\epsilon^3) \\ &= (-8+8) + \epsilon(8a_1 + 4T) \qquad\qquad + \epsilon^2(8a_2 - 6a_1^2 - Ta_1) \qquad\qquad + \mathcal{O}(\epsilon^3) \end{aligned}$$

Now, because the expression on the right must vanish for any value of ϵ, all the coefficients must vanish. First we must have that $(-8+8) = 0$, which checks out. Then the term linear in ϵ must vanish, so

$$(8a_1 + 4T) = 0 \qquad \text{whence} \qquad a_1 = -\tfrac{1}{2}T.$$

And the term quadratic in ϵ must vanish, so

$$(8a_2 - 6a_1^2 - Ta_1) = 0 \qquad \text{whence} \qquad a_2 = \tfrac{3}{4}a_1^2 + \tfrac{1}{8}Ta_1 = \tfrac{1}{8}T^2.$$

The expansion for $x_1(\epsilon)$ is thus

$$x_1(\epsilon) = -2 - \tfrac{1}{2}T\epsilon + \tfrac{1}{8}T^2\epsilon^2 + \mathcal{O}(\epsilon^3)$$

If we set $\epsilon = 1$ and ignore the terms $\mathcal{O}(\epsilon^3)$, we find

$$x_1(1) \approx -2.000399875$$

and comparison to the exact solution of the cubic equation (which is much more difficult to work through) shows that this result is accurate to one part in a billion.

11.3 Derivation of perturbation theory for the energy eigenproblem

Approach

To solve the energy eigenproblem for the Hamiltonian $\hat{H}^{(0)} + \hat{H}^{'}$, where the solution

$$\hat{H}^{(0)}|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle \tag{11.35}$$

is known and where $\hat{H}^{'}$ is "small" compared with $\hat{H}^{(0)}$, (for example the Stark effect, section 18.1)we set

$$\hat{H}(\epsilon) = \hat{H}^{(0)} + \epsilon\hat{H}^{'} \tag{11.36}$$

and then find $|n(\epsilon)\rangle$ and $E_n(\epsilon)$ such that

$$\hat{H}(\epsilon)|n(\epsilon)\rangle = E_n(\epsilon)|n(\epsilon)\rangle \tag{11.37}$$

and

$$\langle n(\epsilon)|n(\epsilon)\rangle = 1. \tag{11.38}$$

Intermediate goal

Find $|\bar{n}(\epsilon)\rangle$ and $E_n(\epsilon)$ such that

$$\hat{H}(\epsilon)|\bar{n}(\epsilon)\rangle = E_n(\epsilon)|\bar{n}(\epsilon)\rangle \tag{11.39}$$

and

$$\langle n^{(0)}|\bar{n}(\epsilon)\rangle = 1. \tag{11.40}$$

Then our final goal will be

$$|n(\epsilon)\rangle = \frac{|\bar{n}(\epsilon)\rangle}{\langle\bar{n}(\epsilon)|\bar{n}(\epsilon)\rangle^{1/2}}. \tag{11.41}$$

Remarkably, it often turns out to be good enough to reach our intermediate goal of finding $|\bar{n}(\epsilon)\rangle$, and one can then invent tricks for extracting information from these unnormalized eigenstates.

Initial assumption

We make the standard perturbation theory guess:

$$\begin{aligned} |\bar{n}(\epsilon)\rangle &= |n^{(0)}\rangle + \epsilon|\bar{n}^{(1)}\rangle + \epsilon^2|\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3) && (11.42)\\ E_n(\epsilon) &= E_n^{(0)} + \epsilon E_n^{(1)} + \epsilon^2 E_n^{(2)} + \mathcal{O}(\epsilon^3) && (11.43) \end{aligned}$$

[Note that the set $\{|\bar{n}^{(1)}\rangle\}$ is not complete, or orthonormal, or any other good thing.]

Consequences of the magnitude choice

The choice $\langle n^{(0)}|\bar{n}(\epsilon)\rangle = 1$ (as opposed the the more usual $\langle\bar{n}(\epsilon)|\bar{n}(\epsilon)\rangle = 1$) gives rise to interesting and useful consequences. First, take the inner product of $|n^{(0)}\rangle$ with equation (11.42)

$$\begin{aligned} \langle n^{(0)}|\bar{n}(\epsilon)\rangle &= \langle n^{(0)}|n^{(0)}\rangle + \epsilon\langle n^{(0)}|\bar{n}^{(1)}\rangle + \epsilon^2\langle n^{(0)}|\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3)\\ 1 &= 1 + \epsilon\langle n^{(0)}|\bar{n}^{(1)}\rangle + \epsilon^2\langle n^{(0)}|\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3) \end{aligned}$$

Because this relationship holds for *all* values of ϵ, the coefficient of each ϵ^m must vanish:

$$\langle n^{(0)}|\bar{n}^{(m)}\rangle = 0 \qquad m = 1, 2, 3, \ldots. \tag{11.44}$$

Whence

$$\begin{aligned}
\langle\bar{n}(\epsilon)|\bar{n}(\epsilon)\rangle &= \left[\langle n^{(0)}| + \epsilon\langle\bar{n}^{(1)}| + \epsilon^2\langle\bar{n}^{(2)}| + \mathcal{O}(\epsilon^3)\right] \\
&\qquad \times \left[|n^{(0)}\rangle + \epsilon|\bar{n}^{(1)}\rangle + \epsilon^2|\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3)\right] \\
&= \langle n^{(0)}|n^{(0)}\rangle + \epsilon\left[\langle\bar{n}^{(1)}|n^{(0)}\rangle + \langle n^{(0)}|\bar{n}^{(1)}\rangle\right] \\
&\qquad + \epsilon^2\left[\langle\bar{n}^{(2)}|n^{(0)}\rangle + \langle\bar{n}^{(1)}|\bar{n}^{(1)}\rangle + \langle n^{(0)}|\bar{n}^{(2)}\rangle\right] + \mathcal{O}(\epsilon^3) \\
&= 1 + \epsilon\left[0+0\right] + \epsilon^2\left[0 + \langle\bar{n}^{(1)}|\bar{n}^{(1)}\rangle + 0\right] + \mathcal{O}(\epsilon^3) \\
&= 1 + \epsilon^2\langle\bar{n}^{(1)}|\bar{n}^{(1)}\rangle + \mathcal{O}(\epsilon^3). \qquad (11.45)
\end{aligned}$$

In other words, while the vector $|\bar{n}(\epsilon)\rangle$ is not exactly normalized, it is "nearly normalized" — the norm differs from 1 by small, second-order terms.

Developing the perturbation expansion

What came before was just warming up. We now go and plug our expansion guesses, equations (11.42) and (11.43) into

$$\hat{H}(\epsilon)|n(\epsilon)\rangle = E_n(\epsilon)|n(\epsilon)\rangle \qquad (11.46)$$

to find

$$\begin{aligned}
&\left[\hat{H}^{(0)} + \epsilon\hat{H}^{'}\right]\left[|n^{(0)}\rangle + \epsilon|\bar{n}^{(1)}\rangle + \epsilon^2|\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3)\right] \qquad (11.47) \\
&\quad = \left[E_n^{(0)} + \epsilon E_n^{(1)} + \epsilon^2 E_n^{(2)} + \mathcal{O}(\epsilon^3)\right]\left[|n^{(0)}\rangle + \epsilon|\bar{n}^{(1)}\rangle + \epsilon^2|\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3)\right].
\end{aligned}$$

Separating out powers of ϵ gives

$$\begin{aligned}
\hat{H}^{(0)}|n^{(0)}\rangle &= E_n^{(0)}|n^{(0)}\rangle \qquad (11.48) \\
\hat{H}^{(0)}|\bar{n}^{(1)}\rangle + \hat{H}^{'}|n^{(0)}\rangle &= E_n^{(1)}|n^{(0)}\rangle + E_n^{(0)}|\bar{n}^{(1)}\rangle \qquad (11.49) \\
\hat{H}^{(0)}|\bar{n}^{(2)}\rangle + \hat{H}^{'}|\bar{n}^{(1)}\rangle &= E_n^{(2)}|n^{(0)}\rangle + E_n^{(1)}|\bar{n}^{(1)}\rangle + E_n^{(0)}|\bar{n}^{(2)}\rangle \qquad (11.50)
\end{aligned}$$

and so forth.

Finding the first-order energy shifts

How do we extract useful information from these expansion equations? Let's focus on what we know and what we want to find. We know $\hat{H}^{(0)}$, $\hat{H}^{'}$, $|n^{(0)}\rangle$, and $E_n^{(0)}$. From equation (11.49) we will find $E_n^{(1)}$ and $|\bar{n}^{(1)}\rangle$. Knowing these, from equation (11.50) we will find $E_n^{(2)}$ and $|\bar{n}^{(2)}\rangle$. And so forth.

To find the energy shifts $E_n^{(1)}$, we multiply equation (11.49) by $\langle n^{(0)}|$ to find

$$\begin{aligned} \langle n^{(0)}|\hat{H}^{(0)}|\bar{n}^{(1)}\rangle + \langle n^{(0)}|\hat{H}^{'}|n^{(0)}\rangle &= E_n^{(1)}\langle n^{(0)}|n^{(0)}\rangle + E_n^{(0)}\langle n^{(0)}|\bar{n}^{(1)}\rangle \\ E_n^{(0)}\langle n^{(0)}|\bar{n}^{(1)}\rangle + \langle n^{(0)}|\hat{H}^{'}|n^{(0)}\rangle &= E_n^{(1)} + E_n^{(0)}\langle n^{(0)}|\bar{n}^{(1)}\rangle \end{aligned} \tag{11.51}$$

Or,

$$E_n^{(1)} = \langle n^{(0)}|\hat{H}^{'}|n^{(0)}\rangle. \tag{11.52}$$

Often you need only these energies, not the states, and you can stop here. But if you do need the states. . .

Finding the first-order state shifts

We will find the state shifts $|\bar{n}^{(1)}\rangle$ by finding all the components of $|\bar{n}^{(1)}\rangle$ in the unperturbed basis $\{|m^{(0)}\rangle\}$.

Multiply equation (11.49) by $\langle m^{(0)}|$ $(m \neq n)$ to find

$$\begin{aligned} \langle m^{(0)}|\hat{H}^{(0)}|\bar{n}^{(1)}\rangle + \langle m^{(0)}|\hat{H}^{'}|n^{(0)}\rangle &= E_n^{(1)}\langle m^{(0)}|n^{(0)}\rangle + E_n^{(0)}\langle m^{(0)}|\bar{n}^{(1)}\rangle \\ E_m^{(0)}\langle m^{(0)}|\bar{n}^{(1)}\rangle + \langle m^{(0)}|\hat{H}^{'}|n^{(0)}\rangle &= 0 + E_n^{(0)}\langle m^{(0)}|\bar{n}^{(1)}\rangle \\ \langle m^{(0)}|\hat{H}^{'}|n^{(0)}\rangle &= (E_n^{(0)} - E_m^{(0)})\langle m^{(0)}|\bar{n}^{(1)}\rangle \end{aligned} \tag{11.53}$$

Now, *if* the state $|n^{(0)}\rangle$ is non-degenerate, then $E_m^{(0)} \neq E_n^{(0)}$ and we can divide both sides to find

$$\langle m^{(0)}|\bar{n}^{(1)}\rangle = \frac{\langle m^{(0)}|\hat{H}^{'}|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} \qquad (m \neq n) \tag{11.54}$$

But we already know, from equation (11.44), that

$$\langle n^{(0)}|\bar{n}^{(1)}\rangle = 0. \tag{11.55}$$

So now all the amplitudes $\langle m^{(0)}|\bar{n}^{(1)}\rangle$ are known, and therefore the vector is known:

$$|\bar{n}^{(1)}\rangle = \sum_m |m^{(0)}\rangle\langle m^{(0)}|\bar{n}^{(1)}\rangle \tag{11.56}$$

In conclusion — if $|n^{(0)}\rangle$ is non-degenerate

$$|\bar{n}^{(1)}\rangle = \sum_{m\neq n} |m^{(0)}\rangle \frac{\langle m^{(0)}|\hat{H}^{'}|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}. \tag{11.57}$$

11.4 Perturbation theory for the energy eigenproblem: Summary of results

Given: Solution for the $\hat{H}^{(0)}$ eigenproblem:

$$\hat{H}^{(0)}|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle \qquad \langle n^{(0)}|n^{(0)}\rangle = 1. \tag{11.58}$$

Find: Solution for the $\hat{H}^{(0)} + \epsilon\hat{H}^{'}$ eigenproblem:

$$(\hat{H}^{(0)} + \epsilon\hat{H}^{'})|n(\epsilon)\rangle = E_n(\epsilon)|n(\epsilon)\rangle \qquad \langle n(\epsilon)|n(\epsilon)\rangle = 1. \tag{11.59}$$

Define the "matrix elements"

$$\langle n^{(0)}|\hat{H}^{'}|m^{(0)}\rangle = H'_{nm}. \tag{11.60}$$

The solutions are (provided $|n^{(0)}\rangle$ is not degenerate):

$$E_n(\epsilon) = E_n^{(0)} + \epsilon H'_{nn} + \epsilon^2 \sum_{m\neq n} \frac{H'_{nm}H'_{mn}}{E_n^{(0)} - E_m^{(0)}} + \mathcal{O}(\epsilon^3) \tag{11.61}$$

$$\begin{aligned}
|n(\epsilon)\rangle = {} & |n^{(0)}\rangle \\
& +\epsilon \sum_{m\neq n} |m^{(0)}\rangle \frac{H'_{mn}}{E_n^{(0)} - E_m^{(0)}} \\
& +\epsilon^2 \left[\sum_{m\neq n}\sum_{\ell\neq n} |m^{(0)}\rangle \frac{H'_{m\ell}H'_{\ell n}}{(E_n^{(0)} - E_m^{(0)})(E_n^{(0)} - E_\ell^{(0)})} \right. \\
& \qquad \left. - \sum_{m\neq n} |m^{(0)}\rangle \frac{H'_{nn}H'_{mn}}{(E_n^{(0)} - E_m^{(0)})^2} - |n^{(0)}\rangle \frac{1}{2} \sum_{m\neq n} \frac{H'_{nm}H'_{mn}}{(E_n^{(0)} - E_m^{(0)})^2} \right] \\
& +\mathcal{O}(\epsilon^3)
\end{aligned} \tag{11.62}$$

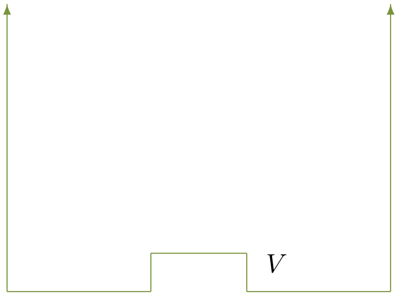
Rules of thumb concerning perturbation theory

- There is no guarantee that the series is convergent, or even asymptotic.
- But experience says "stop at the first non-vanishing energy correction".
- The wavefunctions produced are notoriously poor. How can the energies be good when the wavefunctions are poor? See section 17.3.
- The technique is generally useful for many mathematical problems: classical mechanics, fluid mechanics, etc. Even for solving cubic equations!
- Technique is never guaranteed to succeed, but it is likely to fail (and perhaps fail silently!) if there are degenerate energy states. In this case $E_n^{(0)} = E_m^{(0)}$, so second-order term perhaps diverges, despite the fact that the first-order term $\langle n^{(0)}|\hat{H}'|n^{(0)}\rangle$ looks perfectly fine. (Stark effect in hydrogen.)

Problems

11.2 Square well with a bump

An infinite square well of width L is perturbed by putting in a bit of potential of height V and width a in the middle of the well. Find the first order energy shifts for all the energy eigenstates, and the first order perturbed wavefunction for the ground state (your result will be an infinite series). (Note: Many of the required matrix elements will vanish! Before you integrate, ask yourself whether the integrand is odd.) When $a = L$ the perturbed problem can be solved exactly. Compare the perturbed energies with the exact energies and the perturbed ground state wavefunction with the exact ground state wavefunction.



11.3 Anharmonic oscillator

a. Show that for the simple harmonic oscillator,

$$\langle m|\hat{x}^3|n\rangle = \sqrt{\left(\frac{\hbar}{2m\omega}\right)^3}\ \Big[\sqrt{n(n-1)(n-2)}\,\delta_{m,n-3} + 3\sqrt{n^3}\,\delta_{m,n-1} + 3\sqrt{(n+1)^3}\,\delta_{m,n+1} + \sqrt{(n+1)(n+2)(n+3)}\,\delta_{m,n+3}\Big] . \quad (11.63)$$

b. Recall that the simple harmonic oscillator is always an approximation. The real problem always has a potential $V(x) = \frac{1}{2}kx^2 +$

$bx^3 + cx^4 + \cdots$. The contributions beyond $\frac{1}{2}kx^2$ are called "anharmonic terms". Ignore all the anharmonic terms except for bx^3. Show that to leading order the n^{th} energy eigenvalue changes by

$$-\frac{b^2}{\hbar\omega}\left(\frac{\hbar}{2m\omega}\right)^3(30n^2+30n+11). \tag{11.64}$$

Note that these shifts are not "small" when n is large, in which case it is not appropriate to truncate the perturbation series at leading order. Explain physically why you don't expect the shifts to be small for large n.

11.4 Slightly relativistic simple harmonic oscillator

You know that the concept of potential energy is not applicable in relativistic situations. One consequence of this is that the only fully relativistic quantum theories possible are quantum field theories. However there do exist situations where a particle's motion is "slightly relativistic" (say, $v/c \sim 0.1$) and where the force responds quickly enough to the particle's position that the potential energy concept has approximate validity. For a mass on a spring, this situation hold when the spring's response time is much less than the period.

a. Show that a reasonable approximate Hamiltonian for such a "slightly relativistic SHO" is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\hat{x}^2 - \frac{1}{8c^2m^3}\hat{p}^4. \tag{11.65}$$

b. Show that

$$\langle m|\hat{p}^4|0\rangle = \left(\frac{m\hbar\omega}{2}\right)^2(3\,\delta_{m,0} - 6\sqrt{2}\,\delta_{m,2} + 2\sqrt{6}\,\delta_{m,4}). \tag{11.66}$$

c. Calculate the leading non-vanishing energy shift of the ground state due to this relativistic perturbation.

d. Calculate the leading corrections to the ground state eigenvector $|0\rangle$.

11.5 Two-state systems

The most general Hamiltonian for a two state system (e.g. spin $\frac{1}{2}$, neutral K meson, ammonia molecule) is represented by

$$a_0I + a_1\sigma_1 + a_3\sigma_3 \tag{11.67}$$

where a_0, a_1, and a_3 are real numbers and the σ's are Pauli matrices. (See problem 511.)

a. Assume $a_3 = 0$. Solve the energy eigenproblem.
b. Now assume $a_3 \ll a_0 \approx a_1$. Use perturbation theory to find the leading order shifts in the energy eigenvalues and eigenstates.
c. Find the energy eigenvalues exactly and show that they agree with the perturbation theory results when $a_3 \ll a_0 \approx a_1$.

11.6 **Degenerate perturbation theory in a two-state system**
Consider a two state system with a Hamiltonian represented in some basis by

$$a_0 I + a_1\sigma_1 + a_3\sigma_3. \tag{11.68}$$

We shall call the basis for this representation the "initial basis". This problem shows how to use perturbation theory to solve (approximately) the energy eigenproblem in the case $a_0 \gg a_1 \approx a_3$.

$$\hat{H}^{(0)} = \begin{pmatrix} a_0 & 0 \\ 0 & a_0 \end{pmatrix} \qquad \hat{H}' = \begin{pmatrix} a_3 & a_1 \\ a_1 & -a_3 \end{pmatrix} \tag{11.69}$$

In this case the unperturbed Hamiltonian is degenerate. The initial basis

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \tag{11.70}$$

is a perfectly acceptable energy eigenbasis (both states have energy a_0), but the basis

$$\left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}, \tag{11.71}$$

for example, is just as good.

a. Show that if the non-degenerate formula $E_n^{(1)} = \langle n^{(0)}|\hat{H}'|n^{(0)}\rangle$ were applied (or rather, misapplied) to this problem, then the formula would produce different energy shifts depending upon which basis was used!

Which, if either, are the true energy shifts? The answer comes from equation (11.53), namely

$$(E_n^{(0)} - E_m^{(0)})\langle m^{(0)}|\bar{n}^{(1)}\rangle = \langle m^{(0)}|\hat{H}'|n^{(0)}\rangle \quad \text{whenever } m \neq n. \tag{11.72}$$

This equation was derived from the fundamental assumption that $|n(\epsilon)\rangle$ and $E_n(\epsilon)$ could be expanded in powers of ϵ. If the unperturbed states

$|n^{(0)}\rangle$ and $|m^{(0)}\rangle$ are degenerate, then $E_n^{(0)} = E_m^{(0)}$ and the above equation demands that

$$\langle m^{(0)}|\hat{H}'|n^{(0)}\rangle = 0 \quad \text{whenever } m \neq n \text{ and } E_n^{(0)} = E_m^{(0)}. \tag{11.73}$$

If this does not apply, then the fundamental assumption must be wrong.

And this answers the question of which basis to use! Consistency demands the use of a basis in which the perturbing Hamiltonian is diagonal. (The Hermiticity of $\hat{H}'$ guarantees that such a basis exists.)

b. Without finding this diagonalizing basis, find the representation of $\hat{H}'$ in it.
c. Find the representation of $\hat{H}^{(0)}$ in the diagonalizing basis. (Trick question.)
d. What are the energy eigenvalues of the full Hamiltonian $\hat{H}^{(0)}+\hat{H}'$? (Not "correct to some order in perturbation theory," but the exact eigenvalues!)
e. Still without explicitly producing the diagonalizing basis, show that the states in that basis are exact energy eigenstates of the full Hamiltonian.
f. (Optional) If you're ambitious, you may now go ahead and show that the (normalized) diagonalizing basis vectors are

$$\frac{1}{\sqrt{2}\sqrt{a_1^2+a_3^2-a_3\sqrt{a_1^2+a_3^2}}}\begin{pmatrix} +a_1 \\ -a_3+\sqrt{a_1^2+a_3^2}\end{pmatrix} = \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix},$$

$$\frac{1}{\sqrt{2}\sqrt{a_1^2+a_3^2+a_3\sqrt{a_1^2+a_3^2}}}\begin{pmatrix} -a_1 \\ +a_3+\sqrt{a_1^2+a_3^2}\end{pmatrix} = \begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix},$$

where

$$\tan\theta = \frac{a_1}{a_3+\sqrt{a_1^2+a_3^2}}.$$

Coda: Note the reasoning of degenerate perturbation theory: We expand about the basis that diagonalizes $\hat{H}'$ because expansion about any other basis is immediately self-contradictory, not because this basis is guaranteed to produce a sensible expansion. As usual in perturbation theory, we have no guarantee that this expansion makes sense. We do, however, have a guarantee that any other expansion does not make sense.

Chapter 12

More Dimensions, More Particles

We've been investigating a single, spinless particle moving in one dimension for so long that you might get the misimpression that quantum mechanics is about single particles. As Richard Feynman said: "Mistakes are often made by physics students at first because . . . they work for so long analyzing events involving a single [particle] that they begin to think that the [wavefunction] is somehow associated with the [particle]" rather than with the system.[1]

12.1 More degrees of freedom

Let's think of the process of adding degrees of freedom.

---

[1] Richard P. Feynman, *QED: The Strange Theory of Light and Matter* (Princeton University Press, Princeton, New Jersey, 1985) pages 75–76.

First consider a **spinless particle in one dimension**:

(1) The particle's state is described by a vector $|\psi\rangle$.

(2) The vector has dimension ∞, reflecting the fact that any basis, for example the basis $\{|x\rangle\}$, has ∞ members. (No basis is better than another other basis — for every statement below concerning position there is a parallel statement concerning momentum — but for concreteness we'll discuss only position.)

(3) These basis members are orthonormal,

$$\langle x|x'\rangle = \delta(x-x'), \tag{12.1}$$

and complete

$$\hat{1} = \int_{-\infty}^{+\infty} dx\, |x\rangle\langle x|. \tag{12.2}$$

⟦These two equations may seem recondite, formal, and purely mathematical, but in fact they embody the direct, physical results of measurement experiments: Completeness reflects the fact that when the particle's position is measured, it is found to have a position. Orthonormality reflects the fact that when the particle's position is measured, it is found in only one position. *Statement should be refined. Connection between completeness and interference?*⟧

(4) The state $|\psi\rangle$ is represented (in the position basis) by the numbers $\langle x|\psi\rangle = \psi(x)$. In symbols

$$|\psi\rangle \doteq \langle x|\psi\rangle = \psi(x). \tag{12.3}$$

(5) When position is measured, the probability of measuring a position within a window of width dx about x_0 is

$$|\psi(x_0)|^2\, dx. \tag{12.4}$$

Exercise 12.A. The last sentence would be more compact if I wrote "When the position is measured, the probability of finding the particle within ...". Why didn't I use this more concise wording?

Now consider a **spin-$\frac{1}{2}$ particle in one dimension**:

(1) The particle's state is described by a vector $|\psi\rangle$.

(2) The vector has dimension $\infty \times 2$, reflecting the fact that any basis, for example the basis $\{|x,+\rangle, |x,-\rangle\}$, has $\infty \times 2$ members. (No basis is better than another other basis — for every statement below concerning position plus projection on a vertical axis there is a parallel statement concerning momentum plus projection of a horizontal axis — but for concreteness we'll discuss only position plus projection of a vertical axis.) ⟦For example, the state $|9,+\rangle$ represents a particle at position 9 with spin $+$. The state
$$\tfrac{4}{5}|9,+\rangle - i\tfrac{3}{5}|7,-\rangle] \tag{12.5}$$
represents a particle with amplitude $\frac{4}{5}$ to be at position 9 with spin $+$ and amplitude $-i\frac{3}{5}$ to be at position 7 with spin $-$, but with no amplitude to be at position 9 with spin $-$, and no amplitude to be at position 6 with any spin.⟧

(3) These basis members are orthonormal,
$$\begin{aligned}\langle x,+|x',+\rangle &= \delta(x-x')\\ \langle x,-|x',-\rangle &= \delta(x-x')\\ \langle x,+|x',-\rangle &= 0\\ \langle x,i|x',j\rangle &= \delta(x-x')\delta_{i,j}\end{aligned} \tag{12.6}$$
and complete
$$\begin{aligned}\hat{1} &= \int_{-\infty}^{+\infty} dx\,|x,+\rangle\langle x,+| + \int_{-\infty}^{+\infty} dx\,|x,-\rangle\langle x,-|\\ \hat{1} &= \sum_{i=+,-}\int_{-\infty}^{+\infty} dx\,|x,i\rangle\langle x,i|\end{aligned} \tag{12.7}$$

(4) The state $|\psi\rangle$ is represented (in this basis) by the numbers
$$\begin{pmatrix}\langle x,+|\psi\rangle\\ \langle x,-|\psi\rangle\end{pmatrix} = \begin{pmatrix}\psi_+(x)\\ \psi_-(x)\end{pmatrix}. \tag{12.8}$$
or by
$$\psi(x,i) \tag{12.9}$$
where x takes on continuous values from $-\infty$ to $+\infty$ but i takes on only the two possible values $+$ or $-$. (Some people write this as $\psi_i(x)$ rather than as $\psi(x,i)$, but it is not legitimate to denigrate the variable i to subscript rather than argument just because it happens to be discrete instead of continuous.)

(5) When both spin projection and position are measured, the probability of measuring projection + and position within a window of width dx about x_0 is

$$|\psi_+(x_0)|^2\,dx. \tag{12.10}$$

When position alone is measured, the probability of measuring position within a window of width dx about x_0 is

$$|\psi_+(x_0)|^2\,dx + |\psi_-(x_0)|^2\,dx. \tag{12.11}$$

When spin projection alone is measured, the probability of measuring projection + is

$$\int_{-\infty}^{+\infty} |\psi_+(x)|^2\,dx. \tag{12.12}$$

The proper way of expressing the representation of the state $|\psi\rangle$ in the $\{|x,+\rangle, |x,-\rangle\}$ basis is through the so-called "spinor" above, namely

$$|\psi\rangle \doteq \begin{pmatrix} \psi_+(x) \\ \psi_-(x) \end{pmatrix}.$$

Sometimes you'll see this written instead as

$$|\psi\rangle \doteq \psi_+(x)|+\rangle + \psi_-(x)|-\rangle.$$

Ugh! This is bad notation, because it confuses the state (something like $|\psi\rangle$, a vector) with the representation of a state in a particular basis (something like $\langle x, i|\psi\rangle$, a set of amplitudes). Nevertheless, you'll see it used.

This example represents the way to add degrees of freedom to a description, namely by using a larger basis set. In this case I've merely doubled the size of the basis set, by including spin. I could also add a second dimension by adding the possibility of motion in the y direction, and so forth.

Exercise 12.B. A spin-$\frac{1}{2}$ particle in one dimension is in state (12.5) when both its position and spin are measured. What is the probability of finding the particle at position 9 with spin +? At position 7 with spin −? At position 9 with spin −? At position 6 with any spin?

Consider a **spinless particle in three dimensions**:

(1) The particle's state is described by a vector $|\psi\rangle$.

(2) The vector has dimension ∞^3, reflecting the fact that any basis, for example the basis $\{|x,y,z\rangle\}$ — which is also written as $\{|\vec{r}\rangle\}$ — has ∞^3 members. (No basis is better than another other basis — for every statement below concerning position there is a parallel statement concerning momentum — but for concreteness we'll discuss only position.)

(3) These basis members are orthonormal,

$$\langle x,y,z|x',y',z'\rangle = \delta(x-x')\delta(y-y')\delta(z-z'), \tag{12.13}$$

which is also written as

$$\langle\vec{r}\,|\vec{r}\,'\rangle = \delta(\vec{r}-\vec{r}\,'). \tag{12.14}$$

In addition, the basis members are complete

$$\hat{1} = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz\, |x,y,z\rangle\langle x,y,z|, \tag{12.15}$$

which is also written as

$$\hat{1} = \int_{-\infty}^{+\infty} d^3r\, |\vec{r}\,\rangle\langle\vec{r}\,|. \tag{12.16}$$

(4) The state $|\psi\rangle$ is represented (in the position basis) by the numbers $\langle\vec{r}\,|\psi\rangle = \psi(\vec{r}\,)$ (a complex-valued function of three variables, a vector argument).

(5) When position is measured, the probability of measuring a position within a box of volume d^3r about $\vec{r}_0$ is

$$|\psi(\vec{r}_0)|^2\, d^3r. \tag{12.17}$$

12.1.1 *The symbol for all variables*

When a silver atom moves in three dimensions, the wavefunction takes the form

$$\psi(x,y,z,m_s) \equiv \psi(\underset{\sim}{x}), \tag{12.18}$$

where the single undertilde symbol $\underset{\sim}{x}$ stands for the four variables x,y,z,m_s. [Because the variables x, y, and z are continuous, while the variable m_s is discrete, one sometimes sees the dependence on m_s written as a subscript rather than as an argument: $\psi_{m_s}(x,y,z)$. This is a bad habit: m_s is a variable not a label, and it should not be notated as a second-class variable just because it's discrete.]

Alternatively, you might prefer to use the wavefunction in momentum space, and keep track, not of the spin projection on the z axis, but on the axis rotated from vertical by 27°. In this case, the wavefunction takes the form

$$\tilde{\psi}(p_x, p_y, p_z, m_{27^\circ}). \tag{12.19}$$

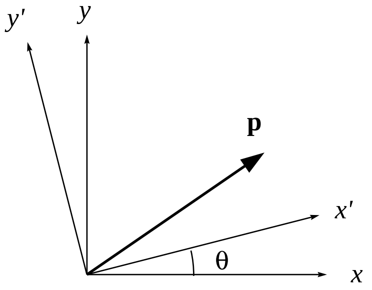
This description is less conventional but just as good as the description in terms of x, y, z, and m_z. But we will still call the set of four variables needed to describe the wavefunction (three of them continuous and one discrete) by the single symbol $\underset{\sim}{x}$.

12.2 Vector operators

So much for states... what about operators?

The general idea of a vector is that it's "something like an arrow". But in what way like an arrow? If you work with the components of a vector, how can the components tell you that they represent something that's "like an arrow"?

Consider the vector momentum $\vec{p}$. If the coordinate axes are x and y, the components of the vector $\vec{p}$ are p_x and p_y. But if the coordinate axes are x' and y', then the components of the vector $\vec{p}$ are $p_{x'}$ and $p_{y'}$. It's the same vector, but it has different components using different coordinate axes.



How are these two sets of coordinates related? It's not hard to show that they're related through

$$\begin{aligned} p_{x'} &= p_x \cos\theta + p_y \sin\theta \\ p_{y'} &= -p_x \sin\theta + p_y \cos\theta \end{aligned} \tag{12.20}$$

(There's a similar but more complicated formula for three-dimensional vectors.)

We use this same formula for change of coordinates under rotation whether it's a position vector or a velocity vector or a momentum vector, despite the fact that position, velocity, and momentum are very different in character. It is in this sense that position, velocity, and momentum are all "like an arrow" and it is in this way that the components of a vector show that the entity behaves "like an arrow".

Now, what is a "vector operator"? In two dimensions, it's a set of two operators that transform under rotation just as the two components of a vector do:

$$\begin{aligned} \hat{p}_{x'} &= \hat{p}_x \cos\theta + \hat{p}_y \sin\theta \\ \hat{p}_{y'} &= -\hat{p}_x \sin\theta + \hat{p}_y \cos\theta \end{aligned} \tag{12.21}$$

(There's a similar but more complicated formula for three-dimensional vector operators.)

Meanwhile, a "scalar operator" is one that doesn't change when the coordinate axes are rotated.

For every vector operator there is a scalar operator

$$\hat{p}^2 = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2. \tag{12.22}$$

12.3 Multiple particles

In section 12.1 we considered adding spin and spatial degrees of freedom for a single particle. But the same scheme works for adding additional particles. (There are peculiarities that apply to the identical particles — see chapter 15 — so in this section we'll consider non-identical particles.)

Consider a system of **two spinless particles** (call them red and green) ambivating **in one dimension**:

(1) The system's state is described by a vector $|\psi\rangle$.

(2) The vector has dimension ∞^2, reflecting the fact that any basis, for example the basis $\{|x_R, x_G\rangle\}$ has ∞^2 members. (No basis is better than another other basis — for every statement below concerning two positions there is a parallel statement concerning two momenta — but for concreteness we'll discuss only position.)

(3) These basis members are orthonormal,

$$\langle x_R, x_G | x'_R, x'_G\rangle = \delta(x_R - x'_R)\delta(x_G - x'_G). \tag{12.23}$$

In addition, the basis members are complete

$$\hat{1} = \int_{-\infty}^{+\infty} dx_R \int_{-\infty}^{+\infty} dx_G\, |x_R, x_G\rangle\langle x_R, x_G|. \tag{12.24}$$

(4) The state $|\psi\rangle$ is represented (in the position basis) by the numbers

$$\langle x_R, x_G|\psi\rangle = \psi(x_R, x_G) \tag{12.25}$$

(a complex-valued function of a two-variable argument).

(5) When the positions of both particles are measured, the probability of finding the red particle within a window of width dx_A about x_A and the green particle within a window of width dx_B about x_B is

$$|\psi(x_A, x_B)|^2\, dx_A\, dx_B. \tag{12.26}$$

Do I need to mention the entirely parallel statements for a system of **two spinless particles** (call them red and green) ambivating **in three dimensions**?

(1) The system's state is described by a vector $|\psi\rangle$.

(2) The vector has dimension ∞^6, reflecting the fact that any basis, for example the basis $\{|\vec{r}_R, \vec{r}_G\rangle\}$ has ∞^6 members. (No basis is better than another other basis — for every statement below concerning two vector positions there is a parallel statement concerning two vector momenta — but for concreteness we'll discuss only position.)

(3) These basis members are orthonormal,

$$\langle \vec{r}_R, \vec{r}_G | \vec{r'}_R, \vec{r'}_G\rangle = \delta^{(3)}(\vec{r}_R - \vec{r'}_R)\delta\ (3)(\vec{r}_G - \vec{r'}_G). \tag{12.27}$$

In addition, the basis members are complete

$$\hat{1} = \int_{-\infty}^{+\infty} d^3r_R \int_{-\infty}^{+\infty} d^3r_G\, |\vec{r}_R, \vec{r}_G\rangle\langle \vec{r}_R, \vec{r}_G|. \tag{12.28}$$

(4) The state $|\psi\rangle$ is represented (in the position basis) by the numbers

$$\langle \vec{r}_R, \vec{r}_G|\psi\rangle = \psi(\vec{r}_R, \vec{r}_G) \tag{12.29}$$

(a complex-valued function of a six-variable argument).

(5) When the positions of both particles are measured, the probability of measuring the red particle within a box of volume d^3r_A about $\vec{r}_A$ and the green particle within a box of volume d^3r_B about $\vec{r}_B$ is

$$|\psi(\vec{r}_A, \vec{r}_B)|^2 \, d^3r_A \, d^3r_B. \tag{12.30}$$

12.4 The phenomena of quantum mechanics

We started (chapter 1) with the phenomena of quantum mechanics: quantization, probability, interference, and entanglement. We used these phenomena to build up the formalism of quantum mechanics: amplitudes, state vectors, operators, etc. (chapter 2).

We've been working at the level of formalism for so long that we're in danger of forgetting the phenomena that underlie the formalism: For example in this chapter we discussed how the formalism of quantum mechanics applies to continuous systems in three dimensions. It's time to return to the level of phenomena and ask how the phenomena of quantum mechanics generalize to continuous systems in three dimensions.

Interference

Interference of a particle — experiments of Tonomura:
A. Tonomura, J. Endo, T. Matsuda, T. Kawasaki, and H. Ezawa, "Demonstration of single-electron buildup of an interference pattern," *American Journal of Physics*, **57** (1989) 117–120.
http://www.hqrd.hitachi.co.jp/em/doubleslit.cfm

Entanglement

How does one describe the state of a single classical particle moving in one dimension? It requires two numbers: a position and a momentum (or a position and a velocity). Two particles moving in one dimension require merely that we specify the state of each particle: four numbers. Similarly specifying the state of three particles require six numbers and N particles require $2N$ numbers. Exactly the same specification counts hold if the particle moves relativistically.

How, in contrast, does one describe the state of a single quantal particle moving in one dimension? A problem arises at the very start, here, because the specification is given through a complex-valued wavefunction $\psi(x)$. Technically the specification requires an infinite number of numbers! Let's approximate the wavefunction through its value on a grid of, say, 100 points. This suggests that a specification requires 200 real numbers, a complex number at each grid point, but one number is taken care of through the overall phase of the wavefunction, and one through normalization. The specification actually requires 198 independent real numbers.

How does one describe the state of two quantal particles moving in one dimension? Now the wavefunction is a function of two variables $\psi(x_A, x_B)$. (This wavefunction might factorize into a function of x_A alone times a function of x_B alone, but it might not. If it does factorize, the two particles are unentangled, if it does not, the two particles are entangled. In the general quantal case a two-particle state is *not* specified by giving the state of each individual particle, because the individual particles might not have states.) The wavefunction of the system is a function of two-dimensional configuration space, so an approximation of the accuracy established previously requires a 100×100 grid of points. Each grid point carries one complex number, and again overall phase and normalization reduce the number of real numbers required by two. For two particles the specification requires $2 \times (100)^2 - 2 = 19998$ independent real numbers.

Similarly, specifying the state of N quantal particles moving in one dimension requires a wavefunction in N-dimensional configuration space which (for a grid of the accuracy we've been using) is specified through $2 \times (100)^N - 2$ independent real numbers.

The specification of a quantal state not only requires more real numbers than the specification of the corresponding classical state, but that number increases exponentially rather than linearly with the number of particles N.

The fact that a quantal state holds more information than a classical state is the fundamental reason that a quantal computer is (in principle) faster than a classical computer, and the basis for much of quantum information theory.

Relativity is different from classical physics, but no more complicated. Quantum mechanics, in contrast, is both *different* from and *richer* than classical physics. You may refer to this richness using terms like "splendor", or “abounding”, or “intricate”, or “ripe with possibilities”. Or you may refer to it using terms like “complicated”, or “messy”, or “full of details likely to trip the innocent”. It’s your choice how to react to this richness, but you can’t deny it.

Chapter 13

Angular Momentum

13.1 Angular momentum in classical mechanics

You remember angular momentum. For a single particle with position $\vec{r}$ and momentum $\vec{p}$, the angular momentum about the origin is

$$\vec{L} = \vec{r} \times \vec{p} = \hat{\imath}(yp_z - zp_y) + \hat{\jmath}(zp_x - xp_z) + \hat{k}(xp_y - yp_x). \tag{13.1}$$

And you remember that, in the absence of external torque, the angular momentum is conserved.

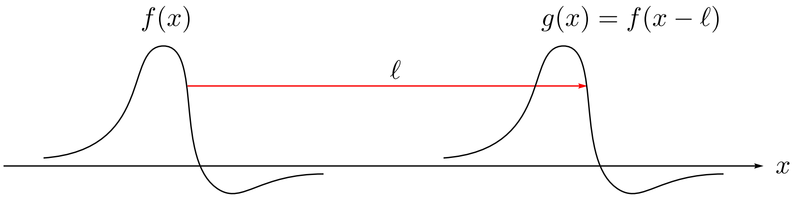
But I find it something of a mystery that angular momentum should be so important. Sure, I can define it and, once defined, I can prove that it's conserved in the absence of external torques. But whatever inspired anyone to define it?

13.2 Angular momentum and rotations

This section shows that angular momentum is intimately connected with rotations and that this connection inspires the definition. The connection can be made in classical mechanics as well as quantum mechanics, but the classical connection has always seemed (to me at least) contrived, whereas the quantal connection seems natural.

We start with a warm-up discussion:

13.2.1 *Linear momentum and translations*



The function $f(x)$ is translated through displacement ℓ to form the function $g(x) = f(x - \ell)$. I ask a purely mathematical question: What is an expression for the translation operator (which is clearly linear) on the space of functions?

$$\mathcal{T}_\ell[f(x)] = g(x) \tag{13.2}$$

To answer this question, we start with translations by a small displacement $\Delta\ell$:

$$\begin{aligned} \mathcal{T}_{\Delta\ell}[f(x)] &= f(x - \Delta\ell) \\ &\approx f(x) - \Delta\ell \frac{df}{dx} \\ &= \left[1 - \Delta\ell \frac{d}{dx}\right] f(x). \end{aligned} \tag{13.3}$$

To translate by a large displacement $\ell = N\Delta\ell$, simply translate by the small displacement N times:

$$\begin{aligned} \mathcal{T}_\ell &= [\mathcal{T}_{\Delta\ell}]^N \\ &= [\mathcal{T}_{\Delta\ell}]^{\ell/\Delta\ell} \\ &\approx \left[1 - \Delta\ell \frac{d}{dx}\right]^{\ell/\Delta\ell}, \end{aligned}$$

which is an approximate expression that becomes better and better as $\Delta\ell$ becomes smaller and smaller.

So what happens in the limit $\Delta\ell \to 0$? If the operator d/dx were a number, say S, we would know exactly what it do:

$$[1 - \Delta\ell\, S]^{\ell/\Delta\ell} = \exp\left\{\frac{\ell}{\Delta\ell} \ln\left[1 - \Delta\ell\, S\right]\right\}$$

but

$$\lim_{\Delta\ell\to 0}\left\{\frac{\ell}{\Delta\ell}\ln\left[1-\Delta\ell\, S\right]\right\} = -\ell S$$

so

$$\lim_{\Delta\ell\to 0}\left[1-\Delta\ell\, S\right]^{\ell/\Delta\ell} = e^{-\ell S}.$$

It is more difficult to perform this reasoning when the number S is replaced by the operator d/dx, but in fact the result still holds:

$$\mathcal{T}_\ell = e^{-\ell\frac{d}{dx}} = \sum_{n=0}^{\infty}\frac{(-\ell)^n}{n!}\frac{d^n}{dx^n}. \tag{13.4}$$

We have answered our purely mathematical question and can start thinking about physics again.

And now that we're thinking about physics, we recognize the representation in position space of the momentum operator:

$$\hat{p} \doteq -i\hbar\frac{d}{dx} \quad \text{or} \quad \frac{d}{dx} \doteq \frac{i\hat{p}}{\hbar}.$$

This inspires the definition of the quantal translation operator as

$$\hat{T}_\ell = e^{-i(\hat{p}/\hbar)\ell} \tag{13.5}$$

because, if

$$|\phi\rangle = \hat{T}_\ell|\psi\rangle,$$

then the wavefunction $\phi(x)$ is just the wavefunction $\psi(x)$ translated by a displacement ℓ:

$$\phi(x) = \psi(x-\ell) = \mathcal{T}_\ell[\psi(x)].$$

Okay, this is all very elegant, but if I really wanted to translate something I'd use a bulldozer. Can this tell us anything practical? It can. Suppose the potential energy function is a constant. Then $[\hat{H}, \hat{T}_\ell] = 0$ holds for any displacement ℓ. Consequently $[\hat{H}, \hat{p}] = 0$, whence momentum is conserved.

We don't have to work out in detail elaborate commutators in a specific representation. From this point of view, the conservation of momentum follows directly from the "homogeneity of space".

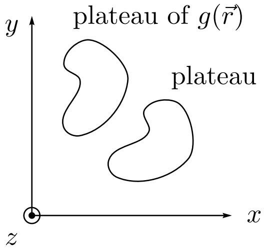
Exercise 13.A. Show that if $[\hat{A}, e^{x\hat{B}}] = 0$ for all x, then $[\hat{A}, \hat{B}] = 0$.

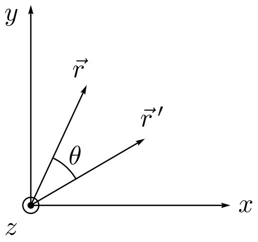
Exercise 13.B. Because $\hat{L}_z$ generates a rotation, any scalar operator $\hat{A}$ must have $[\hat{A}, \hat{L}_z] = 0$. (The same holds for $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_{47^\circ}$.) Verify this explicitly using $\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$ for the scalar operators (a) $\hat{r}^2 = \hat{x}^2 + \hat{y}^2 + \hat{z}^2$ and (b) $\hat{p}^2 = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2$.

Mention crystal momentum here? Exercise?

Now we're done with our warm-up discussion and ready to ask the next question: If linear momentum generates linear displacements, does angular momentum generate angular displacements (that is, rotations)?

13.2.2 *Angular momentum and rotations*





The function $f(\vec{r})$ is rotated through angle θ to form the function $g(\vec{r})$. In the figure above, the functions are indicated by a contour line surrounding a plateau, but everything about the function $f(\vec{r})$, valleys as well as peaks and plateaus, is rotated. The figure shows a rotation is about he z-axis (coming out of the page), but this is not restrictive, because we could just define the z-axis to be parallel to the rotation axis.

In symbols, we say that rotated function is defined through $g(\vec{r}) = f(\vec{r}\,')$, where $\vec{r}$ is the vector resulting from rotating $\vec{r}\,'$, as shown in figure right. We define the rotation operator through

$$g(\vec{r}) = \mathcal{R}_{\theta,\hat{k}}[f(\vec{r})], \tag{13.6}$$

where the subscript indicates a rotation by angle θ about axis $\hat{k}$, the unit vector in the positive z direction.

A few sketches will convince you that, for small rotation angles $\Delta\theta$ about the z-axis, the components of $\vec{r}\,'$ and of $\vec{r}$ are related through

$$\begin{aligned} x' &= x + \Delta\theta\, y \\ y' &= y - \Delta\theta\, x \\ z' &= z. \end{aligned}$$

So under these circumstances

$$\begin{aligned} g(x,y,z) &\approx f(x + \Delta\theta\, y, y - \Delta\theta\, x, z) \\ &\approx f(x,y,z) + \Delta\theta\, y\,\frac{\partial f}{\partial x} - \Delta\theta\, x\,\frac{\partial f}{\partial y} \\ &= f(x,y,z) - \Delta\theta\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) f(x,y,z) \end{aligned}$$

or, in other words,

$$\mathcal{R}_{\theta,\hat{k}}[f(\vec{r})] \approx \left[1 - \Delta\theta\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\right] f(\vec{r}). \tag{13.7}$$

Now follow the same reasoning we used for translations from equations (13.3) to (13.4). The result is

$$\mathcal{R}_{\theta,\hat{k}} = \exp\left\{-\theta\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\right\}. \tag{13.8}$$

Continuing to follow the reasoning we used for translations, we define the quantal operator

$$\hat{R}_{\theta,\hat{k}} = \exp\left\{-i\left[\left(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x\right)/\hbar\right]\theta\right\} = e^{-i(\hat{L}_z/\hbar)\theta}. \tag{13.9}$$

There's nothing special about the unit vector $\hat{k}$. For any rotation about the axis with unit vector $\hat{\alpha}$

$$\hat{R}_{\theta,\hat{\alpha}} = e^{-i(\hat{\vec{L}}\cdot\vec{\alpha}/\hbar)\theta}. \tag{13.10}$$

The operators $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ don't commute. reflecting the fact that rotations about the x-, y- and z-axes don't commute (as you can demonstrate to yourself using a book or a tennis racket). But the operator for the square magnitude of angular momentum,

$$\hat{L}^2 \equiv \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \tag{13.11}$$

is a scalar operator that doesn't change upon rotation, so

$$[\hat{L}^2, \hat{L}_i] = 0 \qquad \text{for } i = x, y, z. \tag{13.12}$$

You can work out these three commutators laboriously using more primitive commutators, but it's clear from inspection once you realize that the operators $\hat{L}_i$ generate rotations.

Similarly, for a Hamiltonian with rotational symmetry,

$$[\hat{H}, \hat{L}_i] = 0 \qquad \text{for } i = x, y, z, \tag{13.13}$$

so all three components of the angular momentum vector are conserved.

13.3 Solution of the angular momentum eigenproblem

We solved the simple harmonic oscillator energy eigenproblem twice: once using a straightforward but laborious differential equation technique, and then again using an operator-factorization technique that was much easier to implement, but which involved unmotivated creative leaps. We'll do the same with the angular momentum eigenproblem, but in the opposite sequence.

Here's the problem:

> *Given Hermitian operators* $\hat{J}_x$, $\hat{J}_y$, $\hat{J}_z$ *obeying*
>
> $$[\hat{J}_x, \hat{J}_y] = i\hbar \hat{J}_z, \qquad \textit{and cyclic permutations} \tag{13.14}$$
>
> *find the eigenvalues and eigenvectors for one such operator, say* $\hat{J}_z$.

Any other component of angular momentum, say $\hat{J}_x$ or $\hat{J}_{42^\circ}$, will have exactly the same eigenvalues, and eigenvectors with the same structure.

Note that we are to solve the problem using *only* the commutation relations — we are not to use, say, the expression for the angular momentum operator in the position basis, nor the relationship between angular momentum and rotation.

Strangely, our first step is to slightly expand the problem. (I warned you that the solution would not take a straightforward, "follow your nose" path.)

> *Define*
>
> $$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \tag{13.15}$$
>
> *and note that*
>
> $$[\hat{J}^2, \hat{J}_i] = 0 \qquad \textit{for } i = x, y, z. \tag{13.16}$$
>
> *Because* $\hat{J}^2$ *and* $\hat{J}_z$ *commute, they have a basis of simultaneous eigenvectors. We expand the problem to find these simultaneous eigenvectors* $|\lambda, \mu\rangle$*, which satisfy*
>
> $$\hat{J}^2 |\lambda, \mu\rangle = \hbar^2 \lambda |\lambda, \mu\rangle \tag{13.17}$$
>
> $$\hat{J}_z |\lambda, \mu\rangle = \hbar \mu |\lambda, \mu\rangle \tag{13.18}$$

Exercise 13.C. Show that λ and μ are dimensionless.

Exercise 13.D. Show that the equations (13.16) follow from the equations (13.14). What is the commutator $[\hat{J}^2, \hat{J}_{28^\circ}]$?

Start off by noting that

$$(\hat{J}_x^2 + \hat{J}_y^2)|\lambda, \mu\rangle = (\hat{J}^2 - \hat{J}_z^2)|\lambda, \mu\rangle = \hbar^2(\lambda - \mu^2)|\lambda, \mu\rangle. \tag{13.19}$$

Now the first operator $(\hat{J}_x^2 + \hat{J}_y^2)$ would be $(\hat{J}_x - i\hat{J}_y)(\hat{J}_x + i\hat{J}_y)$ if $\hat{J}_x$ and $\hat{J}_y$ were numbers. The factorization is not in fact quite that clean, because those operators are not in fact numbers. But we use this factorization to inspire the definitions

$$\hat{J}_- = \hat{J}_x - i\hat{J}_y \quad \text{and} \quad \hat{J}_+ = \hat{J}_x + i\hat{J}_y \tag{13.20}$$

so that

$$\hat{J}_-\hat{J}_+ = \hat{J}_x^2 + \hat{J}_y^2 + i(\hat{J}_x\hat{J}_y - \hat{J}_y\hat{J}_x) = \hat{J}_x^2 + \hat{J}_y^2 + i[\hat{J}_x, \hat{J}_y] = \hat{J}_x^2 + \hat{J}_y^2 - \hbar\hat{J}_z. \tag{13.21}$$

This tells us that

$$\hat{J}_-\hat{J}_+|\lambda, \mu\rangle = (\hbar^2\lambda - \hbar^2\mu^2 - \hbar^2\mu)|\lambda, \mu\rangle = \hbar^2(\lambda - \mu(\mu+1))|\lambda, \mu\rangle. \tag{13.22}$$

We have immediately that

$$\langle\lambda, \mu|\hat{J}_-\hat{J}_+|\lambda, \mu\rangle = \hbar^2(\lambda - \mu(\mu+1)). \tag{13.23}$$

But if we define

$$|\phi\rangle = \hat{J}_+|\lambda, \mu\rangle \quad \text{then} \quad \langle\phi| = \langle\lambda, \mu|\hat{J}_-$$

then equation (13.23) is just the expression for $\langle\phi|\phi\rangle$, and we know that for any vector $\langle\phi|\phi\rangle \geq 0$. Thus

$$\lambda \geq \mu(\mu+1). \tag{13.24}$$

With these preliminaries out of the way, we investigate the operator $\hat{J}_+$. First, its commutation relations:

$$[\hat{J}^2, \hat{J}_+] = 0, \tag{13.25}$$

$$[\hat{J}_z, \hat{J}_+] = [\hat{J}_z, \hat{J}_x] + i[\hat{J}_z, \hat{J}_y] = (i\hbar\hat{J}_y) + i(-i\hbar\hat{J}_x) = \hbar\hat{J}_+. \tag{13.26}$$

Then, use the commutation relations to find the effect of $\hat{J}_+$ on $|\lambda, \mu\rangle$. If we again define $|\phi\rangle = \hat{J}_+|\lambda, \mu\rangle$, then

$$\hat{J}^2|\phi\rangle = \hat{J}^2\hat{J}_+|\lambda, \mu\rangle = \hat{J}_+\hat{J}^2|\lambda, \mu\rangle = \hbar^2\lambda\hat{J}_+|\lambda, \mu\rangle = \hbar^2\lambda|\phi\rangle, \tag{13.27}$$

$$\begin{aligned} \hat{J}_z|\phi\rangle &= \hat{J}_z\hat{J}_+|\lambda, \mu\rangle = (\hat{J}_+\hat{J}_z + \hbar\hat{J}_+)|\lambda, \mu\rangle \\ &= \hbar\mu\hat{J}_+|\lambda, \mu\rangle + \hbar\hat{J}_+|\lambda, \mu\rangle = \hbar(\mu+1)|\phi\rangle. \end{aligned} \tag{13.28}$$

That is, the vector $|\phi\rangle$ is an eigenvector of $\hat{J}^2$ with eigenvalue λ and an eigenvector of $\hat{J}_z$ with eigenvalue $\mu + 1$. In other words,

$$\hat{J}_+|\lambda, \mu\rangle = A|\lambda, \mu + 1\rangle \tag{13.29}$$

where A is a normalization factor to be determined.

To find A, we contrast

$$\langle\phi|\phi\rangle = |A|^2\langle\lambda, \mu|\lambda, \mu\rangle = |A|^2 \tag{13.30}$$

with the result of equation (13.23), namely

$$\langle\phi|\phi\rangle = \langle\lambda, \mu|\hat{J}_-\hat{J}_+|\lambda, \mu\rangle = \hbar^2(\lambda - \mu(\mu + 1)). \tag{13.31}$$

From this we may select $A = \hbar\sqrt{\lambda - \mu(\mu + 1)}$ so that

$$\hat{J}_+|\lambda, \mu\rangle = \hbar\sqrt{\lambda - \mu(\mu + 1)}\,|\lambda, \mu + 1\rangle. \tag{13.32}$$

In short, the operator $\hat{J}_+$ applied to $|\lambda, \mu\rangle$ acts as a *raising operator*: it doesn't change the value of λ, but it increases the value of μ by 1.

Parallel reasoning applied to $\hat{J}_-$ shows that

$$\hat{J}_-|\lambda, \mu\rangle = \hbar\sqrt{\lambda - \mu(\mu - 1)}\,|\lambda, \mu - 1\rangle. \tag{13.33}$$

In short, the operator $\hat{J}_-$ applied to $|\lambda, \mu\rangle$ acts as a *lowering operator*: it doesn't change the value of λ, but it decreases the value of μ by 1.

Exercise 13.E. Execute the "parallel reasoning" that results in equation (13.33).

At first it might appear that we could use these raising or lowering operators to ascend to infinitely high heavens or to dive to infinitely low depths, but that appearance is incorrect. Equation (13.24),

$$\lambda \geq \mu(\mu + 1), \tag{13.34}$$

will necessarily be violated for sufficiently high or sufficiently low values of μ. Instead, there must be some maximum value of μ — call it μ_{max} — such that an attempt to raise $|\lambda, \mu_{\text{max}}\rangle$ results not in a vector proportional to $|\lambda, \mu_{\text{max}} + 1\rangle$, but results instead in 0. It is clear from equation (13.32) that this value of μ satisifies

$$\lambda - \mu_{\text{max}}(\mu_{\text{max}} + 1) = 0. \tag{13.35}$$

And it's equally clear from equation (13.33) that there is a minimum value μ_{min} satisifying

$$\lambda - \mu_{\text{min}}(\mu_{\text{min}} - 1) = 0. \tag{13.36}$$

Solving these two equations simultaneously, we find that

$$\mu_{\max} = -\mu_{\min} \qquad \text{with} \qquad \mu_{\max} \geq 0 \tag{13.37}$$

and that

$$\lambda = \mu_{\max}(\mu_{\max} + 1). \tag{13.38}$$

Exercise 13.F. The simultaneous solution of equations (13.35) and (13.36) results in two possible solutions, namely (13.37) and $\mu_{\min} = \mu_{\max} + 1$. Why do we reject this second solution? Why do we, in equation (13.37), insert the proviso $\mu_{\max} \geq 0$?

But there's more. Because we raise or lower μ by 1 with each application of $\hat{J}_+$ or $\hat{J}_-$, the value of $\mu_{\max}$ must be an integer above $\mu_{\min}$:

$$\begin{aligned} \mu_{\max} &= \mu_{\min} + (\text{an integer}) \\ 2\mu_{\max} &= (\text{an integer}) \\ \mu_{\max} &= \frac{\text{an integer}}{2} \geq 0 \end{aligned} \tag{13.39}$$

Common practice is to call the half-integer $\mu_{\max}$ by the name j, and the half-integer μ by the name m. And common practice is to label the angular momentum state not as $|\lambda, \mu\rangle$ but as $|j, m\rangle$, which contains equivalent information. Using these conventions, the solution to the angular momentum eigenvalue problem is:

The eigenvalues of $\hat{J}^2$ are

$$\hbar^2 j(j+1) \qquad j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \ldots. \tag{13.40}$$

For a given j, the eigenvalues of $\hat{J}_z$ are

$$\hbar m \qquad m = -j, -j+1, \ldots, j-1, j. \tag{13.41}$$

The eigenstates $|j, m\rangle$ are related through the operators

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \qquad \hat{J}_- = \hat{J}_x - i\hat{J}_y \tag{13.42}$$

by

$$\hat{J}_+|j,m\rangle = \hbar\sqrt{j(j+1) - m(m+1)}\,|j, m+1\rangle \tag{13.43}$$

$$\hat{J}_-|j,m\rangle = \hbar\sqrt{j(j+1) - m(m-1)}\,|j, m-1\rangle. \tag{13.44}$$

Exercise 13.G. For a classical rigid body rotating about a fixed axis, the kinetic energy of rotation is $L^2/2I$, where I is the moment of inertia and L is the (magnitude of the) angular momentum. What are the quantal energy eigenvalues of this system?

13.4 Summary of the angular momentum eigenproblem

Given $[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$, and cyclic permutations, the eigenvalues of $\hat{J}^2$ are

$$\hbar^2 j(j+1) \qquad j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \ldots.$$

For a given j, the eigenvalues of $\hat{J}_z$ are

$$\hbar m \qquad m = -j, -j+1, \ldots, j-1, j.$$

The eigenstates $|j, m\rangle$ are related through the operators

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \qquad \hat{J}_- = \hat{J}_x - i\hat{J}_y$$

by

$$\begin{aligned} \hat{J}_+|j,m\rangle &= \hbar\sqrt{j(j+1) - m(m+1)}\,|j, m+1\rangle \\ \hat{J}_-|j,m\rangle &= \hbar\sqrt{j(j+1) - m(m-1)}\,|j, m-1\rangle. \end{aligned}$$

13.5 Angular momentum eigenproblem in the position representation

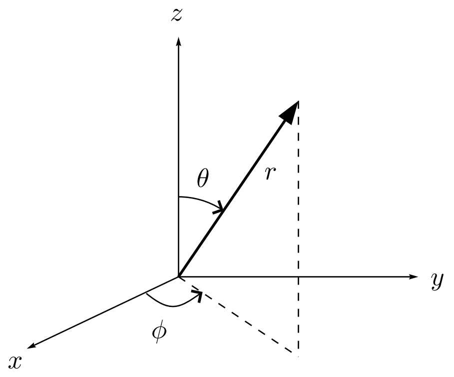
This material is useful in a variety of situations: electromagnetism, gravity, geodesy.

And this material is mathematically intricate: questions like these were first raised by Adrian-Marie Legendre[1] in 1782, and the last elaboration of the structure that I know of was published by Gustav Herglotz in 1962. If we are to cover 180 years of mathematical development in six pages, you can be sure that (1) it's going to be fast-paced and (2) we're going to leave out some of the details.

Setup. Because angular momentum is intimately associated with rotations, you might expect that this problem is most readily solved using not Cartesian coordinates x, y, and z, but spherical coordinates r, θ, and ϕ. And you'd be right.

---

[1] Adrien-Marie Legendre (1752–1833) made contributions throughout mathematics. He originated the "least squares" method of curve fitting. One notable episode from his life is that the French government denied him the pension he had earned when he refused to endorse a government-supported candidate for an honor.



In drawing this diagram, we use the arrow to represent a point, not the position of a particle. A particle generally doesn't have a position, but a geometrical point always does.

You can convert, say, the operator

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \tag{13.45}$$

with its Cartesian position representation

$$\mathcal{L}_z = x\left[-i\hbar\frac{\partial}{\partial y}\right] - y\left[-i\hbar\frac{\partial}{\partial x}\right] \tag{13.46}$$

into spherical coordinates as

$$\mathcal{L}_z = -i\hbar\frac{\partial}{\partial \phi}, \tag{13.47}$$

which makes sense given that $\hat{L}_z$ generates rotations that increase ϕ. It's harder to find and interpret the expressions for $\mathcal{L}_x$ and $\mathcal{L}_y$, but once you do you'll find that the magnitude squared of the angular momentum operator is

$$\mathcal{L}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]. \tag{13.48}$$

Notice that this expression is independent of r, as you might expect for a quantity like angular momentum so intimately associated with rotations. Also makes sense that this is independent of the magnitude r, because if you double x you also double y, so this cancels out.

Exercise 13.H. What is the representation of $\hat{L}_z$ in the momentum basis?

We seek eigenfunctions $y(\theta, \phi)$ and eigenvalues λ such that

$$\mathcal{L}^2 y(\theta, \phi) = \hbar^2 \lambda\, y(\theta, \phi). \tag{13.49}$$

To do this, we need to solve the partial differential equation

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial y(\theta,\phi)}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 y(\theta,\phi)}{\partial\phi^2} = -\lambda y(\theta,\phi). \tag{13.50}$$

Separation of variables. This looks hairy because it is hairy. I'll approach it with a tried-and-true technique called "separation of variables". I will look for solutions that take on the product form

$$y(\theta,\phi) = f(\theta)g(\phi). \tag{13.51}$$

On one hand, there's not yet a guarantee that all or even any of the solutions take this product form. On the other hand, it allows progress to be made. It will turn out (although at this stage it's far from obvious), that *all* of the eigenfunctions do in fact take this form, so at the end we will have found all the eigenfunctions. This seems like extraordinary good luck, and I'd like to solve my problems through skill and intelligence rather than through luck, but better to solve them through luck than not at all.

Now, applying this product form to the eigenequation, we get

$$\frac{g(\phi)}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{df(\theta)}{d\theta}\right) + \frac{f(\theta)}{\sin^2\theta}\frac{d^2g(\phi)}{d\phi^2} = -\lambda f(\theta)g(\phi). \tag{13.52}$$

Multiply both sides by $\sin^2\theta/f(\theta)g(\phi)$ to obtain

$$\frac{\sin\theta}{f(\theta)}\frac{d}{d\theta}\left(\sin\theta\frac{df(\theta)}{d\theta}\right) + \frac{1}{g(\phi)}\frac{d^2g(\phi)}{d\phi^2} = -\lambda\sin^2\theta$$

and write as

$$\frac{\sin\theta}{f(\theta)}\frac{d}{d\theta}\left(\sin\theta\frac{df(\theta)}{d\theta}\right) + \lambda\sin^2\theta = -\frac{1}{g(\phi)}\frac{d^2g(\phi)}{d\phi^2}. \tag{13.53}$$

This is in "separated" form. On the left is a function of θ alone, on the right is a function of ϕ alone. These are independent variables, and the only way the two sides can both be equal for *all* values of θ and ϕ is for both sides to equal the same constant, call it κ. Our one partial differential equation has split into two ordinary differential equations, namely

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{df(\theta)}{d\theta}\right) + \lambda\sin^2\theta f(\theta) = \kappa f(\theta) \tag{13.54}$$

$$\frac{d^2g(\phi)}{d\phi^2} = -\kappa g(\phi) \tag{13.55}$$

The equation in variable ϕ. The second of these equations looks easier, so I'll work on it first. Recall Picard's theorem: there are two linearly independent solutions of that equation for any value of κ. We're looking not just for solutions, but for solutions that come back to themselves when you rotate a full circle, that is, solutions obeying the "full circle" condition

$$g(\phi) = g(\phi + 2\pi). \tag{13.56}$$

For a trial solution, I'll look at $g(\phi) = Ae^{i\alpha\phi}$. This clearly solves the differential equation whenever $\alpha = \pm\sqrt{\kappa}$, and it obeys the full circle condition whenever

$$Ae^{i\alpha\phi} = Ae^{i\alpha(\phi+2\pi)} \qquad \text{or} \qquad e^{i2\pi\alpha} = 1, \tag{13.57}$$

that is, whenever

$$\alpha = m \qquad m = 0, \pm 1, \pm 2, \pm 3, \ldots. \tag{13.58}$$

We have solved the ϕ part of the partial differential equation:

$$g(\phi) = Ae^{im\theta} \qquad m = 0, \pm 1, \pm 2, \pm 3, \ldots. \tag{13.59}$$

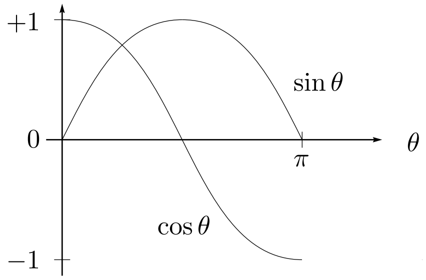
Before going on to the θ part of the problem, we pause and note that $g(\phi)$ is more than just half a solution to the eigenproblem for $\mathcal{L}^2$. It is also a solution to the eigenproblem for $\mathcal{L}_z$, because (see equation 13.47)

$$\mathcal{L}_z e^{im\phi} = -i\hbar\frac{\partial}{\partial\phi}e^{im\phi} = \hbar m\, e^{im\phi}. \tag{13.60}$$

The equation in variable θ. Now we have to go back to the more formidable θ part of the problem, equation (13.54), which now reads

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{df(\theta)}{d\theta}\right) + (\lambda\sin^2\theta - m^2)f(\theta) = 0. \tag{13.61}$$

Seeing all these $\sin\theta$s, you might be tempted to change variable from θ to $\sin\theta$. Bad move.



Because θ ranges from 0 to π, a given value of $\sin\theta$ corresponds to two different angles. On the other hand a given value of $\cos\theta$ corresponds to a single angle, so we change variable from θ to

$$\zeta = \cos\theta. \tag{13.62}$$

(We use the name ζ because it is the value of z on the unit sphere for this angle θ. As θ ranges from 0 to π, the variable ζ ranges from $+1$ to -1. The situation $\zeta = +1$ corresponds to the "north pole" of the spherical coordinate system, that is touching the positive z axis, while the situation $\zeta = -1$ corresponds to the "south pole", that is touching the negative z axis.) In terms of this new variable ζ, equation (13.61) becomes

$$(1-\zeta^2)\frac{d}{d\zeta}\left((1-\zeta^2)\frac{df(\zeta)}{d\zeta}\right) + (\lambda(1-\zeta^2) - m^2)f(\zeta) = 0, \tag{13.63}$$

or

$$(1-\zeta^2)\frac{d^2f(\zeta)}{d\zeta^2} - 2\zeta\frac{df(\zeta)}{d\zeta} + \left(\lambda - \frac{m^2}{1-\zeta^2}\right)f(\zeta) = 0. \tag{13.64}$$

This is called the general Legendre equation.

Power series solution of the Legendre equation. I start out by finding a solution, not for the general Legendre equation, but for the special case $m = 0$:

$$(1-\zeta^2)f''(\zeta) - 2\zeta f'(\zeta) + \lambda f(\zeta) = 0, \tag{13.65}$$

which is called the Legendre equation.

Look for a power series solution:

$$\begin{aligned}
f(\zeta) &= \sum_{k=0}^{\infty} a_k\zeta^k \\
f'(\zeta) &= \sum_{k=0}^{\infty} (k+1)a_{k+1}\zeta^k \\
f''(\zeta) &= \sum_{k=0}^{\infty} (k+2)(k+1)a_{k+2}\zeta^k \\
\zeta f'(\zeta) &= \sum_{k=1}^{\infty} k a_k\zeta^k \\
\zeta^2 f''(\zeta) &= \sum_{k=2}^{\infty} k(k-1)a_k\zeta^k
\end{aligned}$$

When I plug these forms into the Legendre equation, I find that a_0 and a_1 are undetermined — these are the two "adjustable parameters" that enter into the solution of any second-order linear differential equation. But for $k \geq 2$, the equation demands that

$$(k+2)(k+1)a_{k+2} - k(k-1)a_k - 2ka_k + \lambda a_k = 0,$$

or

$$a_{k+2} = \frac{k^2+k-\lambda}{(k+2)(k+1)} a_k \qquad k = 2, 3, 4, \ldots. \tag{13.66}$$

What is the behavior of these coefficients for large values of k? It is $a_{k+2} = a_k$. Such a power series is clearly divergent unless, at some point in the recursion, $a_k = 0$. And this happens if and only if, for some integer k,

$$\lambda = k^2 + k = k(k+1).$$

We have found the eigenvalue condition.

There remains a lot of clean-up to do that I won't detail here. The upshot is that the Legendre equation has normalizable solutions when and only when

$$\lambda = \ell(\ell+1) \quad \text{for} \quad \ell = 0, 1, 2, \ldots. \tag{13.67}$$

For any given ℓ, the solution is a polynomial of order ℓ called a "Legendre polynomial"

$$P_\ell(\zeta). \tag{13.68}$$

If you search the Internet for information about Legendre polynomials (I recommend the "Digital Library of Mathematical Functions") you will find all manner of information: explicit expressions, graphs, integral representations, and more.

Solution of the general Legendre equation. I will describe the solutions of the general Legendre equation without attempting to derive them. The equation has solutions when $\lambda = \ell(\ell+1)$, $\ell = 0, 1, 2, \ldots$, and when $m = -\ell, -\ell+1, \ldots, 0, \ldots, \ell-1, \ell$. These solutions are called the "associated Legendre functions" (not polynomials, because they sometimes involve $\sqrt{1-\zeta^2}$) and are denoted

$$P_\ell^m(\zeta). \tag{13.69}$$

Pulling everything together. The product $f(\theta)g(\phi)$ is called a "spherical harmonic"

$$Y_\ell^m(\theta, \phi) = A\, P_\ell^m(\cos\theta) e^{im\phi}, \tag{13.70}$$

where the normalization constant A is set so that

$$\int_0^\pi d\theta \int_0^{2\pi} d\phi \, \sin\theta \, |Y_\ell^m(\theta,\phi)|^2 = 1. \tag{13.71}$$

These functions are defined for

$$\ell = 0, 1, 2, \ldots \qquad \text{and} \qquad m = -\ell, -\ell+1, \ldots, 0, \ldots, \ell-1, \ell. \tag{13.72}$$

They satisfy

$$\mathcal{L}^2 Y_\ell^m(\theta,\phi) = \hbar^2 \ell(\ell+1)\, Y_\ell^m(\theta,\phi) \tag{13.73}$$

$$\mathcal{L}_z Y_\ell^m(\theta,\phi) = \hbar m \, Y_\ell^m(\theta,\phi). \tag{13.74}$$

You'll notice that these conclusions correspond to the "Summary of the angular momentum eigenproblem" in section 13.4, except that half-integral values of j are omitted.

The spherical harmonics satisfy

$$\left\{ \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \right\} Y_\ell^m(\theta,\phi) = -\ell(\ell+1) Y_\ell^m(\theta,\phi) \tag{13.75}$$

and are complete in the sense that

> *Theorem:* If $f(\zeta,\phi)$ is a differentiable function on the unit sphere, then
>
> $$f(\zeta,\phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell,m} Y_\ell^m(\zeta,\phi) \tag{13.76}$$
>
> where
>
> $$f_{\ell,m} = \int_0^{2\pi} d\phi \int_{-1}^{1} d\zeta \, (Y_\ell^m(\zeta,\phi))^* f(\zeta,\phi). \tag{13.77}$$

The above paragraph is precisely analogous to the Fourier series result that the "trigonometric" functions $e^{i\ell\theta}$ satisfy

$$\left\{ \frac{\partial^2}{\partial\theta^2} \right\} e^{i\ell\theta} = -\ell^2 e^{i\ell\theta} \tag{13.78}$$

and are complete in the sense that

Theorem: If $f(\theta)$ is a differentiable function on the unit circle (i.e. with periodicity 2π), then

$$f(\theta) = \sum_{\ell=-\infty}^{\infty} f_\ell e^{i\ell\theta} \tag{13.79}$$

where

$$f_\ell = \frac{1}{2\pi}\int_0^{2\pi} d\theta\, (e^{i\ell\theta})^* f(\theta). \tag{13.80}$$

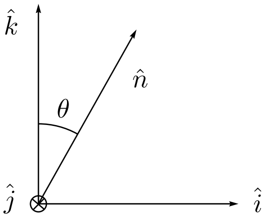
There are a lot of special functions, many of which are used only in very specialized situations. But the spherical harmonics are just as important in three dimensional problems as the trigonometric functions are in two dimensional problems. Spherical harmonics are used in quantum mechanics, in electrostatics, in acoustics, in signal processing, in seismology, and in mapping (to keep track of the deviations of the Earth's shape from spherical). They are as important as sines and cosines. It's worth becoming familiar with them.

Exercise 13.I. Show that the probability density $|Y_\ell^m(\theta,\phi)|^2$ associated with any spherical harmonic is "axially symmetric," that is, independent of rotations about the z axis, that is, independent of ϕ.

13.6 Angular momentum projected onto various axes

Here's a reasonable question: Currently the system is in an angular momentum state with a definite projection of $\hat{\vec{J}}$ on $\hat{k}$, the unit vector in the z direction. What happens when we measure the projection on some other unit vector $\hat{n}$?

In other words: Currently the system is in an angular momentum eigenstate of both $\hat{J}^2$ and $\hat{J}_z$, say $|j, m\,(\hat{k})\rangle$. (Elsewhere in this chapter the projection had been understood to be on the $\hat{k}$ unit vector. In this section our notation makes that understanding explicit.) After measurement of $\hat{J}_\theta \equiv \hat{\vec{J}}\cdot\hat{n}$, the system will be in an angular momentum eigenstate of both $\hat{J}^2$ and $\hat{J}_\theta$, say $|j', m'\,(\hat{n})\rangle$. What is the amplitude $\langle j', m'\,(\hat{n})|j, m\,(\hat{k})\rangle$?



In this figure, the axes are oriented so that $\hat{j}$*, the unit vector in the* y *direction, points into the page.*

The key to solving this problem is to use the angular momentum operator to generate rotations. The unfamiliar state $|j', m'\,(\hat{n})\rangle$ is just the familiar state $|j', m'\,(\hat{k})\rangle$ rotated by an angle θ about the y-axis. In symbols,

$$|j', m'\,(\hat{n})\rangle = e^{-i\hat{J}_y\theta/\hbar}|j', m'\,(\hat{k})\rangle. \tag{13.81}$$

Thus the desired amplitude is just

$$\langle j', m'\,(\hat{n})|j, m\,(\hat{k})\rangle = \langle j', m'\,(\hat{k})|e^{i\hat{J}_y\theta/\hbar}|j, m\,(\hat{k})\rangle. \tag{13.82}$$

It is very clear that the magnitude of $\hat{J}^2$ will not change under rotation, because it is a scalar, so the amplitude above will be zero unless $j' = j$. These amplitudes are conventionally given the symbol

$$d^{(j)}_{m,m'}(\theta) = \langle j, m'\,(\hat{n})|j, m\,(\hat{k})\rangle = \langle j, m'\,(\hat{k})|e^{i\hat{J}_y\theta/\hbar}|j, m\,(\hat{k})\rangle \tag{13.83}$$

and the name "irreducible representations of the rotation group".

The rest of this book considers only states of $\hat{J}_z$, not of $\hat{J}_\theta$, so we drop the explicit axis notation and revert to writing simply

$$d^{(j)}_{m,m'}(\theta) = \langle j, m'|e^{i\hat{J}_y\theta/\hbar}|j, m\rangle. \tag{13.84}$$

How can we evaluate these amplitudes? The obvious way would be to expand $e^{i\hat{J}_y\theta/\hbar}$ in a Taylor series. Then if we knew the values of $\langle j, m'|\hat{J}^n_y|j, m\rangle$ we could evaluate each term of the series. And we could do that by writing $\hat{J}_y$ in terms of raising and lowering operators as $\hat{J}_y = (\hat{J}_+ - \hat{J}_-)/(2i)$. This is a possible scheme but it's difficult. (If you derived equation (11.63) you have an idea of just how difficult it would be.) I'll show you a strategy that is far from obvious but that turns out to be much more straightforward to execute.

The “far from obvious” strategy converts equation (13.84) into a differential equation in θ, then brings our well-developed skills in differential equation solution to bear on the problem. I admit this seems counterintuitive, because you are used to starting with the differential equation and finding the solution, and this strategy seems backwards. But please stick with me.

From (13.84) we see that

$$\begin{aligned}
\frac{d}{d\theta}\left[d^{(j)}_{m,m'}(\theta)\right] &= \langle j,m'|e^{i\hat{J}_y\theta/\hbar}(i\hat{J}_y/\hbar)|j,m\rangle \\
&= \frac{1}{2\hbar}\langle j,m'|e^{i\hat{J}_y\theta/\hbar}(\hat{J}_+ - \hat{J}_-)|j,m\rangle \\
&= +\tfrac{1}{2}\sqrt{j(j+1)-m(m+1)}\,\langle j,m'|e^{i\hat{J}_y\theta/\hbar}|j,m+1\rangle \\
&\quad -\tfrac{1}{2}\sqrt{j(j+1)-m(m-1)}\,\langle j,m'|e^{\hat{J}_y\theta/\hbar}|j,m-1\rangle
\end{aligned}$$

This seems to be, if anything, a step in the wrong direction. But then we recognize the $d(\theta)$ functions on the right-hand side.

$$\begin{aligned}
\frac{d}{d\theta}\left[d^{(j)}_{m,m'}(\theta)\right] &= +\tfrac{1}{2}\sqrt{j(j+1)-m(m+1)}\,d^{(j)}_{m+1,m'}(\theta) \\
&\quad -\tfrac{1}{2}\sqrt{j(j+1)-m(m-1)}\,d^{(j)}_{m-1,m'}(\theta). \qquad (13.85)
\end{aligned}$$

For a given j and m', these are $2j+1$ coupled first-order ODEs, to be solved subject to the initial conditions $d^{(j)}_{m,m'}(0) = \delta_{m,m'}$.

Let’s try this for the simplest case, namely $j = \frac{1}{2}$ and $m' = \frac{1}{2}$. To avoid all those annoying subscripts, I’ll just write $d^{(\frac{1}{2})}_{m,\frac{1}{2}}(\theta)$ as $A_m(\theta)$. Then the equations are: for $m = +\frac{1}{2}$

$$\begin{aligned}
\frac{d}{d\theta}\left[A_{+\frac{1}{2}}(\theta)\right] &= +\tfrac{1}{2}\sqrt{\tfrac{1}{2}(\tfrac{3}{2}) - \tfrac{1}{2}(\tfrac{3}{2})}\,A_{+\frac{3}{2}}(\theta) \\
&\quad -\tfrac{1}{2}\sqrt{\tfrac{1}{2}(\tfrac{3}{2}) - \tfrac{1}{2}(-\tfrac{1}{2})}\,A_{-\frac{1}{2}}(\theta) \\
&= -\tfrac{1}{2}\,A_{-\frac{1}{2}}(\theta) \qquad (13.86)
\end{aligned}$$

while for $m = -\frac{1}{2}$

$$\begin{aligned}
\frac{d}{d\theta}\left[A_{-\frac{1}{2}}(\theta)\right] &= +\tfrac{1}{2}\sqrt{\tfrac{1}{2}(\tfrac{3}{2}) - (-\tfrac{1}{2})(\tfrac{1}{2})}\,A_{+\frac{1}{2}}(\theta) \\
&\quad -\tfrac{1}{2}\sqrt{\tfrac{1}{2}(\tfrac{3}{2}) - (-\tfrac{1}{2})(-\tfrac{3}{2})}\,A_{-\frac{3}{2}}(\theta) \\
&= \tfrac{1}{2}\,A_{+\frac{1}{2}}(\theta) \qquad (13.87)
\end{aligned}$$

Putting these together

$$\frac{d^2}{d\theta^2}\left[A_{+\frac{1}{2}}(\theta)\right] = -\frac{1}{2^2}A_{+\frac{1}{2}}(\theta). \tag{13.88}$$

The ODE for simple harmonic motion! Put together with the initial condition $A_{+\frac{1}{2}}(0) = 1$, this has the immediate solution

$$A_{+\frac{1}{2}}(\theta) = \cos(\theta/2). \tag{13.89}$$

Which, using the definition (13.83), we write out as

$$\langle \tfrac{1}{2}, \tfrac{1}{2}\ (\hat{n})|\tfrac{1}{2}, \tfrac{1}{2}\ (\hat{k})\rangle = \cos(\theta/2) \tag{13.90}$$

a result that we saw many pages ago as equation (2.17):

$$\langle \theta + |z+\rangle = \cos(\theta/2). \tag{13.91}$$

Exercise 13.J. Find the other three equations of (2.17) using the $d^{(j)}_{m,m'}(\theta)$ method.

Problems

13.1 Trivial pursuit

a. Show that if an operator commutes with two components of an angular momentum vector, it commutes with the third as well.

b. If $\hat{J}_x$ and $\hat{J}_z$ are represented by matrices with pure real entries (as is conventionally the case, see problem 13.2), show that $\hat{J}_y$ is represented by a matrix with pure imaginary entries.

13.2 Matrix representations for spin-$\frac{1}{2}$

If we are interested only in a particle's angular momentum, and not in its position, momentum, etc., then for a spin-$\frac{1}{2}$ particle the basis $\{|\frac{1}{2},\frac{1}{2}\rangle, |\frac{1}{2},-\frac{1}{2}\rangle\}$ spans the relevant states. These states are usually denoted simply $\{|\uparrow\rangle, |\downarrow\rangle\}$. Recall that the matrix representation of operator $\hat{A}$ in this basis is

$$\begin{pmatrix} \langle\uparrow|\hat{A}|\uparrow\rangle & \langle\uparrow|\hat{A}|\downarrow\rangle \\ \langle\downarrow|\hat{A}|\uparrow\rangle & \langle\downarrow|\hat{A}|\downarrow\rangle \end{pmatrix}, \tag{13.92}$$

and recall also that this isn't always the easiest way to find a matrix representation.

a. Find matrix representations in the $\{|\uparrow\rangle, |\downarrow\rangle\}$ basis of $\hat{S}_z$, $\hat{S}_+$, $\hat{S}_-$, $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}^2$. Note the reappearance of the Pauli matrices!

b. Find normalized column matrix representations for the eigenstates of $\hat{S}_x$:

$$\hat{S}_x|\rightarrow\rangle = +\frac{\hbar}{2}|\rightarrow\rangle \tag{13.93}$$

$$\hat{S}_x|\leftarrow\rangle = -\frac{\hbar}{2}|\leftarrow\rangle. \tag{13.94}$$

13.3 Rotations and spin-$\frac{1}{2}$

Verify explicitly that

$$|\rightarrow\rangle = e^{-i(\hat{S}_y/\hbar)(+\pi/2)}|\uparrow\rangle, \tag{13.95}$$

$$|\leftarrow\rangle = e^{-i(\hat{S}_y/\hbar)(-\pi/2)}|\uparrow\rangle. \tag{13.96}$$

(Problems 2.9 through 2.11 are relevant here.)

13.4 Spin-1 projection amplitudes

a. (Easy.) Prove that

$$d^{(j)}_{m,m'}(\theta) = [d^{(j)}_{m',m}(-\theta)]^*. \tag{13.97}$$

b. Show that the $d^{(j)}_{m,m'}(\theta)$ with $j = 1$ are

$$\begin{array}{lll}
d^{(1)}_{1,1}(\theta) = +\frac{1}{2}(\cos\theta + 1) & d^{(1)}_{1,0}(\theta) = -\frac{1}{\sqrt{2}}\sin\theta & d^{(1)}_{1,-1}(\theta) = -\frac{1}{2}(\cos\theta - 1) \\
d^{(1)}_{0,1}(\theta) = +\frac{1}{\sqrt{2}}\sin\theta & d^{(1)}_{0,0}(\theta) = \cos\theta & d^{(1)}_{0,-1}(\theta) = -\frac{1}{\sqrt{2}}\sin\theta \\
d^{(1)}_{-1,1}(\theta) = -\frac{1}{2}(\cos\theta - 1) & d^{(1)}_{-1,0}(\theta) = +\frac{1}{\sqrt{2}}\sin\theta & d^{(1)}_{-1,-1}(\theta) = +\frac{1}{2}(\cos\theta + 1)
\end{array}$$

Chapter 14

Central Force Problem and a First Look at Hydrogen

14.1 Examples in nature

One situation we're talking about in this chapter is a point electron moving (or should I say ambivating?) in the vicinity of a point proton, with their interaction described through the potential energy function

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

where r is the magnitude of the separation between the electron and the proton. This situation is a model, called "the Coulomb model", for a hydrogen atom. A real hydrogen atom has a proton of finite size, spin for both the proton and the electron, and relativistic effects for both kinetic energy and the electrodynamic interaction.

But this model is not the *only* situation we're treating in this chapter. A hydrogen atom and a chlorine atom near each other form a molecule where the interaction is not Coulombic but rather more like a Lennard-Jones potential, which again depends only upon the magnitude of the separation. This is also a central force problem.

A proton and a neutron near each other form a nucleus called "the deuteron". They interact via the strong nuclear force, often approximated through the so-called Reid potential energy function, which yet again depends only upon the magnitude of the separation.

A quark and an antiquark near each other form a particle called a meson. This again approximates a central force problem, although in this case relativistic effects dominate and the very idea of "potential energy function" (which implies action-at-a-distance) becomes suspect.

14.2 The classical problem

In all of these cases we can think classically of a six-variable problem: three coordinates $\vec{r}_A$ for the position of particle A, and three coordinates $\vec{r}_B$ for the position of particle B. The energy is

$$\tfrac{1}{2}m_A\dot{\vec{r}}_A^{\,2} + \tfrac{1}{2}m_B\dot{\vec{r}}_B^{\,2} + V(|\vec{r}_B - \vec{r}_A|). \tag{14.1}$$

While the coordinates $\vec{r}_A$ and $\vec{r}_B$ are very natural, they are not the only coordinates possible. Another natural set of six coordinates, just as good as the first set, are the center of mass

$$\vec{R}_{cm} = \frac{m_A\vec{r}_A + m_B\vec{r}_B}{m_A + m_B} \tag{14.2}$$

and the separation between particles

$$\vec{r} = \vec{r}_B - \vec{r}_A. \tag{14.3}$$

In terms of these new coordinates, the energy is

$$\tfrac{1}{2}(m_A + m_B)\dot{\vec{R}}_{cm}^{\,2} + \tfrac{1}{2}\frac{m_A m_B}{m_A + m_B}\dot{\vec{r}}^{\,2} + V(|\vec{r}|). \tag{14.4}$$

Exercise 14.A. Find expressions for $\vec{r}_A$ and $\vec{r}_B$ in terms of $\vec{R}_{cm}$ and $\vec{r}$, then verify the energy expression (14.4).

This new energy expression breaks into two parts: First, a center of mass that moves with constant velocity. We may change reference frame so that our origin is at this center of mass, and in this reference frame $\vec{R}_{cm} = 0$ always, and we needn't ever again consider the motion of the center of mass.

Second, a separation that moves like a particle of mass

$$M = \frac{m_A m_B}{m_A + m_B} \tag{14.5}$$

about a force center at the origin that itself doesn't move. This is called the "reduced mass". For the case of an electron and a proton, where $m_e \ll m_p$,

$$M = \frac{m_e m_p}{m_e + m_p} \approx \frac{m_e m_p}{m_p} = m_e. \tag{14.6}$$

For the case of a quark and an antiquark, each of mass m_q,

$$M = \frac{m_q m_q}{m_q + m_q} = \tfrac{1}{2}m_q. \tag{14.7}$$

Now express classical problem in terms of angular momentum.

14.3 Energy eigenproblem in two dimensions

In one dimension, the energy eigenproblem is

$$-\frac{\hbar^2}{2M}\frac{d^2\eta_n(x)}{dx^2}+V(x)\eta_n(x)=E_n\eta_n(x). \tag{14.8}$$

The generalization to two dimensions is straightforward:

$$-\frac{\hbar^2}{2M}\left[\frac{\partial^2\eta_{\underset{\sim}{n}}(x,y)}{\partial x^2}+\frac{\partial^2\eta_{\underset{\sim}{n}}(x,y)}{\partial y^2}\right]+V(x,y)\eta_{\underset{\sim}{n}}(x,y)=E_{\underset{\sim}{n}}\eta_{\underset{\sim}{n}}(x,y). \tag{14.9}$$

(In one dimension the index n stands for a single integer. In equation (14.36) we will see that in two dimensions the index $\underset{\sim}{n}$ stands for two integers. This is why we use the label $\underset{\sim}{n}$ rather than n.) The part in square brackets is called "the Laplacian of $\eta_{\underset{\sim}{n}}(x,y)$" and represented by the symbol "∇^2" as follows

$$\left[\frac{\partial^2 f(x,y)}{\partial x^2}+\frac{\partial^2 f(x,y)}{\partial y^2}\right]\equiv\nabla^2 f(x,y). \tag{14.10}$$

Thus the "mathematical form" of the energy eigenproblem is

$$\nabla^2\eta_{\underset{\sim}{n}}(\vec{r})+\frac{2M}{\hbar^2}[E_{\underset{\sim}{n}}-V(\vec{r})]\eta_{\underset{\sim}{n}}(\vec{r})=0. \tag{14.11}$$

Suppose $V(x,y)$ is a "central potential" — that is, a function of distance from the origin r only. Then it makes sense to use polar coordinates r and θ rather than Cartesian coordinates x and y. What is the expression for the Laplacian in polar coordinates? This can be uncovered through the chain rule, and it's pretty hard to do. Fortunately, you can look up the answer:

$$\nabla^2 f(\vec{r})=\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f(r,\theta)}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 f(r,\theta)}{\partial\theta^2}\right]. \tag{14.12}$$

Thus, the partial differential equation to be solved is

$$\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\eta_{\underset{\sim}{n}}(r,\theta)}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2\eta_{\underset{\sim}{n}}(r,\theta)}{\partial\theta^2}\right]+\frac{2M}{\hbar^2}[E_{\underset{\sim}{n}}-V(r)]\eta_{\underset{\sim}{n}}(r,\theta)=0 \tag{14.13}$$

or

$$\left\{\frac{\partial^2}{\partial\theta^2}+r\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right)+\frac{2M}{\hbar^2}r^2[E_{\underset{\sim}{n}}-V(r)]\right\}\eta_{\underset{\sim}{n}}(r,\theta)=0. \tag{14.14}$$

For convenience, we wrap up all the r dependence into one piece by defining the "radial linear operator"

$$\mathcal{Q}_{\underset{\sim}{n}}(r)\equiv r\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right)+\frac{2M}{\hbar^2}r^2[E_{\underset{\sim}{n}}-V(r)] \tag{14.15}$$

and write the above as

$$\left\{ \frac{\partial^2}{\partial \theta^2} + \mathcal{Q}_{\underset{\sim}{n}}(r) \right\} \eta_{\underset{\sim}{n}}(r, \theta) = 0. \tag{14.16}$$

This is a linear partial differential equation, so we cast around for solutions knowing that a linear combination of solutions will also be a solution, and hoping that we will cast our net wide enough to catch all the members of a basis. We cast around using the technique of "separation of variables", namely by looking for solutions of the form

$$\eta_{\underset{\sim}{n}}(r, \theta) = R(r)\Theta(\theta). \tag{14.17}$$

Plugging this form into the PDE gives

$$\begin{aligned} R(r)\Theta''(\theta) + \Theta(\theta)\mathcal{Q}_{\underset{\sim}{n}}(r)R(r) &= 0 \\ \frac{\Theta''(\theta)}{\Theta(\theta)} + \frac{\mathcal{Q}_{\underset{\sim}{n}}(r)R(r)}{R(r)} &= 0 \end{aligned} \tag{14.18}$$

Through the usual separation-of-variables argument, we recognize that if a function of θ alone plus a function of r alone sum to zero, where θ and r are independent variables, then both functions must be equal to a constant:

$$\frac{\mathcal{Q}_n(r)R(r)}{R(r)} = -\frac{\Theta''(\theta)}{\Theta(\theta)} = \text{const}. \tag{14.19}$$

First, look at the angular part:

$$\Theta''(\theta) = -\text{const}\,\Theta(\theta). \tag{14.20}$$

This is the differential equation for a mass on a spring! The two linearly independent solutions are

$$\Theta(\theta) = e^{+i\sqrt{\text{const}}\,\theta} \quad \text{or} \quad \Theta(\theta) = e^{-i\sqrt{\text{const}}\,\theta}. \tag{14.21}$$

Now, the boundary condition for this ODE is just that the function must come back to itself if θ increases by 2π:

$$\Theta(\theta) = \Theta(2\pi + \theta). \tag{14.22}$$

If you think about this for a second, you'll see that this means $\sqrt{\text{const}}$ must be an integer

$$\sqrt{\text{const}} = \ell \quad \text{where} \quad \ell = 0, \pm 1, \pm 2, \ldots. \tag{14.23}$$

In summary, the solution to the angular problem is

$$\Theta(\theta) = e^{i\ell\theta} \quad \text{where} \quad \ell = 0, \pm 1, \pm 2, \ldots. \tag{14.24}$$

Now examine the radial part of the problem:

$$\frac{\mathcal{Q}_{\underset{\sim}{n}}(r)R(r)}{R(r)} = \text{const} = \ell^2. \tag{14.25}$$

Write this as

$$\begin{aligned}
\mathcal{Q}_{\underset{\sim}{n}}(r)R(r) - \ell^2 R(r) &= 0 \\
\left\{ r\frac{d}{dr}\left(r\frac{d}{dr}\right) + \frac{2M}{\hbar^2} r^2 [E_{\underset{\sim}{n}} - V(r)] - \ell^2 \right\} R(r) &= 0 \\
\left\{ \frac{1}{r}\frac{d}{dr}\left(r\frac{d}{dr}\right) + \frac{2M}{\hbar^2} [E_{\underset{\sim}{n}} - V(r)] - \frac{\ell^2}{r^2} \right\} R(r) &= 0 \\
\left\{ \frac{1}{r}\frac{d}{dr}\left(r\frac{d}{dr}\right) + \frac{2M}{\hbar^2} \left[E_{\underset{\sim}{n}} - V(r) - \frac{\hbar^2}{2M}\frac{\ell^2}{r^2}\right] \right\} R(r) &= 0 \qquad (14.26)
\end{aligned}$$

Compare this differential equation with another one-variable differential equation, namely the ODE for the energy eigenvalue problem in one dimension:

$$\left\{ \frac{d^2}{dx^2} + \frac{2M}{\hbar^2}\left[E - V(x)\right] \right\} \eta(x) = 0. \tag{14.27}$$

The parts to the right are rather similar, but the parts to the left — the derivatives — are rather different. In addition, the one-dimensional energy eigenfunction satisfies the normalization

$$\int_{-\infty}^{+\infty} |\eta(x)|^2 \, dx = 1, \tag{14.28}$$

whereas the two-dimensional energy eigenfunction satisfies the normalization

$$\begin{aligned}
\int |\eta(x,y)|^2 \, dx\, dy &= 1 \\
\int_0^\infty dr \int_0^{2\pi} r \, d\theta \, |R(r)e^{i\ell\theta}|^2 &= 1 \\
2\pi \int_0^\infty dr \, r |R(r)|^2 &= 1. \qquad (14.29)
\end{aligned}$$

This suggests that the true analog of the one-dimensional $\eta(x)$ is not $R(r)$, but rather

$$u(r) = \sqrt{r}R(r). \tag{14.30}$$

Furthermore,

$$\text{if} \quad u(r) = \sqrt{r}R(r), \quad \text{then} \quad \frac{1}{r}\frac{d}{dr}\left(rR'(r)\right) = \frac{1}{\sqrt{r}}\left(u''(r) + \frac{1}{4}\frac{u(r)}{r^2}\right). \tag{14.31}$$

Using this change of function, the radial equation (14.26) becomes

$$\left\{\frac{d^2}{dr^2} + \frac{1}{4}\frac{1}{r^2} + \frac{2M}{\hbar^2}\left[E - V(r) - \frac{\hbar^2}{2M}\frac{\ell^2}{r^2}\right]\right\}u(r) = 0$$

$$\left\{\frac{d^2}{dr^2} + \frac{2M}{\hbar^2}\left[E - V(r) - \frac{\hbar^2}{2M}\left(\ell^2 - \frac{1}{4}\right)\frac{1}{r^2}\right]\right\}u(r) = 0. \tag{14.32}$$

In this form, the radial equation is exactly like a one-dimensional energy eigenproblem, except that where the one-dimensional problem has the function $V(x)$, the radial problem has the function $V(r) + \hbar^2(\ell^2 - \frac{1}{4})/(2Mr^2)$. These two functions play parallel mathematical roles in the two problems. To emphasize these similar roles, we define an "effective potential energy function" for the radial problem, namely

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2(\ell^2 - \frac{1}{4})}{2M}\frac{1}{r^2}. \tag{14.33}$$

Don't read too much into the term "effective potential energy". No actual potential energy function will depend upon $\hbar$, still less upon the separation constant ℓ! I'm not saying that $V_{\text{eff}}(r)$ *is* a potential energy function, merely that it plays the mathematical role of one in solving this one-dimensional eigenproblem.

Now that the radial equation (14.32) is in exact correspondence with the one-dimensional equation (14.27), we can solve this eigenproblem using any technique that works for the one-dimensional problem. The resulting eigenfunctions and eigenvalues will, of course, depend upon the value of the separation constant ℓ, because the effective potential depends upon the value of ℓ. And as always, for each ℓ there will be many eigenvalues and eigenfunctions, which we will label by index $n = 1, 2, 3, \ldots$ calling them $u_{n,\ell}(r)$ with eigenvalue $E_{n,\ell}$.

Finally, note that the effective potential energy for $\ell = +5$ is the same as the effective potential energy for $\ell = -5$. Thus the eigenfunctions $u_{n,+5}(r)$ and eigenvalues $E_{n,+5}$ will be identical to the eigenfunctions $u_{n,-5}(r)$ and eigenvalues $E_{n,-5}$.

This is a really charming result. We haven't yet specified the potential energy function $V(r)$, so we can't yet determine, say, $E_{7,+5}$ *or* $E_{7,-5}$. Yet we know that these two energy eigenvalues will be equal! Whenever there are two different eigenfunctions, in this case

$$\frac{u_{n,+5}(r)}{\sqrt{r}}e^{+i5\theta} \quad \text{and} \quad \frac{u_{n,+5}(r)}{\sqrt{r}}e^{-i5\theta},$$

attached to the same eigenvalue, the eigenfunctions are said to be *degenerate*. I don't know how such a disparaging term came to be attached to such a charming result, but it has been. ⟦Consider better placement of this remark.⟧

Did we catch all the solutions? It's not obvious, but we did.

Summary:

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

$$-\frac{\hbar^2}{2M}\nabla^2\eta(\vec{r}) + V(r)\eta(\vec{r}) = E\eta(\vec{r}), \tag{14.34}$$

first solve the radial energy eigenproblem

$$-\frac{\hbar^2}{2M}\frac{d^2u(r)}{dr^2} + \left[V(r) + \frac{\hbar^2(\ell^2 - \frac{1}{4})}{2M}\frac{1}{r^2}\right]u(r) = Eu(r) \tag{14.35}$$

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

$$\eta_{n,\ell}(r,\theta) = \frac{u_{n,\ell}(r)}{\sqrt{r}}e^{i\ell\theta} \quad \text{with energy } E_{n,\ell}. \tag{14.36}$$

Notice that the two different solutions with $u_{n,\ell}(r)$ and with $u_{n,-\ell}(r)$ are (except for $\ell = 0$) degenerate.

Exercise 14.B. Show that if you didn't like complex numbers you could select a set of energy eigenfunctions that are pure real.

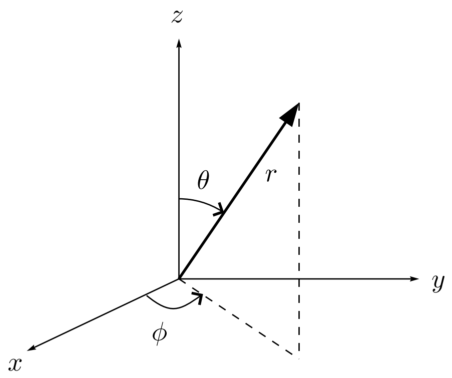
Reflection:

So we've reduced the two-dimensional problem to a one-dimensional problem. How did this miracle occur? Two things happened:

- The original eigenvalue problem was of the form
$$\{\text{angular operator} + \text{radial operator}\}\eta_n(r,\theta) = 0. \tag{14.37}$$
- There was an angular operator eigenbasis $\{\Phi_\ell(\theta)\}$ such that
$$\{\text{angular operator}\}\Phi_\ell(\theta) = \text{number } \Phi_\ell(\theta). \tag{14.38}$$

14.4 Energy eigenproblem in three dimensions

Can we get the same miracle to occur in three dimensions?



In fact, the result is parallel to the two-dimensional result:

Summary:

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

$$-\frac{\hbar^2}{2M}\nabla^2\eta(\vec{r})+V(r)\eta(\vec{r})=E\eta(\vec{r}), \tag{14.39}$$

first solve the radial energy eigenproblem

$$-\frac{\hbar^2}{2M}\frac{d^2u(r)}{dr^2}+\left[V(r)+\frac{\hbar^2\ell(\ell+1)}{2M}\frac{1}{r^2}\right]u(r)=Eu(r) \tag{14.40}$$

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

$$\eta_{n,\ell,m}(r,\theta,\phi)=\frac{u_{n,\ell}(r)}{r}Y_\ell^m(\theta,\phi)\quad \text{with energy } E_{n,\ell}, \tag{14.41}$$

where m takes on the $2\ell+1$ values $-\ell$, $-\ell+1$, …, 0, …, $\ell-1$, ℓ. Notice that the $2\ell+1$ different solutions for a given n and ℓ, but with different m, are degenerate.

Because of spherical symmetry, the operators $\hat{H}$, $\hat{L}^2$, and $\hat{L}_z$ all commute. We seek a simultaneous eigenbasis for *all three* operators.

The energy eigenproblem is

$$-\frac{\hbar^2}{2M}\nabla^2\eta(\vec{r}) + V(r)\eta(\vec{r}) = E\eta(\vec{r}), \tag{14.42}$$

and the Laplacian in spherical coordinates is

$$\begin{aligned}\nabla^2 &= \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\\ &= \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\mathcal{L}^2}{\hbar^2}\right] \qquad (14.43)\end{aligned}$$

where we have recognized the "angular momentum squared" operator defined at equation (13.48).

The energy eigenproblem is then

$$\left\{\nabla^2 + \frac{2M}{\hbar^2}[E - V(r)]\right\}\eta(\vec{r}) = 0, \tag{14.44}$$

or

$$\left\{-\frac{\mathcal{L}^2}{\hbar^2} + \frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{2M}{\hbar^2}r^2[E - V(r)]\right\}\eta(\vec{r}) = 0. \tag{14.45}$$

It will not surprise you that we tackle this equation using separation of variables: we search for solutions of the form $R(r)y(\theta,\phi)$

$$R(r)\left\{-\frac{\mathcal{L}^2}{\hbar^2}\right\}y(\theta,\phi) + y(\theta,\phi)\left\{\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{2M}{\hbar^2}r^2[E - V(r)]\right\}R(r) = 0,$$

and then

$$-\frac{\mathcal{L}^2 y(\theta,\phi)}{\hbar^2\, y(\theta,\phi)} + \frac{1}{R(r)}\left\{\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{2M}{\hbar^2}r^2[E - V(r)]\right\}R(r) = 0.$$

Because this is a function of angle alone plus a function of radius alone that always sums to zero, both functions must be constant: name it $-\lambda$ for the angular part and $+\lambda$ for the radial part.

$$\mathcal{L}^2 y(\theta,\phi) = \lambda y(\theta,\phi) \tag{14.46}$$

$$\left\{\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{2M}{\hbar^2}r^2[E - V(r)]\right\}R(r) = \lambda R(r). \tag{14.47}$$

We have already solved the angular part of the problem, equation (14.46). Back at equation (13.73) we found that the eigenvalues were

$$\lambda = \ell(\ell+1) \qquad \text{for } \ell = 0, 1, 2, 3, \ldots \tag{14.48}$$

and the eigenfunctions were the spherical harmonics

$$Y_\ell^m(\theta,\phi) \text{ with } m=-\ell,-\ell+1,\ldots,0,\ldots,\ell-1,\ell. \tag{14.49}$$

Now return to the radial problem (14.47), which becomes

$$\left\{\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)+\frac{2M}{\hbar^2}\left[E_{\underset{\sim}{n}}-V(r)-\frac{\hbar^2}{2M}\frac{\ell(\ell+1)}{r^2}\right]\right\}R_{\underset{\sim}{n}}(r)=0. \tag{14.50}$$

Note that the differential equation is independent of m, so the solution must also be independent of m.

The energy eigenfunction satisfies the normalization

$$\begin{aligned}\int |\eta(x,y,z)|^2\,dx\,dy\,dz &= 1\\ \int_0^\infty dr\int_0^\pi d\theta\int_0^{2\pi} d\phi\, r^2\sin\theta\,|R_{n,\ell}(r)Y_\ell^m(\theta,\phi)|^2 &= 1\\ \int_0^\infty dr\, r^2|R_{n,\ell}(r)|^2 &= 1.\end{aligned} \tag{14.51}$$

This suggests that the true analog to a one-dimensional wavefunction is $u_{n,\ell}(r)=rR_{n,\ell}(r)$, and sure enough $u_{n,\ell}(r)$ satisfies the equation

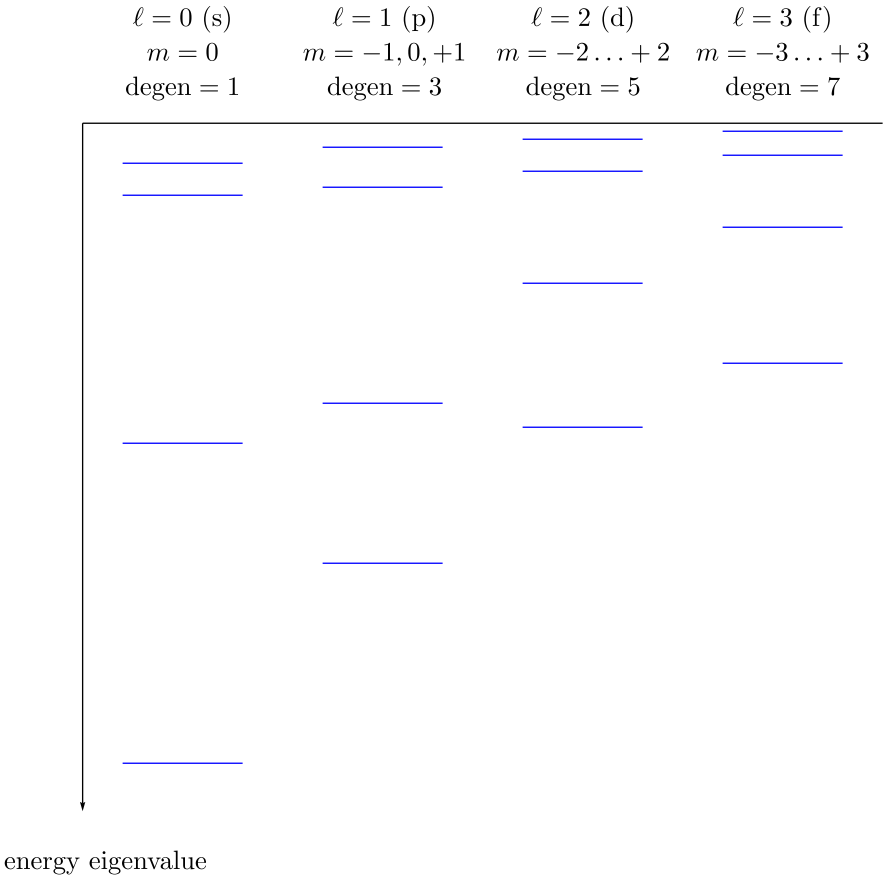
$$\left\{-\frac{\hbar^2}{2M}\frac{d^2}{dr^2}+\left[V(r)+\frac{\hbar^2\ell(\ell+1)}{2M}\frac{1}{r^2}\right]\right\}u_{n,\ell}(r)=E_{n,\ell}u_{n,\ell}(r). \tag{14.52}$$

14.5 Qualitative character of energy solutions

So we need to solve this one-dimensional eigenproblem an infinite number of times: for $\ell=0$, for $\ell=1$, for $\ell=2$, and so on. Each solution will produce an infinite number of eigenvalues: for $\ell=0$, they are $E_{1,0}$, $E_{2,0}$, $E_{3,0}$, …; for $\ell=1$, they are $E_{1,1}$, $E_{2,1}$, $E_{3,1}$, …; for $\ell=2$, they are $E_{1,2}$, $E_{2,2}$, $E_{3,2}$, …; and so forth. Now, look at the effective potential energy function that appears within square brackets in equation (14.52). You can see that, at every point r, it increases with ℓ. It seems reasonable, then, that for any value of n, $E_{n,\ell}$ increases with ℓ. (This is actually a theorem.)

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

In summary, the energy eigenvalues for some generic three-dimensional radially symmetric potential will look sort of like this:



---

[1] William B. Jensen, “The origin of the s, p, d, f orbital labels” *Journal of Chemical Education* **84** (5) (May 2007) 757–758.

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

Exercise 14.C. Carry out a parallel qualitative discussion for the energy eigenproblem if the potential energy function is the "Lennard-Jones" or "6-12" potential

$$V(r) = \frac{A}{r^{12}} - \frac{B}{r^6}. \tag{14.53}$$

14.6 Bound state energy eigenproblem for Coulombic potentials

Problem: Given a (reduced) mass M and a Coulombic potential energy $V(r) = -k/r$, find the negative values $E_{n,\ell}$ such that the corresponding solutions $U_{n,\ell}(r)$ of

$$\left\{-\frac{\hbar^2}{2M}\frac{d^2}{dr^2} + \left[-\frac{k}{r} + \frac{\hbar^2\ell(\ell+1)}{2Mr^2}\right]\right\} U_{n,\ell}(r) = E_{n,\ell}U_{n,\ell}(r) \tag{14.54}$$

are normalizable wavefunctions

$$\int_0^\infty |U_{n,\ell}(r)|^2\, dr = 1. \tag{14.55}$$

Strategy: Same as for the simple harmonic oscillator eigenproblem:

(1) Convert to dimensionless variable.
(2) Remove asymptotic behavior of solutions.
(3) Find non-asymptotic behavior using the series method.
(4) Invoke normalization to terminate the series as a polynomial.

1. Convert to dimensionless variable: Only one length can be constructed from M, k, and $\hbar$. It is

$$a = \frac{\hbar^2}{kM}. \tag{14.56}$$

For the hydrogen problem

$$M = \frac{m_p m_e}{m_p + m_e} \approx m_e \quad \text{and} \quad k = \frac{e^2}{4\pi\epsilon_0},$$

so this length is approximately

$$\frac{\hbar^2}{m_e}\frac{4\pi\epsilon_0}{e^2} \equiv a_0 \equiv \text{“the Bohr radius”} = 0.0529 \text{ nm}. \tag{14.57}$$

Convert to the dimensionless variable

$$\tilde{r} = \frac{r}{a} \tag{14.58}$$

and the dimensionless wavefunction

$$u_{n,\ell}(\tilde{r}) = \sqrt{a}\, U_{n,\ell}(a\tilde{r}). \tag{14.59}$$

The resulting eigenproblem is

$$\left\{-\frac{d^2}{d\tilde{r}^2} - \frac{2}{\tilde{r}} + \frac{\ell(\ell+1)}{\tilde{r}^2}\right\} u_{n,\ell}(\tilde{r}) = \frac{E_{n,\ell}}{k^2M/2\hbar^2} u_{n,\ell}(\tilde{r}) \tag{14.60}$$

with

$$\int_0^\infty |u_{n,\ell}(\tilde{r})|^2\, d\tilde{r} = 1. \tag{14.61}$$

It’s clear that the energy

$$\frac{k^2M}{2\hbar^2} \tag{14.62}$$

is the characteristic energy for this problem. For hydrogen, its value is approximately

$$\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \frac{m_e}{2\hbar^2} \equiv \text{Ry} \equiv \text{“the Rydberg energy”} = 13.6 \text{ eV}.$$

Thus it is reasonable, for brevity, to define the dimensionless energy parameter

$$\mathcal{E}_{n,\ell} = \frac{E_{n,\ell}}{k^2M/2\hbar^2}. \tag{14.63}$$

Furthermore, for the bound state problem $\mathcal{E}_{n,\ell}$ is negative so we define

$$b_{n,\ell}^2 = -\mathcal{E}_{n,\ell} \tag{14.64}$$

and the eigenproblem becomes

$$\left\{\frac{d^2}{d\tilde{r}^2} + \frac{2}{\tilde{r}} - \frac{\ell(\ell+1)}{\tilde{r}^2} - b_{n,\ell}^2\right\} u_{n,\ell}(\tilde{r}) = 0 \tag{14.65}$$

with

$$\int_0^\infty |u_{n,\ell}(\tilde{r})|^2\, d\tilde{r} = 1. \tag{14.66}$$

2. Remove asymptotic behavior of solutions:

> *Note:* In this section we will show that as $\tilde{r} \to 0$,
>
> $$u_{n,\ell}(\tilde{r}) \approx \tilde{r}^{\ell+1}, \tag{14.67}$$
>
> and that as $\tilde{r} \to \infty$,
>
> $$u_{n,\ell}(\tilde{r}) \approx e^{-b_{n,\ell}\tilde{r}}, \tag{14.68}$$
>
> so we will set
>
> $$u_{n,\ell}(\tilde{r}) = \tilde{r}^{\ell+1} e^{-b_{n,\ell}\tilde{r}} v_{n,\ell}(\tilde{r}) \tag{14.69}$$
>
> and then solve an ODE for $v_{n,\ell}(\tilde{r})$. As far as rigor is concerned we could have just pulled the change-of-function (14.69) out of a hat. Thus this section is motivational and doesn't need to be rigorous.

Because equation (14.65) has problems (or, formally, a "regular singular point") at $\tilde{r} = 0$, it pays to find the asymptotic behavior when $\tilde{r} \to 0$ as well as when $\tilde{r} \to \infty$.

2A. Find asymptotic behavior as $\tilde{r} \to 0$: The ODE is

$$\left\{ \frac{d^2}{d\tilde{r}^2} + \left[\frac{2}{\tilde{r}} - \frac{\ell(\ell+1)}{\tilde{r}^2} - b_{n,\ell}^2 \right] \right\} u_{n,\ell}(\tilde{r}) = 0. \tag{14.70}$$

As $\tilde{r} \to 0$ the term in square brackets is dominated (unless $\ell = 0$) by $-\ell(\ell+1)/\tilde{r}^2$. The equation

$$\left\{ \frac{d^2}{d\tilde{r}^2} - \frac{\ell(\ell+1)}{\tilde{r}^2} \right\} u(\tilde{r}) = 0 \tag{14.71}$$

is solved by

$$u(\tilde{r}) = A\tilde{r}^{\ell+1} + B\tilde{r}^{-\ell}. \tag{14.72}$$

However, it's not healthy to keep factors like $\tilde{r}^{-\ell}$ around, because

$$\int_0^{\tilde{r}_0} \tilde{r}^{-2\ell}\, d\tilde{r} = \frac{1}{-2\ell+1} \left[\frac{1}{\tilde{r}^{2\ell-1}} \right]_0^{\tilde{r}_0} = \infty \qquad [\text{for } \ell > \tfrac{1}{2}], \tag{14.73}$$

so wavefunctions with $\tilde{r}^{-\ell}$ prefactors tend to be unnormalizable. (Here $\tilde{r}_0$ is just any positive number.) Thus the wavefunction must behave as

$$u(\tilde{r}) \approx A\tilde{r}^{\ell+1} \tag{14.74}$$

as $\tilde{r} \to 0$.

Our arguments have relied upon $\ell \neq 0$, but it turns out that by stupid good luck the result (14.74) applies when $\ell = 0$ as well. However, it's rather hard to prove this, and since this section is really just motivation anyway, I'll not pursue the matter.

2B. Find asymptotic behavior as $\tilde{r} \to \infty$: In this case, the square bracket term in equation (14.70) is dominated by $-b_{n,\ell}^2$, so the approximate ODE is

$$\left\{ \frac{d^2}{d\tilde{r}^2} - b_{n,\ell}^2 \right\} u_{n,\ell}(\tilde{r}) = 0 \tag{14.75}$$

with solutions

$$u_{n,\ell}(\tilde{r}) = Ae^{-b_{n,\ell}\tilde{r}} + Be^{+b_{n,\ell}\tilde{r}}. \tag{14.76}$$

Clearly, normalization requires that $B = 0$, so the wavefunction has the expected exponential cutoff for large $\tilde{r}$.

In this way, we have justified the definition of $v_{n,\ell}(\tilde{r})$ in equation (14.69). Plugging (14.69) into ODE (14.65), we find that $v_{n,\ell}(\tilde{r})$ satisfies the ODE

$$\left\{ \tilde{r}\frac{d^2}{d\tilde{r}^2} + 2[\ell + 1 - b_{n,\ell}\tilde{r}]\frac{d}{d\tilde{r}} - 2[b_{n,\ell}\ell + b_{n,\ell} - 1] \right\} v_{n,\ell}(\tilde{r}) = 0 \tag{14.77}$$

3. Find non-asymptotic behavior using the series method: We try out the solution

$$v_{n,\ell}(\tilde{r}) = \sum_{k=0}^{\infty} a_k \tilde{r}^k \tag{14.78}$$

and readily find that

$$a_{k+1} = \frac{2b_{n,\ell}(k + \ell + 1) - 2}{(k+1)(k + 2\ell + 2)} a_k \qquad k = 0, 1, 2, \ldots \tag{14.79}$$

(Note that because k and ℓ are both non-negative, the denominator never vanishes.)

4. Invoke normalization to terminate the series as a polynomial: If the a_k coefficient never vanishes, then

$$\frac{a_{k+1}}{a_k} \to \frac{2b_{n,\ell}}{k} \qquad \text{as} \qquad k \to \infty. \tag{14.80}$$

As in the SHO, this leads to $v(\tilde{r}) \approx e^{2b_{n,\ell}\tilde{r}}$ as $\tilde{r} \to \infty$, which is pure disaster. To avoid catastrophe, we must truncate the series as a kth order polynomial by demanding

$$b_{n,\ell} = \frac{1}{k + \ell + 1} \qquad k = 0, 1, 2, \ldots \tag{14.81}$$

Thus $b_{n,\ell}$ is always the reciprocal of the integer

$$n = k + \ell + 1 \tag{14.82}$$

and

$$\mathcal{E}_{n,\ell} = -b_{n,\ell}^2 = -\frac{1}{n^2} \qquad n = 1, 2, 3, \ldots. \tag{14.83}$$

We have found the permissible bound state energies!

What are the eigenfunctions? The solution $v_{n,\ell}(\tilde{r})$ that is a polynomial of order $k = n - \ell - 1$ has a name: it is the Laguerre[2] polynomial

$$L_{n-\ell-1}^{2\ell+1}((2/n)\tilde{r}). \tag{14.84}$$

It would be nicer to have a more direct notation like our own $v_{n,\ell}(\tilde{r})$, but Laguerre died before quantum mechanics was born, so he could not have known how to make his notation convenient for the quantum mechanical Coulomb problem. The Laguerre polynomials are just one more class of special functions not worth knowing much about.

All together, the energy eigenfunctions are

$$\eta_{n,\ell,m}(\tilde{r}, \theta, \phi) = [\text{constant}]\tilde{r}^\ell e^{-\tilde{r}/n} L_{n-\ell-1}^{2\ell+1}((2/n)\tilde{r}) Y_\ell^m(\theta, \phi). \tag{14.85}$$

Degeneracy

Recall that each $v_{n,\ell}(\tilde{r})$ already has an associated $2\ell + 1$-fold degeneracy. In addition, each ℓ gives rise to an infinite number of eigenvalues:

$$\mathcal{E}_{n,\ell} = -\frac{1}{(k + \ell + 1)^2} \qquad k = 0, 1, 2, \ldots. \tag{14.86}$$

In tabular form

$$\begin{array}{llllllll}
\ell = 0 & \text{gives} & n = & 1, & 2, & 3, & 4, & \ldots \\
\ell = 1 & \text{gives} & n = & & 2, & 3, & 4, & \ldots \\
\ell = 2 & \text{gives} & n = & & & 3, & 4, & \ldots \\
\vdots & & & & & & &
\end{array}$$

So...

---

[2] Edmond Laguerre (1834–1886), French artillery officer and mathematician, made contributions to analysis and especially geometry.

$$\begin{array}{llllllll}
\ell = 0 & \text{(degeneracy 1)} & \text{gives} & \mathcal{E}_{n,\ell} = & -1, & -\dfrac{1}{2^2}, & -\dfrac{1}{3^2}, & -\dfrac{1}{4^2}, \quad \ldots \\
\ell = 1 & \text{(degeneracy 3)} & \text{gives} & \mathcal{E}_{n,\ell} = & & -\dfrac{1}{2^2}, & -\dfrac{1}{3^2}, & -\dfrac{1}{4^2}, \quad \ldots \\
\ell = 2 & \text{(degeneracy 5)} & \text{gives} & \mathcal{E}_{n,\ell} = & & & -\dfrac{1}{3^2}, & -\dfrac{1}{4^2}, \quad \ldots \\
\vdots & & & & & & &
\end{array}$$

Eigenenergies of $-1/n^2$ are associated with n different values of ℓ, namely $\ell = 0, 1, \ldots, n-1$. The total degeneracy is thus

$$\sum_{\ell=0}^{n-1} (2\ell + 1) = n^2. \qquad (14.87)$$

14.7 Summary of the bound state energy eigenproblem for a Coulombic potential

A complete set of energy eigenfunctions is $\eta_{n,\ell,m}(r, \theta, \phi)$

where $n = 1, 2, 3, \ldots$
and for each n $\ell = 0, 1, 2, \ldots, n-1$
and for each n and ℓ $m = -\ell, -\ell+1, \ldots, \ell-1, \ell.$

This wavefunction represents a state of energy

$$E_n = -\frac{k^2 M / 2\hbar^2}{n^2},$$

independent of ℓ and m. Thus energy E_n has an n^2-fold degeneracy. In particular, for hydrogen this eigenenergy is nearly

$$E_n = -\frac{\text{Ry}}{n^2}, \qquad \text{Ry} = 13.6 \text{ eV}.$$

In addition, the wavefunction $\eta_{n,\ell,m}(r, \theta, \phi)$ represents a state with an angular momentum squared of $\hbar^2 \ell(\ell + 1)$ and an angular momentum z component of $\hbar m$.

[I recommend that you memorize this summary…it's the sort of thing that frequently comes up on GREs and physics oral exams.]

14.8 Hydrogen atom fine structure

The solution to the Coulomb problem that we've just produced is a magnificent achievement, but it is not a solution to the hydrogen atom problem. The Coulomb problem is a model for the hydrogen atom: highly accurate but not perfect. It ignores collisions, electronic and nuclear spin, the finite size of the proton, relativity, and other factors. These factors account for the "fine structure" of the hydrogen atom.

One element of the fine structure, the only element we'll discuss here, is the relativistic correction to the kinetic energy.

Recall that a classical free relativistic particle of mass m has

$$E^2 - (pc)^2 = (mc^2)^2. \tag{14.88}$$

Thus the classical kinetic energy is

$$\text{KE} = E - mc^2 = \sqrt{(mc^2)^2 + (pc)^2} - mc^2 \tag{14.89}$$

It's hard to see how to convert this into a quantal operator, because in quantum mechanics we treat momentum as an operator $\hat{p}$, and it's hard to know how to deal with the square root of an operator. Instead, for an approximate treatment, we expand the square root in a power series expansion. Recall that

$$(1+\epsilon)^n = 1 + n\epsilon + \tfrac{1}{2}n(n-1)\epsilon^2 + \cdots\,, \tag{14.90}$$

so

$$\begin{aligned}
\text{KE} &= \sqrt{(mc^2)^2 + (pc)^2} - mc^2 \\
&= mc^2\left[\sqrt{1 + \left(\frac{pc}{mc^2}\right)^2} - 1\right] \\
&= mc^2\left[1 + \frac{1}{2}\left(\frac{p}{mc}\right)^2 + \frac{1}{2}\frac{1}{2}\left(-\frac{1}{2}\right)\left(\frac{p}{mc}\right)^4 + \cdots - 1\right] \\
&= \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \cdots
\end{aligned} \tag{14.91}$$

This is *not* a fully relativistic treatment of hydrogen, because it treats relativistic effects on the kinetic energy only approximately, and treats relativistic effects on the potential energy not at all. But it's a start.

First *estimate* the size of this relativistic effect in hydrogen.

$$\frac{p^4}{8m^3c^2} = \frac{1}{2}\left(\frac{p^2}{2m_e}\right)^2 \frac{1}{m_ec^2}. \tag{14.92}$$

The term in parentheses is about the non-relativistic kinetic energy of hydrogen, and we've seen that this is about a Ry. Meanwhile $m_e c^2$ is the energy equivalent of the mass of an electron. At this gross level of approximation, the $\frac{1}{2}$ is a fine point so we just write

$$\frac{p^4}{8m^3c^2} \sim \frac{\text{Ry}^2}{m_e c^2}. \tag{14.93}$$

The size of this relativistic effect, relative to the size of course structure, is about

$$\frac{\text{Ry}^2/m_e c^2}{\text{Ry}} = \frac{\text{Ry}}{m_e c^2} = \frac{13.6 \text{ eV}}{511\,000 \text{ eV}} \approx \frac{25}{1\,000\,000}. \tag{14.94}$$

This correction of about 25 parts per million is small but measurable. The problem calls out for perturbation theory.

Quantal Hamiltonian set up for perturbation theory:

$$\hat{H} = \left[\frac{\hat{p}^2}{2m_e} + V(\hat{\vec{r}})\right] - \frac{\hat{p}^4}{8m^3c^2} = \left[\hat{H}^{(0)}\right] + \hat{H}^{'} \tag{14.95}$$

Do we need to use degenerate perturbation theory? Because $\hat{H}^{'}$ is spherically symmetric, the basis needed as a point of departure for perturbation theory is exactly the one we've been using.

Energy correction:

$$\langle \hat{H}^{'} \rangle_{n\ell m} = -\frac{1}{8m^3c^2} \langle \hat{p}^4 \rangle_{n\ell m} \tag{14.96}$$

Direct approach:

$$\langle \hat{p}^4 \rangle_{n\ell m} = \hbar^4 \int d^3r \, \eta^*_{n\ell m}(\vec{r}) \nabla^2 \left(\nabla^2 \eta_{n\ell m}(\vec{r})\right) \tag{14.97}$$

where the triple integral runs over all space. Remember expression (14.43) for the Laplacian. Do you want to apply this expression not once, but twice, followed by a triple integral? You could do it if you had to, but this direct approach is a lot of work. Isn't there an easier way?

Indirect approach:

$$\begin{aligned}
\hat{H}^{(0)} &= \frac{p^2}{2m_e} + V(\hat{\vec{r}}) \\
\hat{p}^2 &= 2m_e \left(\hat{H}^{(0)} - V(\hat{\vec{r}})\right) \\
\hat{p}^4 &= 4m_e^2 \left((\hat{H}^{(0)})^2 - \hat{H}^{(0)} V(\hat{\vec{r}}) - V(\hat{\vec{r}})\hat{H}^{(0)} + V^2(\hat{\vec{r}})\right)
\end{aligned} \tag{14.98}$$

Sandwich this operator between $|n\ell m\rangle$, keeping in mind that

$$\hat{H}^{(0)}|n\ell m\rangle = -\frac{\mathrm{Ry}}{n^2}|n\ell m\rangle$$

and, because $\hat{H}^{(0)}$ is Hermitian,

$$\langle n\ell m|\hat{H}^{(0)} = -\frac{\mathrm{Ry}}{n^2}\langle n\ell m|.$$

This gives

$$\langle \hat{p}^4\rangle_{n\ell m} = 4m_e^2\left(\frac{\mathrm{Ry}^2}{n^4} - 2\frac{\mathrm{Ry}}{n^2}\left\langle\frac{k}{r}\right\rangle_{n\ell m} + \left\langle\frac{k^2}{r^2}\right\rangle_{n\ell m}\right). \tag{14.99}$$

Exercise 14.D. Does $\left\langle\frac{1}{r}\right\rangle = \frac{1}{\langle r\rangle}$?

The two mean values above are far easier to work out the two Laplacians and one triple integral in the form (14.97). (For one thing, they involve only single integrals over r rather than a triple integral over $\vec{r}$.) The first is worked out indirectly at equation (14.113). Or, you may look them up.[3] The results are

$$\left\langle\frac{1}{r}\right\rangle_{n\ell m} = \frac{1}{a_0 n^2} \quad \text{and} \quad \left\langle\frac{1}{r^2}\right\rangle_{n\ell m} = \frac{1}{a_0^2 n^3(\ell+\frac{1}{2})}. \tag{14.100}$$

Pulling all these pieces together, the energy shifts are

$$E_{n\ell}^{(1)} = \frac{\mathrm{Ry}^2}{2m_e c^2}\left[\frac{3}{n^4} - \frac{4}{n^3(\ell+\frac{1}{2})}\right]. \tag{14.101}$$

In this more refined approximation, the accidental degeneracy is removed: the energy depends on ℓ as well as n.

This is not the end of the story for hydrogen, because there are additional defects in the Coulomb model: For example, there is a contribution to the true Hamiltonian called "spin-orbit coupling" that involves the interaction of the electron's intrinsic ("spin") magnetic moment with the magnetic moment due to the electron's motion ("orbit"). The spin-orbit effect turns out to be of about the same size as this relativistic effect.

This relativistic effect plus spin-orbit coupling are together called the "fine structure". They result in energy shifts on the order of 50 ζeV (corresponding to frequencies of about 10 GHz, in the microwave regime).

---

[3] E.U. Condon and G.H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, Cambridge, UK, 1935) page 117.

An even smaller effect is "spin-spin coupling", involving the interaction of the electron's magnetic moment and the proton's magnetic moment. This is called "hyperfine structure" and result in energy shifts on the order of 5 ζeV (corresponding to frequencies of about 1400 MHz, in the television regime). The famous 21 cm radiation used to map our galaxy comes from hyperfine structure.

Finely there is the "Lamb shift" at 1057 MHz, due to correcting the electrostatic potential $-k/r$ with a fully relativistic treatment of the quantal electromagnetic field acting between the proton and the electron.

14.9 Problems

14.1 **Positronium**

The "atom" positronium is a bound state of an electron and a positron. Find the allowed energies for positronium.

14.2 **Operator factorization solution of the Coulomb problem**

The bound state energy eigenvalues of the hydrogen atom can be found using the operator factorization method. In reduced units, the radial wave equation is

$$\left[-\frac{d^2}{d\tilde{r}^2} + \frac{\ell(\ell+1)}{\tilde{r}^2} - \frac{2}{\tilde{r}}\right] u_{n,\ell}(\tilde{r}) \equiv h_\ell \; u_{n,\ell}(\tilde{r}) = \mathcal{E}_{n,\ell} \; u_{n,\ell}(\tilde{r}). \quad (14.102)$$

Introduce the operators

$$D_\pm^{(\ell)} \equiv \frac{d}{d\tilde{r}} \mp \frac{\ell}{\tilde{r}} \pm \frac{1}{\ell} \quad (14.103)$$

and show that

$$D_-^{(\ell+1)} \; D_+^{(\ell+1)} = -h_\ell - \frac{1}{(\ell+1)^2}, \quad D_+^{(\ell)} D_-^{(\ell)} = -h_\ell - \frac{1}{\ell^2}. \quad (14.104)$$

From this, conclude that

$$h_{\ell+1} \; D_+^{(\ell+1)} \; u_{n,\ell}(\tilde{r}) = \mathcal{E}_{n,\ell} \; D_+^{(\ell+1)} \; u_{n,\ell}(\tilde{r}) \quad (14.105)$$

whence

$$D_+^{(\ell+1)} u_{n,\ell}(\tilde{r}) \;\propto\; u_{n,\ell+1}(\tilde{r}) \quad (14.106)$$

and $\mathcal{E}_{n,\ell}$ is independent of ℓ.

Argue that for every $\mathcal{E}_{n,\ell} < 0$ there is a maximum ℓ. (*Clue:* Examine the effective potential for radial motion.) Call this value ℓ_{max}, and set $n = \ell_{\text{max}} + 1$ to show that

$$\mathcal{E}_{n,\ell} = -\frac{1}{n^2}, \quad \ell = 0, \ldots, n-1. \quad (14.107)$$

14.3 A non-Coulombic central force

The central potential

$$V(r) = -\frac{k}{r} + \frac{c}{r^2} \tag{14.108}$$

is a model (albeit a poor one) for the interaction of the two atoms in a diatomic molecule. (Arnold Sommerfeld called this the "rotating oscillator" potential: see his *Atomic Structure and Spectral Lines*, 3rd ed., 1922, appendix 17.) Steven A. Klein (class of 1989) investigated this potential and found that its energy eigenproblem could be solved exactly.

a. Sketch the potential, assuming that k and c are both positive.

b. Following the method of section 14.6, convert the radial equation of the energy eigenproblem into

$$\left[-\frac{d^2}{d\tilde{r}^2} - \frac{2}{\tilde{r}} + \frac{\gamma + \ell(\ell+1)}{\tilde{r}^2}\right] u_{n,\ell}(\tilde{r}) = \mathcal{E}_{n,\ell}\, u_{n,\ell}(\tilde{r}). \tag{14.109}$$

where $\gamma = 2cM/\hbar^2$ and where $\tilde{r}$, $\mathcal{E}_{n,\ell}$, and $u_{n,\ell}(\tilde{r})$ are to be identified.

c. Find two values of x such that $x(x+1) = \gamma + \ell(\ell+1)$. Select whichever one will be most convenient for later use.

d. Convince yourself that the solution described in section 14.6 does not depend upon ℓ being an integer, and conclude that the energy eigenvalues are

$$\mathcal{E}_{n,\ell} = \frac{-1}{[n - \ell + \frac{1}{2}(-1 + \sqrt{(2\ell+1)^2 + 4\gamma})]^2} \tag{14.110}$$

where $n = 1, 2, 3, \ldots$ and where for each n, ℓ can take on values $\ell = 0, 1, 2, \ldots, n-1$.

e. Verify that this energy spectrum reduces to the Coulomb limit when $c = 0$.

14.4 The quantum mechanical virial theorem

a. Argue that, in an energy eigenstate $|\eta(t)\rangle$, the mean value $\langle \hat{\vec{r}} \cdot \hat{\vec{p}} \rangle$ does not change with time.

b. Hence conclude that $\langle \eta(t)|[\hat{\vec{r}} \cdot \hat{\vec{p}}, \hat{H}]|\eta(t)\rangle = 0$.

c. Show that $[\hat{\vec{r}} \cdot \hat{\vec{p}}, \hat{p}^2] = 2i\hbar\,\hat{p}^2$, while $[\hat{\vec{r}} \cdot \hat{\vec{p}}, V(\hat{\vec{r}})] = -i\hbar\,\hat{\vec{r}} \cdot \nabla V(\hat{\vec{r}})$, where $V(\vec{r})$ is any scalar function of the vector $\vec{r}$. (*Clue:* For the second commutator, use an explicit position basis representation.)

d. Suppose the Hamiltonian is

$$\hat{H} = \frac{1}{2m}\hat{p}^2 + V(\hat{r}) = \hat{T} + \hat{V}. \tag{14.111}$$

Define the force function $\vec{F}(\vec{r}) = -\nabla V(\vec{r})$ and the force operator $\hat{\vec{F}} = \vec{F}(\hat{\vec{r}})$. Conclude that, for an energy eigenstate,

$$2\langle \hat{T} \rangle = -\langle \hat{\vec{r}} \cdot \hat{\vec{F}} \rangle. \tag{14.112}$$

This is the "virial theorem."

e. If $V(\vec{r}) = C/r^n$, show that

$$2\langle \hat{T} \rangle = -n\langle \hat{V} \rangle \tag{14.113}$$

for any energy eigenstate, and that

$$\langle \hat{T} \rangle = \frac{n}{n-2}E, \qquad \langle \hat{V} \rangle = \frac{-2}{n-2}E, \tag{14.114}$$

for the energy eigenstate with energy E.

14.5 **Research project**

Discuss the motion of wavepackets in a Coulombic potential. Does the mean value of $\hat{\vec{r}}$ follow the classical Kepler ellipse? Is it even restricted to a plane? Does the wavepacket spread out in time (as with the force-free particle) or remain compact (as with the simple harmonic oscillator)?

Chapter 15

Identical Particles

Please review section 6.2, "Wavefunction: Two particles in one or three dimensions", on page 176. In that section we talked about two different particles, say an electron and a neutron. We set up a grid, discussed bin amplitudes $\psi_{i,j}$, and talked about the limit as the width of each bin shrank to zero.

15.1 Two identical particles

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

And here is the situation when one particle is found in bin 8, the other in bin 5:

No difference, of course. . . that's the *meaning* of "identical". And of course this holds not only for bins 5 and 8, but for any pair of bins i and j, even if $i = j$. (If the two particles don't interact, it is perfectly plausible for both of them to occupy the same bin at the same time.)

What does this mean for the state of a system with two identical particles? Suppose that, by hook or by crook, we come up with a set of bin amplitudes $\psi_{i,j}$ that describes the state of the system. Then the set of amplitudes $\phi_{i,j} = \psi_{j,i}$ describes that state just as well as the original set $\psi_{i,j}$. Does this mean that $\phi_{i,j} = \psi_{i,j}$? Not at all. Remember global phase freedom (pages 75 and **??**): If every bin amplitude is multiplied by the same "overall phase factor" — a complex number with magnitude unity — then the resulting set of amplitudes describes the state just as well as the original set did. Calling that overall phase factor s, we conclude that $\phi_{i,j} = s\psi_{i,j}$.

But, because $\phi_{i,j} = \psi_{j,i}$, the original set of amplitudes must satisfy $\psi_{j,i} = s\psi_{i,j}$. The variable name s comes from "swap": when we swap subscripts, we introduce a factor of s. The quantity s is a number…not a function of i or j. For example, the same value of s must work for $\psi_{8,5} = s\psi_{5,8}$, for $\psi_{7,3} = s\psi_{3,7}$, for $\psi_{5,8} = s\psi_{8,5}$, …. Wait. What was that last one? Put together the first and last examples:

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

Clearly, $s^2 = 1$, so s can't be any old complex number with magnitude unity: it can be only $s = +1$ or $s = -1$.

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

$$\psi(x_A, x_B) = +\psi(x_B, x_A) \quad \text{or} \quad \psi(x_A, x_B) = -\psi(x_B, x_A). \tag{15.1}$$

The first kind of wavefunction is called "symmetric under coordinate swapping", the second is called "antisymmetric under coordinate swapping". This requirement for symmetry or antisymmetry under coordinate swapping is called the Pauli[1] principle. It holds for all quantal states, not just energy eigenstates. It holds for interacting as well as for non-interacting

---

[1] Wolfgang Pauli (1900–1958), Vienna-born Swiss physicist, was one of the founders of quantum mechanics. In 1924 he proposed the "exclusion principle", ancestor of today's symmetry/antisymmetry requirement; in 1926 he produced the first solution for the energy eigenproblem for atomic hydrogen; in 1930 he proposed the existence of the neutrino, a prediction confirmed experimentally in 1956; in 1934 he and "Viki" Weisskopf discovered how to make sense of relativistic quantum mechanics by realizing that the solutions to relativistic quantal equations do not give an amplitude for a single particle to have a position (technically, a wavefunction), but rather an amplitude for an additional particle to be created at a position or for an existing particle to be annihilated at a position (technically, a creation or annihilation operator). He originated the insult, applied to ideas that cannot be tested, that they are "not even wrong".

identical particles. It holds for wavefunctions in both momentum and position representations (see problem 15.2). And it has a number of surprising consequences, both within the domain of quantum mechanics and atomic physics, as we will soon see, but also within the domain of statistical mechanics.[2]

It might distress you to see variables like x_A: doesn't x_A mean the position of particle "A" while x_B means the position of particle "B"? So doesn't this terminology label the particles as "A" and "B", which would violate our initial requirement that the particles be identical? The answer is that this terminology does *not* label one particle "A" and the other particle "B". Instead, it labels one *point* "A" and the other *point* "B". Look back to the figures on page 369: the numbers 5 and 8 label *bins*, not *particles*, so when these bins shrink to zero the variables x_A and x_B apply to *points*, not *particles*. That's why I like to call these wavefunctions "(anti)symmetric under swap of coordinates". But you'll hear people using terms like "(anti)symmetric under particle swapping" or "…under particle interchange" or "…under particle exchange".

What if the two particles are in three-dimensional space, and what if they have spin? In that case, the swap applies to *all* the coordinates: using the undertilde notation[3] of equation (12.18),

$$\psi(\underset{\sim}{x}_A, \underset{\sim}{x}_B) = +\psi(\underset{\sim}{x}_B, \underset{\sim}{x}_A) \quad \text{or} \quad \psi(\underset{\sim}{x}_A, \underset{\sim}{x}_B) = -\psi(\underset{\sim}{x}_B, \underset{\sim}{x}_A). \tag{15.2}$$

15.2 Three or more identical particles

What if there are three identical particles? The wavefunction is $\psi(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C)$ and you can swap either the first and second coordinates, or the second and third coordinates, or the first and third coordinates. This section will show that the wavefunction must be either symmetric under each of these three swaps or else antisymmetric under each of these three swaps.

---

[2] For example, particles with wavefunctions symmetric under coordinate swapping can undergo a phase transition called "Bose-Einstein condensation", whereas those with wavefunctions antisymmetric cannot. See, for example, R.K. Pathria and Paul Beale, *Statistical Mechanics*.

[3] So that the symbol $\underset{\sim}{x}$ represents whatever is needed to specify the state: For a spinless particle in one dimension $\underset{\sim}{x}$ represents the coordinate x or, if you are working in momentum space, the coordinate p. For a particle with spin moving in three dimensions, $\underset{\sim}{x}$ represents (x, y, z, m_z), or perhaps (p_x, p_y, p_z, m_x).

Any swap must produce a wavefunction representing the same state, so it can introduce at most a constant phase factor. We call that factor $s_{1,2}$ for swapping the first and second coordinates, $s_{1,3}$ for swapping the first and third coordinates, and $s_{2,3}$ for swapping the second and third coordinates. In other words

$$\begin{aligned}
&\psi(\underset{\sim}{x}_A,\underset{\sim}{x}_B,\underset{\sim}{x}_C)\\
&= s_{1,2}\psi(\underset{\sim}{x}_B,\underset{\sim}{x}_A,\underset{\sim}{x}_C)\\
&= s_{1,3}\psi(\underset{\sim}{x}_C,\underset{\sim}{x}_B,\underset{\sim}{x}_A)\\
&= s_{2,3}\psi(\underset{\sim}{x}_A,\underset{\sim}{x}_C,\underset{\sim}{x}_B).
\end{aligned}$$

The "swap then swap back" argument above shows that each of the three s factors must be either $+1$ or -1. We gain more information through repeated swappings that return ultimately to the initial sequence. For example

$$\begin{aligned}
&\psi(\underset{\sim}{x}_A,\underset{\sim}{x}_B,\underset{\sim}{x}_C) \quad \llbracket\text{swap the first and second coordinates giving...}\rrbracket\\
&= s_{1,2}\psi(\underset{\sim}{x}_B,\underset{\sim}{x}_A,\underset{\sim}{x}_C) \quad \llbracket\text{swap the second and third coordinates giving...}\rrbracket\\
&= s_{1,2}s_{2,3}\psi(\underset{\sim}{x}_B,\underset{\sim}{x}_C,\underset{\sim}{x}_A) \quad \llbracket\text{swap the first and third coordinates giving ...}\rrbracket\\
&= s_{1,2}s_{2,3}s_{1,3}\psi(\underset{\sim}{x}_A,\underset{\sim}{x}_C,\underset{\sim}{x}_B) \quad \llbracket\text{swap the second and third coordinates giving...}\rrbracket\\
&= s_{1,2}s_{2,3}s_{1,3}s_{2,3}\psi(\underset{\sim}{x}_A,\underset{\sim}{x}_B,\underset{\sim}{x}_C)
\end{aligned}$$

We already know that $(s_{2,3})^2 = 1$, so this argument reveals that $s_{1,2}s_{1,3} = 1$, i.e., these two phase factors are either both $+1$ or both -1. There are four possibilities:

$$\begin{array}{llll}
\text{A:} & s_{1,2} = +1; & s_{1,3} = +1; & s_{2,3} = +1\\
\text{B:} & s_{1,2} = +1; & s_{1,3} = +1; & s_{2,3} = -1\\
\text{C:} & s_{1,2} = -1; & s_{1,3} = -1; & s_{2,3} = +1\\
\text{D:} & s_{1,2} = -1; & s_{1,3} = -1; & s_{2,3} = -1
\end{array}$$

Furthermore, we can go from $\psi(\underset{\sim}{x}_A,\underset{\sim}{x}_B,\underset{\sim}{x}_C)$ to $\psi(\underset{\sim}{x}_B,\underset{\sim}{x}_C,\underset{\sim}{x}_A)$ via two different swapping routes:

$$\begin{aligned}
&\psi(\underset{\sim}{x}_A,\underset{\sim}{x}_B,\underset{\sim}{x}_C) \quad \llbracket\text{swap the first and second coordinates giving...}\rrbracket\\
&= s_{1,2}\psi(\underset{\sim}{x}_B,\underset{\sim}{x}_A,\underset{\sim}{x}_C) \quad \llbracket\text{swap the second and third coordinates giving...}\rrbracket\\
&= s_{1,2}s_{2,3}\psi(\underset{\sim}{x}_B,\underset{\sim}{x}_C,\underset{\sim}{x}_A)
\end{aligned}$$

or

$$\begin{aligned}
&\psi(\underset{\sim}{x}_A,\underset{\sim}{x}_B,\underset{\sim}{x}_C) \quad \llbracket\text{swap the first and third coordinates giving...}\rrbracket\\
&= s_{1,3}\psi(\underset{\sim}{x}_C,\underset{\sim}{x}_B,\underset{\sim}{x}_A) \quad \llbracket\text{swap the first and second coordinates giving...}\rrbracket\\
&= s_{1,3}s_{1,2}\psi(\underset{\sim}{x}_B,\underset{\sim}{x}_C,\underset{\sim}{x}_A)
\end{aligned}$$

The conclusion is that $s_{2,3} = s_{1,3}$, so possibilities B and C above are ruled out. A wavefunction for three identical particles must be either symmetric under all swaps or else antisymmetric under all swaps.

Exercise 15.A. Four or more particles. Show that the same result applies for wavefunctions of four identical particles by applying the above argument to clusters of three coordinates. There are four clusters: first, second, and third; first, second, and fourth; first, third, and fourth; second, third, and fourth. Argue that because the clusters overlap, the wavefunction must be either completely symmetric or completely antisymmetric. Generalize your argument to five or more identical particles.

In conclusion, a wavefunction for any number of identical particles must be either "completely symmetric" (every swap introduces a phase factor of +1) or else "completely antisymmetric" (every swap introduces a phase factor of −1). This is called the "exchange symmetry" of the wavefunction.

15.1 **How many swaps?**
If there are two particles, there is one possible swap. If there are three particles, there are three possible swaps. Show that for four particles there are six possible swaps and that for N particles there are $N(N-1)/2$ possible swaps.

15.2 **Pauli principle in the momentum representation**
Show that the momentum wavefunction as the same interchange symmetry as the position wavefunction (i.e., symmetric or antisymmetric). How about the energy coefficients? (Exactly what does that last question mean?)

15.3 **Conservation of exchange symmetry**
Show that exchange symmetry is conserved: If the system starts out in a symmetric state it will remain symmetric at all times in the future, and similarly for antisymmetric.

15.3 Bosons and fermions

Given what we've uncovered so far, I would guess that a collection of neutrons could start out in a symmetric state (in which case they would be in a symmetric state for all time) or else they could start out in an antisymmetric state (in which case they would be in an antisymmetric state for all time). In fact, however, this is not the case. For suppose you had a collection of five neutrons in a symmetric state and a different collection of two neutrons in an antisymmetric state. Just by changing which collection is under consideration, you could consider this as one collection of seven neutrons. That collection of seven neutrons would have to be either completely symmetric or completely antisymmetric, and it wouldn't be if the five were in a symmetric state and the two in an antisymmetric state.

So the exchange symmetry has nothing to do with history or with what you consider to be the extent of the collection, but instead depends only on the type of particle. Neutrons, protons, electrons, carbon-13 nuclei (in their ground state), ^{3}He atoms (in their ground state), and sigma baryons are always antisymmetric under swapping — they are called "fermions".[4] Photons, alpha particles, carbon-12 nuclei (in their ground state), ^{4}He atoms (in their ground state), and pi mesons are always symmetric under swapping — they are called "bosons".[5]

Furthermore, all bosons have integral spin and all fermions have half-integral spin. There is a mathematical result in relativistic quantum field theory called "the spin-statistics theorem" that sheds some light on this astounding fact.[6]

---

[4] Enrico Fermi (1901–1954) of Italy excelled in both experimental and theoretical physics. He directed the building of the first nuclear reactor and produced the first theory of the weak interaction. The Fermi surface in the physics of metals was named in his honor. He elucidated the statistics of what are now called fermions in 1926. He produced so many thoughtful conceptual and estimation problems that such problems are today called "Fermi problems". I never met him (he died before I was born) but I have met several of his students, and all of them speak of him in that rare tone reserved for someone who is not just a great scientist and a great teacher and a great leader, but also a great human being.

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

[6] See Ian Duck and E.C.G. Sudarshan, *Pauli and the Spin-Statistics Theorem* (World Scientific, Singapore, 1997), and the review of this book by A.S. Wightman in *American Journal of Physics* **67** (August 1999) 742–746.

15.4 Symmetrization and antisymmetrization

Given the importance of wavefunctions symmetric or antisymmetric under coordinate swaps, it makes sense to investigate the mathematics of such "permutation symmetry". This section treats systems of two or three particles; the generalization to systems of four or more particles is straightforward.

Start with any two-variable garden-variety function $f(\underset{\sim}{x}_A, \underset{\sim}{x}_B)$, not necessarily symmetric or antisymmetric. Can that function be used as a "seed" to build a symmetric or antisymmetric function? It can. The function

$$s(\underset{\sim}{x}_A, \underset{\sim}{x}_B) = f(\underset{\sim}{x}_A, \underset{\sim}{x}_B) + f(\underset{\sim}{x}_B, \underset{\sim}{x}_A) \tag{15.3}$$

is symmetric under swapping while the function

$$a(\underset{\sim}{x}_A, \underset{\sim}{x}_B) = f(\underset{\sim}{x}_A, \underset{\sim}{x}_B) - f(\underset{\sim}{x}_B, \underset{\sim}{x}_A) \tag{15.4}$$

is antisymmetric. If you don't believe me, try it out:

$$\begin{aligned} s(5,2) &= f(5,2) + f(2,5) \\ s(2,5) &= f(2,5) + f(5,2) \end{aligned}$$

so clearly $s(5,2) = s(2,5)$. Meanwhile

$$\begin{aligned} a(5,2) &= f(5,2) - f(2,5) \\ a(2,5) &= f(2,5) - f(5,2) \end{aligned}$$

so just as clearly $a(5,2) = -a(2,5)$.

Can this be generalized to three variables? Start with a three-variable garden-variety function $f(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C)$. The function

$$\begin{aligned} s(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C) = \;& f(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C) \\ & +f(\underset{\sim}{x}_A, \underset{\sim}{x}_C, \underset{\sim}{x}_B) \\ & +f(\underset{\sim}{x}_C, \underset{\sim}{x}_A, \underset{\sim}{x}_B) \\ & +f(\underset{\sim}{x}_C, \underset{\sim}{x}_B, \underset{\sim}{x}_A) \\ & +f(\underset{\sim}{x}_B, \underset{\sim}{x}_C, \underset{\sim}{x}_A) \\ & +f(\underset{\sim}{x}_B, \underset{\sim}{x}_A, \underset{\sim}{x}_C) \end{aligned} \tag{15.5}$$

is completely symmetric while the function

$$\begin{aligned} a(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C) = \;& f(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C) \\ & -f(\underset{\sim}{x}_A, \underset{\sim}{x}_C, \underset{\sim}{x}_B) \\ & +f(\underset{\sim}{x}_C, \underset{\sim}{x}_A, \underset{\sim}{x}_B) \\ & -f(\underset{\sim}{x}_C, \underset{\sim}{x}_B, \underset{\sim}{x}_A) \\ & +f(\underset{\sim}{x}_B, \underset{\sim}{x}_C, \underset{\sim}{x}_A) \\ & -f(\underset{\sim}{x}_B, \underset{\sim}{x}_A, \underset{\sim}{x}_C) \end{aligned} \tag{15.6}$$

is completely antisymmetric. Once again, if you don't believe me I invite you to try it out with $\underset{\sim}{x}_A = 5$, $\underset{\sim}{x}_B = 2$, and $\underset{\sim}{x}_C = 7$.

⟦These $6 = 3!$ permutations are listed in the sequence called[7] "plain changes" or "the Johnson-Trotter sequence". This sequence has the admirable property that each permutation differs from its predecessor by a single swap of adjacent letters.⟧

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

$$f(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C) = f_1(\underset{\sim}{x}_A) f_2(\underset{\sim}{x}_B) f_3(\underset{\sim}{x}_C), \tag{15.7}$$

in which case you may think of the symmetrization/antisymmetrization machinery as being the sum over all permutations of the coordinates $\underset{\sim}{x}_A$, $\underset{\sim}{x}_B$, and $\underset{\sim}{x}_C$, as above, or as the sum over all permutations of the functions $f_1(\underset{\sim}{x})$, $f_2(\underset{\sim}{x})$, and $f_3(\underset{\sim}{x})$: the function

$$\begin{aligned} s(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C) = \quad & f_1(\underset{\sim}{x}_A) f_2(\underset{\sim}{x}_B) f_3(\underset{\sim}{x}_C) \\ & + f_1(\underset{\sim}{x}_A) f_3(\underset{\sim}{x}_B) f_2(\underset{\sim}{x}_C) \\ & + f_3(\underset{\sim}{x}_A) f_1(\underset{\sim}{x}_B) f_2(\underset{\sim}{x}_C) \\ & + f_3(\underset{\sim}{x}_A) f_2(\underset{\sim}{x}_B) f_1(\underset{\sim}{x}_C) \\ & + f_2(\underset{\sim}{x}_A) f_3(\underset{\sim}{x}_B) f_1(\underset{\sim}{x}_C) \\ & + f_2(\underset{\sim}{x}_A) f_1(\underset{\sim}{x}_B) f_3(\underset{\sim}{x}_C) \end{aligned} \tag{15.8}$$

is completely symmetric while the function

$$\begin{aligned} a(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C) = \quad & f_1(\underset{\sim}{x}_A) f_2(\underset{\sim}{x}_B) f_3(\underset{\sim}{x}_C) \\ & - f_1(\underset{\sim}{x}_A) f_3(\underset{\sim}{x}_B) f_2(\underset{\sim}{x}_C) \\ & + f_3(\underset{\sim}{x}_A) f_1(\underset{\sim}{x}_B) f_2(\underset{\sim}{x}_C) \\ & - f_3(\underset{\sim}{x}_A) f_2(\underset{\sim}{x}_B) f_1(\underset{\sim}{x}_C) \\ & + f_2(\underset{\sim}{x}_A) f_3(\underset{\sim}{x}_B) f_1(\underset{\sim}{x}_C) \\ & - f_2(\underset{\sim}{x}_A) f_1(\underset{\sim}{x}_B) f_3(\underset{\sim}{x}_C) \end{aligned} \tag{15.9}$$

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

$$a(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C) = \begin{vmatrix} f_1(\underset{\sim}{x}_A) & f_2(\underset{\sim}{x}_A) & f_3(\underset{\sim}{x}_A) \\ f_1(\underset{\sim}{x}_B) & f_2(\underset{\sim}{x}_B) & f_3(\underset{\sim}{x}_B) \\ f_1(\underset{\sim}{x}_C) & f_2(\underset{\sim}{x}_C) & f_3(\underset{\sim}{x}_C) \end{vmatrix}, \tag{15.10}$$

---

[7] Donald Knuth, *The Art of Computer Programming*, volume 4A, "Combinatorial Algorithms, Part 1" (Addison-Wesley, Boston, 1997) section 7.2.1.2, "Generating all permutations".

and call it the "Slater[8] determinant". I personally think this terminology confuses the issue (the expression works only if the seed function is a product of one-variable functions, it suppresses the delightful and useful "plain changes" sequence of permutations, plus I never liked determinants[9] to begin with), but it's widely used.

15.4 **Symmetrizing and antisymmetrizing the already symmetric**
If the seed $f(x_A, x_B, x_C)$ happens to be completely symmetric to begin with, what are the symmetrized and antisymmetrized functions? What if the seed happens to be antisymmetric to begin with?

15.5 **Two variables versus three variables**
Show that any two-variable function can be represented as a sum of a symmetric and an antisymmetric function. Can any three-variable function be represented as a sum of a completely symmetric and a completely antisymmetric function?

15.5 Consequences of the Pauli principle

Does the requirement of symmetry or antisymmetry under coordinate swapping have any consequences? Here's an immediate one for fermions: Take both $\underset{\sim}{x}_A = \underset{\sim}{x}$ and $\underset{\sim}{x}_B = \underset{\sim}{x}$. Now when these coordinates are swapped, you get back to where you started:

$$\psi(\underset{\sim}{x}, \underset{\sim}{x}) = -\psi(\underset{\sim}{x}, \underset{\sim}{x}) \quad \text{so} \quad \psi(\underset{\sim}{x}, \underset{\sim}{x}) = 0. \tag{15.11}$$

Thus, the probability density for two identical fermions to have all the same coordinates is zero.

And here's a consequence for both bosons and fermions. Think about space only, no spin. The (unnormalized) seed function

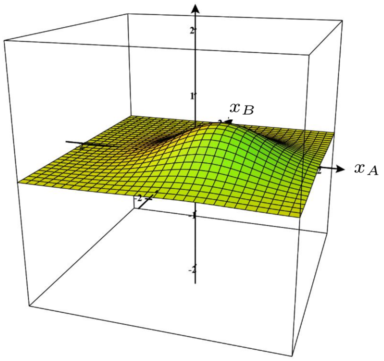
$$f(x_A, x_B) = e^{-[(x_A-0.5\sigma)^2+(x_B+0.3\sigma)^2]/2\sigma^2}$$

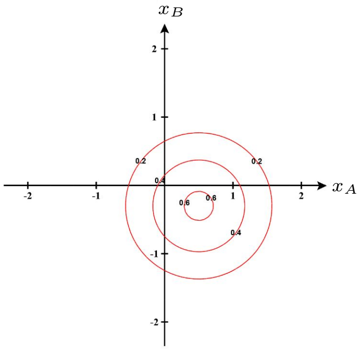
has a maximum when $x_A = 0.5\sigma$ and when $x_B = -0.3\sigma$. This shows up as one hump in the two-variable plots below (drawn taking $\sigma = 1$), which show the normalized probability density proportional to $|f(x_A, x_B)|^2$.

---

[8] John C. Slater (1900–1976), American theoretical physicist who made major contributions to our understanding of atoms, molecules, and solids. Also important as a teacher, textbook author, and administrator.

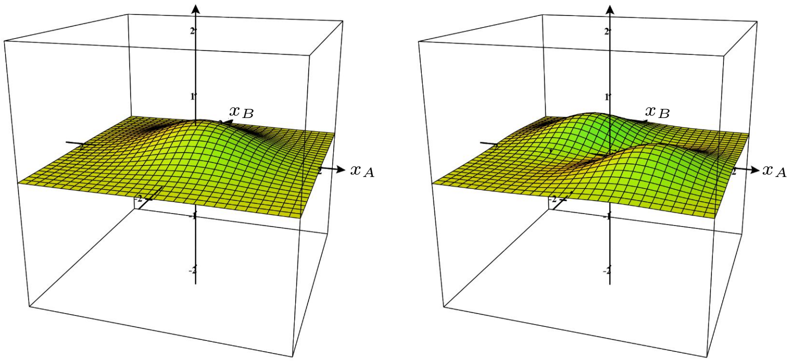
[9] I am not alone. See Sheldon Axler, "Down with determinants!" *American Mathematical Monthly* **102** (February 1995) 139–154.



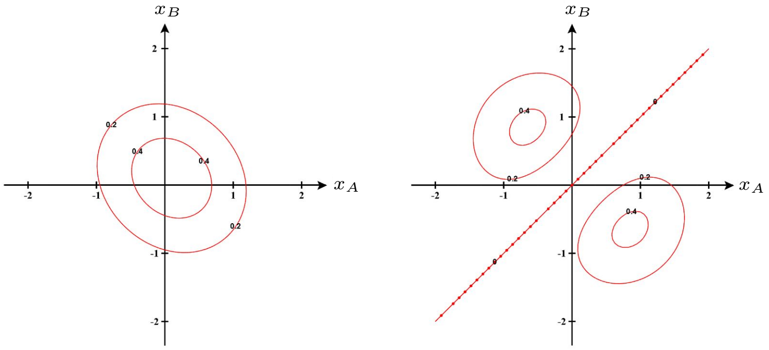


Depending on your background and preferences, you might find it easier to read either the surface plot on the left or the contour plot on the right: both depict the same two-variable function. (And both were drawn using Paul Seeburger's applet CalcPlot3D.)

But what of the symmetric and antisymmetric combinations generated from this seed? Here are surface plots of the normalized probability densities associated with the symmetric (left) and antisymmetric (right) combinations:



And here are the corresponding contour plots:



The seed function has no special properties on the $x_A = x_B$ diagonal axis. But, as required by equation (15.11), the antisymmetric combination vanishes there. And the symmetric combination is high there!

The "vanishing on diagonal requirement" and this particular example are but two facets of the more general rule of thumb that:

> In a symmetric spatial wavefunction, the particles tend to huddle together.
>
> In an antisymmetric spatial wavefunction, the particles tend to spread apart.

This rule is *not* a theorem and you can find counterexamples,[10] but such exceptions are rare.

In everyday experience, when two people tend to huddle together or spread apart, it's for emotional reasons. In everyday experience, when two particles tend to huddle together or spread apart, it's because they're attracted to or repelled from each other through a force. This quantal case is vastly different. The huddling or spreading is of course not caused by emotions and it's also not caused by a force — it occurs for identical particles even when they don't interact. The cause is instead the symmetry/antisymmetry requirement: not a force like a hammer blow, but a piece of mathematics!

Therefore it's difficult to come up with terms for the behavior of identical particles that don't suggest either emotions or forces ascribed to particles: congregate, avoid; gregarious, loner; attract, repel; flock, scatter. "Huddle together" and "spread apart" are the best terms I've been able to devise, but you might be able to find better ones.

Exercise 15.B. Does the "huddle together/spread apart" rule of thumb hold for wavefunctions in momentum space?

Problem

15.6 **Symmetric and antisymmetric combinations: infinite square well**

Two identical particles ambivate in a one-dimensional infinite square well. Take as a seed function the product of energy eigenstates

---

[10] See D.F. Styer, "On the separation of identical particles in quantum mechanics" *European Journal of Physics* **41** (14 October 2020) 065402.

$\eta_2(x_A)\eta_3(x_B)$. Use your favorite graphics package to plot the probability densities associated with the symmetric and antisymmetric combinations generated from this seed. Does the "huddle together/spread apart" rule hold?

15.6 Consequences of the Pauli principle for product states

A commonly encountered special case comes when the many-particle seed function is a product of one-particle functions — we glanced at this special case in equation (15.7). What happens if two of these one-particle functions are the same? Nothing special happens for the symmetrization case. But the answer for antisymmetrization is cute. It pops out of equation (15.9): If $f_1(\underset{\sim}{x}) = f_2(\underset{\sim}{x})$, then the last line cancels the first line, the second cancels the fifth, and the fourth cancels the third. The antisymmetric combination vanishes everywhere!

Unlike the "huddle together/spread apart" rule of thumb, this result *is* a theorem: the antisymmetric combination vanishes if any two of the one-particle functions are the same. It is a partner to the $\underset{\sim}{x}_A = \underset{\sim}{x}_B$ theorem of equation (15.11): just as the two particles can't have the same coordinates, so their wavefunction can't be built from the same one-particle functions.

A fascinating but more specialized result concerns the root-mean-square separation between the two identical particles ambivating in one dimension:

$$s_{\text{rms}} \equiv \Big[\langle (x_A - x_B)^2\rangle\Big]^{1/2}, \tag{15.12}$$

whence

$$s_{\text{rms}}^2 = \langle x_A^2\rangle + \langle x_B^2\rangle - 2\langle x_A x_B\rangle. \tag{15.13}$$

If the one-particle seed functions $f_1(x)$ and $f_2(x)$ are normalized and orthogonal, then the unsymmetrized wavefunction is

$$f_1(x_A)f_2(x_B), \tag{15.14}$$

the symmetrized wavefunction is

$$\tfrac{1}{\sqrt{2}}\Big[f_1(x_A)f_2(x_B) + f_2(x_A)f_1(x_B)\Big], \tag{15.15}$$

and the antisymmetrized wavefunction is

$$\tfrac{1}{\sqrt{2}}\Big[f_1(x_A)f_2(x_B) - f_2(x_A)f_1(x_B)\Big]. \tag{15.16}$$

Exercise 15.C. Verify the normalization constants in equations (15.15) and (15.16).

We now calculate the rms separations for these three wavefunctions in turn. For the unsymmetrized wavefunction (15.14)

$$\begin{aligned}\langle x_A^2\rangle &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_1^*(x_A)f_2^*(x_B)\,x_A^2\,f_1(x_A)f_2(x_B)\,dx_A\,dx_B\\ &= \int_{-\infty}^{+\infty} f_1^*(x_A)\,x_A^2\,f_1(x_A)\,dx_A\int_{-\infty}^{+\infty} f_2^*(x_B)f_2(x_B)\,dx_B\\ &= \langle x^2\rangle_1,\end{aligned}$$

where $\langle x^2\rangle_1$ represents the mean value of x^2 in the one-particle state $f_1(x)$. Similarly

$$\langle x_B^2\rangle = \langle x^2\rangle_2.$$

And

$$\begin{aligned}\langle x_A x_B\rangle &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_1^*(x_A)f_2^*(x_B)\,x_A x_B\,f_1(x_A)f_2(x_B)\,dx_A\,dx_B\\ &= \int_{-\infty}^{+\infty} f_1^*(x_A)\,x_A\,f_1(x_A)\,dx_A\int_{-\infty}^{+\infty} f_2^*(x_B)\,x_B\,f_2(x_B)\,dx_B\\ &= \langle x\rangle_1\langle x\rangle_2.\end{aligned}$$

Thus for the unsymmetrized wavefunction (15.14),

$$s_{\text{rms}}^2 = \langle x^2\rangle_1 + \langle x^2\rangle_2 - 2\langle x\rangle_1\langle x\rangle_2. \tag{15.17}$$

We can do the calculations for both the symmetrized (15.15) and the antisymmetrized (15.16) two-particle wavefunctions at once:

$$\begin{aligned}
\langle x_A^2\rangle = \frac{1}{2}\Bigg[&\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_1^*(x_A)f_2^*(x_B)\,x_A^2\,f_1(x_A)f_2(x_B)\,dx_A\,dx_B\\
&\pm\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_1^*(x_A)f_2^*(x_B)\,x_A^2\,f_2(x_A)f_1(x_B)\,dx_A\,dx_B\\
&\pm\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_2^*(x_A)f_1^*(x_B)\,x_A^2\,f_1(x_A)f_2(x_B)\,dx_A\,dx_B\\
&+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_2^*(x_A)f_1^*(x_B)\,x_A^2\,f_2(x_A)f_1(x_B)\,dx_A\,dx_B\Bigg]\\
= \frac{1}{2}\Bigg[&\int_{-\infty}^{+\infty} f_1^*(x_A)\,x_A^2\,f_1(x_A)\,dx_A\int_{-\infty}^{+\infty} f_2^*(x_B)f_2(x_B)\,dx_B\\
&\pm\int_{-\infty}^{+\infty} f_1^*(x_A)\,x_A^2\,f_2(x_A)\,dx_A\int_{-\infty}^{+\infty} f_2^*(x_B)f_1(x_B)\,dx_B\\
&\pm\int_{-\infty}^{+\infty} f_2^*(x_A)\,x_A^2\,f_1(x_A)\,dx_A\int_{-\infty}^{+\infty} f_1^*(x_B)f_2(x_B)\,dx_B\\
&+\int_{-\infty}^{+\infty} f_2^*(x_A)\,x_A^2\,f_2(x_A)\,dx_A\int_{-\infty}^{+\infty} f_1^*(x_B)f_1(x_B)\,dx_B\Bigg]\\
= \tfrac{1}{2}[&\langle x^2\rangle_1 + \langle x^2\rangle_2].
\end{aligned}$$

Of course, $\langle x_B^2\rangle$ has the same value.

Finally

$$
\begin{aligned}
\langle x_A x_B \rangle = \frac{1}{2} \Bigg[& \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_1^*(x_A) f_2^*(x_B)\, x_A x_B\, f_1(x_A) f_2(x_B)\, dx_A\, dx_B \\
& \pm \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_1^*(x_A) f_2^*(x_B)\, x_A x_B\, f_2(x_A) f_1(x_B)\, dx_A\, dx_B \\
& \pm \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_2^*(x_A) f_1^*(x_B)\, x_A x_B\, f_1(x_A) f_2(x_B)\, dx_A\, dx_B \\
& + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_2^*(x_A) f_1^*(x_B)\, x_A x_B\, f_2(x_A) f_1(x_B)\, dx_A\, dx_B \Bigg] \\
= \frac{1}{2} \Bigg[& \int_{-\infty}^{+\infty} f_1^*(x_A)\, x_A\, f_1(x_A)\, dx_A \int_{-\infty}^{+\infty} f_2^*(x_B)\, x_B\, f_2(x_B)\, dx_B \\
& \pm \int_{-\infty}^{+\infty} f_1^*(x_A)\, x_A\, f_2(x_A)\, dx_A \int_{-\infty}^{+\infty} f_2^*(x_B)\, x_B\, f_1(x_B)\, dx_B \\
& \pm \int_{-\infty}^{+\infty} f_2^*(x_A)\, x_A\, f_1(x_A)\, dx_A \int_{-\infty}^{+\infty} f_1^*(x_B)\, x_B\, f_2(x_B)\, dx_B \\
& + \int_{-\infty}^{+\infty} f_2^*(x_A)\, x_A\, f_2(x_A)\, dx_A \int_{-\infty}^{+\infty} f_1^*(x_B)\, x_B\, f_1(x_B)\, dx_B \Bigg] \\
= {} & \langle x \rangle_1 \langle x \rangle_2 \pm \langle 1|x|2\rangle \langle 2|x|1\rangle \\
= {} & \langle x \rangle_1 \langle x \rangle_2 \pm |\langle 2|x|1\rangle|^2,
\end{aligned}
$$

where

$$\langle 2|x|1\rangle \equiv \int_{-\infty}^{+\infty} f_2^*(x)\, x\, f_1(x)\, dx.$$

Thus for the symmetrized (15.15) or antisymmetrized wavefunction (15.14),

$$s_{\text{rms}}^2 = \langle x^2 \rangle_1 + \langle x^2 \rangle_2 - 2\langle x \rangle_1 \langle x \rangle_2 \mp 2|\langle 2|x|1\rangle|^2. \tag{15.18}$$

This result is generally in accord with our "huddle together/spread apart" rule of thumb, because

rms separation for symmetrized
≤ rms separation for unsymmetrized
≤ rms separation for antisymmetrized.

On the other hand, if it should happen that $|\langle 2|x|1\rangle|$ vanishes, then all three rms separations are exactly the same.

From California, to the New York island

There is an electron in California with wavefunction $\phi_{\rm CA}(\vec{x})$, and an electron in New York with wavefunction $\phi_{\rm NY}(\vec{x})$. (Neglect spin for simplicity.) Do I really need to treat them as a system of two electrons with wavefunction

$$\psi(\vec{r}_A,\vec{r}_B) = \tfrac{1}{\sqrt{2}}[\phi_{\rm CA}(\vec{x}_A)\phi_{\rm NY}(\vec{x}_B) - \phi_{\rm NY}(\vec{x}_A)\phi_{\rm CA}(\vec{x}_B)] \text{ ?} \qquad (15.19)$$

No, and this section explains why.

What is the probability density for finding an electron at point $\vec{x}$, regardless of the position of the other electron? It is

$$\rho(\vec{x}) = \int |\psi(\vec{x},\vec{x}_B)|^2 \, d^3x_B + \int |\psi(\vec{x}_A,\vec{x})|^2 \, d^3x_A, \qquad (15.20)$$

where the integrals run over all space. Plugging in our expression (15.19) for $\psi(\vec{r}_A,\vec{r}_B)$ shows that this probability density equals exactly

$$|\phi_{\rm NY}(\vec{x})|^2 + |\phi_{\rm CA}(\vec{x})|^2 - 2\Re e\left\{\phi^*_{\rm CA}(\vec{x})\phi_{\rm NY}(\vec{x}) \int \phi^*_{\rm NY}(\vec{x}_A)\phi_{\rm CA}(\vec{x}_A)\, d^3x_A\right\}. \qquad (15.21)$$

But, because the two electrons are so far apart, it is an excellent approximation (called "no overlap") that

$$\phi_{\rm CA}(\vec{x})\phi_{\rm NY}(\vec{x}) = 0 \quad \text{for all } \vec{x}. \qquad (15.22)$$

In this excellent approximation, the right-most term of equation (15.21), the "interference term", vanishes. Furthermore, for points $\vec{x}$ in California, $|\phi_{\rm NY}(\vec{x})|^2 = 0$, again to an excellent approximation. Thus if you're in California the probability density is

$$|\phi_{\rm CA}(\vec{x})|^2, \qquad (15.23)$$

which is exactly the conclusion you would have drawn without all this New York rigamarole.

Problem

15.7 Mean separation in the infinite square well

Two noninteracting particles are in an infinite square well of width L. The associated one-body energy eigenstates are $\eta_n(x)$ and $\eta_m(x)$, where

$$\eta_n(x) = \sqrt{\frac{2}{L}}\,\sin\left(n\pi\frac{x}{L}\right).$$

Calculate the root-mean-square separation if these are

a. two non-identical particles, one in state $\eta_n(x_A)$ and the other in state $\eta_m(x_B)$

b. two identical bosons, in state

$$\tfrac{1}{\sqrt{2}}\left[\eta_n(x_A)\eta_m(x_B) + \eta_m(x_A)\eta_n(x_B)\right]$$

c. two identical fermions, in state

$$\tfrac{1}{\sqrt{2}}\left[\eta_n(x_A)\eta_m(x_B) - \eta_m(x_A)\eta_n(x_B)\right]$$

Do your results always adhere to our rule of thumb that "symmetric means huddled together; antisymmetric means spread apart"?

15.7 A basis for three identical particles

15.7.1 *Three-particle states built from one-particle levels*

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

$$\eta_1(\underset{\sim}{x}), \eta_2(\underset{\sim}{x}), \eta_3(\underset{\sim}{x}), \ldots, \eta_M(\underset{\sim}{x}). \tag{15.24}$$

Now **three non-identical particles**, each with the same mass, ambivate subject to the same potential energy. If they don't interact with each other, you can see what the energy eigenstates are: state $\eta_3(\underset{\sim}{x}_A)\eta_8(\underset{\sim}{x}_B)\eta_2(\underset{\sim}{x}_C)$, for example, has energy $E_3 + E_8 + E_2$. There's necessarily a degeneracy, as defined on page **??**, because the different state $\eta_8(\underset{\sim}{x}_A)\eta_3(\underset{\sim}{x}_B)\eta_2(\underset{\sim}{x}_C)$ has the same energy. If the three particles *do* interact, then these states are not energy states, but they *do* constitute a basis. Any state can be represented as a linear combination of these basis members. These states are normalized. I could go on, but the picture is clear: the fact that there are three particles rather than one is unimportant; this basis has all the properties you expect of a basis.[11]

---

[11] Notice that if the three particles don't interact, it's perfectly okay for two or even three of them to have the same position. Only for particles that repel, with infinite potential energy when the separation vanishes, is it true that "two particles cannot occupy the same place at the same time".

I list some members of this basis of product single-particle wavefunctions.

$$\begin{array}{lll}
\eta_1(\underset{\sim}{x}_A)\eta_1(\underset{\sim}{x}_B)\eta_1(\underset{\sim}{x}_C) & |1,1,1\rangle & E_1+E_1+E_1 \\
\eta_1(\underset{\sim}{x}_A)\eta_1(\underset{\sim}{x}_B)\eta_2(\underset{\sim}{x}_C) & |1,1,2\rangle & E_1+E_1+E_2 \\
\eta_1(\underset{\sim}{x}_A)\eta_2(\underset{\sim}{x}_B)\eta_1(\underset{\sim}{x}_C) & |1,2,1\rangle & E_1+E_2+E_1 \\
\eta_2(\underset{\sim}{x}_A)\eta_1(\underset{\sim}{x}_B)\eta_1(\underset{\sim}{x}_C) & |2,1,1\rangle & E_2+E_1+E_1 \\
\eta_1(\underset{\sim}{x}_A)\eta_1(\underset{\sim}{x}_B)\eta_3(\underset{\sim}{x}_C) & |1,1,3\rangle & E_1+E_1+E_3 \\
\quad\vdots & \quad\vdots & \quad\vdots \\
\eta_1(\underset{\sim}{x}_A)\eta_4(\underset{\sim}{x}_B)\eta_3(\underset{\sim}{x}_C) & |1,4,3\rangle & E_1+E_4+E_3 \\
\eta_4(\underset{\sim}{x}_A)\eta_3(\underset{\sim}{x}_B)\eta_1(\underset{\sim}{x}_C) & |4,3,1\rangle & E_4+E_3+E_1 \\
\quad\vdots & \quad\vdots & \quad\vdots \\
\eta_2(\underset{\sim}{x}_A)\eta_7(\underset{\sim}{x}_B)\eta_3(\underset{\sim}{x}_C) & |2,7,3\rangle & E_2+E_7+E_3 \\
\eta_7(\underset{\sim}{x}_A)\eta_3(\underset{\sim}{x}_B)\eta_2(\underset{\sim}{x}_C) & |7,3,2\rangle & E_7+E_3+E_2 \\
\quad\vdots & \quad\vdots & \quad\vdots \\
\eta_M(\underset{\sim}{x}_A)\eta_M(\underset{\sim}{x}_B)\eta_M(\underset{\sim}{x}_C) & |M,M,M\rangle & E_M+E_M+E_M
\end{array}$$

The left column gives the conventional name of the product wavefunction. It's tiring to write these long names, we we abbreviate them as shown in the center column. The right column shows the energy of the state if the three particles don't interact.

A few remarks: (1) There are M^3 members in the basis. (2) Sequence matters: the state $|4,5,1\rangle$ is different from the state $|1,5,4\rangle$. (3) This is a basis of product wavefunctions, but that doesn't mean that every state is a product state, because an arbitrary state is a sum of basis members.

To keep in mind the distinction between this basis for the three-particle system (with M^3 members) and the basis for the one-particle system from which it is built (with M members), we often call the three-particle basis members "states" and the one-particle basis members "levels". The levels are the building blocks out of which states are constructed.[12]

---

[12] Some people, particularly chemists referring to atomic systems, use the term "orbital" rather than "level". This term unfortunately suggests a circular Bohr orbit. An electron with an energy does *not* execute a circular Bohr orbit at constant speed. Instead it ambivates without position or velocity.

15.7.2 *Building a symmetric basis.*

Any wavefunction can be expressed as a sum over the above basis,

$$\psi(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C) = \sum_{r=1}^{M}\sum_{s=1}^{M}\sum_{t=1}^{M} c_{r,s,t}\eta_r(\underset{\sim}{x}_A)\eta_s(\underset{\sim}{x}_B)\eta_t(\underset{\sim}{x}_C) = \sum_{r,s,t} c_{r,s,t}|r,s,t\rangle,$$

but if we have three identical bosons, we're not interested in any wavefunction, we're interested only in symmetric wavefunctions. To build a symmetric wavefunction, we execute the symmetrization process (15.5) on $\psi(\underset{\sim}{x}_A, \underset{\sim}{x}_B, \underset{\sim}{x}_C)$. Doing so, we conclude that this symmetric wavefunction can be expressed as a sum over the symmetrization of each member of the basis. As a result, if we go through and symmetrize each member of the basis for three non-identical particles (the one on page 387), we will produce a basis for symmetric states.

The symmetrization of

$$\eta_r(\underset{\sim}{x}_A)\eta_s(\underset{\sim}{x}_B)\eta_t(\underset{\sim}{x}_C) \quad \text{also known as} \quad |r,s,t\rangle$$

can be executed with the process at equation (15.8). We represent this symmetrization as

$$\hat{\mathcal{S}}|r,s,t\rangle = A_s\left(|r,s,t\rangle + |r,t,s\rangle + |t,r,s\rangle + |t,s,r\rangle + |s,t,r\rangle + |s,r,t\rangle\right)$$

where A_s is a normalization constant.

Let's execute this process starting with $|1,1,1\rangle$. This symmetrizes to itself:

$$\hat{\mathcal{S}}|1,1,1\rangle = |1,1,1\rangle.$$

Next comes $|1,1,2\rangle$:

$$\begin{aligned}\hat{\mathcal{S}}|1,1,2\rangle &= A_s\left(|1,1,2\rangle + |1,2,1\rangle + |2,1,1\rangle + |2,1,1\rangle + |1,2,1\rangle + |1,1,2\rangle\right)\\ &= 2A_s\left(|1,1,2\rangle + |1,2,1\rangle + |2,1,1\rangle\right).\end{aligned}$$

It's clear, now, that

$$\hat{\mathcal{S}}|1,1,2\rangle = \hat{\mathcal{S}}|1,2,1\rangle = \hat{\mathcal{S}}|2,1,1\rangle,$$

so we must discard two of these three states from our symmetric basis. In fact, it's clear that all states built through symmetrizing any three given levels are the same state. For example

$$\hat{\mathcal{S}}|3,9,2\rangle = \hat{\mathcal{S}}|3,2,9\rangle = \hat{\mathcal{S}}|2,3,9\rangle = \hat{\mathcal{S}}|2,9,3\rangle = \hat{\mathcal{S}}|9,2,3\rangle = \hat{\mathcal{S}}|9,3,2\rangle,$$

and we must discard five of these six states from our symmetric basis.

We are left with a basis for symmetric functions

$\lvert 1,1,1\rangle$	$E_1+E_1+E_1$
$\hat{\mathcal{S}}\lvert 1,1,2\rangle$	$E_1+E_1+E_2$
~~$\hat{\mathcal{S}}\lvert 1,2,1\rangle$~~	~~$E_1+E_2+E_1$~~
~~$\hat{\mathcal{S}}\lvert 2,1,1\rangle$~~	~~$E_2+E_1+E_1$~~
$\hat{\mathcal{S}}\lvert 1,1,3\rangle$	$E_1+E_1+E_3$
$\vdots$	$\vdots$
$\hat{\mathcal{S}}\lvert 1,4,3\rangle$	$E_1+E_4+E_3$
~~$\hat{\mathcal{S}}\lvert 4,3,1\rangle$~~	~~$E_4+E_3+E_1$~~
$\vdots$	$\vdots$
$\hat{\mathcal{S}}\lvert 2,7,3\rangle$	$E_2+E_7+E_3$
~~$\hat{\mathcal{S}}\lvert 7,3,2\rangle$~~	~~$E_7+E_3+E_2$~~
$\vdots$	$\vdots$
$\lvert M,M,M\rangle$	$E_M+E_M+E_M$

A few remarks: (1) There are

$$\frac{M(M+1)(M+2)}{3!}$$

members in the basis. (2) Sequence doesn't matter: the state $\hat{\mathcal{S}}|4,5,1\rangle$ is the same as the state $\hat{\mathcal{S}}|1,5,4\rangle$. (3) This is a basis of symmetrizations of products of levels, but that doesn't mean that every state is a symmetrization of products of levels because an arbitrary state is a sum of basis members.

You'll notice that in this table (unlike the table on page 387) I don't write out the conventional name of the wavefunction. That's because these names are long . . . for example one of them is

$$\begin{aligned}\tfrac{1}{\sqrt{3!}}\big[&\eta_2(\underset{\sim}{x}_A)\eta_7(\underset{\sim}{x}_B)\eta_3(\underset{\sim}{x}_C)+\eta_2(\underset{\sim}{x}_A)\eta_3(\underset{\sim}{x}_B)\eta_7(\underset{\sim}{x}_C)\\ &+\eta_3(\underset{\sim}{x}_A)\eta_2(\underset{\sim}{x}_B)\eta_7(\underset{\sim}{x}_C)+\eta_3(\underset{\sim}{x}_A)\eta_7(\underset{\sim}{x}_B)\eta_2(\underset{\sim}{x}_C)\\ &+\eta_7(\underset{\sim}{x}_A)\eta_3(\underset{\sim}{x}_B)\eta_2(\underset{\sim}{x}_C)+\eta_7(\underset{\sim}{x}_A)\eta_2(\underset{\sim}{x}_B)\eta_3(\underset{\sim}{x}_C)\big]\,.\end{aligned}$$

On the other hand to specify these basis states we need only list the three levels that go into building it (the three "building blocks" that go into making it). [This was *not* the case for three non-identical particles.] Consequently one often speaks of this state as "a particle in level 2, a particle in level 7, and a particle in level 3". This phrase is *not* correct: If a particle *were* in level 7, then it could be distinguished as "the particle in level 7" and hence would not be identical to the other two particles. The correct statement is that the *system* is in the symmetric state given above, and that the *individual particles* do not have states. On the other hand, the correct statement is a mouthful and you may use the “balls in buckets” picture as shorthand — as long as you say it but don’t think it.

15.7.3 *Building an antisymmetric basis.*

We can build a basis of states, each of which is antisymmetric, in a parallel manner by antisymmetrizing each member of the basis for non-identical particles and discarding duplicates.

The antisymmetrization of

$$\eta_r(\underset{\sim}{x}_A)\eta_s(\underset{\sim}{x}_B)\eta_t(\underset{\sim}{x}_C) \quad \text{also known as} \quad |r,s,t\rangle$$

can be executed with the process at equation (15.9). We represent this antisymmetrization as

$$\hat{\mathcal{A}}|r,s,t\rangle = A_a\left(|r,s,t\rangle - |r,t,s\rangle + |t,r,s\rangle - |t,s,r\rangle + |s,t,r\rangle - |s,r,t\rangle\right)$$

where A_a is again a normalization constant.

Let’s execute this process starting with $|1,1,1\rangle$. This antisymmetrizes to zero:

$$\hat{\mathcal{A}}|1,1,1\rangle = 0.$$

Same with $|1,1,2\rangle$:

$$\begin{aligned}\hat{\mathcal{A}}|1,1,2\rangle &= A_a\left(|1,1,2\rangle - |1,2,1\rangle + |2,1,1\rangle - |2,1,1\rangle + |1,2,1\rangle - |1,1,2\rangle\right)\\ &= 0.\end{aligned}$$

It’s clear, in fact, that *any* basis member with two indices the same will antisymmetrize to zero. (This reflects the theorem on page 381 that “the antisymmetric combination vanishes if any two of the one-particle functions are the same.”) The only way to avoid antisymmetrization to zero is for all of the level indices to differ. Furthermore

$$\hat{\mathcal{A}}|r,s,t\rangle = -\hat{\mathcal{A}}|r,t,s\rangle = \hat{\mathcal{A}}|t,r,s\rangle = -\hat{\mathcal{A}}|t,s,r\rangle = \hat{\mathcal{A}}|s,t,r\rangle = -\hat{\mathcal{A}}|s,r,t\rangle$$

so the six distinct basis members $|2,7,3\rangle$, $|7,3,2\rangle$, $|3,7,2\rangle$, etc. all antisymmetrize to the same state.

We are left with a basis for antisymmetric functions

~~$\hat{\mathcal{A}}\vert 1,1,1\rangle$~~	~~$E_1 + E_1 + E_1$~~
~~$\hat{\mathcal{A}}\vert 1,1,2\rangle$~~	~~$E_1 + E_1 + E_2$~~
~~$\hat{\mathcal{A}}\vert 1,2,1\rangle$~~	~~$E_1 + E_2 + E_1$~~
~~$\hat{\mathcal{A}}\vert 2,1,1\rangle$~~	~~$E_2 + E_1 + E_1$~~
~~$\hat{\mathcal{A}}\vert 1,1,3\rangle$~~	~~$E_1 + E_1 + E_3$~~
$\vdots$	$\vdots$
$\hat{\mathcal{A}}\vert 1,4,3\rangle$	$E_1 + E_4 + E_3$
~~$\hat{\mathcal{A}}\vert 4,3,1\rangle$~~	~~$E_4 + E_3 + E_1$~~
$\vdots$	$\vdots$
$\hat{\mathcal{A}}\vert 2,7,3\rangle$	$E_2 + E_7 + E_3$
~~$\hat{\mathcal{A}}\vert 7,3,2\rangle$~~	~~$E_7 + E_3 + E_2$~~
$\vdots$	$\vdots$
~~$\hat{\mathcal{A}}\vert M,M,M\rangle$~~	~~$E_M + E_M + E_M$~~

A few remarks: (1) There are

$$\frac{M(M-1)(M-2)}{3!}$$

members in the basis. (2) Sequence doesn't matter: the expression $\hat{\mathcal{A}}|4,5,1\rangle$ is the negative of the express $\hat{\mathcal{A}}|1,5,4\rangle$, but they represent the same state. (3) This is a basis of antisymmetrizations of products of levels, but that doesn't mean that every state is an antisymmetrization of products of levels because an arbitrary state is a sum of basis members.

Once again these states have long expressions like

$$\begin{aligned}\tfrac{1}{\sqrt{3!}}\big[&\eta_2(\underset{\sim}{x}_A)\eta_7(\underset{\sim}{x}_B)\eta_3(\underset{\sim}{x}_C) - \eta_2(\underset{\sim}{x}_A)\eta_3(\underset{\sim}{x}_B)\eta_7(\underset{\sim}{x}_C)\\ &+ \eta_3(\underset{\sim}{x}_A)\eta_2(\underset{\sim}{x}_B)\eta_7(\underset{\sim}{x}_C) - \eta_3(\underset{\sim}{x}_A)\eta_7(\underset{\sim}{x}_B)\eta_2(\underset{\sim}{x}_C)\\ &+ \eta_7(\underset{\sim}{x}_A)\eta_3(\underset{\sim}{x}_B)\eta_2(\underset{\sim}{x}_C) - \eta_7(\underset{\sim}{x}_A)\eta_2(\underset{\sim}{x}_B)\eta_3(\underset{\sim}{x}_C)\big]\,.\end{aligned}$$

but to specify the three-particle state we need only list the one-particle building blocks ("levels") used in its construction. This results in almost the same "balls in buckets" picture that we drew for symmetric wavefunctions, but with the additional restriction that any bucket can contain only one or zero balls. Once again you may use the "balls in buckets" picture as a shorthand, as long as you keep in mind that it conceals a considerably more intricate process of building and antisymmetrizing.

Generalizations. It is easy to generalize this procedure for building antisymmetric and symmetric many-particle basis states out of one-particle levels for any number of particles. The only special case is for two particles, where the symmetric basis has $M(M+1)/2$ members and the antisymmetric basis has $M(M-1)/2$ members. Putting these two bases together results in a full basis of M^2 members. This reflects the fact that any function of two variables can be written as the sum of an antisymmetric and a symmetric function. The same is *not* true for systems of three or more particles.

If there are N particles, the symmetric basis has

$$\binom{M+N-1}{N} = \frac{(M+N-1)!}{N!(M-1)!}. \tag{15.25}$$

members, the antisymmetric basis has

$$\binom{M}{N} = \frac{M!}{N!(M-N)!} \tag{15.26}$$

members.

15.7.4 *The occupation number representation*

We have seen that in order to specify a member of the symmetric or the antisymmetric basis that we have just produced, it is not necessary to specify the order of the one-particle level building blocks. For example $\hat{A}|4,9,7\rangle$ represents the same state as $\hat{A}|4,7,9\rangle$, so there's no need to pay attention to the order in which the 4, 7, and 9 appear. This observation permits the "occupation number" representation of such states, in which we specify the basis state simply by listing the one-particle levels that are used as building blocks to make up that state. Or, equivalently but more commonly, we specify the basis state by listing the number n_r of one-body levels of each type r that are used as building blocks. (And, of course, we must also specify whether we're considering the symmetric or the antisymmetric basis.) Thus, for example:

level r:	1	2	3	4	5	6	$\cdots$	M
$\hat{S}\|3,4,4\rangle$ has n_r:	0	0	1	2	0	0	$\cdots$	0
$\hat{A}\|1,3,4\rangle$ has n_r:	1	0	1	1	0	0	$\cdots$	0

The second line in this table means that the state $\hat{S}|3,4,4\rangle$ is built by starting with the three levels $\eta_3(\underset{\sim}{x}_A)$, $\eta_3(\underset{\sim}{x}_B)$, and $\eta_4(\underset{\sim}{x}_C)$, multiplying them together, and then symmetrizing. Sometimes you will hear this state described by the phrase "there is one particle in level 3 and two particles in level 4", but that can't be literally true... the three particles are identical, and if they could be assigned to distinct levels they would not be identical! Phrases such as the one above[13] invoke the "balls in buckets" picture of N-particle quantal wavefunctions: The state $\hat{S}|3,4,4\rangle$ is pictured as one ball in bucket number 3 and two balls in bucket number 4. It is all right to use this picture and this phraseology, as long as you don't believe it. Always keep in mind that it is a shorthand for a more elaborate process of building up states from levels by multiplication and symmetrization.

The very term "occupation number" for n_r is a poor one, because it so strongly suggests the balls-in-buckets picture: "Particles A and B are in level 3, particle C is in level 4." If this were correct, then particle A could not be identical with particle C — they are distinguished by being in different levels. (Just as a fast baseball cannot be identical with a slow baseball of the same construction — they are distinguished by having different speeds.) The fact is, the individual particles don't have labels and they *don't have states.* Instead, the system as a whole has a state. That state is built by taking one level 3 and two levels 4, multiplying them and then symmetrizing them.

A more accurate picture than the "balls in buckets" picture is: You have a stack of bricks of type 1, a stack of bricks of type 2, ..., a stack of bricks of type M. Build a state by taking one brick from stack 3, and two bricks from stack 4.

The balls in buckets picture is easy to work with, but gives the misimpression that a particle is in a particular level, and the state of the system is given by listing the state (level) of each individual particle. No. The system is in a particular non-product state, and the particles themselves *don't have states (or levels).*

A somewhat better (yet still imperfect) name for n_r is "occupancy". If you can think of a better name, please let the world know!

To summarize the occupation number representation: a member of the symmetric basis is specified by the list

$$n_r, \qquad \text{for} \qquad r = 1, 2, \ldots M, \qquad \text{where} \qquad n_r \text{ is } 0, 1, 2, \ldots, \tag{15.27}$$

---

[13] For example, phrases like "the level is filled" or "the level is empty" or "the level is half-filled".

and a member of the antisymmetric basis is specified by the list

$$n_r, \quad \text{for} \quad r = 1, 2, \ldots M, \quad \text{where} \quad n_r \text{ is 0 or 1.} \tag{15.28}$$

The total number of particles in such a state is

$$N = \sum_{r=1}^{M} n_r, \tag{15.29}$$

and, if the particles don't interact, the energy of the state is

$$E = \sum_{r=1}^{M} n_r E_r. \tag{15.30}$$

15.8 **Problem:** Count the members of the antisymmetric and antisymmetric bases for N particles rather than three. (Continue to use M levels.) Does your expression have the proper limits when $N = 1$ and when $N = M$?

15.9 **Problem:** Find the normalization constant for $\hat{\mathcal{S}}|7, 3, 7\rangle$.

15.10 **Problem:** Any two-variable function may be written as a sum of a symmetric and an antisymmetric function. Consequently the union of the symmetric basis and the antisymmetric basis is a basis for the set of all two-variable functions. Show that neither of these statements is true for functions of three variables.

15.11 **Building basis states for three particles**
Suppose you had three particles and three "building block" levels (say the orthonormal levels $\eta_1(x)$, $\eta_3(x)$, and $\eta_7(x)$). Construct normalized three-particle basis states for the case of

a. three non-identical particles
b. three identical bosons
c. three identical fermions

How many states are there in each basis? Repeat for three particles with four one-particle levels, but in this case simply count and don't write down all the three-particle states.

15.8 Spin plus space, two electrons

Electrons are spin-half fermions. Two of them ambivate subject to the same potential. Energy doesn't depend on spin. Pretend the two electrons don't interact. (Perhaps a better name for this section would be "Spin plus space, two noninteracting spin-$\frac{1}{2}$ fermions", but yikes, how long do you want this section's title to be? Should I add "non-relativistic" and "ignoring collisions" and "ignoring radiation"?)

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

Well, what two-particle states can we build from the one-particle spatial levels with, say, $n = 1$ and $n = 3$? (Once you see how to do it for $n = 1$ and $n = 3$, you can readily generalize to any two values of n.) These correspond to four levels:

$$\eta_1(\vec{x})\chi_+, \tag{15.31}$$

$$\eta_1(\vec{x})\chi_-, \tag{15.32}$$

$$\eta_3(\vec{x})\chi_+, \tag{15.33}$$

$$\eta_3(\vec{x})\chi_-. \tag{15.34}$$

What states mixing $n = 1$ with $n = 3$ can be built from these four levels?

The antisymmetric combination of (15.31) with itself vanishes. The antisymmetric combination of (15.31) with (15.32) is a combination of $n = 1$ with $n = 1$, not of $n = 1$ with $n = 3$. The (unnormalzed) antisymmetric combination of (15.31) with (15.33) is

$$\eta_1(\vec{x}_A)\chi_+(A)\eta_3(\vec{x}_B)\chi_+(B) - \eta_3(\vec{x}_A)\chi_+(A)\eta_1(\vec{x}_B)\chi_+(B). \tag{15.35}$$

The antisymmetric combination of (15.31) with (15.34) is

$$\eta_1(\vec{x}_A)\chi_+(A)\eta_3(\vec{x}_B)\chi_-(B) - \eta_3(\vec{x}_A)\chi_-(A)\eta_1(\vec{x}_B)\chi_+(B). \tag{15.36}$$

The antisymmetric combination of (15.32) with (15.33) is

$$\eta_1(\vec{x}_A)\chi_-(A)\eta_3(\vec{x}_B)\chi_+(B) - \eta_3(\vec{x}_A)\chi_+(A)\eta_1(\vec{x}_B)\chi_-(B). \tag{15.37}$$

The antisymmetric combination of (15.32) with (15.34) is

$$\eta_1(\vec{x}_A)\chi_-(A)\eta_3(\vec{x}_B)\chi_-(B) - \eta_3(\vec{x}_A)\chi_-(A)\eta_1(\vec{x}_B)\chi_-(B). \tag{15.38}$$

Finally, the antisymmetric combination of (15.33) with (15.34) is a combination of $n = 3$ with $n = 3$, not of $n = 1$ with $n = 3$.

All four of these states are energy eigenstates with energy $E_1 + E_3$. State (15.35) factorizes into a convenient space-times-spin form:

$$\begin{aligned} &\eta_1(\vec{x}_A)\chi_+(A)\eta_3(\vec{x}_B)\chi_+(B) - \eta_3(\vec{x}_A)\chi_+(A)\eta_1(\vec{x}_B)\chi_+(B) \\ &= \left[\eta_1(\vec{x}_A)\eta_3(\vec{x}_B) - \eta_3(\vec{x}_A)\eta_1(\vec{x}_B)\right]\chi_+(A)\chi_+(B). \end{aligned} \tag{15.39}$$

The space part of the wavefunction is antisymmetric under coordinate swap. The spin part is symmetric. Thus the total wavefunction is antisymmetric.

Before proceeding I confess that I'm sick and tired of writing all these ηs and χs and As and Bs that convey no information. I *always* write the η in front of the χ. I *always* write the As in front of the Bs. You'll *never* confuse an η with a χ, because the ηs are labeled 1, 3 while the χs are labeled $+$, $-$. Dirac introduced a notation (see page 58) that takes all this for granted, so that neither you nor I have to write the same thing out over and over again. This notation usually replaces $+$ with $\uparrow$ and $-$ with $\downarrow$ (see page 80). In this notation, equation (15.39) is written

$$|1\uparrow, 3\uparrow\rangle - |3\uparrow, 1\uparrow\rangle = \Big[|1,3\rangle - |3,1\rangle\Big]\,|\uparrow\uparrow\rangle. \tag{15.40}$$

In this new notation the states (15.35) through (15.38) are written

$$\Big[|1,3\rangle - |3,1\rangle\Big]\,|\uparrow\uparrow\rangle \tag{15.41}$$

$$|1\uparrow, 3\downarrow\rangle - |3\downarrow, 1\uparrow\rangle \tag{15.42}$$

$$|1\downarrow, 3\uparrow\rangle - |3\uparrow, 1\downarrow\rangle \tag{15.43}$$

$$\Big[|1,3\rangle - |3,1\rangle\Big]\,|\downarrow\downarrow\rangle. \tag{15.44}$$

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

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

Because all four states (15.41) through (15.44) have the same energy, namely $E_1 + E_3$, I can make linear combinations of the states to form other equally good energy states. Can I make a combination of states (15.42) and (15.43) that *does* factorize into space times spin? Nothing ventured, nothing gained. Let's try it:

$$\alpha\left[|1\uparrow,3\downarrow\rangle-|3\downarrow,1\uparrow\rangle\right]+\beta\left[|1\downarrow,3\uparrow\rangle-|3\uparrow,1\downarrow\rangle\right]$$
$$=|1,3\rangle\left[\alpha|{\uparrow\downarrow}\rangle+\beta|{\downarrow\uparrow}\rangle\right]-|3,1\rangle\left[\alpha|{\downarrow\uparrow}\rangle+\beta|{\uparrow\downarrow}\rangle\right].$$

This will factorize only if the left term in square brackets is proportional to the right term in square brackets:

$$\left[\alpha|{\uparrow\downarrow}\rangle+\beta|{\downarrow\uparrow}\rangle\right]=c\left[\beta|{\uparrow\downarrow}\rangle+\alpha|{\downarrow\uparrow}\rangle\right],$$

that is only if

$$\alpha=c\beta\qquad\text{and}\qquad\beta=c\alpha.$$

Combining these two equations results in $c=\pm1$. If $c=+1$ then the combination results in the state

$$\left[|1,3\rangle-|3,1\rangle\right]\alpha\left[|{\uparrow\downarrow}\rangle+|{\downarrow\uparrow}\rangle\right], \tag{15.45}$$

whereas when $c=-1$ the result is

$$\left[|1,3\rangle+|3,1\rangle\right]\alpha\left[|{\uparrow\downarrow}\rangle-|{\downarrow\uparrow}\rangle\right]. \tag{15.46}$$

Putting all this together and, for the sake of good form, insuring normalized states, we find that the two-electron energy states in equations (15.41) through (15.44) can be recast as

$$\left[\tfrac{1}{\sqrt{2}}(|1,3\rangle-|3,1\rangle)\right]|{\uparrow\uparrow}\rangle \tag{15.47}$$

$$\left[\tfrac{1}{\sqrt{2}}(|1,3\rangle-|3,1\rangle)\right]\left[\tfrac{1}{\sqrt{2}}(|{\uparrow\downarrow}\rangle+|{\downarrow\uparrow}\rangle)\right] \tag{15.48}$$

$$\left[\tfrac{1}{\sqrt{2}}(|1,3\rangle-|3,1\rangle)\right]|{\downarrow\downarrow}\rangle \tag{15.49}$$

$$\left[\tfrac{1}{\sqrt{2}}(|1,3\rangle+|3,1\rangle)\right]\left[\tfrac{1}{\sqrt{2}}(|{\uparrow\downarrow}\rangle-|{\downarrow\uparrow}\rangle)\right]. \tag{15.50}$$

The first three of these states have spatial wavefunctions antisymmetric under coordinate swaps and spin wavefunctions symmetric under coordinate swaps — these are called "ortho states" or "a triplet". The last one has a symmetric spatial wavefunction and an antisymmetric spin wavefunction — these are called "para states" or "a singlet". Our discussion in section 15.5, "Consequences of the Pauli principle", demonstrates that in ortho states, the two electrons tend to spread apart in space; in para states, they tend to huddle together.

I write out the singlet spin state

$$\frac{1}{\sqrt{2}}\left[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right] \tag{15.51}$$

using the verbose terminology

$$\frac{1}{\sqrt{2}}\left[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)\right] \tag{15.52}$$

to make it absolutely clear that coordinate A is associated with both spin $+$ and spin $-$, as is coordinate B. It is impossible to say that "one electron has spin up and the other has spin down".

This abstract machinery might seem purely formal, but in fact it has tangible experimental consequences. In the sample problem below, the machinery suggests that the ground state of the hydrogen atom is two-fold degenerate, while the ground state of the helium atom is non-degenerate. And this prediction is borne out by experiment!

15.8.1 *Sample Problem: Ground state degeneracy for one and two electrons*

A certain potential energy function has two spatial energy eigenstates: $\eta_1(\vec{x})$ with energy E_1 and $\eta_2(\vec{x})$ with a higher energy E_2. These energies are independent of spin.

a. A single electron (spin-$\frac{1}{2}$) ambivates in this potential. Write out the four energy eigenstates and the energy eigenvalue associated with each. What is the ground state degeneracy?
b. Two non-interacting electrons ambivate in this same potential. Write out the six energy eigenstates and the energy eigenvalue associated with each. What is the ground state degeneracy?

Solution: **(a)** For the single electron:

energy eigenstate	energy eigenvalue
$\eta_1(\vec{x})\chi_+$	E_1
$\eta_1(\vec{x})\chi_-$	E_1
$\eta_2(\vec{x})\chi_+$	E_2
$\eta_2(\vec{x})\chi_-$	E_2

The first two states listed are both ground states, so the ground state is two-fold degenerate.

(b) For the two electrons, we build states from levels just as we did in this section. The first line below is the antisymmetrized combination of $\eta_1(\vec{x})\chi_+$ with $\eta_1(\vec{x})\chi_-$. This state has energy $2E_1$. The next four lines are built up exactly as equations (15.47) through (15.50) were. Each of these four states has energy $E_1 + E_2$. The last line is the antisymmetrized combination of $\eta_2(\vec{x})\chi_+$ with $\eta_2(\vec{x})\chi_-$. This state has energy $2E_2$.

$$\eta_1(\vec{x}_A)\eta_1(\vec{x}_B)\tfrac{1}{\sqrt{2}}\left[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)\right]$$
$$\tfrac{1}{\sqrt{2}}\left[\eta_1(\vec{x}_A)\eta_2(\vec{x}_B) - \eta_2(\vec{x}_A)\eta_1(\vec{x}_B)\right]\left[\chi_+(A)\chi_+(B)\right]$$
$$\tfrac{1}{\sqrt{2}}\left[\eta_1(\vec{x}_A)\eta_2(\vec{x}_B) - \eta_2(\vec{x}_A)\eta_1(\vec{x}_B)\right]\tfrac{1}{\sqrt{2}}\left[\chi_+(A)\chi_-(B) + \chi_-(A)\chi_+(B)\right]$$
$$\tfrac{1}{\sqrt{2}}\left[\eta_1(\vec{x}_A)\eta_2(\vec{x}_B) - \eta_2(\vec{x}_A)\eta_1(\vec{x}_B)\right]\left[\chi_-(A)\chi_-(B)\right]$$
$$\tfrac{1}{\sqrt{2}}\left[\eta_1(\vec{x}_A)\eta_2(\vec{x}_B) + \eta_2(\vec{x}_A)\eta_1(\vec{x}_B)\right]\tfrac{1}{\sqrt{2}}\left[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)\right]$$
$$\eta_2(\vec{x}_A)\eta_2(\vec{x}_B)\tfrac{1}{\sqrt{2}}\left[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)\right].$$

The ground state of the two-electron system is the first state listed: it is non-degenerate.

Problems

15.12 Combining a spatial one-particle level with itself

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

15.13 Change of basis through abstract rotation

Show that, in retrospect, the process of building states (15.48) and (15.50) from states (15.42) and (15.43) is nothing but a "45° rotation" in the style of equation (**??**).

15.14 Normalization of singlet spin state

Justify the normalization constant $\frac{1}{\sqrt{2}}$ that enters in moving from equation (15.46) to equation (15.50). Compare this singlet spin state to the entangled state (2.37). (Indeed, one way to produce an entangled pair of electrons is to start in a singlet state and then draw the two electrons apart.)

15.15 Ortho and para accounting

Show that in our case with $M/2$ spatial energy levels, the two-electron energy basis has $\frac{1}{2}M(M-1)$ members, of which

$$\tfrac{3}{2}(M/2)[(M/2)-1] \text{ are ortho}$$

(antisymmetric in space and symmetric in spin) and

$$\tfrac{1}{2}(M/2)[(M/2)+1] \text{ are para}$$

(symmetric in space and antisymmetric in spin).

15.16 **Intersystem crossing**

A one-electron system has a ground level $\eta_g(\vec{x})$ and an excited level $\eta_e(\vec{x})$, for a total of four basis levels:

$$\eta_g(\vec{x})\chi_+, \quad \eta_g(\vec{x})\chi_-, \quad \eta_e(\vec{x})\chi_+, \quad \eta_e(\vec{x})\chi_-.$$

A basis for two-electron states is then the six states:

$$\begin{aligned}
&\eta_g(\vec{x}_A)\eta_g(\vec{x}_B)\tfrac{1}{\sqrt{2}}\left[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)\right]\\
&\tfrac{1}{\sqrt{2}}\left[\eta_g(\vec{x}_A)\eta_e(\vec{x}_B) - \eta_e(\vec{x}_A)\eta_g(\vec{x}_B)\right]\left[\chi_+(A)\chi_+(B)\right]\\
&\tfrac{1}{\sqrt{2}}\left[\eta_g(\vec{x}_A)\eta_e(\vec{x}_B) - \eta_e(\vec{x}_A)\eta_g(\vec{x}_B)\right]\tfrac{1}{\sqrt{2}}\left[\chi_+(A)\chi_-(B) + \chi_-(A)\chi_+(B)\right]\\
&\tfrac{1}{\sqrt{2}}\left[\eta_g(\vec{x}_A)\eta_e(\vec{x}_B) - \eta_e(\vec{x}_A)\eta_g(\vec{x}_B)\right]\left[\chi_-(A)\chi_-(B)\right]\\
&\tfrac{1}{\sqrt{2}}\left[\eta_g(\vec{x}_A)\eta_e(\vec{x}_B) + \eta_e(\vec{x}_A)\eta_g(\vec{x}_B)\right]\tfrac{1}{\sqrt{2}}\left[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)\right]\\
&\eta_e(\vec{x}_A)\eta_e(\vec{x}_B)\tfrac{1}{\sqrt{2}}\left[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)\right].
\end{aligned}$$

A transition from the second state listed above to the first is called an "intersystem crossing". One sometimes reads, in association with the diagram below, that in an intersystem crossing "the spin of the excited electron is reversed". In five paragraphs or fewer, explain why this phrase is inaccurate, perhaps even grotesque, and suggest a replacement.

15.9 Spin plus space, three electrons, ground state

Three electrons are in the situation described in the first paragraph of section 15.8 (energy independent of spin, electrons don't interact). The full listing of energy eigenstates has been done, but it's an accounting nightmare, so I ask a simpler question: What is the ground state?

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

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

or the antisymmetrized combination of the three levels

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

The two states so generated are degenerate:[14] both have energy $2E_1 + E_2$.

Write out the first state in detail. It is

$$\begin{aligned}\tfrac{1}{\sqrt{6}}[\quad &\eta_1(\vec{x}_A)\chi_+(A)\,\eta_1(\vec{x}_B)\chi_-(B)\,\eta_2(\vec{x}_C)\chi_+(C)\\ -&\eta_1(\vec{x}_A)\chi_+(A)\,\eta_2(\vec{x}_B)\chi_+(B)\,\eta_1(\vec{x}_C)\chi_-(C)\\ +&\eta_2(\vec{x}_A)\chi_+(A)\,\eta_1(\vec{x}_B)\chi_+(B)\,\eta_1(\vec{x}_C)\chi_-(C)\\ -&\eta_2(\vec{x}_A)\chi_+(A)\,\eta_1(\vec{x}_B)\chi_-(B)\,\eta_1(\vec{x}_C)\chi_+(C)\\ +&\eta_1(\vec{x}_A)\chi_-(A)\,\eta_2(\vec{x}_B)\chi_+(B)\,\eta_1(\vec{x}_C)\chi_+(C)\\ -&\eta_1(\vec{x}_A)\chi_-(A)\,\eta_1(\vec{x}_B)\chi_+(B)\,\eta_2(\vec{x}_C)\chi_+(C)\ ].\end{aligned} \tag{15.53}$$

This morass is another good argument for the abbreviated Dirac notation introduced on page 396. I'm not concerned with normalization for the moment, so I'll write this first state as

$$\begin{aligned}&|1\uparrow, 1\downarrow, 2\uparrow\rangle\\ -&|1\uparrow, 2\uparrow, 1\downarrow\rangle\\ +&|2\uparrow, 1\uparrow, 1\downarrow\rangle\\ -&|2\uparrow, 1\downarrow, 1\uparrow\rangle\\ +&|1\downarrow, 2\uparrow, 1\uparrow\rangle\\ -&|1\downarrow, 1\uparrow, 2\uparrow\rangle\end{aligned} \tag{15.54}$$

and the second one (with $2\downarrow$ replacing $2\uparrow$) as

$$\begin{aligned}&|1\uparrow, 1\downarrow, 2\downarrow\rangle\\ -&|1\uparrow, 2\downarrow, 1\downarrow\rangle\\ +&|2\downarrow, 1\uparrow, 1\downarrow\rangle\\ -&|2\downarrow, 1\downarrow, 1\uparrow\rangle\\ +&|1\downarrow, 2\downarrow, 1\uparrow\rangle\\ -&|1\downarrow, 1\uparrow, 2\downarrow\rangle.\end{aligned} \tag{15.55}$$

Both of these states are antisymmetric, but neither factorizes into a neat "space part times spin part". If, following the approach used with two electrons, you attempt to find a linear combination of these two that does so factorize, you will fail: see problem 15.17. The ground state wavefunction cannot be made to factor into a space part times a spin part.

---

[14] See the definition on page **??** and problem 6.**??** on page **??**.

Problems

15.17 **A doomed attempt** (essential problem)
Any linear combination of state (15.54) with state (15.55) has the form

$$\begin{aligned} &|1,1,2\rangle \Big[\alpha|{\uparrow\downarrow\uparrow}\rangle + \beta|{\uparrow\downarrow\downarrow}\rangle\Big] \\ -&|1,2,1\rangle \Big[\alpha|{\uparrow\uparrow\downarrow}\rangle + \beta|{\uparrow\downarrow\downarrow}\rangle\Big] \\ +&|2,1,1\rangle \Big[\alpha|{\uparrow\uparrow\downarrow}\rangle + \beta|{\downarrow\uparrow\downarrow}\rangle\Big] \\ -&|2,1,1\rangle \Big[\alpha|{\uparrow\downarrow\uparrow}\rangle + \beta|{\downarrow\downarrow\uparrow}\rangle\Big] \\ +&|1,2,1\rangle \Big[\alpha|{\downarrow\uparrow\uparrow}\rangle + \beta|{\downarrow\downarrow\uparrow}\rangle\Big] \\ -&|1,1,2\rangle \Big[\alpha|{\downarrow\uparrow\uparrow}\rangle + \beta|{\downarrow\uparrow\downarrow}\rangle\Big]. \end{aligned} \tag{15.56}$$

Show that this form can never be factorized into a space part times a spin part.

15.18 **Two-electron ions**
Apply the techniques of Griffiths, section 7.2, "Ground State of Helium," to the H^- and Li^+ ions. Each of these ions has two electrons, like helium, but nuclear charges $Z = 1$ and $Z = 3$, respectively. For each ion find the effective (partially shielded) nuclear charge and determine the best upper bound on the ground state energy.

15.19 **The meaning of two-particle wavefunctions** (Old)

a. The wavefunction $\psi(x_A, x_B)$ describes two non-identical particles in one dimension. Does

$$\int_{-\infty}^{\infty} dx_A \int_{-\infty}^{\infty} dx_B \; |\psi(x_A, x_B)|^2 \tag{15.57}$$

equal one (the usual normalization) or two (the number of particles)? Write integral expressions for:

   i. The probability of finding particle A between x_1 and x_2 and particle B between x_3 and x_4.
   ii. The probability of finding particle A between x_1 and x_2, regardless of where particle B is.

b. The wavefunction $\psi(x_A, x_B)$ describes two identical particles in one dimension. Does

$$\int_{-\infty}^{\infty} dx_A \int_{-\infty}^{\infty} dx_B \; |\psi(x_A, x_B)|^2 \qquad (15.58)$$

equal one or two? Assuming that $x_1 < x_2 < x_3 < x_4$, write integral expressions for:

   i. The probability of finding one particle between x_1 and x_2 and the other between x_3 and x_4.
   ii. The probability of finding a particle between x_1 and x_2.

c. Look up the definition of "configuration space" in a classical mechanics book. Does the wavefunction inhabit configuration space or conventional three-dimensional position space? For discussion: Does your answer have any bearing upon the question of whether the wavefunction is "physically real" or a "mathematical convenience"? Does it affect your thoughts concerning measurement and the "collapse of the wavepacket"?

15.20 **Symmetrization and antisymmetrization (mathematical)** (Old)

a. Show that any two-variable function can be written as the sum of a symmetric function and an antisymmetric function.
b. Show that this is *not* true for functions of three variables. (*Clue:* Try the counterexample $f(x, y, z) = g(x)$.)
c. There is a function of three variables that is:
   i. Antisymmetric under interchange of the first and second variables: $f(x, y, z) = -f(y, x, z)$.
   ii. Symmetric under interchange of the second and third variables: $f(x, y, z) = f(x, z, y)$.
   iii. Symmetric under interchange of the first and third variables: $f(x, y, z) = f(z, y, x)$.

   Find this function and show that it is unique.

15.21 **Questions** (recommended problem)
Update your list of quantum mechanics questions that you started at problem 1.13 on page 56. Write down new questions and, if you have uncovered answers to any of your old questions, write them down briefly.

Chapter 16

A First Look at Helium

Helium: two electrons and one nucleus. The three-body problem! But wait, the three-body problem hasn't been solved exactly even in classical mechanics, there's no hope for an exact solution in quantum mechanics. Does this mean we give up? No. If you give up on a problem you can't solve exactly, you give up on life.[1] Instead, we look for approximate solutions.

If we take account of the Coulomb forces, but ignore things like the finite size of the nucleus, nuclear motion, relativistic motion of the electron, spin-orbit effects, and so forth, the Hamiltonian for two electrons and one nucleus is

$$\begin{aligned}\hat{H} &\doteq \left[-\frac{\hbar^2}{2m_e}\nabla_A^2 - \frac{2e^2}{4\pi\epsilon_0}\frac{1}{r_A}\right] + \left[-\frac{\hbar^2}{2m_e}\nabla_B^2 - \frac{2e^2}{4\pi\epsilon_0}\frac{1}{r_B}\right] + \left[\frac{e^2}{4\pi\epsilon_0}\frac{1}{|\vec{r}_A - \vec{r}_B|}\right] \\ &= \qquad \underbrace{\widehat{\mathrm{KE}}_A + \hat{U}_{nA}}_{\equiv \hat{H}_A} \qquad + \qquad \underbrace{\widehat{\mathrm{KE}}_B + \hat{U}_{nB}}_{\equiv \hat{H}_B} \qquad + \qquad \hat{U}_{AB}\end{aligned}$$

Recall that in using the subscripts "A" and "B" we are not labeling the *electrons* as "electron A" and "electron B": the electrons are identical and can't be labeled. Instead we are labeling the *points in space* where an electron might exist as "point A" and "point B".

We look for eigenstates of the partial Hamiltonian $\hat{H}_A + \hat{H}_B$. These are not eigenstates of the full Hamiltonian, but they *are* a basis, and they can be used as a place to start.

---

[1] Can't find the exact perfect apartment to rent? Can't find the exact perfect candidate to vote for? Can't find the exact perfect friend? Of course you can't find any of these things. But we get on with our lives accepting imperfections because we realize that the alternatives (homelessness, political corruption, friendlessness) are worse.

One-particle levels

We begin by finding the one-particle levels for the Hamiltonian $\hat{H}_A$ alone. We combine these with levels for $\hat{H}_B$ alone, and antisymmetrize the result.

The problem $\hat{H}_A$ is just the Hydrogen atom Coulomb problem with two changes: First, the nuclear mass is $4m_p$ instead of m_p. At our level of approximation ("ignore nuclear motion") this has no effect. Second, the nuclear charge is $2e$ instead of e. Remembering that the Rydberg energy is

$$\text{Ry} = \frac{m_e}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2,$$

this change means that the energy eigenvalues for $\hat{H}_A$ are

$$E^{(A)}_{n_A} = -\frac{4\,\text{Ry}}{n_A^2} \qquad \text{where } n_A = 1, 2, 3, \ldots.$$

Similarly, the energy eigenstates for $\hat{H}_A$ are represented by familiar functions like

$$\eta_{n\ell m}(\vec{r})|\!\uparrow\rangle \qquad \text{or} \qquad \eta_{n\ell m}(\vec{r})\chi_+.$$

Soon we will need to keep track of $\hat{H}_A$ versus $\hat{H}_B$. A notation like $\eta_{n\ell m}(\vec{r}_A)|\!\uparrow\rangle$ is fine for the space part of the eigenstate, but leaves the spin part ambiguous. We will instead use notation like

$$\eta_{n\ell m}(A)\chi_+(A)$$

to mean the same thing.

⟦Notice that the eigenstates don't *have* to take on the factorized form of "space part"×"spin part" — for example

$$\tfrac{1}{\sqrt{2}}[\eta_{200}(\vec{r})\chi_+ + \eta_{210}(\vec{r})\chi_-]$$

is a perfectly good eigenstate — but that the factorized form is particularly convenient for working with. (If we were to consider spin-orbit coupling, then the eigenstates could *not* take the factorized form.)⟧

Antisymmetrization

This is the situation of section 15.8, "Spin plus space, two electrons". You will remember from that section that a pair of position levels come together through the antisymmetrization process to form a singlet and a triplet as in equations (15.47) through (15.50).

The ground state

The ground levels of $\hat{H}_A$ and of $\hat{H}_B$ are both doubly degenerate due to spin. So if you had distinguishable particles, the ground state of $\hat{H}_A + \hat{H}_B$ would be four-fold degenerate:

$$\begin{array}{c}\text{distinguishable}\\ \eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_+(B)\\ \eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_-(B)\\ \eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_+(B)\\ \eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_-(B)\end{array}$$

But if you have identical fermions, the triplet (equations 15.47 through 15.49) vanishes and the singlet (equation 15.50) becomes (see problem 15.12)

$$\eta_{100}(A)\eta_{100}(B)\tfrac{1}{\sqrt{2}}[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)]. \tag{16.1}$$

Hence the Hamiltonian $\hat{H}_A + \hat{H}_B$ has a non-degenerate ground state.

It's common to hear things like "In the ground state of Helium, one electron is in one-body level $|100\rangle$ with spin up and the other is in one-body level $|100\rangle$ with spin down." This claim is false. The equation makes it clear that "In the ground state of Helium, one electron is in one-body level $|100\rangle$, the other is in one-body level $|100\rangle$, and the spins are not in a product state." If the first phrase *were* correct, then you *would* be able to distinguish the two electrons, and they *would not* be identical. But it's not correct.

States built from one ground level

Now build a state by combining the ground level of one Hamiltonian with $|n\ell m\rangle$ from the other. If you had distinguishable particles, this "combination" means a simple multiplication, and there would be eight states (all with the same energy):

distinguishable

$$\begin{gathered}
\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_+(B)\\
\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_-(B)\\
\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_+(B)\\
\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_-(B)\\
\eta_{n\ell m}(A)\chi_+(A)\eta_{100}(B)\chi_+(B)\\
\eta_{n\ell m}(A)\chi_+(A)\eta_{100}(B)\chi_-(B)\\
\eta_{n\ell m}(A)\chi_-(A)\eta_{100}(B)\chi_+(B)\\
\eta_{n\ell m}(A)\chi_-(A)\eta_{100}(B)\chi_-(B)
\end{gathered}$$

But if you have identical fermions, the "combination" means a multiplication followed by an antisymmetrization, and we've seen that they antisymmetrize to a triplet and a singlet

$$\begin{aligned}
&\tfrac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B)-\eta_{n\ell m}(A)\eta_{100}(B)]\chi_+(A)\chi_+(B)\\
&\tfrac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B)-\eta_{n\ell m}(A)\eta_{100}(B)]\tfrac{1}{\sqrt{2}}[\chi_+(A)\chi_-(B)+\chi_-(A)\chi_+(B)]\\
&\tfrac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B)-\eta_{n\ell m}(A)\eta_{100}(B)]\chi_-(A)\chi_-(B)\\
&\tfrac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B)+\eta_{n\ell m}(A)\eta_{100}(B)]\tfrac{1}{\sqrt{2}}[\chi_+(A)\chi_-(B)-\chi_-(A)\chi_+(B)].
\end{aligned}$$

The first three basis states are called a "triplet" (with "space antisymmetric, spin symmetric"). The last basis state is called a "singlet" (with "space symmetric, spin antisymmetric"). This particular basis has three nice properties: (1) Every member of the basis factorizes into a spatial part times a spin part. (2) Every member of the basis factorizes into a symmetric part times an antisymmetric part. (3) All three members of the triplet have identical spatial parts.

The third point means that when we take account of electron-electron repulsion through perturbation theory, we will necessarily find that all three members of any triplet remain degenerate even when the effects of the sub-Hamiltonian $\hat{U}_{AB}$ are considered.

States built from two excited levels

What happens if we carry out the above process but combining an excited level of one sub-Hamiltonian (say $\eta_{200}(A)$) with an arbitrary level of the other sub-Hamiltonian (say $\eta_{n\ell m}(B)$)?

The process goes on in a straightforward way, but it turns out that the resulting eigenenergies are are always so high that the atom is unstable: it decays rapidly to a positive helium atom plus an ejected electron. Such electrons are called “Auger electrons” (pronounced “oh-jey” because Pierre Victor Auger was French) and Auger electron spectroscopy is an important analytical technique in surface and materials science.

Strange names

So all stable energy states for Helium are built from a ground level (1s) plus another level. If the other level is itself a 1s level, then the two levels come together and then antisymmetrize to the singlet (16.1). This basis member is given the name 1^1S, pronounced “one singlet S”, after the “other level” 1s.

If the other level is anything else, say a 3p level, then the ground level plus the other level come together and then antisymmetrize to a triplet plus a singlet as show on page 408. The singlet is called 3^1P (“three singlet P”) and the triplet is called 3^3P (“three triplet P”).

16.22 **Electron-electron repulsion**
The electron-electron repulsion term $\hat{U}_{AB}$ is defined in the equation on page 405. Write down expressions for the mean value of $\hat{U}_{AB}$ in the states 2^1S, 2^3S, and 2^1P. (That is, set up the integrals in terms of the levels $\eta_{n\ell m}(\vec{r})$. Do *not* evaluate the integrals.) *Bonus:* Argue that the mean value for 2^1P is greater than the mean value for 2^1S.

Chapter 17

Breather

Why do we need a breather at this point?

There are no new principles, but lots of applications. The applications will shed light on the principles and the principles will shed light on the applications. I will not attempt to fool you: the applications will be hard. For example, the three-body problem has not been solved in classical mechanics. In the richer, more intricate, world of quantum mechanics, we will not solve it either.

You know from solving problems in classical mechanics that you should think first, before plunging into a hard problem. You know, for example, that if you use the appropriate variables, select the most appropriate coordinate system, or use a symmetry – that you can save untold amounts of labor. (See, for example, George Pólya, *How to Solve it* (Doubleday, Garden City, NY, 1957). Sanjoy Mahajan, *Street-Fighting Mathematics* (MIT Press, Cambridge, MA, 2010).) This rule holds even more so in the more complex world of quantum mechanics.

And that's the role of this chapter. We'll take a breather, pull back from the details, and organize ourselves for facing the difficult problems that lie before us.

Henry David Thoreau, *Walden* (1854): "I went to the woods because I wished to live deliberately, to front only the essential facts of life, and see if I could not learn what it had to teach, and not, when I came to die, discover that I had not lived."

17.1 What's ahead?

At this point, we have encountered all the principles of non-relativistic quantum mechanics.

That doesn't mean we have no more to do. Applying these known principles to various systems not only gives practical results (such as the laser), it also tests and strengthens our understanding of the principles. Where would you like to go with our newfound knowledge?

One obvious direction is a better understanding of atoms. Start with hydrogen. We have an exact solution for the Coulomb problem, but as we've already mentioned (section 14.8, "Hydrogen atom fine structure") the Coulomb problem is not a perfect model for a physical hydrogen atom. Plus we need to understand what happens when hydrogen is placed in an external electric or magnetic field. (The "Stark effect" or "Zeeman effect", respectively. The latter is more easily implemented in the laboratory and was historically important in the development of quantum mechanics.) We need to understand "collisions", when a hydrogen atom gets close to an electron, or a proton, or another hydrogen atom ("scattering theory"). Most important, we need to understand how a hydrogen atom emits and absorbs light or any other form of electromagnetic radiation.

We can continue with helium. Our model was decidedly crude, in that the two electrons were assumed to attract the nucleus but not to repel each other! We need a better model for helium, and once we have it we'll want to understand the fine structure of helium, the Stark and Zeeman effects in helium, scattering theory with helium, and the emission and absorption of light by helium.

The obvious path is to larger and larger atoms. Then molecules. We can start with simple molecules like diatomic hydrogen, but we'll need to build up to water, and benzine, and hydrocarbon polymers, and proteins and DNA. For all these systems, after we have a basic understanding we might want to move on to the fine structure, the Stark and Zeeman effects, collisions, and the interaction with radiation.

Finally, for the ultimate in complexity, we could explore membranes, solids (both crystalline and amorphous), and liquids. We will encounter, and need to explain, everyday phenomena like the hardness and shininess and electrical conductivity of metals, but also exotic phenomena like superconductivity, superfluidity, and phase transitions.

Along the way, we could investigate quantum information processing.

Naturally, as we move on to these more complex systems, we will need more powerful mathematical tools. To perturbation theory we will need to add the variational method, the Hartree-Fock mean field approximation for multi-electronic atoms, perturbation theory for the time evolution problem, density functional theory,[1] and more.

Or perhaps you want to explore in the opposite direction: instead of bigger and bigger things, you might want to investigate smaller and smaller things. Moving down from atomic-size systems we could examine first atomic nuclei, then the constituents of nuclei like neutrons and protons, then the constituents of neutrons and protons like quarks and gluons. Very early on in this path we will realize that relativistic effects are important and we will have to not just apply non-relativistic quantum mechanics, but develop a new relativistically correct version of quantum mechanics. Because relativistic particles must interact through fields rather than through instantaneous "action at a distance", the relativistically correct quantum mechanics we develop will necessarily be a quantum field theory.

In this path, also, we will need to develop competent mathematical tools such as diagrammatic perturbation theory and the renormalization group.

But there are explorations to perform even if we remain in the domain of a single non-relativistic particle ambivating in one dimension subject to a static potential energy function: questions such as the classical limit, and chaotic behavior. Here again new mathematical tools are required, including the WKB (or quasiclassical) approximation.

One thing is certain: If you choose to continue in quantum mechanics, your life will never be boring.

17.2 Scaled variables

Here's the energy eigenproblem for the hydrogen atom (at the level of approximation ignoring collisions, radiation, nuclear mass, nuclear size, spin,

---

[1] When the 1998 Nobel Prize in Chemistry was awarded to the physicist Kalter Kohn and the mathematician John Pople for their development of computational techniques in quantum mechanics, I heard some chemists grumble that chemistry Nobel laureates should have taken at least *one* undergraduate chemistry course.

magnetic effects, relativity, and the quantum character of the electromagnetic field):

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)-\frac{e^2}{4\pi\epsilon_0}\frac{1}{r}\right]\eta(\vec{r})=E\eta(\vec{r}). \tag{17.1}$$

This section uses dimensional analysis to find the characteristic length and characteristic energy for this problem, then uses scaled variables to express this equation in a more natural and more easily-worked-with form.

Whatever result comes out of this energy eigenequation, whether the result be a length, or an energy, or anything else, the result can only depend on three parameters: $\hbar$, m, and $e^2/4\pi\epsilon_0$. These parameters have the following dimensions:

	dimensions	base dimensions (mass, length, time)
$\hbar$	[Energy×T]	[ML^2/T]
m	[M]	[M]
$e^2/4\pi\epsilon_0$	[Energy×L]	[ML^3/T^2]

How can we build a quantity with the dimensions of length from these three parameters? Well, the quantity will have to involve $\hbar$ and $e^2/4\pi\epsilon_0$, because these are the only parameters that include the dimensions of length, but we'll have to get rid those dimensions of time. We can do that by squaring the first and dividing by the third:

quantity	dimensions
$\dfrac{\hbar^2}{e^2/4\pi\epsilon_0}$	[ML]

And now there's only one way to get rid of the dimension of mass (without reintroducing a dimension of time), namely dividing this quantity by m:

quantity	dimensions
$\dfrac{\hbar^2}{m\,e^2/4\pi\epsilon_0}$	[L]

We have uncovered the one and only way to combine these three parameters to produce a quantity with the dimensions of length. We define the *Bohr radius*

$$a_0 \equiv \frac{\hbar^2}{m\,e^2/4\pi\epsilon_0} \approx 0.05 \text{ nm}. \tag{17.2}$$

This quantity sets the typical scale for any length in a hydrogen atom. For example, if I ask for the mean distance from the nucleus to an electron in energy eigenstate $\eta_{5,4,-3}(\vec{r})$ the answer will be some pure (dimensionless) number times a_0. If I ask for the uncertainty in $\hat{x}$ of an electron in state $\eta_{2,1,0}(\vec{r})$ the answer will be some pure number times a_0.

Is there a characteristic energy? Yes, it is given through $e^2/4\pi\epsilon_0$ divided by a_0. The characteristic energy is

$$\mathcal{E}_0 \equiv \frac{m\,(e^2/4\pi\epsilon_0)^2}{\hbar^2} = 2\text{Ry}. \tag{17.3}$$

This characteristic energy doesn't have its own name, because we just call it twice the Rydberg energy (the minimum energy required to ionize a hydrogen atom). It plays the same role for energies that a_0 plays for lengths: Any energy value concerning hydrogen will be a pure number times $\mathcal{E}_0$.

Now is the time to introduce scaled variables. Whenever I specify a length, I specify that length in terms of some other length. For example, when I say the Eiffel tower is 324 meters tall, I mean that the ratio of the height of the Eiffel tower to the length of the prototype meter bar — that bar stored in a vault in Sèvres, France — is 324.

Now, what is the relevance of the prototype meter bar to atomic phenomena? None! Instead of measuring atomic lengths relative to the prototype meter, it makes more sense to measure them relative to something atomic, namely to the Bohr radius. I define the dimensionless "scaled length" $\tilde{x}$ as

$$\tilde{x} \equiv \frac{x}{a_0}, \tag{17.4}$$

and it's my preference to measure atomic lengths using this standard, rather than using the prototype meter bar as a standard.

So, what is the energy eigenproblem (17.1) written in terms of scaled lengths? For any function $f(x)$, the chain rule of calculus tells us that

$$\frac{\partial f(x)}{\partial x} = \frac{\partial f(\tilde{x})}{\partial \tilde{x}}\frac{\partial \tilde{x}}{\partial x} = \frac{\partial f(\tilde{x})}{\partial \tilde{x}}\frac{1}{a_0}$$

and consequently that

$$\frac{\partial^2 f(x)}{\partial x^2} = \frac{\partial^2 f(\tilde{x})}{\partial \tilde{x}^2}\frac{1}{a_0^2}.$$

Consequently the energy eigenproblem (17.1) is

$$\left[-\frac{\hbar^2}{2m}\frac{1}{a_0^2}\left(\frac{\partial^2}{\partial \tilde{x}^2} + \frac{\partial^2}{\partial \tilde{y}^2} + \frac{\partial^2}{\partial \tilde{z}^2}\right) - \frac{e^2}{4\pi\epsilon_0}\frac{1}{a_0\tilde{r}}\right]\eta(\vec{\tilde{r}}) = E\eta(\vec{\tilde{r}}), \tag{17.5}$$

which seems like a nightmare, until you realize that

$$\frac{\hbar^2}{m}\frac{1}{a_0^2} = \frac{e^2}{4\pi\epsilon_0}\frac{1}{a_0} = \mathcal{E}_0.$$

The eigenproblem (17.5) is thus

$$\left[-\frac{1}{2}\left(\frac{\partial^2}{\partial\tilde{x}^2}+\frac{\partial^2}{\partial\tilde{y}^2}+\frac{\partial^2}{\partial\tilde{z}^2}\right)-\frac{1}{\tilde{r}}\right]\eta(\vec{\tilde{r}}) = \frac{E}{\mathcal{E}_0}\eta(\vec{\tilde{r}}). \tag{17.6}$$

Defining the dimensionless "scaled energy"

$$\tilde{E} \equiv \frac{E}{\mathcal{E}_0}, \tag{17.7}$$

we see immediately that the energy eigenproblem, expressed in scaled variables, is

$$\left[-\frac{1}{2}\left(\frac{\partial^2}{\partial\tilde{x}^2}+\frac{\partial^2}{\partial\tilde{y}^2}+\frac{\partial^2}{\partial\tilde{z}^2}\right)-\frac{1}{\tilde{r}}\right]\eta(\vec{\tilde{r}}) = \tilde{E}\eta(\vec{\tilde{r}}) \tag{17.8}$$

or

$$\left[-\frac{1}{2}\tilde{\nabla}^2-\frac{1}{\tilde{r}}\right]\eta(\vec{\tilde{r}}) = \tilde{E}\eta(\vec{\tilde{r}}). \tag{17.9}$$

Whoa! It's considerably easier to work with the energy eigenproblem written in this form than it is to work with form (17.1) — there are no $\hbar$s and $e^2/4\pi\epsilon_0$s to keep track of (and to lose through algebra errors).

The only problem is that there are so many tildes to write down. People get tired of writing tildes, so they just omit them, with the understanding that they are now working with scaled variables rather than traditional variables, and the energy eigenproblem becomes

$$\left[-\frac{1}{2}\nabla^2-\frac{1}{r}\right]\eta(\vec{r}) = E\eta(\vec{r}). \tag{17.10}$$

I like to call this process "using scaled variables". Others call it "measuring length and energy in atomic units". Still others say that we get equation (17.10) from (17.1) by

$$\text{"setting } \hbar = m = \frac{e^2}{4\pi\epsilon_0} = 1\text{"}.$$

This last phrase is particularly opaque, because taken literally it's absurd. So you must not take it literally: it's a code phrase for the more interesting process of converting to scaled variables and then dropping the tildes.

One last point. Some people call this system not "atomic units" but "natural units". While these units are indeed the natural system for solving problems in atomic physics, they not the natural units for solving problems in nuclear physics, or in stellar physics, or in cosmology. And they are particularly unnatural and inappropriate for measuring the heights of towers.

17.3 Variational method for the ground state energy

Imagine a gymnasium full of fruits

$$\text{smallest fruit} \;\leq\; \text{smallest cantaloupe.}$$

Similarly

$$\text{ground state energy} \;\leq \langle\psi|\hat{H}|\psi\rangle \text{ for any } |\psi\rangle.$$

So try out a bunch of states, turn the crank, find the smallest. Very mechanical.

For example, to estimate the ground state energy of a quartic oscillator $V(x) = \alpha x^4$, you could use as trial wavefunctions the Gaussians

$$\psi(x) = \frac{1}{\sqrt[4]{\pi}\sqrt{\sigma}} e^{-x^2/2\sigma^2}.$$

Turn the crank to find $\langle\psi|\hat{H}|\psi\rangle$, then minimize to find which value of σ minimizes that mean value.

Two things to remember: First, it's a mathematical technique useful in many fields, not just in quantum mechanics. Second, it seems merely mechanical, but in fact it relies on picking good trial wavefunctions: you have to gain an intuitive understanding of how the real wavefunction is going to behave, then pick trial wavefunctions capable of mimicing that behavior. In the words of Forman S. Acton[2]: "In the hands of a Feynman the [variational] technique works like a Latin charm; with ordinary mortals the result is a mixed bag."

Sample problem: Variational estimate for the ground state energy of a quartic oscillator
The trial wavefunction

$$\psi(x) = \frac{1}{\sqrt[4]{\pi}\sqrt{\sigma}} e^{-x^2/2\sigma^2}$$

---

[2] *Numerical Methods that Work* (Harper & Row, New York, 1970) page 252.

is normalized. (If you don't know this, you should verify it.) We look for

$$\begin{aligned}
\langle\psi|\hat{H}|\psi\rangle &= \frac{1}{\sqrt{\pi}\sigma}\int_{-\infty}^{+\infty} e^{-x^2/2\sigma^2}\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \alpha x^4\right] e^{-x^2/2\sigma^2}\,dx \\
&= -\frac{\hbar^2}{2m}\frac{1}{\sqrt{\pi}\sigma}\int_{-\infty}^{+\infty} e^{-x^2/2\sigma^2}\left(-\frac{1}{\sigma^2}\right)\left(1-\frac{x^2}{\sigma^2}\right) e^{-x^2/2\sigma^2}\,dx \\
&\qquad + \alpha\frac{1}{\sqrt{\pi}\sigma}\int_{-\infty}^{+\infty} e^{-x^2/2\sigma^2} x^4 e^{-x^2/2\sigma^2}\,dx \\
&= \frac{\hbar^2}{2m}\frac{1}{\sqrt{\pi}\sigma^3}\int_{-\infty}^{+\infty}\left(1-\frac{x^2}{\sigma^2}\right) e^{-x^2/\sigma^2}\,dx \\
&\qquad + \alpha\frac{1}{\sqrt{\pi}\sigma}\int_{-\infty}^{+\infty} x^4 e^{-x^2/\sigma^2}\,dx \\
&= \frac{\hbar^2}{2m}\frac{1}{\sqrt{\pi}\sigma^2}\int_{-\infty}^{+\infty}\left(1-\tilde{x}^2\right) e^{-\tilde{x}^2}\,d\tilde{x} + \alpha\frac{\sigma^4}{\sqrt{\pi}}\int_{-\infty}^{+\infty}\tilde{x}^4 e^{-\tilde{x}^2}\,d\tilde{x}.
\end{aligned}$$

Already, even before evaluating the integrals, we can see that both integrals are numbers independent of the trial wavefunction width σ. Thus the expected kinetic energy, on the left, decreases with σ while the expected potential energy, on the right, increases with σ. Does this make sense to you?

When you work out (or look up) the integrals, you find

$$\langle\psi|\hat{H}|\psi\rangle = \frac{\hbar^2}{2m}\frac{1}{\sqrt{\pi}\sigma^2}\left(\sqrt{\pi}-\tfrac{1}{2}\sqrt{\pi}\right) + \alpha\frac{\sigma^4}{\sqrt{\pi}}\left(\frac{3}{4}\sqrt{\pi}\right) = \frac{\hbar^2}{2m}\frac{1}{2\sigma^2} + \alpha\frac{3\sigma^4}{4}.$$

If you minimize this energy with respect to σ, you will find that the minimum value (which is, hence, the best upper bound for the ground state energy) is

$$\left(9\frac{\hbar^2}{2m}\right)^{2/3}\frac{\alpha^{1/3}}{4}.$$

Problem: Show that the width of the minimum-energy wavefunction is

$$\sigma = \left(\frac{\hbar^2}{2m}\frac{1}{3\alpha}\right)^{1/6}.$$

Added: You can use the variational technique for other states as well: for example, in one dimensional systems, the first excited state is less than or equal to $\langle\psi|\hat{H}|\psi\rangle$ for all states $|\psi\rangle$ with a single node.

17.4 Sum over paths/histories/trajectories

We have wandered so far from our original ideas about amplitude combining in series and in parallel that it's easy to "lose the forest behind the trees" and forget where we started.

17.4.1 *Action in classical mechanics*

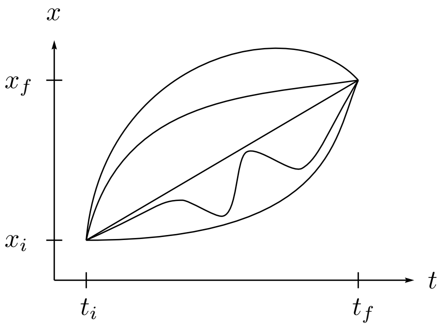
Classical mechanics can be expressed in many different formulations. The most familiar is the Newtonian formulation, encapsulated in the famous formula $\sum \vec{F} = m\vec{a}$. But there is also a Lagrangian formulation, a Hamiltonian formulation, a Poisson bracket formulation, and others. These different formulations differ dramatically in familiarity, in ease of use, in ability to extent to relativistic and field-theoretic situations, in elegance, and in philosophical "feel", but all of them result in the same answer to any given problem. So you might prefer one formulation to another, but you cannot say that one formulation is right and another is wrong — they're all right.

Perhaps the most remarkable formulations is the "principle of least action". (I express it here for one particle moving in one dimension, but it readily generalizes to higher dimensions and multiple particles.) Suppose a particle of mass m, subject to a potential energy function $V(x)$, moves from initial position x_i at time t_i to final position x_f at time t_f. What path $x(t)$ does it take? The Newtonian formulation of classical mechanics says to set up and solve the differential equation

$$m\frac{d^2x(t)}{dt^2} = -\frac{\partial V(x)}{\partial x}, \tag{17.11}$$

subject to the boundary conditions $x(t_i) = x_i$ and $x(t_f) = x_f$. This solution will be the real trajectory taken by the particle.

The principle of least action says instead to consider all the possible trajectories $x(t)$ tracing from the given initial place and time to the final place and time. One trajectory has uniform speed. Others are fast at the beginning and slow at the ending. Still others are slow at the beginning and fast at the ending. Some overshoot the final mark and have to return. Some jitter back and forth before reaching their goal. (Do not consider trajectories going backward in time.)

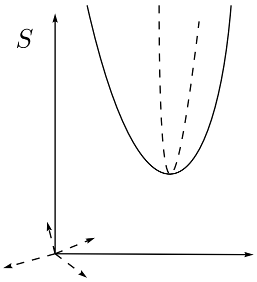


For each trajectory at every time find the kinetic energy and subtract the potential energy, then integrate that difference with respect to time. The result for any given trajectory $x(t)$ is called the "action"

$$S\{x(t)\} = \int_{t_i}^{t_f} \left[\tfrac{1}{2} m \left(\frac{dx(t)}{dt} \right)^2 - V(x(t)) \right] dt. \tag{17.12}$$

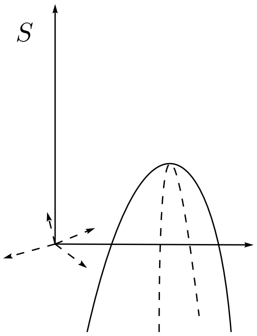
The real trajectory taken by the particle will be the one with the smallest action. Hence the name "principle of least action".

The graph below pictorializes the situation. The vertical axis represents action. The horizontal axes represent the space of various trajectories that lead from x_i at t_i to x_f at t_f. Because of the great variety of such trajectories, these are represented by one solid axis within the plane of the page and numerous dashed axes that symbolize the additional parameters that would specify various aspects of the trajectory. The real trajectory is the one that minimizes the action over all possible trajectories that move forward in time.

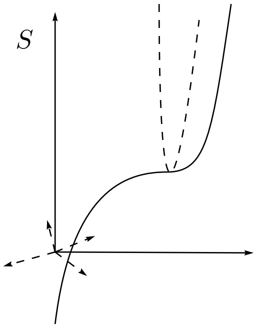


This formulation is appealing in that most of us have a good intuitive feel for minimization, because we have spent most of our lives attempting (perhaps unconsciously) to minimize cost and effort and travel time. On the other hand, this formulation has a philosophical feel like magic: How is the particle supposed to "know" the action of the paths it hasn't taken?

I have to confess now that recently, below equation (17.12), I told you a little lie. Although usually the real trajectory is the one that minimizes the action, occasionally it is the trajectory that *maximizes* the action



And very rarely the real trajectory neither minimizes nor maximizes the action, but instead lies at a point of inflection



For these reasons the "principle of least action" is more properly called the "principle of stationary action".

How can it be that minimizing action or maximizing action are both as good? Anyone running a factory attempts to minimize costs; no factory manager would ever say "minimize the costs or maximize the costs, it's all the same to me".

The resolution to these two conundrums ("How can the particle know the action of paths not taken?" and "How can maximization be just as good as minimization?") lies in quantum mechanics.

17.4.2 *Action in quantum mechanics*

The picture implicit in our "three desirable rules for amplitude" on page 60 is that we will list all possible paths from the initial to the final state, assign an amplitude to each path, and sum the amplitudes over all possible paths. The situations we considered then had two or three possible paths. The situation we consider now has an infinite number of paths, only five of which are sketched on page 420. What amplitude should be assigned to each path?

The answer turns out to be that the amplitude for trajectory $x(t)$ is

$$Ae^{iS\{x(t)\}/\hbar} \tag{17.13}$$

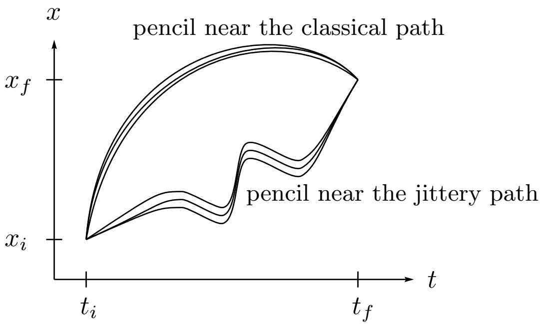
where A is a normalization constant, the same for each possible path, and $S\{x(t)\}$ is the classical action for this particular path. Our rule for combining amplitudes in parallel tells us that the amplitude to go from x_i at t_i to x_f at t_f, called the "propagator", must be

$$K(x_f, t_f; x_i, t_i) = \sum_{\text{all paths}} Ae^{iS\{x(t)\}/\hbar}. \tag{17.14}$$

Obviously, to turn this idea into a useful tool, we must first solve the technical problems of determining the normalization constant A and figuring out how to sum over an infinite number of paths ("path integration"). And once those technical problems are solved we need to prove that this formulation of quantum mechanics is correct (i.e., that it gives the same results as the Schrödinger equation). We will need to ask about what happens if the initial and final states are *not* states of definite position, but instead states of definite momentum, or arbitrary states. We will need to generalize this formulation to particles with spin. These questions are answered in R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals*, emended edition (Dover Publications, Mineola, NY, 2010). This introduction investigates only two questions.

We said on page 60 that "If an action[3] takes place through several successive stages, the amplitude for that action is the product of the amplitudes for each stage." Does equation (17.13) for path amplitude reflect this rule? It does, because a path from the initial state x_i, t_i to the final state x_f, t_f passes through some middle state x_m, t_m. Because the action from initial to final is the sum of the action from initial to middle plus the action from middle to final, the amplitude for going from initial to final is the product of the amplitude for going from initial to middle times the amplitude for going from middle to final.

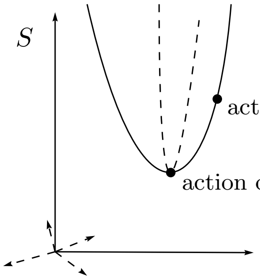
How can this "sum over histories" formulation possibly have a classical limit? Every path, from the classical path to weird jittery paths to paths that go to Mars and back, enters into the sum with the same magnitude, just with different phases. Doesn't that mean they're all equally important, and none of them will drop out in classical situations? The resolution to this conundrum comes through considering not individual paths, but small clusters of paths called pencils.



At the classical path, the "action as a function of path" graph is flat, so nearby paths have almost the same action, and hence almost the same path amplitude. When those amplitudes are summed over the pencil, the amplitudes interfere constructively and add up to a substantial sum amplitude and hence a substantial probability of travel on the pencil near the classical path.

---

[3] Don't confuse this everyday use of the word "action" with the mathematical function $S\{x(t)\}$, also called "action".



But at the jittery path, the "action as a function of path" graph is sloped, so nearby paths have quite a different action, and hence the phase of path amplitude differs dramatically from one path to another within the same pencil. The amplitudes of paths within this pencil all have the same magnitude, but they have wildly varying phases. When those amplitudes are summed over the pencil, the amplitudes interfere destructively and cancel out to a near-zero sum amplitude. Hence there is negligible probability of travel on the pencil near the jittery path.

We have seen how the classical limit emerges from the sum over histories formulation, but we've seen even more. We've seen that the key to a classical limit is having a pencil of trajectories with nearly identical actions. It doesn't care whether that pencil is a minimum, or a maximum, or a point of inflection. This is why the classical principle is not a "principle of least action" but in fact a "principle of stationary action". This is why classical mechanics seems to be saying "minimize the action or maximize the action, it's all the same to me".

And we've also seen how the classical particle can take a single path without "knowing" the actions of other paths: the quantal particle does indeed have an amplitude to take any path.

I will be the first to acknowledge that we have entered a territory that is not only unfamiliar and far from common sense, but also intricate and complex. But that complexity does *not* arise from the fundamentals of quantum mechanics, which are just the three simple rules for amplitude presented on page 60. Instead, the complexity arises from using those simple rules over and over so that the simple rules generate complex and, frankly, fantastic situations. Quantum mechanics is like the game of chess, where simple rules are applied over and over again to produce a complex and subtle game.

17.5 Problems

17.1 Quantal recurrence in the infinite square well

a. Find the period as a function of energy for a classical particle of mass m in an infinite square well of width L.

b. Show that any wavefunction, regardless of energy, in the same infinite square well is periodic in time with a period

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

(This part can be solved knowing only the energy eigenvalues.)

c. What happens after one-half of this time has passed? (This part requires some knowledge of the energy eigenfunctions.)

d. For a macroscopic particle of mass $m = 1$ kg moving in a macroscopic square well of length $L = 1$ m, what is the numerical revival time in seconds? Compare to the age of the universe. What's going on? (*Clue:* The classical period arrives when the position $x(t)$ and momentum $p(t)$ come back to their original values. The quantal revival time arrives when the mean position $\langle \hat{x} \rangle_t$ and mean momentum $\langle \hat{p} \rangle_t$ and the indeterminacy in position $(\Delta x)_t$ and the indeterminacy in momentum $(\Delta p)_t$ and the position-momentum correlation $\langle \hat{x}\hat{p} \rangle_t$ and indeed the entire wavefunction come back to their original values.)

[*Note:* This problem raises deep questions about the character of quantum mechanics and of its classical limit. See D.F. Styer, "Quantum revivals versus classical periodicity in the infinite square well," *American Journal of Physics* **69** (January 2001) 56–62.]

17.2 Quantal recurrence in the Coulomb problem

Show that in the Coulomb problem, any quantal state consisting of a superposition of two or more bound energy eigenstates with principal quantal numbers n_1, n_2, …, n_r evolves in time with a period of

$$\frac{h}{\text{Ry}} N^2,$$

where Ry is the Rydberg energy and the integer N is the least common multiple of n_1, n_2, …, n_r.

17.3 Atomic units

The Schrödinger equation for the Coulomb problem is

$$i\hbar\frac{\partial\Psi(x,y,z,t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi(x,y,z,t) - \frac{e^2}{4\pi\epsilon_0}\frac{1}{r}\Psi(x,y,z,t).$$

It is clear that the answer to any physical problem can depend only on the three parameters $\hbar$, m, and $e^2/4\pi\epsilon_0$. In section 17.2, we used these ideas to show that any problem that asked for a length had to have an answer which was a dimensionless number times the characteristic length, the so-called Bohr radius

$$a_0 = \frac{4\pi\epsilon_0}{e^2}\frac{\hbar^2}{m}.$$

a. Show that there is only one characteristic energy, i.e. only one way to combine the three parameters to produce a quantity with the dimensions of energy. (Section 17.2 found *one* way to perform this combination, but I want you to prove that this is the *only* way. *Clue:* Instead of the conventional base dimensions of length, mass, and time, use the unconventional base dimensions of length, mass, and energy.)

b. Find the characteristic time τ_0. What is its numerical value in terms of femtoseconds?

c. *Bonus:* Show that, in the Bohr model, the period of the innermost orbit is $2\pi\tau_0$. What is the period of the nth orbit?

d. Estimate the number of heartbeats made in a lifetime by a typical person. If each Bohr model orbit corresponds to a heartbeat, how many "lifetimes of hydrogen" pass in a second?

e. Write the time-dependent Schrödinger equation in terms of the scaled variables

$$\tilde{r} = \frac{r}{a_0} \qquad \text{“lengths measured in atomic units”}$$

and

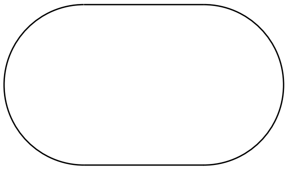
$$\tilde{t} = \frac{t}{\tau_0} \qquad \text{“time measured in atomic units”.}$$

Be sure to use the dimensionless wavefunction

$$\widetilde{\Psi}(\tilde{x},\tilde{y},\tilde{z},\tilde{t}) = (a_0)^{3/2}\Psi(x,y,z,t).$$

17.4 **Scaling in the stadium problem**

The "stadium" problem is often used as a model chaotic system, in both classical and quantum mechanics. [See E.J. Heller, "Bound-State Eigenfunctions of Classically Chaotic Hamiltonian Systems: Scars of Periodic Orbits" *Phys. Rev. Lett.*, **53**, 1515–1518 (1984); S. Tomsovic and E.J. Heller, "Long-Time Semiclassical Dynamics of Chaos: The Stadium Billiard" *Phys. Rev.* E, **47**, 282–299 (1993); E.J. Heller and S. Tomsovic, "Postmodern Quantum Mechanics" *Physics Today*, **46** (7), 38–46 (July 1993).] This is a two-dimensional infinite well shaped as a rectangle with semi-circular caps on opposite ends. Suppose one stadium has the same shape but is exactly three times as large as another. Show that in the larger stadium, wavepackets move just as they do in the smaller stadium, but nine times more slowly. (The initial wavepacket is of course also enlarged three times.) And show that the energy eigenvalues of the larger stadium are one-ninth the energy eigenvalues of the smaller stadium.



17.5 **Variational principle for the harmonic oscillator**

Find the best bound on the ground state energy of the one-dimensional harmonic oscillator using a trial wavefunction of form

$$\psi(x) = \frac{A}{x^2 + b^2},$$

where A is determined through normalization and b is an adjustable parameter. (*Clue:* Put the integrals within $\langle H \rangle$ into dimensionless form so that they are independent of A and b, and are "just numbers": call them C_K and C_P. Solve the problem in terms of these numbers, then evaluate the integrals only at the end.)

17.6 **Solving the Coulomb problem through operator factorization**

Griffiths (section 4.2) finds the bound state energy eigenvalues for the Coulomb problem using power series solutions of the Schrödinger equation. Here is another way, based on operator factorization (ladder operators). In atomic units, the radial wave equation is

$$\left[-\frac{1}{2}\frac{d^2}{dr^2}+\frac{\ell(\ell+1)}{2r^2}-\frac{1}{r}\right]u_{n,\ell}(r)\equiv h_\ell u_{n,\ell}(r)=\epsilon_{n,\ell}u_{n,\ell}(r)$$

where $u_{n,\ell}(r)$ is r times the radial wavefunction. Introduce the operators

$$D_\pm^{(\ell)}\equiv\frac{d}{dr}\mp\frac{\ell}{r}\pm\frac{1}{\ell}.$$

a. Show that

$$D_+^{(\ell)}D_-^{(\ell)}=-2h_\ell-\frac{1}{\ell^2}.$$

and that

$$D_-^{(\ell+1)}D_+^{(\ell+1)}=-2h_\ell-\frac{1}{(\ell+1)^2}$$

b. Conclude that

$$h_{\ell+1}D_+^{(\ell+1)}=D_+^{(\ell+1)}h_\ell,$$

and apply this operator equation to $u_{n,\ell}(r)$ to show that

$$D_+^{(\ell+1)}u_{n,\ell}(r)\propto u_{n,\ell+1}(r)$$

and that $\epsilon_{n,\ell}$ is independent of ℓ.

c. Argue that for every $\epsilon_{n,\ell}<0$ there is a maximum ℓ. (*Clue:* Examine the effective potential for radial motion.) Call this ℓ value ℓ_n.

d. Define $n=\ell_n+1$ and show that

$$\epsilon_{n,\ell}=-\frac{1}{2n^2}\quad\text{where}\quad\ell=0,\ldots,n-1.$$

(One can also continue this game to find the energy eigenfunctions.)

Chapter 18

Hydrogen

Recall the structure of states summarized in section 14.7.

18.1 The Stark effect

The unpeturbed Hamiltonian, as represented in the position basis, is

$$\hat{H}^{(0)} \doteq -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0}\frac{1}{r}. \tag{18.1}$$

An electric field of magnitude E is applied, and we name the direction of the electric field the z direction. The perturbing Hamiltonian, again represented in the position basis, is

$$\hat{H}' \doteq eEz = eEr\cos\theta. \tag{18.2}$$

Perturbation theory for the energy eigenvalues tells us that, provided the unperturbed energy state $|n^{(0)}\rangle$ is non-degenerate,

$$E_n = E_n^{(0)} + \langle n^{(0)}|\hat{H}'|n^{(0)}\rangle + \sum_{m\neq n}\frac{|\langle m^{(0)}|\hat{H}'|n^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}} + \cdots . \tag{18.3}$$

Let us apply perturbation theory to the ground state $|n,\ell,m\rangle = |1,0,0\rangle$. This state is non-degenerate, so equation (18.3) applies without question. A moment's thought will convince you that $\langle 1,0,0|\hat{H}'|1,0,0\rangle =$

$eE\langle 1,0,0|\hat{z}|1,0,0\rangle = 0$, so the result is

$$\begin{aligned} E_1 &= E_1^{(0)} + \sum_{n=2}^{\infty}\sum_{\ell=0}^{n-1}\sum_{m=-\ell}^{+\ell} \frac{|\langle n,\ell,m|\hat{H}'|1,0,0\rangle|^2}{E_1^{(0)} - E_n^{(0)}} + \cdots \\ &= -\mathrm{Ry} + \sum_{n=2}^{\infty}\sum_{\ell=0}^{n-1}\sum_{m=-\ell}^{+\ell} \frac{|eE\langle n,\ell,m|\hat{z}|1,0,0\rangle|^2}{-\mathrm{Ry} + \mathrm{Ry}/n^2} + \cdots \\ &= -\mathrm{Ry} - \frac{e^2E^2}{\mathrm{Ry}} \sum_{n=2}^{\infty}\sum_{\ell=0}^{n-1}\sum_{m=-\ell}^{+\ell} \frac{|\langle n,\ell,m|\hat{z}|1,0,0\rangle|^2}{1 - 1/n^2} + \cdots . \qquad (18.4) \end{aligned}$$

It would take a lot of work to evaluate the sum here, but one thing is clear: that sum is just some quantity with the dimensions [length^2], and independent of the field strength E. So when the electric field is turned on, the ground state energy decreases from the zero-field energy of $-\mathrm{Ry}$, quadratically with E. Without even evaluating the sum, we get a lot of important information.

Well, that went well. What if we apply perturbation theory to the first excited state $|2,0,0\rangle$? My first thought is that, once again $\langle 2,0,0|\hat{H}'|2,0,0\rangle = eE\langle 2,0,0|\hat{z}|2,0,0\rangle = 0$, so we'll need to go on to second-order perturbation theory, and hence we'll again find a quadratic Stark effect. The same argument holds for the excited state $|2,1,+1\rangle$, the state $|7,5,-3\rangle$ and indeed for *any* energy state.

But that quick and easy argument is wrong. In making it we've forgotten that the equation 18.3 applies only to non-degenerate energy states.[1] The first excited state is four-fold degenerate: the states $|2,0,0\rangle$, $|2,1,+1\rangle$, $|2,1,0\rangle$, and $|2,1,-1\rangle$ all have the same energy, namely $-\mathrm{Ry}/2^2$. If we were to try to evaluate the sum, we'd have to look at terms like

$$\frac{|\langle 2,1,0|\hat{H}'|2,0,0\rangle|^2}{E_{2,0,0} - E_{2,1,0}} = \frac{|\langle 2,1,0|\hat{H}'|2,0,0\rangle|^2}{0},$$

which equals infinity! In our attempt to "get a lot of important information without actually evaluating the sum" we have missed the fact that the sum diverges.

There's only one escape from this trap. We can avoid infinities by making sure that, whenever we have a zero in the denominator, we also have a zero in the numerator. (Author's note to self: Change chapter 11

---

[1] This is a favorite trick question in physics oral exams.

to show this more rigorously.) That is, we can't perform the perturbation theory expansion using the basis

$$\{|2,0,0\rangle, |2,1,+1\rangle, |2,1,0\rangle, |2,1,-1\rangle\}$$

but we *can* perform it using some new basis, a linear combination of these states, such that in this new basis the matrix elements of $\hat{H}'$ vanish *except on the diagonal.* In other words, we must diagonalize the 4×4 matrix of $\hat{H}'$, and perform the perturbation expansion using that new basis rather than the initial basis.

The process, in other words, requires three stages: First find the matrix of $\hat{H}'$, then diagonalize it, and finally perform the expansion.

Start by finding the 4×4 matrix in the initial basis. Each matrix element will have the form

$$\begin{aligned} \langle a|\hat{H}'|b\rangle &= eE\langle a|\hat{z}|b\rangle \\ &= eE\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta \int_0^{\infty} r^2\, dr\, \eta_a^*(r,\theta,\phi)\, r\cos\theta\, \eta_b(r,\theta,\phi) \end{aligned} \tag{18.5}$$

and they will be arrayed in a matrix like this:

$$\begin{array}{cc} \begin{matrix} \langle 200| & \langle 211| & \langle 210| & \langle 21\bar{1}| \end{matrix} & \\ \left[\begin{matrix} & & & \\ & & & \\ & & & \\ & & & \end{matrix}\right] & \begin{matrix} |200\rangle \\ |211\rangle \\ |210\rangle \\ |21\bar{1}\rangle \end{matrix} \end{array}$$

(Here the m value of -1 is shown as $\bar{1}$ because otherwise it messes up the spacing.)

You might think that there are 16 matrix elements to calculate, that each one is a triple integral, and that the best way to start off is by going to a bar and getting drunk. Courage! The operator is Hermitian, so the subdiagonal elements are the complex conjugates of the corresponding superdiagonal elements — there are only 10 matrix elements to calculate.

The diagonal elements are all proportional to the mean values of $\hat{z}$, and these means vanish for *any* of the traditional Coulomb problem eigenstates $|n,\ell,m\rangle$.

$$\begin{array}{cc} \begin{matrix} \langle 200| & \langle 211| & \langle 210| & \langle 21\bar{1}| \end{matrix} & \\ \left[\begin{matrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{matrix}\right] & \begin{matrix} |200\rangle \\ |211\rangle \\ |210\rangle \\ |21\bar{1}\rangle \end{matrix} \end{array}$$

Remember what the wavefunctions look like:

$$\begin{aligned}
|2,0,0\rangle &\doteq R_{2,0}(r)Y_0^0(\theta,\phi) \sim 1 \\
|2,1,+1\rangle &\doteq R_{2,1}(r)Y_1^{+1}(\theta,\phi) \sim \sin\theta\, e^{+i\phi} \\
|2,1,0\rangle &\doteq R_{2,1}(r)Y_1^{0}(\theta,\phi) \sim \cos\theta \\
|2,1,-1\rangle &\doteq R_{2,1}(r)Y_1^{-1}(\theta,\phi) \sim \sin\theta\, e^{-i\phi}
\end{aligned}$$

where $\sim$ means that I've written down the angular dependence but not the radial dependence.

The leftmost matrix element on the top row is

$$\begin{aligned}
&\langle 2,1,+1|\hat{H}'|2,0,0\rangle \\
&= eE\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta \int_0^{\infty} r^2\, dr\, R_{2,1}(r)Y_1^{+1*}(\theta,\phi)\, r\cos\theta\, R_{2,0}(r)Y_0^0(\theta,\phi).
\end{aligned}$$

There are three integrals here: r, θ, and ϕ. To do the r integral I would have to look up the expressions for $R_{2,1}(r)$ and $R_{2,0}(r)$, and then do a gnarly integral. To do the θ integral I would have to look up the spherical harmonics and then do an integral not quite so gnarly as the r integral. But to do the ϕ integral is straightforward: The function $Y_1^{+1*}(\theta,\phi)$ contributes an $e^{-i\phi}$ and that's it. The ϕ integral is

$$\int_0^{2\pi} d\phi\, e^{-i\phi}$$

and this integral is easy to do. . . it's zero.

$$\begin{array}{ccccc}
\langle 200| & \langle 211| & \langle 210| & \langle 21\bar{1}| & \\
0 & 0 & & & |200\rangle \\
0 & 0 & & & |211\rangle \\
 & & 0 & & |210\rangle \\
 & & & 0 & |21\bar{1}\rangle
\end{array}$$

It's a good thing we put off doing the difficult r and θ integrals, because if we had sweated away working them out, and then found that all we did with those hard-won results was to multiply them by zero, then we'd really need to visit that bar. When I was a child, my Protestant-work-ethic parents told me that when faced with two tasks, I should always "be a man" and do the difficult one first. I'm telling you to do the opposite, because doing the easy task might make you realize that you don't have to do the difficult one.

If you look at the two other matrix elements on the superdiagonal,

$$\langle 2,1,0|\hat{H}'|2,1,+1\rangle \quad \text{and} \quad \langle 2,1,-1|\hat{H}'|2,1,0\rangle,$$

you'll recognize instantly that for each of these two the ϕ integral is

$$\int_0^{2\pi} d\phi\, e^{+i\phi} = 0.$$

The same holds for $\langle 2,1,-1|\hat{H}'|2,0,0\rangle$, so the matrix is shaping up as

$$\begin{array}{cc} \begin{array}{cccc} \langle 200| & \langle 211| & \langle 210| & \langle 21\bar{1}| \end{array} & \\ \left[\begin{array}{cccc} 0 & 0 & & 0 \\ 0 & 0 & 0 & \\ & 0 & 0 & 0 \\ 0 & & 0 & 0 \end{array}\right] & \begin{array}{c} |200\rangle \\ |211\rangle \\ |210\rangle \\ |21\bar{1}\rangle \end{array} \end{array}$$

and we have just two more elements to calculate.

The matrix element

$$\langle 2,1,-1|\hat{H}'|2,1,1\rangle \sim \int_0^{2\pi} d\phi\, e^{+2i\phi} = 0,$$

so the only hard integral we have to do is

$$\langle 2,1,0|\hat{H}'|2,0,0\rangle = eE\langle 2,1,0|\hat{z}|2,0,0\rangle.$$

The matrix element $\langle 2,1,0|\hat{z}|2,0,0\rangle$ is a length, and any length for the Coulomb problem must turn out to be a dimensionless number times the Bohr radius

$$\langle 2,1,0|\hat{H}'|2,0,0\rangle = eE\langle 2,1,0|\hat{z}|2,0,0\rangle = eE(\text{number})a_0. \tag{18.6}$$

The only thing that remains to do is to find that dimensionless number. I ask you to do this yourself in problem 18.1 (part a). The answer is -3. Thus the matrix is

$$\begin{array}{ccc} & \begin{array}{cccc} \langle 200| & \langle 211| & \langle 210| & \langle 21\bar{1}| \end{array} & \\ -eEa_0 & \left[\begin{array}{cccc} 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] & \begin{array}{c} |200\rangle \\ |211\rangle \\ |210\rangle \\ |21\bar{1}\rangle \end{array} \end{array}$$

and we are done with the first stage of our three-stage problem.

You will be tempted to rush immediately into the problem of diagonalizing this matrix, but "fools rush in where angels fear to tread" (Alexander Pope). If you think about it for an instant, you'll realize that it will be a lot easier to do the problem if we rearrange the sequence of basis vectors so that the matrix reads

$$-eEa_0 \begin{array}{cccccc} \langle 200| & \langle 210| & \langle 211| & \langle 21\bar{1}| & & \\ \left[\begin{matrix} 0 \\ 3 \\ 0 \\ 0 \end{matrix}\right. & \begin{matrix} 3 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0 \\ 0 \end{matrix} & \left.\begin{matrix} 0 \\ 0 \\ 0 \\ 0 \end{matrix}\right] & \begin{matrix} |200\rangle \\ |210\rangle \\ |211\rangle \\ |21\bar{1}\rangle \end{matrix} \end{array}$$

Now we start the second stage, diagonalizing the matrix. First, find the eigenvalues:

$$\begin{aligned} 0 &= \det|\mathsf{M} - \lambda\mathsf{I}| \\ &= \det \begin{vmatrix} -\lambda & 3 & 0 & 0 \\ 3 & -\lambda & 0 & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & 0 & -\lambda \end{vmatrix} \\ &= -\lambda \det \begin{vmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} - 3 \det \begin{vmatrix} 3 & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} \\ &= \lambda^4 - 3^2\lambda^2 \\ &= \lambda^2(\lambda^2 - 3^2) \end{aligned}$$

Normally, it's hard to solve a quartic equation, but in this case we can just read off the four solutions:

$$\lambda = +3, -3, 0, 0.$$

The eigenvectors associated with $\lambda = 0$ and $\lambda = 0$ are clearly

$$|2, 1, +1\rangle \quad \text{and} \quad |2, 1, -1\rangle.$$

The eigenvector associated with $\lambda = 3$ will be a linear combination

$$x|2, 0, 0\rangle + y|2, 1, 0\rangle$$

where

$$\begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 3 \begin{pmatrix} x \\ y \end{pmatrix}.$$

Any $x = y$ is a solution, but I choose the normalized solution so that the eigenvector with eigenvalue 3 is

$$\tfrac{1}{\sqrt{2}} \left(|2, 0, 0\rangle + |2, 1, 0\rangle\right).$$

The parallel process for $\lambda = -3$ reveals the eigenvector

$$\tfrac{1}{\sqrt{2}}\left(-|2,0,0\rangle + |2,1,0\rangle\right).$$

⟦Why, you will ask, do I use this eigenvector rather than

$$\tfrac{1}{\sqrt{2}}\left(|2,0,0\rangle - |2,1,0\rangle\right),$$

which is also an eigenvector but which I can write down with fewer pen strokes? The answer is simple personal preference. The version I use is the same one used for geometrical vectors in a plane, and where the change of basis is a 45° rotation. This helps me remember that, even in this recondite and abstruse situation, the process of matrix diagonalization does not change the physical situation, it merely changes the basis vectors we select to help us describe the physical situation.⟧

To summarize, in the basis

$$\left\{\tfrac{1}{\sqrt{2}}\left(|2,0,0\rangle + |2,1,0\rangle\right), \tfrac{1}{\sqrt{2}}\left(-|2,0,0\rangle + |2,1,0\rangle\right), |2,1,+1\rangle, |2,1,-1\rangle\right\}$$

the matrix representation of the operator $\hat{H}'$ is

$$-eEa_0 \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

And now, for the final stage, executing perturbation theory starting from this new basis, which I'll call $\{|a\rangle, |b\rangle, |c\rangle, |d\rangle\}$. The energy value associated with $|a\rangle$ is

$$E_2 = E_2^{(0)} + \langle a|\hat{H}'|a\rangle + \sum_m \frac{|\langle a|\hat{H}'|m\rangle|^2}{E_a^{(0)} - E_m^{(0)}} + \cdots$$

The first correction we already know: it is $\langle a|\hat{H}'|a\rangle = -3eEa_0$. The second correction — the sum — contains terms like

$$\frac{|\langle a|\hat{H}'|b\rangle|^2}{E_a^{(0)} - E_b^{(0)}} = \frac{0}{0}$$

and

$$\frac{|\langle a|\hat{H}'|c\rangle|^2}{E_a^{(0)} - E_c^{(0)}} = \frac{0}{0}$$

and

$$\frac{|\langle a|\hat{H}'|1,0,0\rangle|^2}{E_a^{(0)} - E_{1,0,0}^{(0)}} = \frac{\text{something}}{-\frac{3}{4}\text{Ry}}$$

but it contains *no* terms where a number is divided by zero. I will follow the usual rule-of-thumb for perturbation theory, which is to stop at the first non-zero correction and ignore the sum altogether.

Similarly, the leading energy correction associated with $|b\rangle$ is $\langle b|\hat{H}'|b\rangle = 3eEa_0$.

The first-order corrections for $|c\rangle$ and $|d\rangle$ vanish, so these states will be subject to a quadratic Stark effect, just like the ground state. I could work them out if I really needed to, but instead I will quote and follow the age-old dictum (modified from "The Lay of the Last Minstrel" by Walter Scott):

> Breathes there the man, with soul so dead,
> Who never to himself hath said
> "To hell with it, I'm going to bed."

18.1 The Stark effect

a. Find the numerical factor in equation (18.6).

b. The "good" energy eigenstates for the $n = 2$ Stark effect — the states that one should use as unperturbed states in perturbation theory — are

$$\begin{gathered} |2,1,+1\rangle \\ |2,1,-1\rangle \\ \tfrac{1}{\sqrt{2}}(+|2,0,0\rangle + |2,1,0\rangle) \\ \tfrac{1}{\sqrt{2}}(-|2,0,0\rangle + |2,1,0\rangle) \end{gathered}$$

Find the mean position $\langle \vec{r} \rangle$ in each of these states.

c. The mean position is zero in state $|2,0,0\rangle$ and zero in state $|2,1,0\rangle$, yet it is non-zero in state $(|2,0,0\rangle + |2,1,0\rangle)/\sqrt{2}$. This might seem like a contradiction: After all, if the mean position vanishes for two probability densities, then it vanishes for the sum of the two. What great principle of quantum mechanics allows this fact to escape the curse of contradiction? (Answer in one sentence.)

d. (**Bonus.**) Describe these four states qualitatively and explain why they are the "good" states for use in the Stark effect.

e. Consider the Stark effect for the $n = 3$ states of hydrogen. There are initially nine degenerate states. Construct a 9×9 matrix representing the perturbing Hamiltonian. (*Clue:* Before actually working any integrals, use a selection rule to determine the sequence of basis members that will produce a block diagonal matrix.)

f. Find the eigenvalues and degeneracies.

18.2 **Bonus**

In the previous problem, on the Stark effect, we had to calculate a lot of matrix elements of the form

$$\int_0^\infty r^2\, R_{n,\ell}(r)\, r\, R_{n',\ell'}(r)\, dr.$$

This was possible but (to put it mildly) tedious. Can you think of some easy way to do integrals of this form? Could the operator factorization technique (problem 17.6) give us any assistance? Can you derive any inspiration from our proof of Kramers' relation (problem below)?

18.3 **Kramers' relation**

Kramers' relation states that for any energy eigenstate $\eta_{n\ell m}(\vec{r})$ of the Coulomb problem, the expected values of r^s, r^{s-1}, and r^{s-2} are related through

$$\frac{s+1}{n^2}\langle r^s\rangle - (2s+1)a_0\langle r^{s-1}\rangle + \frac{s}{4}[(2\ell+1)^2 - s^2]a_0^2\langle r^{s-2}\rangle = 0.$$

a. Prove Kramers' relation. *Clues:* Use atomic units. Start with the radial equation in form

$$u''(r) = \left[\frac{\ell(\ell+1)}{r^2} - \frac{2}{r} + \frac{1}{n^2}\right] u(r),$$

and use it to express

$$\int_0^\infty u(r) r^s u''(r)\, dr$$

in terms of $\langle r^s\rangle$, $\langle r^{s-1}\rangle$, and $\langle r^{s-2}\rangle$. Then perform that integral by parts to find an integral involving $u'(r)$ as the highest derivative. Show that

$$\int_0^\infty u(r) r^s u'(r)\, dr = -\frac{s}{2}\langle r^{s-1}\rangle$$

and that

$$\int_0^\infty u'(r) r^s u'(r)\, dr = -\frac{2}{s+1}\int_0^\infty u''(r) r^{s+1} u'(r)\, dr.$$

b. Use Kramers' relation with $s = 0$, $s = 1$, $s = 2$, and $s = 3$ to find formulas for $\langle r^{-1}\rangle$, $\langle r\rangle$, $\langle r^2\rangle$, and $\langle r^3\rangle$. Note that you could continue indefinitely to find $\langle r^s\rangle$ for any positive power.

c. However, you can't use this chain to work downward. Try it for $s = -1$, and show that you get a relation between $\langle r^{-2}\rangle$ and $\langle r^{-3}\rangle$, but not either quantity by itself.

Chapter 19

Helium

The helium problem is a "three-body problem". This problem has never been solved exactly even in classical mechanics, and it is hopeless to expect an exact solution in the richer and more intricate regime of quantum mechanics. Does this mean we should give up? Of course not. Most physics problems cannot be solved exactly, but some can be solved approximately well enough to compare theory to experiment, which is itself imperfect. (In the same way, most problems you have with your parents, or with your boy/girlfriend, cannot be solved perfectly. But they can often be solved well enough to continue your relationship.)

19.1 Ground state energy of helium

The role of theory

Jacov Ilich Frenkel (also Yakov Ilich Frenkel or Iakov Ilich Frenkel; 1894–1952) was a prolific physicist. Among other things he coined the term "phonon". In a review article on the theory of metals (quoted by M.E. Fisher in "The Nature of Critical Points", Boulder lectures, 1965) he said:

> The more complicated the system considered, the more simplified must its theoretical description be. One cannot demand that a theoretical description of a complicated atom, and all the more of a molecule or a crystal, have the same degree of accuracy as of the theory of the simplest hydrogen atom. Incidentally, such a requirement is not only impossible to fulfill but also essentially useless. . . . An exact calculation of the constants characterizing the simplest

> physical system has essential significance as a test on the correctness of the basic principles of the theory. However, once it passes this test brilliantly there is no sense in subjecting it to further tests as applied to more complicated systems. The most ideal theory cannot pass such tests, owing to the practically unsurmountable mathematical difficulties unavoidably encountered in applications to complicated systems. In this case all that is demanded of the theory is a correct interpretation of the general character of the quantities and laws pertaining to such a system. The theoretical physicist is in this respect like a cartoonist, who must depict the original, not in all details like a photographic camera, but simplify and schematize it in a way as to disclose and emphasize the most characteristic features. Photographic accuracy can and should be required only of the description of the simplest system. A good theory of complicated systems should represent only a good "caricature" of these systems, exaggerating the properties that are most difficult, and purposely ignoring all the remaining inessential properties.

Which case is the ground state of He?

1) Fundamental test of symmetrization postulate.
2) Test to see whether QM breaks down for complex systems (Anthony J. Leggett).
3) Refinements can involve new physical ideas.
4) Physical effects other than ground state energy.

Experiment

$E_g = -78.975$ eV.

Theory

(Summarizing Griffiths 5.2.1 and 7.2.) If we take account of the Coulomb forces, but ignore things like the finite size of the nucleus, nuclear motion, relativistic motion of the electron, spin-orbit effects, and so forth, the Hamiltonian for two electrons and one nucleus is

$$\hat{H} = \hat{H}_A + \hat{H}_B + \hat{U}_{AB} \tag{19.1}$$

where

$$\hat{U}_{AB} = \frac{e^2}{4\pi\epsilon_0}\frac{1}{|\vec{r}_A - \vec{r}_B|}. \tag{19.2}$$

The ground state wavefunction for H is

$$\eta_{100}(\vec{r}) = \frac{1}{\sqrt{\pi}a_0^{3/2}}e^{-r/a_0}. \tag{19.3}$$

But if the nucleus had charge $+Ze$, this would be

$$\eta_{100}(\vec{r}) = \frac{Z^{3/2}}{\sqrt{\pi}a_0^{3/2}}e^{-Zr/a_0}. \tag{19.4}$$

So the $\hat{U}_{AB} = 0$ ground state is

$$\eta_{100}(\vec{r}_A)\eta_{100}(\vec{r}_B) = \frac{Z^3}{\pi a_0^3}e^{-Z(r_A+r_B)/a_0} \quad \text{with} \quad Z = 2. \tag{19.5}$$

This state gives a ground state energy of $E_g = -8(\text{Ry}) = -109$ eV.

Turning on the electron-electron repulsion, perturbation theory finds $\langle\hat{U}_{AB}\rangle$ and jacks up E_g to -75 eV.

The variational method uses the same wavefunction as above, but considers Z not as 2 but as an adjustable parameter. Interpretation: "shielding" — expect $1 < Z_{\text{min}} < 2$. And in fact minimizing $\langle H\rangle$ with over this class of trial wavefunctions gives $Z_{\text{min}} = 1.69$ and $E_g = -77.5$ eV. (Sure enough, an overestimate.) Griffiths stops here and suggests that the rest of the work is humdrum.

Further theory

Review: A. Hibbert, *Rept. Prog. Phys.* **38** (1975) 1222–1225.

Hylleraas (1929): Trial wavefunction of form (atomic units)

$$\psi(\vec{r}_A, \vec{r}_B) = e^{-Z(r_A+r_B)}\sum c_{nlm}(Z(r_A+r_B))^n(Z(r_A-r_B))^{2l}(Z|\vec{r}_A-\vec{r}_B|)^m.$$

[I won't go into all the reasons why he picked this trial wavefunction, but… ask why only even powers $2l$.] Using Z and six terms in sum as variational parameters, he got an energy good to 2 parts in 10,000.

This is a good energy. Is there any point in doing better? Yes. Although it gives you a good energy, it gives you a poor wavefunction: Think of a $d = 2$ landscape with a hidden valley — e.g. a crater, an absolute minimum. The $d = 2$ landscape represents two variational parameters — by coincidence, the exact wavefunction has the form that you guessed. If you tried just one variational parameter, you'd be walking a line in this landscape. The line could be quite far from the valley bottom while giving very good elevation estimates for the valley bottom, because the valley is flat at the bottom. [Sketch.]

In fact, you can show that *no* wavefunction of this form, no matter how many terms you pick, can satisfy the Schrödinger Equation — even if you picked an infinite number of terms, you'd never hit the wavefunction right on!

Is there any reason to get the wavefunction right? Yes! For example if you wanted to calculate Stark or Zeeman effect, or spin-orbit, or whatever, you'd need those wavefunctions for doing perturbation theory!

Kinoshita (1959): One of the "great fiddlers of physics". Trial wavefunction of form (atomic units)

$$\psi(\vec{r}_A, \vec{r}_B) = e^{-Z(r_A+r_B)} \sum c_{nlm}(Z(r_A+r_B))^n \left(Z \frac{r_A - r_B}{|\vec{r}_A - \vec{r}_B|} \right)^{2l} \left(\frac{|\vec{r}_A - \vec{r}_B|}{r_A + r_B} \right)^m .$$

He showed that this could satisfy the Schrödinger Equation exactly if sum were infinite. Used 80 terms for accuracy 1 part in 100,000.

Pekeris (1962): A different trial wavefunction guaranteed to get the correct form when both electrons are far from nucleus. Used 1078 terms, added fine structure and hyperfine structure, got accuracy 1 part in 10^9.

Schwartz (1962): Added terms like $[Z(r_A + r_B)]^{n/2}$ …not smooth. Got better energies with 189 terms!

Frankowski and Pekeris (1966): Introduced terms like $\ln^k(Z(r_A + r_B))$ …not smooth. 246 terms, accuracy 1 part in 10^{12}.

Kato: (See Drake, page 155.) Looked at condition for two electrons close, both far from nucleus. In this case it's like H atom, wavefunction must have cusp. Allow electrons to show this cusp.

State of art: Gordon W.F. Drake, ed. *Atomic, Molecular, and Optical Physics Handbook* page 163. [Reference QC173.A827 1996]

New frontiers: experiment. S.D. Bergeson, *et al.*, "Measurement of the He ground state Lamb shift", *Phys. Rev. Lett.* **80** (1998) 3475–3478.

New frontiers: theory. S.P. Goldman, "Uncoupling correlated calculations in atomic physics: Very high accuracy and ease," *Phys. Rev.* A **57** (1998) 677–680. 8066 terms, 1 part in 10^{18}.

New frontiers: Lithium, metallic Hydrogen.

Sometimes people get the impression that variational calculations are dry and mechanical: simply add more parameters to your trial wavefunction, and your results will improve (or at least, they can't get worse). The history of the Helium ground state calculation shows how wrong this impression is. Progress is made by deep thinking about the character of the true wavefunction (What is the character when both electrons are far from the nucleus and far from each other? What is the character when both electrons are far from the nucleus and close to each other?) and then choosing trail wavefunctions that can display (or at least mimic) those characteristics of the true wavefunction.

Chapter 20

Atoms

20.1 Addition of angular momenta

We often have occasion to add angular momenta. For example, an electron might have orbital angular momentum with respect to the nucleus, but also spin angular momentum. What is the total angular momentum?

Or again, there might be two electrons in an atom, each with orbital angular momentum. What is the total orbital angular momentum of the two electrons?

Or again, there might be an electron with orbital angular momentum relative to the nucleus, but the nucleus moves relative to some origin. What is the total angular momentum of the electron relative to the origin?

This section demonstrates how to perform such additions through a specific example, namely adding angular momentum A with $\ell_A = 1$ to angular momentum B with $\ell_B = 2$. (For the moment, assume that these angular momenta belong to non-identical particles. If the two particles are identical — as in the second example above — then there is an additional requirement that the sum wavefunction be symmetric or antisymmetric under swapping/interchange/exchange.)

First, recall the states for a single angular momentum: There are no states with values of $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$, and $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$ simultaneously, reflecting such facts as that $\hat{L}_x$ and $\hat{L}_z$ do not commute. However, because $\hat{L}^2$ and $\hat{L}_z$ *do* commute, there *are* states (in fact, a basis of states) that have values of $\hat{L}^2$ and $\hat{L}_z$ simultaneously.

For angular momentum A, with $\ell_A = 1$, these basis states are

$$\begin{aligned}&|1,+1\rangle\\&|1,\quad 0\rangle\\&|1,-1\rangle\end{aligned}$$

where

$$\hat{L}_A^2|\ell_A, m_A\rangle = \hbar^2\ell_A(\ell_A+1)|\ell_A, m_A\rangle = \hbar^2(1)(2)|\ell_A, m_A\rangle$$

and

$$\hat{L}_{A,z}|\ell_A, m_A\rangle = \hbar m_A|\ell_A, m_A\rangle.$$

These states are called the "$\ell_A = 1$ triplet".

For angular momentum B, with $\ell_B = 2$, these basis states are

$$\begin{aligned}&|2,+2\rangle\\&|2,+1\rangle\\&|2,\quad 0\rangle\\&|2,-1\rangle\\&|2,-2\rangle\end{aligned}$$

where

$$\hat{L}_B^2|\ell_B, m_B\rangle = \hbar^2\ell_B(\ell_B+1)|\ell_B, m_B\rangle = \hbar^2(2)(3)|\ell_B, m_B\rangle$$

and

$$\hat{L}_{B,z}|\ell_B, m_B\rangle = \hbar m_B|\ell_B, m_B\rangle.$$

These states are called the "$\ell_B = 2$ quintet".

Now, what sort of states can we have for the sum of these two angular momenta? The relevant total angular momentum operator is

$$\hat{\vec{J}} = \hat{\vec{L}}_A + \hat{\vec{L}}_B$$

so

$$\hat{J}_z = \hat{L}_{A,z} + \hat{L}_{B,z}$$

but

$$\hat{J}^2 \neq \hat{L}_A^2 + \hat{L}_B^2.$$

We can ask for states with values of $\hat{J}^2$ and $\hat{J}_z$ simultaneously, but such states will *not* necessarily have values of $\hat{L}_{A,z}$ and $\hat{L}_{B,z}$, because $\hat{J}^2$ and $\hat{L}_{A,z}$ do not commute (see problem 201, "Angular momentum commutators").

For the same reason, we can ask for states with values of $\hat{L}_{A,z}$ and $\hat{L}_{B,z}$ simultaneously, but such states will *not* necessarily have values of $\hat{J}^2$.

For most problems, there are two bases that are natural and useful. The first is consists of states like $|\ell_A, m_A\rangle|\ell_B, m_B\rangle$ — simple product states of the bases we discussed above. The second basis consists of states like $|j, m_J\rangle$. To find how these are connected, we list states in the first basis according to their associated[1] value of m_J:

$\vert\ell_A, m_A\rangle\vert\ell_B, m_B\rangle$			m_J		
$\vert 1,+1\rangle\vert 2,+2\rangle$			$+3$		
$\vert 1,+1\rangle\vert 2,+1\rangle$	$\vert 1,0\rangle\vert 2,+2\rangle$		$+2$	$+2$	
$\vert 1,+1\rangle\vert 2,\ \ 0\rangle$	$\vert 1,0\rangle\vert 2,+1\rangle$	$\vert 1,-1\rangle\vert 2,+2\rangle$	$+1$	$+1$	$+1$
$\vert 1,+1\rangle\vert 2,-1\rangle$	$\vert 1,0\rangle\vert 2,\ \ 0\rangle$	$\vert 1,-1\rangle\vert 2,+1\rangle$	0	0	0
$\vert 1,+1\rangle\vert 2,-2\rangle$	$\vert 1,0\rangle\vert 2,-1\rangle$	$\vert 1,-1\rangle\vert 2,\ \ 0\rangle$	-1	-1	-1
	$\vert 1,0\rangle\vert 2,-2\rangle$	$\vert 1,-1\rangle\vert 2,-1\rangle$		-2	-2
		$\vert 1,-1\rangle\vert 2,-2\rangle$			-3

These values of m_J fall into a natural structure:

There is a heptet of seven states with $m_J = +3, +2, +1, 0, -1, -2, -3$. This heptet must be associated with $j = 3$.

There is a quintet of five states with $m_J = +2, +1, 0, -1, -2$. This quintet must be associated with $j = 2$.

There is a triplet of three states with $m_J = +1, 0, -1$. This triplet must be associated with $j = 1$.

So now we know what the values of j are! If you think about this problem for general values of ℓ_A and ℓ_B, you will see immediately that the values of j run from $\ell_A + \ell_B$ to $|\ell_A - \ell_B|$. Often, this is all that's needed.[2] But sometimes you need more. Sometimes you need to express total-angular-momentum states like $|j, m_J\rangle$ in terms of in individual-angular-momentum states like $|\ell_A, m_A\rangle|\ell_B, m_B\rangle$.

The basic set-up of our problem comes through the table below:

---

[1] While the state $|\ell_A, m_A\rangle|\ell_B, m_B\rangle$ doesn't have a value of j, it does have a value of m_J, namely $m_J = m_A + m_B$.

[2] In particular, many GRE questions that appear on their face to be deep and difficult only go this far.

$\|\ell_A, m_A\rangle\|\ell_B, m_B\rangle$			$\|j, m_J\rangle$		
$\|1,+1\rangle_A\|2,+2\rangle_B$			$\|3,+3\rangle_J$		
$\|1,+1\rangle_A\|2,+1\rangle_B$	$\|1,0\rangle_A\|2,+2\rangle_B$		$\|3,+2\rangle_J$	$\|2,+2\rangle_J$	
$\|1,+1\rangle_A\|2,\ \ 0\rangle_B$	$\|1,0\rangle_A\|2,+1\rangle_B$	$\|1,-1\rangle_A\|2,+2\rangle_B$	$\|3,+1\rangle_J$	$\|2,+1\rangle_J$	$\|1,+1\rangle_J$
$\|1,+1\rangle_A\|2,-1\rangle_B$	$\|1,0\rangle_A\|2,\ \ 0\rangle_B$	$\|1,-1\rangle_A\|2,+1\rangle_B$	$\|3,\ \ 0\rangle_J$	$\|2,\ \ 0\rangle_J$	$\|1,\ \ 0\rangle_J$
$\|1,+1\rangle_A\|2,-2\rangle_B$	$\|1,0\rangle_A\|2,-1\rangle_B$	$\|1,-1\rangle_A\|2,\ \ 0\rangle_B$	$\|3,-1\rangle_J$	$\|2,-1\rangle_J$	$\|1,-1\rangle_J$
	$\|1,0\rangle_A\|2,-2\rangle_B$	$\|1,-1\rangle_A\|2,-1\rangle_B$	$\|3,-2\rangle_J$	$\|2,-2\rangle_J$	
		$\|1,-1\rangle_A\|2,-2\rangle_B$	$\|3,-3\rangle_J$		

Note that we have labeled states like

$$|\ell_A, m_A\rangle|\ell_B, m_B\rangle \text{ as } |\ell_A, m_A\rangle_A|\ell_B, m_B\rangle_B$$

and states like

$$|j, m_J\rangle \text{ as } |j, m_J\rangle_J.$$

Otherwise we might confuse the state $|2,+1\rangle_B$ on the left side of the second row with the completely different state $|2,+1\rangle_J$ on the right side of the of the third row. (Some authors solve this notation vexation by writing the states of total angular momentum as $|j, m_J, \ell_A, \ell_B\rangle$, taking advantage of the fact that ℓ_A and ℓ_B are the same for all states on the right — and for all states on the left, for that matter. This means every state on the right would be written as $|j, m_J, 1, 2\rangle$. For me, it rapidly grows frustrating to tack a "1,2" on to the end of every such state.)

The second line of this table means that the state $|3,+2\rangle_J$ is some linear combination of the states $|1,+1\rangle_A|2,+1\rangle_B$ and $|1,0\rangle_A|2,+2\rangle_B$. Similarly for the state $|2,+2\rangle_J$. ⟦This is the *meaning* of the assertion made earlier that in the state $|3,+2\rangle_J$ there is no value for m_A: The state $|3,+2\rangle_J$ is a superposition of a state with $m_A = +1$ and a state with $m_A = 0$, but the state $|3,+2\rangle_J$ itself has no value for m_A.⟧ Similarly, the state $|1,+1\rangle_A|2,+1\rangle_B$ is a linear combination of states $|3,+2\rangle_J$ and $|2,+2\rangle_J$

But what linear combination? We start with the first line of the table. Because there's only one state on each side, we write

$$|3,+3\rangle_J = |1,+1\rangle_A|2,+2\rangle_B. \tag{20.1}$$

(We could have inserted an overall phase factor of magnitude one, such as $|3,+3\rangle_J = -|1,+1\rangle_A|2,+2\rangle_B$ or $|3,+3\rangle_J = i|1,+1\rangle_A|2,+2\rangle_B$ or even $|3,+3\rangle_J = -\sqrt{i}\,|1,+1\rangle_A|2,+2\rangle_B$. But this insertion would have only made our lives difficult for no reason.)

Now, to find an expression for $|3,+2\rangle$, apply the lowering operator

$$\hat{J}_- = \hat{L}_{A,-} + \hat{L}_{B,-}$$

to both sides of equation (20.1). Remembering that

$$\hat{J}_-|j,m\rangle = \hbar\sqrt{j(j+1)-m(m-1)}\,|j,m-1\rangle,$$

this lowering gives

$$\begin{aligned}
\hat{J}_-|3,+3\rangle_J &= \left[\hat{L}_{A,-}|1,+1\rangle_A\right]|2,+2\rangle_B \\
&\qquad + |1,+1\rangle_A\left[\hat{L}_{B,-}|2,+2\rangle_B\right] \\
\hbar\sqrt{3(4)-3(2)}\,|3,+2\rangle_J &= \left[\hbar\sqrt{1(2)-1(0)}\,|1,0\rangle_A\right]|2,+2\rangle_B \\
&\qquad + |1,+1\rangle_A\left[\hbar\sqrt{2(3)-2(1)}\,|2,+1\rangle_B\right] \\
\sqrt{6}\,|3,+2\rangle_J &= \left[\sqrt{2}\,|1,0\rangle_A\right]|2,+2\rangle_B + |1,+1\rangle_A\left[\sqrt{4}\,|2,+1\rangle_B\right] \\
|3,+2\rangle_J &= \sqrt{\frac{1}{3}}\,|1,0\rangle_A|2,+2\rangle_B + \sqrt{\frac{2}{3}}\,|1,+1\rangle_A|2,+1\rangle_B.
\end{aligned} \tag{20.2}$$

Before, we knew only that if the system were in state $|3,+2\rangle_J$ and we measured m_A, the result might be 0 or it might be +1. Now we know that the probability of obtaining the result 0 is $\frac{1}{3}$, while the probability of obtaining the result +1 is $\frac{2}{3}$.

You can continue this process: lower $|3,+2\rangle_J$ to find an expression for $|3,+1\rangle_J$, lower $|3,+1\rangle_J$ to find an expression for $|3,0\rangle_J$, and so forth. When you get to $|3,-2\rangle_J$, you should lower it to find

$$|3,-3\rangle_J = |1,-1\rangle_A|2,-2\rangle_B,$$

and if that's not the result you get, then you made an error somewhere in this long chain.

Now we know how to find expressions for the entire heptet $|3,m\rangle_J$, with m ranging from +3 to −3. But what about the quintet $|2,m\rangle_J$, with m ranging from +2 to −2? If we knew the top member $|2,+2\rangle_J$, we could lower away to find the rest of the quintet. But how do we find this starting point?

The trick to use here is orthogonality. We know that

$$|2,+2\rangle_J = \alpha|1,0\rangle_A|2,+2\rangle_B + \beta|1,+1\rangle_A|2,+1\rangle_B,$$

where α and β are to be determined, and that

$$|3,+2\rangle_J = \sqrt{\frac{1}{3}}\,|1,0\rangle_A|2,+2\rangle_B + \sqrt{\frac{2}{3}}\,|1,+1\rangle_A|2,+1\rangle_B$$

and that

$$\langle 3,+2|2,+2\rangle_J = 0.$$

We use the orthogonality to find the expansion coefficients α and β:

$$\begin{aligned}
0 &= \langle 3,+2|2,+2\rangle_J \\
&= \left[\sqrt{\frac{1}{3}}\,{}_A\langle 1,0|_B\langle 2,+2| + \sqrt{\frac{2}{3}}\,{}_A\langle 1,+1|_B\langle 2,+1|\right]\left[\alpha|1,0\rangle_A|2,+2\rangle_B + \beta|1,+1\rangle_A|2,+1\rangle_B\right] \\
&= \sqrt{\frac{1}{3}}\,\alpha\langle 1,0|1,0\rangle_A\langle 2,+2|2,+2\rangle_B + \sqrt{\frac{1}{3}}\,\beta\langle 1,0|1,+1\rangle_A\langle 2,+2|2,+1\rangle_B \\
&\qquad + \sqrt{\frac{2}{3}}\,\alpha\langle 1,+1|1,0\rangle_A\langle 2,+1|2,+2\rangle_B + \sqrt{\frac{2}{3}}\,\beta\langle 1,+1|1,+1\rangle_A\langle 2,+1|2,+1\rangle_B \\
&= \sqrt{\frac{1}{3}}\,\alpha(1) + \sqrt{\frac{1}{3}}\,\beta(0) + \sqrt{\frac{2}{3}}\,\alpha(0) + \sqrt{\frac{2}{3}}\,\beta(1) \\
&= \alpha\sqrt{\frac{1}{3}} + \beta\sqrt{\frac{2}{3}}.
\end{aligned}$$

There are, of course, many solutions to this equation, but you can read off a normalized solution, namely

$$\alpha = \sqrt{\frac{2}{3}}, \qquad \beta = -\sqrt{\frac{1}{3}}$$

so that

$$|2,+2\rangle_J = \sqrt{\frac{2}{3}}\,|1,0\rangle_A|2,+2\rangle_B - \sqrt{\frac{1}{3}}\,|1,+1\rangle_A|2,+1\rangle_B. \tag{20.3}$$

You could have taken $|2,+2\rangle_J$ to be the negative of the expression above (or i times the expression above, or $\sqrt{i}$ times the expression above, etc.) I recommend against this: life is hard enough on its own, don't go out of your way to deliberately make difficulties for yourself.

Once the expression for $|2,+2\rangle_J$ is known, we can lower m_J from $+2$ all the way to -2 to find expressions for the entire $j = 2$ quintet.

And then one can find the expression for $|1,+1\rangle_J$ by demanding that it be orthogonal to $|3,+1\rangle_J$ and $|2,+1\rangle_J$. And once that's found we can lower to find expressions for the entire $j = 1$ triplet.

In summary, the states of these two angular momenta, $\ell_A = 1$ and $\ell_B = 2$, fall in a Hilbert space with a fifteen-member basis. While there are, of course, an infinite number of bases, the most natural and most useful bases are (1) the states of definite individual angular momenta (the 15 states like $|\ell_A, m_A\rangle|\ell_B, m_B\rangle$) or (2) the states of definite total angular momentum (the 15 states like $|j, m_J\rangle_J$). We now know (in principle) how to express states of the second basis in terms of states in the first basis. The coefficients, like $\sqrt{1/3}$ and $\sqrt{2/3}$ in equation (20.2), or $\sqrt{2/3}$ and $-\sqrt{1/3}$ in equation (20.3) that implement this change of basis are called Clebsch-Gordon coefficients.[3]

As you can see, it takes a lot of work to compute Clebsch-Gordon coefficients, but fortunately you don't have to do it. There are published tables of Clebsch-Gordon coefficients. Griffiths explains how to use them.

Problem

20.1 **Angular momentum commutators**
Show that

$$[\hat{J}^2, \hat{L}_{A,z}] = 2i\hbar(\hat{L}_{A,x}\hat{L}_{B,y} - \hat{L}_{A,y}\hat{L}_{B,x}).$$

Without performing any new calculation, find $[\hat{J}^2, \hat{L}_{B,z}]$.

20.2 Hartree-Fock approximation

For atom with atomic number Z.

(1) Guess some spherically-symmetric potential energy function that interpolates between

$$\text{for small } r, \quad V(r) \approx -\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r} \tag{20.4}$$

and

$$\text{for large } r, \quad V(r) \approx -\frac{1}{4\pi\epsilon_0}\frac{e^2}{r}. \tag{20.5}$$

---

[3] Alfred Clebsch (1833–1872) and Paul Gordan (1837–1912) were German mathematicians who recognized the importance of these coefficients in the purely mathematical context of invariant theory in about 1868, years before quantum mechanics was discovered. Gordan went on to serve as thesis advisor for Emmy Noether.

(2) Using all the tricks we've learned about spherically-symmetric potential energy functions, solve (numerically) the energy eigenproblem for the lowest $Z/2$ one-body energy levels. (If Z is odd, round up.)

(3) Use the antisymmetrization machinery to combine those levels into the Z-body ground state.

(4) From the quantal probability density for electrons in configuration space, deduce an electrostatic charge density in position space.

(5) Average that charge density over angle to make it spherically symmetric.

(6) From this spherically-symmetric charge density, use the shell theorem of electrostatics to deduce a spherically-symmetric potential energy function.

(7) Go to step **(2)**

You'll notice that this process never ends. In practice, you repeat until either you've earned a Ph.D. or you can't stand it any longer.

This is a "mean-field approximation". An electron is assumed to interact with the mean (average) of all the other electrons. Even if you go through this process an infinite number of times, you will never get the fine points of two electrons interacting far from the nucleus and from the other electrons.

Nevertheless, even two or three cycles through this algorithm can produce results in close accord with experiment. This has always surprised me and I think if I understood it I'd discover something valuable about quantum mechanics.

20.3 Atomic ground states

In addition to the process described above, you have to worry about spin, and about orbital angular momentum and (when you go on to Hamiltonians more accurate than the above) their interaction.

Friedrich Hund[4] did many such perturbation calculations and noticed regularities that he codified into "Hund's rules". Griffith talks about them.

---

[4] German physicist (1896–1997) who applied quantum mechanics to atoms and molecules, and who discovered quantum tunneling.

The some aspects of an electronic state are described using a particular notation — called a "term symbol" —- which you should know about. A state will have a particular orbital angular momentum L, spin angular momentum S, and total angular momentum J. (It will also have values of L_z, S_z, and J_z, but they are not recorded in this notation.) You would think these three numbers would be presented as three numbers, but no: they are conventionally presented as the term symbol

$$ {}^{2S+1}L_J. \tag{20.6} $$

By further convention, S and J are given as numbers, while L is presented as a letter using the S, P, D, F encoding. (In this notation, capital not lower case letters are used. Please don't ask why.) The ground state of carbon, for example, happens to have $S = 1$, $L = 1$, and $J = 0$; it is described as a ${}^3P_0$ state. One last convention: the spin number is written as a number, but pronounced as a degeneracy. The ground state of carbon is pronounced "triplet pee zero". The ground state of sodium, ${}^2S_{1/2}$, is pronounced "doublet ess one-half".

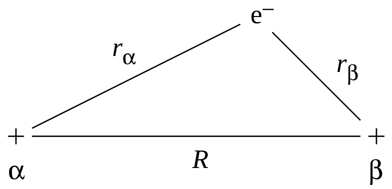
The people who write the physics GRE have fallen into the misconception that this term symbol notation tells us something important about nature, rather than about human convention. I recommend that you review the above paragraph the night before you take the GRE.

Chapter 21

Molecules

21.1 The hydrogen molecule ion

The hydrogen molecule ion[1] is two protons an a single electron. . . H_2^+. If we had managed to successfully solve the helium atom problem we would also have solved this one, because it's just three particles interacting through $1/r^2$ forces. However, you know that this problem has not been exactly solved even in the classical limit. Thus we don't even look for an exact solution: we look for the approximation most applicable to the case of two particles much more massive than the third.



If we take account of the Coulomb forces, but ignore things like the finite size of the nucleus, relativistic motion of the electron, spin-orbit effects, and so forth, the Hamiltonian for one electron and two protons (α and β) is

$$\hat{H} = \widehat{\mathrm{KE}}_\alpha + \widehat{\mathrm{KE}}_\beta + \widehat{\mathrm{KE}}_e + \hat{U}_{\alpha\beta} + \hat{U}_{\alpha e} + \hat{U}_{\beta e} \tag{21.1}$$

This is, of course, also the Hamiltonian for the helium atom, or for any three-body problem with pair interactions. Now comes the approximation suitable for the hydrogen molecule ion (but *not* appropriate for the helium

---

[1] Technically the hydrogen molecule cation.

atom): Assume that the two protons are so massive that they are fixed, and the interaction between them is treated classically. In equations, this approximation demands

$$\widehat{\mathrm{KE}}_\alpha = 0; \qquad \widehat{\mathrm{KE}}_\beta = 0; \qquad \hat{U}_{\alpha\beta} = U_{\alpha\beta} = \frac{e^2}{4\pi\epsilon_0}\frac{1}{R}. \tag{21.2}$$

The remaining, quantum mechanical, piece of the full Hamiltonian is the electronic Hamiltonian

$$\hat{H}_e = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0}\left(\frac{1}{r_\alpha} + \frac{1}{r_\beta}\right). \tag{21.3}$$

This approximation is called the "Born-Oppenheimer" approximation.

What shall we do with the electronic Hamiltonian? It would be nice to have an analytic solution of the energy eigenproblem. Then we could do precise comparisons between these results and the experimental spectrum of the hydrogen molecule ion, and build on them to study the hydrogen molecule, in exactly the same way that we built on our exact solution for He^+ to get an approximate solution for He. This goal is hopelessly beyond our reach. [Check out Gordon W.F. Drake, editor, *Atomic, Molecular, and Optical Physics Handbook* (AIP Press, Woodbury, NY, 1996) Reference QC173.A827 1996. There's a chapter on high-precision calculations for helium, but no chapter on high-precision calculations for the hydrogen molecule ion.] Instead of giving up, we might instead look for an exact solution to the ground state problem. This goal is also beyond our reach. Instead of giving up, we use the variational method to look for an approximate ground state.

Before doing so, however, we notice one exact symmetry of the electronic Hamiltonian that will guide us in our search for approximate solutions. The Hamiltonian is symmetric under the interchange of symbols α and β or, what is the same thing, symmetric under inversion about the point midway between the two nuclei. Any discussion of parity (see, for example, Gordon Baym *Lectures on Quantum Mechanics* pages 99–101) shows that this means the energy eigenfunctions can always be chosen either odd or even under the interchange of α and β.

Where will we find a variational trial wavefunction? If nucleus β did not exist, the ground state wavefunction would be the hydrogen ground state wavefunction centered on nucleus α:

$$\eta_\alpha(\vec{r}) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r_\alpha/a_0} \equiv |\alpha\rangle. \tag{21.4}$$

Similarly if nucleus α did not exist, the ground state wavefunction would be

$$\eta_\beta(\vec{r}) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r_\beta/a_0} \equiv |\beta\rangle. \tag{21.5}$$

We take as our trial wavefunction a linear combination of these two wavefunctions. This trial wavefunction is called a “linear combination of atomic orbitals” or “LCAO”. So the trial wavefunction is

$$\psi(\vec{r}) = A\eta_\alpha(\vec{r}) + B\eta_\beta(\vec{r}). \tag{21.6}$$

At first glance, it seems that the variational parameters are the complex numbers A and B, for a total of four real parameters. However, one parameter is taken up through normalization, and one through overall phase. Furthermore, because of parity the swapping of α and β can result in at most a change in sign, whence $B = \pm A$. Thus our trial wavefunction is

$$\psi(\vec{r}) = A_\pm[\eta_\alpha(\vec{r}) \pm \eta_\beta(\vec{r})], \tag{21.7}$$

where $A_\pm$ is the normalization constant, selected to be real and positive. (The notation $A_\pm$ reflects the fact that depending on whether we take the + sign or the − sign, we will get a different normalization constant.)

This might seem like a letdown. We have discussed exquisitely precise variational wavefunction involving hundreds or even thousands of real parameters. Here the only variational parameter is the binary choice: + sign or − sign! Compute $\langle \hat{H}_e \rangle$ both ways and see which is lower! You don’t even have to take a derivative at the end! Clearly this is a first attempt and more accurate calculations are possible. Rather than give in to despair, however, let’s recognize the limitations and forge on to see what we can discover. At the very least what we learn here will guide us in selecting better trial wavefunctions for our next attempt.

There are only two steps: normalize the wavefunction and evaluate $\langle \hat{H}_e \rangle$. However, these steps can be done through a frontal assault (which is likely to get hopelessly bogged down in algebraic details) or through a more subtle approach recognizing that we already know quite a lot about the functions $\eta_\alpha(\vec{r})$ and $\eta_\beta(\vec{r})$, and using this knowledge to our advantage. Let’s use the second approach.

Normalization demands that

$$\begin{aligned} 1 &= |A_\pm|^2(\langle\alpha| \pm \langle\beta|)(|\alpha\rangle \pm |\beta\rangle) \\ &= |A_\pm|^2(\langle\alpha|\alpha\rangle \pm \langle\alpha|\beta\rangle \pm \langle\beta|\alpha\rangle + \langle\beta|\beta\rangle) \\ &= 2|A_\pm|^2(1 \pm \langle\alpha|\beta\rangle) \end{aligned}$$

where in the last step we have used the normalization of $|\alpha\rangle$ and $|\beta\rangle$. The integral $\langle\alpha|\beta\rangle$ is not easy to calculate, so we set it aside for later by naming it the *overlap integral*

$$I(R) \equiv \langle\alpha|\beta\rangle = \int \eta_\alpha(\vec{r})\eta_\beta(\vec{r})\, d^3r. \tag{21.8}$$

In terms of this integral, we can select the normalization to be

$$A_\pm = \frac{1}{\sqrt{2(1 \pm I(R))}}. \tag{21.9}$$

Evaluating the electronic Hamiltonian in the trial wavefunction gives

$$\begin{aligned}\langle\hat{H}_e\rangle &= \frac{(\langle\alpha| \pm \langle\beta|)\hat{H}_e(|\alpha\rangle \pm |\beta\rangle)}{2(1 \pm I(R))} \\ &= \frac{\langle\alpha|\hat{H}_e|\alpha\rangle \pm \langle\alpha|\hat{H}_e|\beta\rangle \pm \langle\beta|\hat{H}_e|\alpha\rangle + \langle\beta|\hat{H}_e|\beta\rangle}{2(1 \pm I(R))} \\ &= \frac{\langle\alpha|\hat{H}_e|\alpha\rangle \pm \langle\beta|\hat{H}_e|\alpha\rangle}{1 \pm I(R)}\end{aligned} \tag{21.10}$$

But we have already done large parts of these two integrals:

$$\begin{aligned}\hat{H}_e|\alpha\rangle &= \left[\widehat{\mathrm{KE}} - \frac{e^2}{4\pi\epsilon_0}\frac{1}{r_\alpha} - \frac{e^2}{4\pi\epsilon_0}\frac{1}{r_\beta}\right]|\alpha\rangle \\ &= \left[\widehat{\mathrm{KE}} - \frac{e^2}{4\pi\epsilon_0}\frac{1}{r_\alpha}\right]|\alpha\rangle - \frac{e^2}{4\pi\epsilon_0}\frac{1}{r_\beta}|\alpha\rangle \\ &= -\mathrm{Ry}\,|\alpha\rangle - 2\,\mathrm{Ry}\,a_0\frac{1}{r_\beta}|\alpha\rangle \\ &= -\mathrm{Ry}\left[|\alpha\rangle + 2\frac{a_0}{r_\beta}|\alpha\rangle\right]\end{aligned} \tag{21.11}$$

whence

$$\langle\alpha|\hat{H}_e|\alpha\rangle = -\mathrm{Ry}\left[1 + 2\left\langle\alpha\left|\frac{a_0}{r_\beta}\right|\alpha\right\rangle\right] \tag{21.12}$$

$$\langle\beta|\hat{H}_e|\alpha\rangle = -\mathrm{Ry}\left[\langle\beta|\alpha\rangle + 2\left\langle\beta\left|\frac{a_0}{r_\beta}\right|\alpha\right\rangle\right]. \tag{21.13}$$

On the right-hand side we recognize the overlap integral, $I(R) = \langle\beta|\alpha\rangle$, and two new (dimensionless) integrals, which are called the *direct integral*

$$D(R) \equiv \left\langle\alpha\left|\frac{a_0}{r_\beta}\right|\alpha\right\rangle \tag{21.14}$$

and the *exchange integral*

$$X(R) \equiv \left\langle \beta \left| \frac{a_0}{r_\beta} \right| \alpha \right\rangle. \tag{21.15}$$

These two integrals are not easy to work out (I will assign them as homework) but once we do them (plus the overlap integral) we can find the mean value of the electronic Hamiltonian in the trial wavefunction. It is

$$\begin{aligned} \langle \hat{H}_e \rangle &= -\text{Ry}\frac{1 + 2D(R) \pm I(R) \pm 2X(R)}{1 \pm I(R)} \\ &= -\text{Ry}\left[1 + 2\frac{D(R) \pm X(R)}{1 \pm I(R)}\right]. \end{aligned} \tag{21.16}$$

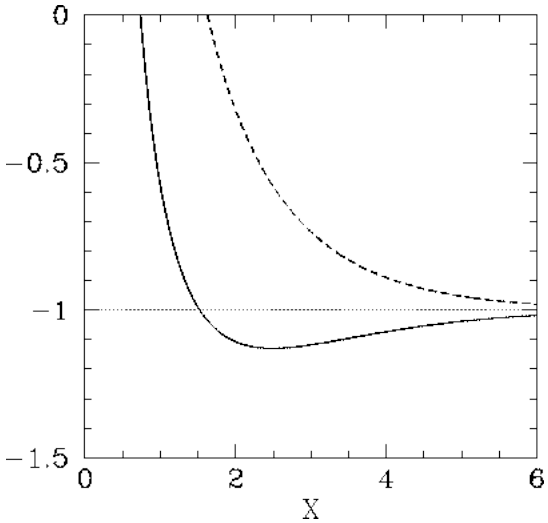
This, remember, is only the electronic part of the Hamiltonian. In the Born-Oppenheimer approximation the nuclear part has no kinetic energy and Coulombic potential energy

$$\frac{e^2}{4\pi\epsilon_0}\frac{1}{R} = 2\,\text{Ry}\frac{a_0}{R}, \tag{21.17}$$

so the upper bound on the total ground state energy is

$$\text{Ry}\left[2\frac{a_0}{R} - 1 - 2\frac{D(R) \pm X(R)}{1 \pm I(R)}\right]. \tag{21.18}$$

What are the results?



Here the dashed line represents −, the solid line represents +. X means R/a_0, and the vertical axis is energy in Ry. [When $R \to \infty$, the system is a hydrogen atom (ground state energy −Ry) and a clamped proton far away (ground state energy 0).]

21.1.1 *Why is + lower energy than −?*

21.1.2 *Understanding the integrals*

How can we understand these integrals? This section uses scaled units.

First, all three integrals are always positive.

The overlap integral: $I(R) = \langle\beta|\alpha\rangle$.
When $R \to \infty$, $I(R)$ approaches zero, exponentially quickly.
When $R = 0$, $I(R) = 1$.

The direct integral: $D(R) = \langle\alpha|1/r_\beta|\alpha\rangle$.
When $R \to \infty$, $D(R) \to 1/R$.
When $R = 0$, $D(R) = \langle 1/r\rangle = 1$.

The exchange integral: $X(R) = \langle\beta|1/r_\beta|\alpha\rangle$.
When $R \to \infty$, $X(R)$ approaches zero even faster than $I(R)$ does.
When $R = 0$, $X(R) = \langle 1/r\rangle = 1$.

Do the analytic expressions bear these limits out?

$$I(R) = e^{-R}\left(1 + R + \tfrac{1}{3}R^2\right) \tag{21.19}$$

$$D(R) = \frac{1}{R} - \left(1 + \frac{1}{R}\right)e^{-2R} \tag{21.20}$$

$$X(R) = e^{-R}(1 + R) \tag{21.21}$$

First conclusion: For R positive, $I(R) > X(R)$. Check.

For $R \to \infty$, $I(R)$ and $X(R)$ go to zero exponentially, while $D(R) \to 1/R$. Check.

For $R \to 0$,

$$I(R) \to 1 - \tfrac{1}{6}R^2 + \mathcal{O}(R^3) \tag{21.22}$$

$$X(R) \to 1 - \tfrac{1}{2}R^2 + \mathcal{O}(R^3) \tag{21.23}$$

Check, check. But what of $D(R)$? As $R \to 0$, you might say $D(R) \to \infty - (1 + \infty)1$, and the infinities cancel, so you're left with $D(R) \to -1$, but of course that's silly. . . we've already said that $D(R)$ is positive. We need to do the limit with some care.

$$\begin{aligned}
D(R) &= \frac{1}{R} - \left(1 + \frac{1}{R}\right) e^{-2R} \\
&= \frac{1}{R} - \left(1 + \frac{1}{R}\right) \left(1 + (-2R) + \tfrac{1}{2}(-2R)^2 + \tfrac{1}{6}(-2R)^3 + \mathcal{O}((-2R)^4)\right) \\
&= \frac{1}{R} - \left(1 + \frac{1}{R}\right) \left(1 - 2R + 2R^2 - \tfrac{4}{3}R^3 + \mathcal{O}(R^4)\right) \\
&= \frac{1}{R} - \left(1 - 2R + 2R^2 - \tfrac{4}{3}R^3 + \mathcal{O}(R^4)\right) \\
&\quad -\frac{1}{R} \left(1 - 2R + 2R^2 - \tfrac{4}{3}R^3 + \mathcal{O}(R^4)\right) \\
&= \frac{1}{R} - \left(1 - 2R + 2R^2 - \tfrac{4}{3}R^3 + \mathcal{O}(R^4)\right) \\
&\qquad - \left(\frac{1}{R} - 2 + 2R - \tfrac{4}{3}R^2 + \mathcal{O}(R^3)\right) \\
&= -\left(-1 + \tfrac{2}{3}R^2 + \mathcal{O}(R^3)\right) \\
&= 1 - \tfrac{2}{3}R^2 + \mathcal{O}(R^3). \qquad (21.24)
\end{aligned}$$

All three integrals start at 1 when $R = 0$. As R increases they all take off with zero slope, but drop quadratically: $I(R)$ is highest, then $X(R)$, and $D(R)$ lowest. But at some point $D(R)$ crosses the other two. While all three approach zero as $R \to \infty$, $D(R)$ does so much more slowly than the other two.

21.1.3 *Why is H_2^+ hard?*

Obviously not Pauli exclusion! But if you plot the various contributions, you see that it's classical nuclear repulsion, not "Heisenberg hardness".

21.2 Problems

21.1 **The hydrogen molecule ion: Evaluation of integrals**
Evaluate the direct and exchange integrals $D(R)$ and $X(R)$. (*Clue:* Remember that $\sqrt{x^2} = |x|$.) Plot as a function of R the overlap integral, $I(R)$, as well as $D(R)$ and $X(R)$.

21.2 **The hydrogen molecule ion: Thinking about integrals**
For the hydrogen molecule ion, find and plot the mean values of nuclear potential energy, total electronic energy, kinetic electronic energy, and potential electronic energy for the state $\psi_+(\vec{r})$, as functions of R. Do these plots shed any light on our initial question of "Why is stuff hard?" (We gave possible answers of "repulsion hardness," "Heisenberg hardness," and "Pauli hardness.") *Bonus:* The hydrogen molecule ion cannot display Pauli hardness, because it has only one quantal particle. Can you generalize this discussion to the neutral hydrogen molecule?

21.3 **Improved variational wavefunction**
Everett Schlawin ('09) suggested using "shielded" subwavefunctions like equation (19.4) in place of the subwavefunctions (21.4) and (21.5) that go into making trial wavefunction (21.7). Then there would be a variational parameter Z in addition to the binary choice of + or −. I haven't tried this, but through the usual variational argument, it can't be worse than what we've tried so far! (That is, the *results* can't be worse. The *amount of labor* involved can be far, far worse.) Execute this suggestion. Show that this trial wavefunction results in the exact helium ion ground state energy in the case $R = 0$.

21.3 The hydrogen molecule

When we discussed the helium atom, we had available an exact solution (that is, exact ignoring fine and hyperfine structure) of the helium ion problem. We used the one-body levels of the helium ion problem as building blocks for the two-body helium atom problem. Then we added electron-electron repulsion. You will recall, for example, that the helium atom ground state had the form (where "level" refers to a solution of the one-body helium ion problem)

$$\text{(two electrons in ground level)} \times \text{(spin singlet)} \tag{21.25}$$

while the helium atom first excited state had the form

$$\text{(one electron in ground level, one in first excited level)} \times \text{(spin triplet).} \tag{21.26}$$

We will attempt the same strategy for the hydrogen molecule, but we face a roadblock at the very first step — we lack an exact solution to the hydrogen molecule ion problem! Using LCAO, we have a candidate for a ground state, namely

$$\psi_+(\vec{r}) = A_+[\eta_\alpha(\vec{r}) + \eta_\beta(\vec{r})]. \tag{21.27}$$

21.4 Can we do better?

Try out our LCAO upper bound for the electronic ground state energy (21.16) at $R = 0$: The result is -3 Ry. But for $R = 0$ this is just the Helium ion, for which the exact ground state energy is -4 Ry. Sure enough, the variational method produces an upper bound, but it's a poor one.

We've seen before that the trick to getting good variational bounds is to figure out the qualitative character of the true wavefunction and select a trial wavefunction that mimics that character. Friedrich Hund, Robert Mulliken, John C. Slater, and John Lennard-Jones started out by dreaming up a trial wavefunction that could mimic the character of the true wavefunction at $R = 0$. Their techniques evolved into what is today called the "molecular orbital method". This is only one of several choices of trial wavefunction. Others are called "valance bond theory" or "the Hückel method" or "the extended Hückel method".

Story about Roald Hoffmann.

All these are primitive, but in synthetic chemistry, you don't need the spectrum, you don't need the ground state energy, all you need to know is which structure has lower energy, and that's the one you'll synthesize.

Today, chemists are much more likely to use a completely different approach, called "density-functional theory". This was developed by the physicist Walter Kohn and made readily accessible through the computer program GAUSSIAN written by the mathematician John Pople. When Kohn and Pople won the Nobel Prize in Chemistry in 1998, I heard some chemists grumble that Chemistry Nobel laureates should have taken at least one chemistry course.

Chapter 22

WKB: The Quasiclassical Approximation

When I started learning quantum mechanics, I worked a lot of integrals and diagonalized a lot of matrices. But I also vaguely wondered "Why is quantum mechanics true?". For example, why *can't* a particle simultaneously have a position and a momentum? Eventually I realized that I had the question backwards. The real question is "We know that interference and entanglement exist. Why don't we notice them in daily life?" The Heisenberg indeterminacy principle, for example, answers the question[1] "When is the classical approximation adequate?" That is, the real question concerns the classical limit of quantum mechanics. Ehrenfest's Theorem shows that classical mechanics *can* be the limit of quantum mechanics, but not that it *has to* be. Research on this topic is vast and continues under the name "decoherence". We approach the topic through the quasiclassical approximation.

The WKB technique finds approximate solutions to the energy eigenproblem in one dimension. It is named for three physicists who independently discovered it: the German Gregor Wentzel, the Dutchman Hendrik Kramers, and the Frenchman Léon Brillouin. In the Netherlands it is known as the KWB approximation, in France as BWK, and in Britain as JWKB (adding a tribute to the English mathematician Sir Harold Jeffreys, who in fact discovered the approximation three years before Wentzel, Kramers, and Brillouin did). In Russia it is known as the quasiclassical approximation, the name that I prefer.

---

[1] Reference to the Bethe papers

The fact that this approximation was discovered independently four times suggests, correctly, that the idea is pretty straightforward.[2] Focus on a region where the potential energy function $V(x)$ is *constant*. Within that region the eigenfunction of energy E is, when $E > V$, given by

$$\eta(x) = Re^{\pm ikx} \quad \text{where} \quad \hbar k = \sqrt{2m(E-V)}. \tag{22.1}$$

The plus sign indicates positive momentum, the minus sign negative momentum, and the general solution is of course a linear combination of the two. The wavefunction is sinusoidal oscillatory, with constant wavelength

$$\lambda = \frac{2\pi\hbar}{\sqrt{2m(E-V)}} \tag{22.2}$$

and constant amplitude R. Now suppose that $V(x)$ is *not* constant, but that it varies slowly over the length λ. Then my guess would be that $\eta(x)$ is almost sinusoidal, but the wavelength and amplitude vary slowly with x. That is, I would seek oscillatory solutions like

$$\eta(x) = R(x)e^{\pm ik(x)x} \quad \text{where} \quad \hbar k(x) = \sqrt{2m(E-V(x))}. \tag{22.3}$$

On the other hand, if the potential energy function $V(x)$ is constant but $E < V$, then the energy eigenfunction is

$$\eta(x) = Re^{\pm x/d} \quad \text{where} \quad d = \frac{\hbar}{\sqrt{2m(V-E)}}. \tag{22.4}$$

Here d is the characteristic exponential decay length: If one walks in the direction of decreasing function, then the function diminishes by a factor of $1/e$ (about $1/3$) every time one steps a distance d. However when $V(x)$ is *not* constant, but varies slowly over the length of d, then my guess would be that $\eta(x)$ is almost exponential, but the decay length and amplitude vary slowly with x. That is, I would seek solutions like

$$\eta(x) = R(x)e^{\pm x/d(x)} \quad \text{where} \quad d(x) = \frac{\hbar}{\sqrt{2m(V(x)-E)}}. \tag{22.5}$$

What we have said so far reinforces the qualitative expectations for energy eigenfunction sketching established in section 9.1.

There is one place where this entire scheme is guaranteed to fail. If $E = V(x)$ then $\lambda(x) = d(x) = \infty$, and no potential energy function varies "slowly on the scale of infinity". The proper handling of these so-called

---

[2] The same basic idea can be used in many similar situations: to light moving in a medium where the index of diffraction varies slowly, for example, or to waves on a string of slowly-varying density.

"classical turning points" is the most difficult facet of deriving the quasi-classical approximation. However we will find that once the derivation is done the final result is easy to state and to use.

If you apply these ideas to two- or three-dimensional problems, you find that the classical turning points are now lines (in two dimensions) or surfaces (in three dimensions). The matching program at turning points becomes a matching program over lines or surfaces (called in this context "caustics") and the results are neither easy to state nor simple to use. They are connected with classical chaos, and, remarkably, with the theory of the rainbow. Such are the nimble abstractions of mathematics. We will not pursue these avenues in this book.

22.1 Polar form for the energy eigenproblem

Define the "classical momentum"

$$p_c(x) = \sqrt{2m(E - V(x))}. \tag{22.6}$$

This is the magnitude of the momentum that a classical particle of energy E would have if it were located at x. Of course, whenever $E < V(x)$, that is within a "classically prohibited region" (where a classical particle of energy E would never be), $p_c(x)$ is pure imaginary.

The energy eigenproblem equation

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

can be compactly written in terms of $p_c(x)$ as

$$\frac{d^2\eta(x)}{dx^2} = -\frac{p_c^2(x)}{\hbar^2}\eta(x). \tag{22.7}$$

We have already begun discussing (equations 22.3 and 22.5) the energy eigenfunction $\eta(x)$ in polar form, that is as a real-valued amplitude function $R(x)$ in addition to a real-valued phase function $\phi(x)$:

$$\eta(x) = R(x)e^{i\phi(x)} \tag{22.8}$$

(compare equation 6.27). To continue our discussion, we write the energy eigenproblem (22.7) in terms of $R(x)$ and $\phi(x)$. Using a prime to denote differentiation with respect to x,

$$\frac{d\eta}{dx} = [R' + iR\phi']e^{i\phi} \tag{22.9}$$

$$\frac{d^2\eta}{dx^2} = [R'' + 2iR'\phi' + iR\phi'' - R(\phi')^2]e^{i\phi}. \tag{22.10}$$

Whence energy eigenproblem (22.7) becomes

$$R'' + 2iR'\phi' + iR\phi'' - R(\phi')^2 = -\frac{p_c^2}{\hbar^2}R. \tag{22.11}$$

This complex equation is equivalent to two pure real equations, one for the real part and the other for the complex part. The real part is

$$R'' - R(\phi')^2 = -\frac{p_c^2}{\hbar^2}R \quad \text{or} \quad R'' = R\left[(\phi')^2 - \frac{p_c^2}{\hbar^2}\right], \tag{22.12}$$

while the imaginary part is

$$2R'\phi' + R\phi'' = 0 \quad \text{or} \quad \left[R^2\phi'\right]' = 0. \tag{22.13}$$

We have made no approximations: these two equations are equivalent to the original energy eigenproblem.

Furthermore, the second equation is readily solved to find

$$R^2\phi' = \tilde{C}^2 \quad \text{or} \quad R = \frac{\tilde{C}}{\sqrt{\phi'}}, \tag{22.14}$$

where $\tilde{C}$ is a constant.

Exercise 22.A. What are the dimensions of $\tilde{C}$? Show that it must be either pure real or pure imaginary.

22.1 **Energy eigenproblem in terms of phase**
Show that the phase $\phi(x)$ for an energy eigenstate obeys the non-linear differential equation

$$\phi'''\phi' - \tfrac{3}{2}(\phi'')^2 + 2(\phi')^4 - 2\frac{p_c^2(x)}{\hbar^2}(\phi')^2 = 0. \tag{22.15}$$

If the phase can be found through solving this equation (it usually can't be), then the magnitude can be found through equation (22.14).

22.2 Far from classical turning points

In contrast, the *real* part of the polar form of the energy eigenproblem, namely equation (22.13), usually cannot be solved. The quasiclassical approximation is that the magnitude $R(x)$ varies slowly enough that R'' is negligible in that equation. (To be precise, the magnitude $|R''/R|$ is small compared to $(\phi')^2$, *and* small compared to $(p_c(x)/\hbar)^2$.) When this assumption holds,

$$(\phi')^2 = \frac{p_c^2}{\hbar^2} \quad \text{or} \quad \frac{d\phi}{dx} = \pm\frac{p_c(x)}{\hbar}, \tag{22.16}$$

and consequently

$$\phi(x) = \pm\frac{1}{\hbar}\int p_c(x)\,dx, \tag{22.17}$$

where the expression is left as an indefinite integral, without a set constant of integration. This establishes the phase, and then equation (22.14) gives the magnitude, so all together

$$\eta(x) = \frac{C}{\sqrt{p_c(x)}} e^{\pm\frac{i}{\hbar}\int p_c(x)\,dx}, \tag{22.18}$$

where $C = \tilde{C}\sqrt{\pm\hbar}$. Furthermore any constant of integration can be absorbed into the constant C, which may now be complex.

For any value of E there are two linearly independent (approximate) solutions, one with the + sign and one with the − sign, and the general solution is a linear combination of the two.

In the classically allowed region, where $p_c(x)$ is real, equation (22.18) is the most convenient expression for the approximate energy eigenfunction. In the classically prohibited region, where $p_c(x)$ is imaginary, it is more convenient to use the equivalent

$$\eta(x) = \frac{C}{\sqrt{|p_c(x)|}} e^{\pm\frac{1}{\hbar}\int |p_c(x)|\,dx}. \tag{22.19}$$

As mentioned in the paragraph below equation (22.5), this approximation is guaranteed to fail when $E = V(x)$, that is where $p_c(x) = 0$ (the "classical turning point"), and this failure is demonstrated through the division by zero at classical turning points for both equations (22.18) and (22.19).

Note that within the classically allowed region, for either of these two solutions, the probability density is

$$|\eta(x)|^2 = \frac{|C|^2}{p_c(x)}, \tag{22.20}$$

which is the quantitative formulation of our principle, already determined on page 255, that the probability density for the quantal particle is *small* where the classical particle would be *fast*.

22.2 **Alternative derivation of quasiclassical approximation**
There are several ways to derive the quasiclassical approximate wavefunction (22.18). Here is an alternative to the derivation in the text that uses an expansion in terms of $\hbar$.

Inspired by the free-particle solution $\eta(x) = Ae^{\pm ipx/\hbar}$, write

$$\eta(x) = e^{if(x)/\hbar},$$

where $f(x)$ is some *complex* function. *Any* non-zero function can be written in this form.

a. Show that the energy eigenproblem is

$$i\hbar f''(x) - (f'(x))^2 + p_c^2(x) = 0.$$

b. Write $f(x)$ as a power series in $\hbar$

$$f(x) = f_0(x) + \hbar f_1(x) + \hbar^2 f_2(x) + \cdots,$$

plug into the energy eigenproblem, and collect like powers of $\hbar$ (dimensional analysis!) to show that

$$(f_0')^2 = p_c^2, \quad if_0'' = 2f_0'f_1', \quad if_1'' = 2f_0'f_2' + (f_1')^2, \quad \text{etc.}$$

c. Solve for $f_0(x)$ and $f_1(x)$ to rederive equation (22.18).

(This derivation is in principle superior to the "one-shot" derivation in the text, because it would be possible to solve for $f_2(x)$, for $f_3(x)$, etc., each time making the approximation more accurate. I do not personally known of anyone who has actually followed this possibility.)

22.3 The connection region

We have formula (22.18) accurate within the classically allowed region, and formula (22.18) accurate within the classically prohibited region. But we lack a formula accurate within the connection region near the classical turning point, where the quasiclassical approximation fails. The job of this section is to find a formula accurate in this region.

The classical turning point is x_R.
GRAPH with blue horizontal line marked E.
olive line slanted from SW to NE

$$V(x) = V(x_R) - F(x - x_R) = E - F(x - x_R).$$

(The slope is called $-F$ because F is the classical force experienced in the connection region, in this case a negative number.)
Vertical dashed line x_R
arrow to right of dashed line $\bar{x} = x - x_R$

Graph with qualitative $\eta(x)$ sketch?

$$-\frac{\hbar^2}{2m}\frac{d^2\eta}{dx^2} + V(x)\eta(x) = E\eta(x) \tag{22.21}$$

$$-\frac{\hbar^2}{2m}\frac{d^2\eta}{dx^2} + [E - F(x - x_R)]\eta(x) = E\eta(x) \tag{22.22}$$

In terms of the new variable $\bar{x}$

$$-\frac{\hbar^2}{2m}\frac{d^2\eta}{d\bar{x}^2} - F\bar{x}\eta(\bar{x}) = 0 \tag{22.23}$$

There are only two parameters: $\hbar^2/2m$ and F. What is the characteristic length for this problem?

quantity	dimensions
$\hbar^2/2m$	[mass][length]^4/[time]^2
F	[mass][length]/[time]^2

Clearly the characteristic length is

$$x_c = \left(\frac{\hbar^2/2m}{-F}\right)^{1/3}. \tag{22.24}$$

Defining the scaled variable

$$\tilde{x} = \bar{x}/x_c \tag{22.25}$$

we have

$$-\frac{d^2\eta}{d\tilde{x}^2} + \tilde{x}\eta(\tilde{x}) = 0. \tag{22.26}$$

This is called "Airy's equation", and the solutions are called Airy[3] functions. The two linearly-independent Airy functions are denoted $\mathrm{Ai}(\tilde{x})$ and

---

[3] George Biddell Airy (1801–1892), English astronomer and mathematician, found the density of the Earth, established the theory of the rainbow, refined the prime meridian at Greenwich, and tested the pre-relativistic ether drag hypothesis, among other activities. He encountered Richarda Smith during a walking tour of Derbyshire, and proposed marriage to her two days later.

Bi($\tilde{x}$). These functions have been studied expensively, and the results are summarized in the "Digital Library of Mathematical Functions". Here is some information quoted from that source:

Integral representations:

$$\text{Ai}(x) = \frac{1}{\pi}\int_0^\infty \cos(t^3/3 + xt)\,dt \tag{22.27}$$

$$\text{Bi}(x) = \frac{1}{\pi}\int_0^\infty \left[e^{-t^3/3+xt} + \sin(t^3/3 + xt)\right]\,dt \tag{22.28}$$

Asymptotic forms accurate when $1 \ll x$:

$$\text{Ai}(x) \sim \frac{1}{2\sqrt{\pi}x^{1/4}}e^{-(2/3)x^{3/2}} \tag{22.29}$$

$$\text{Bi}(x) \sim \frac{1}{\sqrt{\pi}x^{1/4}}e^{(2/3)x^{3/2}} \tag{22.30}$$

Asymptotic forms accurate when $x \ll -1$:

$$\text{Ai}(x) \sim \frac{1}{\sqrt{\pi}(-x)^{1/4}}\sin\left[\tfrac{2}{3}(-x)^{3/2} + \tfrac{\pi}{4}\right] \tag{22.31}$$

$$\text{Bi}(x) \sim \frac{1}{\sqrt{\pi}(-x)^{1/4}}\cos\left[\tfrac{2}{3}(-x)^{3/2} + \tfrac{\pi}{4}\right] \tag{22.32}$$

End of section on Airy functions.

22.4 Patching

Equation (22.18) is approximately correct within the classically allowed region; equation (22.18) is approximately correct within the classically prohibited region. But both equations result in nonsense (division by zero) at the classical turning point. Equation (XXX) is approximately correct near the classical turning point. Our task now is to patch all three equations together.

22.5 Why is WKB the "quasiclassical" approximation?

The approximation works when the de Broglie wavelength h/p is much less than the characteristic length L_c of variations in the potential energy function:

$$\begin{aligned} h/p &\ll L_c \\ p &\gg h/L_c. \end{aligned} \tag{22.33}$$

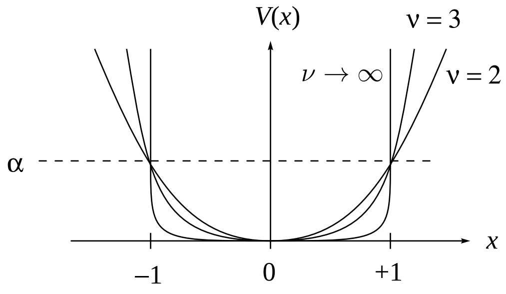
That is, it works for large — classical — values of momentum. Remember that when I say large I don't mean large on a human scale (say by comparing the momentum of a gnat to the momentum of a semi-truck), I mean large on the scale of h/L_C. So the momentum could be very small on a human scale yet the WKB approximation would still work very well.

22.6 The "power law" potential

While the quasiclassical approximation is difficult to derive, it is straightforward to apply. This section applies the approximation to the so-called "power law" potential energy function,

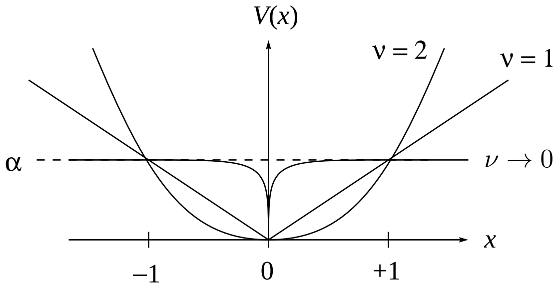
$$V(x) = \alpha|x|^{\nu}. \tag{22.34}$$

When $\nu = 2$ this is just the simple harmonic oscillator, which we have studied extensively. When $\nu > 2$ this potential traces out successively steeper potential wells as ν increases:



In the limit $\nu \to \infty$, the power law potential approaches an infinite square well.

Meanwhile, when $\nu < 2$ this potential traces out successively flatter potential wells as ν decreases:

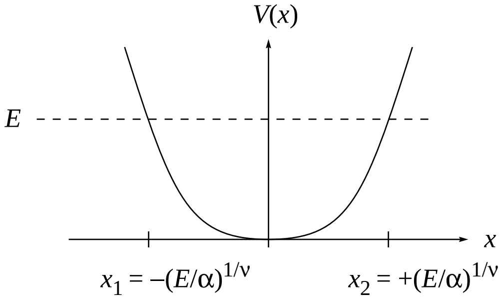


In the limit $\nu \to 0$, the power law potential approaches the flat potential $V(x) = \alpha$.

I don't know of any physical system that obeys the power law potential (except for the special cases $\nu = 0$, $\nu = 2$, and $\nu \to \infty$), but it's a good idea to understand quantum mechanics even in cases where it doesn't reflect any physical system.

To apply the quasiclassical approximation, locate the classical turning points at

$$x_1 = -(E/\alpha)^{1/\nu} \quad \text{and} \quad x_2 = +(E/\alpha)^{1/\nu}, \tag{22.35}$$



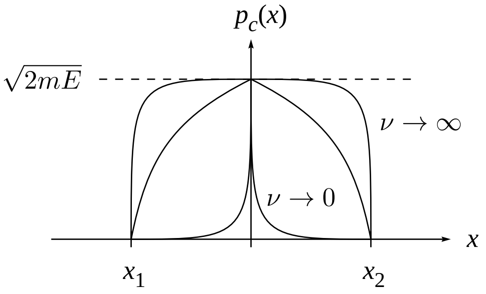
and then perform the integration

$$\int_{x_1}^{x_2} p_c(x)\,dx = (n - \tfrac{1}{2})\pi\hbar \tag{22.36}$$

where

$$p_c(x) = \sqrt{2m(E - V(x))} = \sqrt{2m(E - \alpha|x|^{1/\nu})}. \tag{22.37}$$

It's always a good idea to sketch the integrand before executing the integral, and that's what I do here:



So

$$\begin{aligned}\int_{x_1}^{x_2} p_c(x)\,dx &= \int_{x_1}^{x_2} \sqrt{2m(E-V(x))}\,dx \\ &= \sqrt{2m}\int_{-(E/\alpha)^{1/\nu}}^{+(E/\alpha)^{1/\nu}} \sqrt{E-\alpha|x|^\nu}\,dx \\ &= 2\sqrt{2m}\int_{0}^{+(E/\alpha)^{1/\nu}} \sqrt{E-\alpha x^\nu}\,dx.\end{aligned}$$

How should one execute this integral? I prefer to integrate over dimensionless variables, so as to separate the physical operation of setting up an integral from the mathematical operation of executing that integral. For that reason I define the dimensionless variable u through

$$\begin{aligned}\alpha x^\nu &= E u^\nu, \\ x &= \left(\frac{E}{\alpha}\right)^{1/\nu} u, \\ u &= \left(\frac{\alpha}{E}\right)^{1/\nu} x.\end{aligned}$$

Changing the integral to this variable

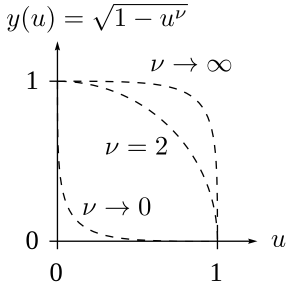
$$\begin{aligned}\int_{x_1}^{x_2} p_c(x)\,dx &= 2\sqrt{2m}\int_0^1 \sqrt{E-Eu^\nu}\left(\frac{E}{\alpha}\right)^{1/\nu} dx \\ &= 2\sqrt{2mE}\left(\frac{E}{\alpha}\right)^{1/\nu}\int_0^1 \sqrt{1-u^\nu}\,dx \\ &= \frac{(8m)^{1/2}}{\alpha^{1/\nu}} E^{(2+\nu)/2\nu}\int_0^1 \sqrt{1-u^\nu}\,dx\end{aligned}$$

where the integral here is a numerical function of ν independent of m or E or α. Let's call it

$$I(\nu) = \int_0^1 \sqrt{1-u^\nu}\,dx. \tag{22.38}$$

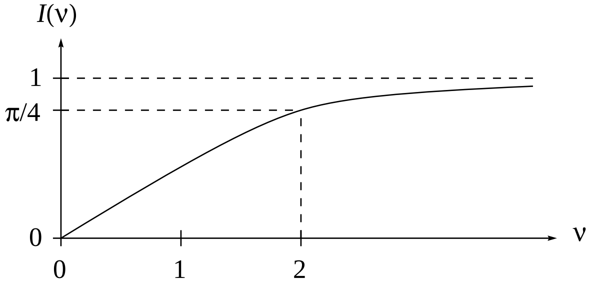
If you try to evaluate this integal in terms of polynomials or trig functions or anything familiar, you will fail. This is a function of ν all right, but we're going to have to uncover its properties on our own without recourse to familiar functions.

Let's start by graphing the integrand.



$I(\nu)$ is the area under the curve. You could produce a table of values through numerical integration, but let's uncover its properties first. It's clear from the graph that $I(0) = 0$, that as $\nu \to \infty$, $I(\nu) \to 1$, and that $I(\nu)$ increases monotonically.

When $\nu = 2$, the integrand y is $y = \sqrt{1 - u^2}$ so $u^2 + y^2 = 1\ldots$ the integrand traces out a quarter circle of radius 1. The area under this curve is of course $\pi/4$. So my first thought is that the function $I(\nu)$ looks like this:



But I want to investigate one detail further: What is the behavior of $I(\nu)$ for small values of ν? To find this, I need to understand the behavior of u^ν for small values of ν.

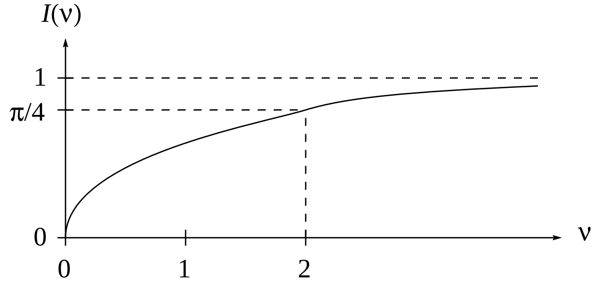
$$\begin{aligned}
e^x &= 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{3!}x^3 + \cdots \\
u^\nu = e^{\nu \ln u} &= 1 + \nu \ln u + \tfrac{1}{2}\nu^2 \ln^2 u + \tfrac{1}{6}\nu^3 \ln^3 u + \cdots \\
1 - u^\nu &= -\nu \ln u - \tfrac{1}{2}\nu^2 \ln^2 u - \tfrac{1}{6}\nu^3 \ln^3 u + \cdots \\
\sqrt{1-u^\nu} &\approx \sqrt{\nu}\sqrt{-\ln u}
\end{aligned}$$

At first glance it looks very bad to see that negative sign under the square root radical, but then you remember that when $0 < u < 1$, $\ln u$ is negative, so it's a good thing that the negative sign is there!

For small values of ν,

$$I(\nu) \approx \sqrt{\nu}\int_0^1 \sqrt{-\ln u}\, du = \sqrt{\nu}\,(\text{some positive number}). \tag{22.39}$$

Even without knowing the value of that postive number, you know that $I(\nu)$ takes off from $\nu = 0$ with infinite slope, like this:



⟦You don't really need the value of "some positive number", but if you're insatiably curious, use the substitution $v = -\ln u$ to find

$$\int_0^1 \sqrt{-\ln u}\, du = \int_\infty^0 \sqrt{v}(-e^{-v})\, dv = \int_0^\infty v^{1/2} e^{-v}\, dv = \Gamma(\tfrac{3}{2}) = \frac{\sqrt{\pi}}{2},$$

so for small values of ν,

$$I(\nu) \approx \frac{\sqrt{\pi}}{2}\sqrt{\nu}.\ ⟧$$

A formal analysis shows that our integral $I(\nu)$ can be expressed in terms of gamma functions as

$$I(\nu) = \frac{\sqrt{\pi}}{2+\nu}\frac{\Gamma(\frac{1}{\nu})}{\Gamma(\frac{1}{\nu}+\frac{1}{2})},$$

but the graph actually tells you more than this formal expression does. When I was an undergraduate only a very few special functions (for example the Γ function) had been laboriously worked out numerically and tabulated, so it was important to express your integral of interest in terms of one of those few that had been worked out. Now numerical integration is a breeze (your phone is more powerful than the single computer we had on campus when I was an undergraduate), so it's more important to be able to tease information out of the function as we've done here.

In summary, the energy eigenvalues obtained through the quasiclassical approximation

$$(n - \tfrac{1}{2})\pi\hbar = \frac{(8m)^{1/2}}{\alpha^{1/\nu}} E^{(2+\nu)/2\nu} I(\nu)$$

are

$$E_n = \left[\frac{\alpha^{1/\nu}}{(8m)^{1/2} I(\nu)} (n - \tfrac{1}{2})\pi\hbar \right]^{2\nu/(2+\nu)} \qquad n = 1, 2, 3, \ldots. \tag{22.40}$$

You could spend a lot of time probing this equation to find out what it tells us about quantum mechanics. (You could also spend a lot of time looking at the quasiclassical wavefunctions.) I'll content myself with examining the energy eigenvalues for the three special cases $\nu = 2$, $\nu \to \infty$, and $\nu \to 0$.

When $\nu = 2$ the power-law potential $V(x) = \alpha x^2$ becomes the simple harmonic oscillator $V(x) = \frac{1}{2} m\omega^2 x^2$. Equation (22.40) becomes

$$\begin{aligned} E_n &= \frac{\alpha^{1/2}}{(8m)^{1/2} I(2)} (n - \tfrac{1}{2})\pi\hbar \\ &= \frac{(\frac{1}{2} m\omega^2)^{1/2}}{(8m)^{1/2} \pi/4} (n - \tfrac{1}{2})\pi\hbar \\ &= (n - \tfrac{1}{2})\hbar\omega \qquad n = 1, 2, 3, \ldots. \end{aligned} \tag{22.41}$$

The exact eigenvalues are of course

$$E_n = (n + \tfrac{1}{2})\hbar\omega \qquad n = 0, 1, 2, 3, \ldots.$$

For the simple harmonic oscillator, the quasiclassical energy eigenvalues are exactly correct. ⟦The energy eigenfunctions are not.⟧

When $\nu \to \infty$ the power-law potential becomes an infinite square well of width $L = 2$. Equation (22.40) becomes

$$\begin{aligned} E_n &= \left[\frac{\alpha^{1/\infty}}{(8m)^{1/2} I(\infty)} (n - \tfrac{1}{2})\pi\hbar \right]^2 \\ &= \left[\frac{1}{(8m)^{1/2}} (n - \tfrac{1}{2})\pi\hbar \right]^2 \\ &= \frac{\pi^2\hbar^2}{8m}(n - \tfrac{1}{2})^2. \end{aligned} \tag{22.42}$$

The exact eigenvalues are (when $L = 2$)

$$E_n = \frac{\pi^2\hbar^2}{2mL^2} n^2 = \frac{\pi^2\hbar^2}{8m} n^2.$$

Not bad for an approximation.

When $\nu \to 0$ the power-law potential becomes the flat, constant potential $V(x) = \alpha$. This "free particle" potential admits no bound states. How will the quasiclassical approximation deal with this?

$$\begin{aligned} E_n &= \left[\frac{\alpha^{1/\nu}}{(8m)^{1/2} I(\nu)} (n - \tfrac{1}{2})\pi\hbar \right]^{2\nu/(2+\nu)} \\ &= \frac{\alpha^{2/(2+\nu)}}{(8m)^{\nu/(2+\nu)}\, I(\nu)^{2\nu/(2+\nu)}} \left[(n - \tfrac{1}{2})\pi\hbar \right]^{2\nu/(2+\nu)} \\ &\to \frac{\alpha}{(8m)^0\, I(\nu)^\nu} \left[(n - \tfrac{1}{2})\pi\hbar \right]^0 \\ &\to \frac{\alpha}{I(\nu)^\nu}. \end{aligned}$$

But what is $I(\nu)^\nu$ for small ν? We've already seen at equation (22.39) that it is $\sqrt{\nu}^{\,\nu}$ (some positive number)$^\nu$. The right part goes to 1, but

$$\sqrt{\nu}^{\,\nu} = \nu^{\nu/2} = e^{\nu \ln \nu/2} \to e^0 = 1.$$

Thus as $\nu \to 0$,

$$E_n \to \alpha \qquad \text{for all values of } n. \tag{22.43}$$

Chapter 23

The Interaction of Matter and Radiation

Two questions:

(1) Our theorem says atoms stay in excited energy state forever!

(2) Absorb light of only one frequency ... what, will absorb light of wavelength 471.3428 nm but not 471.3427 nm?

Strangely, we start our quest to solve these problems by figuring out how to solve differential equations.

23.1 Perturbation Theory for the Time Evolution Problem

By now, you have realized that quantum mechanics is an art of approximations. I make no apologies for this: After all, physics is an art of approximations. (The classical "three-body problem" has never been solved exactly, and never will be.) Indeed, life is an art of approximations. (If you're waiting for the perfect boyfriend or girlfriend before making a commitment, you'll be waiting for a long time — and for some, that long wait is a poor solution to the problem of life.)

Furthermore, much of the fun and creativity of theoretical physics comes from finding applicable approximations. If theoretical physics were nothing but turning a mathematical crank to mechanically grind out solutions, it would not be exciting. I do not apologize for the fact that, to do theoretical physics, you have to think!

23.2 Setup

Here's our problem:

Solve the initial value problem for the Hamiltonian

$$\hat{H}(t) = \hat{H}^{(0)} + \hat{H}'(t) \tag{23.1}$$

given the solution $\{|n\rangle\}$ of the unperturbed energy eigenproblem

$$\hat{H}^{(0)}|n\rangle = E_n|n\rangle. \tag{23.2}$$

Here we're thinking of $\hat{H}'(t)$ as being in some sense "small" compared to the unperturbed Hamiltonian $\hat{H}^{(0)}$. One common example is a burst of light shining on an atom. Note also that it doesn't make sense to solve the energy eigenproblem for $\hat{H}(t)$, because this Hamiltonian depends upon time, so it doesn't have stationary state solutions!

We solve this problem by expanding the solution $|\psi(t)\rangle$ in the basis $\{|n\rangle\}$:

$$|\psi(t)\rangle = \sum_n C_n(t)|n\rangle \quad \text{where} \quad C_n(t) = \langle n|\psi(t)\rangle. \tag{23.3}$$

Once we know the $C_n(t)$, we'll know the solution $|\psi(t)\rangle$. Now, the state vector evolves according to

$$\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle \tag{23.4}$$

so the expansion coefficients evolve according to

$$\begin{aligned}
\frac{dC_n(t)}{dt} &= -\frac{i}{\hbar}\langle n|\hat{H}|\psi(t)\rangle \\
&= -\frac{i}{\hbar}\sum_m \langle n|\hat{H}|m\rangle C_m(t) \\
&= -\frac{i}{\hbar}\sum_m \left[\langle n|\hat{H}^{(0)}|m\rangle + \langle n|\hat{H}'|m\rangle\right] C_m(t) \\
&= -\frac{i}{\hbar}\sum_m \left[E_m\delta_{m,n} + H'_{n,m}\right] C_m(t) \\
&= -\frac{i}{\hbar}\left[E_nC_n(t) + \sum_m H'_{n,m}C_m(t)\right] \qquad (23.5)
\end{aligned}$$

This result is exact: we have yet to make any approximation.

Now, if $\hat{H}'(t)$ vanished, the solutions would be

$$C_n(t) = C_n(0)e^{-(i/\hbar)E_nt}, \tag{23.6}$$

which motivates us to define new variables $c_n(t)$ through

$$C_n(t) = c_n(t)e^{-(i/\hbar)E_nt}. \tag{23.7}$$

Because the "bulk of the time evolution" comes through the $e^{-(i/\hbar)E_n t}$ term, the $c_n(t)$ presumably have "less time dependence" than the $C_n(t)$. In other words, we expect the $c_n(t)$ to vary slowly with time.

Plugging this definition into the time evolution equation (23.5) gives

$$\begin{aligned}
&\frac{dc_n(t)}{dt} e^{-(i/\hbar)E_n t} + c_n(t) \left(-(i/\hbar)E_n\right) e^{-(i/\hbar)E_n t} \qquad (23.8)\\
&= -\frac{i}{\hbar}\left[E_n c_n(t) e^{-(i/\hbar)E_n t} + \sum_m H'_{n,m} c_m(t) e^{-(i/\hbar)E_m t}\right]
\end{aligned}$$

or

$$\frac{dc_n(t)}{dt} = -\frac{i}{\hbar}\sum_m H'_{n,m} c_m(t) e^{+(i/\hbar)(E_n - E_m)t}. \qquad (23.9)$$

Once again, this equation is exact. Its formal solution, given the initial values $c_n(0)$, is

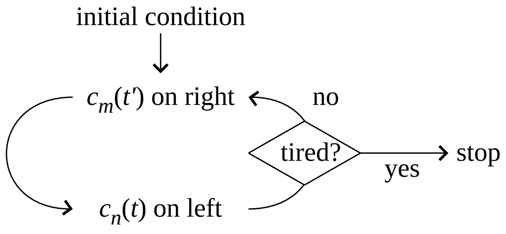
$$c_n(t) = c_n(0) - \frac{i}{\hbar}\sum_m \int_0^t H'_{n,m}(t') c_m(t') e^{+(i/\hbar)(E_n - E_m)t'} dt'. \qquad (23.10)$$

This set of equations (one for each basis member) is exact, but at first glance seems useless. The unknown quantities $c_n(t)$ are present on the left, but also the right-hand sides.

We make progress using our idea that the coefficients $c_n(t)$ are changing *slowly.* In a very crude approximation, we can think that they're not changing at all. So on the right-hand side of equation (23.10) we plug in not functions, but the constants $c_m(t') = c_m(0)$, namely the given initial conditions.

Having made that approximation, we can now perform the integrations and produce, on the left-hand side of equation (23.10), functions of time $c_n(t)$. These coefficients aren't exact, because they were based on the crude approximation that the coefficients were constant in time, but they're likely to be better approximations than we started off with.

Now, armed with these more accurate coefficients, we can plug these into the right-hand side of equation (23.10), perform the integration, and produce yet more accurate coefficients on the left-hand side. This process can be repeated over and over, for as long as our stamina lasts.



There is actually a theorem assuring us that this process will converge!

> *Theorem* (Picard[1]) If the matrix elements $H'_{n,m}(t)$ are continuous in time and bounded, and if the basis is finite, then this method converges to the correct solution.

The theorem does not tell us how many iterations will be needed to reach a desired accuracy. In practice, one usually stops upon reaching the first non-zero correction.

In particular, if the initial state is some eigenstate $|a\rangle$ of the unperturbed Hamiltonian $\hat{H}^{(0)}$, then to first order

$$c_n(t) = -\frac{i}{\hbar}\int_0^t H'_{n,a}(t')e^{+(i/\hbar)(E_n-E_a)t'}\,dt' \quad \text{for } n \neq a \qquad (23.11)$$

$$c_a(t) = 1 - \frac{i}{\hbar}\int_0^t H'_{a,a}(t')\,dt'$$

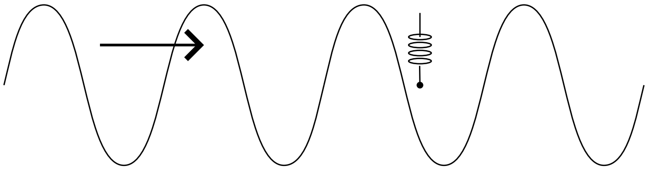
If the system is in energy state $|a\rangle$ at time zero, then the probability of finding it in energy state $|b\rangle$ at time t, through the influence of perturbation $\hat{H}'(t)$, is called the *transition probability*

$$P_{a\to b}(t) = |C_b(t)|^2 = |c_b(t)|^2. \qquad (23.12)$$

Example: An electron bound to an atom is approximated by a one-dimensional simple harmonic oscillator of natural frequency ω_0. The oscillator is in its ground state $|0\rangle$ and then exposed to light of electric field amplitude E_0 and frequency ω for time t. (The light is polarized in the direction of the oscillations.) What is probability (in first-order perturbation theory) of ending up in state $|b\rangle$?

---

[1] Émile Picard (1856–1941) made immense contributions to complex analysis and to the theory of differential equations. He wrote one of the first textbooks concerning the theory of relativity, and married the daughter of Charles Hermite.



Solution part A — What is the Hamiltonian? If it were a classical particle of charge $-e$ exposed to electric field $E_0 \sin \omega t$, it would experience a force $-eE_0 \sin \omega t$ and hence have a potential energy of $eE_0 x \sin \omega t$. (We can ignore the spatial variation of electric field because the electron is constrained to move only up and down — that's our "one dimensional" assumption. We can ignore magnetic field for the same reason.)

The quantal Hamiltonian is then

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega_0^2}{2}\hat{x}^2 + eE_0\hat{x}\sin \omega t. \tag{23.13}$$

We identify the first two terms as the time-independent Hamiltonian $\hat{H}^{(0)}$ and the last term as the perturbation $\hat{H}'(t)$.

Solution part B — Apply perturbation theory. The matrix element is

$$H_{n,0}(t) = \langle n|\hat{H}'(t)|0\rangle = eE_0 \sin \omega t \, \langle n|\hat{x}|0\rangle = eE_0 \sin \omega t \sqrt{\frac{\hbar}{2m\omega_0}}\delta_{n,1}. \tag{23.14}$$

(Remember your raising and lowering operators! See equation (G.31).) Invoking equations (23.11), we obtain

$$c_n(t) = 0 \quad \text{for } n \neq 0, 1 \tag{23.15}$$

$$c_1(t) = -\frac{i}{\hbar} eE_0 \sqrt{\frac{\hbar}{2m\omega_0}} \int_0^t \sin \omega t' e^{i\omega_0 t'} \, dt' \tag{23.16}$$

$$c_0(t) = 1 \tag{23.17}$$

We will eventually need to perform the time integral in equation (23.16), but even before doing so the main qualitative features are clear: First, probability is *not* conserved within first order perturbation theory. The probability of remaining in the ground state is 1, but the probability of transition to the first excited state is finite! Second, to first order transitions go only to the first excited state. This is an example of a *selection rule.*

The time integral in equation (23.16) will be evaluated at equation (23.30). For now, let's just call it $I(t)$. In terms of this integral, the transition probabilities are

$$P_{0\to b}(t) = 0 \quad \text{for } b \neq 0, 1 \tag{23.18}$$

$$P_{0\to 1}(t) = \frac{e^2 E_0^2}{2m\hbar\omega_0} I(t) I^*(t) \tag{23.19}$$

$$P_{0\to 0}(t) = 1 \tag{23.20}$$

23.3 Light absorption

How do atoms absorb light?

More specifically, if an electron in atomic energy eigenstate $|a\rangle$ (usually but not always the ground state) is exposed to a beam of monochromatic, polarized light for time t, what is the probability of it ending up in atomic energy eigenstate $|b\rangle$? We answer this question to first order in time-dependent perturbation theory.

First, we need to find the effect of light on the electron. We'll treat the light classically — that is, we'll ignore the quantization of the electromagnetic field (quantum electrodynamics) that gives rise to the concept of photons. Consider the light wave (polarized in the $\hat{k}$ direction, with frequency ω) as an electric field

$$\vec{E}(\vec{r}, t) = E_0 \hat{k} \sin(\vec{k} \cdot \vec{r} - \omega t). \tag{23.21}$$

Presumably, the absorption of light by the atom will result in some sort of diminution of the light beam's electric field, but we'll ignore that. (A powerful beam from a laser will be somewhat diminished when some of the light is absorbed by a single atom, but not a great deal.) The light beam has a magnetic field as well as an electric field, but the magnetic field amplitude is $B_0 = E_0/c$, so the electric force is on the order of eE_0 while the magnetic force is on the order of $evB_0 = e(v/c)E_0$. Since the electron moves at non-relativistic speeds, $v/c \ll 1$ and we can ignore the magnetic effect. Finally, the electric field at one side of the atom differs from the electric field at the other side of the atom, but the atom is so small compared to the wavelength of light (atom: about 0.1 nm; wavelength of violet light: about 400 nm) that we can safely ignore this also.

Using these approximations, the force experienced by an electron due to the light beam is

$$\vec{F}(t) = -eE_0 \hat{k} \sin(\omega t), \tag{23.22}$$

so the associated potential energy is

$$U(t) = eE_0 z \sin(\omega t). \tag{23.23}$$

Turning this classical potential energy into a quantal operator gives

$$\hat{H}'(t) = eE_0 \hat{z} \sin(\omega t). \tag{23.24}$$

(Note that the hat $\hat{k}$ in equation (23.22) signifies unit vector, whereas the hat $\hat{z}$ in equation (23.24) signifies quantal operator. I'm sorry for any confusion. . . there just aren't enough symbols in the world to represent everything unambiguously!)

Now that we have the quantal operator for the perturbation, we can turn to the time-dependent perturbation theory result (23.11). (Is it legitimate to use perturbation theory in this case? See the problem.)

For all of the atomic energy states $|a\rangle$ we've considered in this book,

$$H'_{a,a}(t) = \langle a|H'(t)|a\rangle = eE_0\langle a|\hat{z}|a\rangle \sin(\omega t) = 0, \tag{23.25}$$

whence $c_a(t) = 1$ and $P_{a\to a} = 1$. Most of the atoms *don't* make transitions.

But what about those that do? For these we need to find the matrix elements

$$H'_{b,a}(t) = \langle b|H'(t)|a\rangle = eE_0\langle b|\hat{z}|a\rangle \sin(\omega t). \tag{23.26}$$

These are just the $z_{b,a}$ matrix elements that we calculated for the Stark effect. (And after all, what we're considering here is just the Stark effect with an oscillating electric field.) The transition amplitudes are

$$c_b(t) = -\frac{i}{\hbar} eE_0\langle b|\hat{z}|a\rangle \int_0^t \sin(\omega t') e^{+(i/\hbar)(E_b-E_a)t'}\,dt'. \tag{23.27}$$

It is convenient (and conventional!) to follow the lead of Einstein's $\Delta E = \hbar\omega$ and define

$$E_b - E_a = \hbar\omega_0. \tag{23.28}$$

The time integral is then

$$\begin{aligned}
&\int_0^t \sin(\omega t') e^{i\omega_0 t'}\,dt' \\
&= \int_0^t \frac{e^{+i\omega t'} - e^{-i\omega t'}}{2i} e^{i\omega_0 t'}\,dt' \\
&= \frac{1}{2i}\left[\int_0^t e^{i(\omega_0+\omega)t'}\,dt' - \int_0^t e^{i(\omega_0-\omega)t'}\,dt'\right] \\
&= \frac{1}{2i}\left[\frac{e^{i(\omega_0+\omega)t'}}{i(\omega_0+\omega)} - \frac{e^{i(\omega_0-\omega)t'}}{i(\omega_0-\omega)}\right]_0^t \\
&= -\frac{1}{2}\left[\frac{e^{i(\omega_0+\omega)t} - 1}{\omega_0+\omega} - \frac{e^{i(\omega_0-\omega)t} - 1}{\omega_0-\omega}\right]
\end{aligned} \tag{23.29}$$

Enrico Fermi thought about this expression and realized that in most cases it would not be substantial (as reflected in the fact that $P_{a\to a} = 1$). The numerators are complex numbers in magnitude between 0 and 2. For light, we're thinking of frequencies ω near ZZZ. The only case when this expression is big, is when $\omega \approx \omega_0$, and when that's true only the right-hand part is big. So it's legitimate to ignore the left-hand part and write

$$\begin{aligned}
&\int_0^t \sin(\omega t')e^{i\omega_0 t'}\,dt' \\
&\approx -\frac{1}{2}\left[-\frac{e^{i(\omega_0-\omega)t}-1}{\omega_0-\omega}\right] \\
&= \frac{1}{2}\left[e^{i(\omega_0-\omega)t/2}\frac{e^{i(\omega_0-\omega)t/2}-e^{-i(\omega_0-\omega)t/2}}{\omega_0-\omega}\right] \\
&= \frac{1}{2}\left[e^{i(\omega_0-\omega)t/2}\frac{2i\sin((\omega_0-\omega)t/2)}{\omega_0-\omega}\right] \\
&= ie^{i(\omega_0-\omega)t/2}\frac{\sin((\omega_0-\omega)t/2)}{\omega_0-\omega} \\
&= ie^{-i(\omega-\omega_0)t/2}\frac{\sin((\omega-\omega_0)t/2)}{\omega-\omega_0}. \qquad (23.30)
\end{aligned}$$

Plugging this approximation for the integral into equation (23.27) produces

$$c_b(t) = \frac{eE_0\langle b|\hat{z}|a\rangle}{\hbar}e^{-i(\omega-\omega_0)t/2}\frac{\sin((\omega-\omega_0)t/2)}{\omega-\omega_0}. \qquad (23.31)$$

The transition probability is then

$$P_{a\to b} = \frac{e^2E_0^2|\langle b|\hat{z}|a\rangle|^2}{\hbar^2}\frac{\sin^2((\omega-\omega_0)t/2)}{(\omega-\omega_0)^2}. \qquad (23.32)$$

This rule, like all rules,[2] has limits on its applicability: we've already mentioned that it applies when the wavelength of light is much larger than an atom, when the light can be treated classically, when $\omega \approx \omega_0$, etc. Most importantly, it applies only when the transition probability is small, because when that probability is large the whole basis of perturbation theory breaks down. You might think that with all these restrictions, it's not a very important result. You'd be wrong. In fact Fermi used it so often that he called it "the golden rule."

---

[2] A father needs to leave his child at home for a short time. Concerned for his child's safety, he issues the sensible rule "Don't leave home while I'm away." While the father is away, the home catches fire. Should the child violate the rule?

Physical implications of Fermi's golden rule

We have derived Fermi's golden rule, but that's only the start and not the end of our quest to answer the question of "How do atoms absorb light?". What does Fermi's golden rule say about nature? First, we'll think of the formula as a function of frequency ω for fixed time t, then we'll think of the formula as a function of time t at fixed frequency ω.

Write the transition probability as

$$P_{a\to b} = A\frac{\sin^2((\omega-\omega_0)t/2)}{(\omega-\omega_0)^2} \tag{23.33}$$

where the value of A is independent of both frequency and time. Clearly, this expression is always positive or zero (good thing!) and is symmetric about the natural transition frequency ω_0. The expression is always less then the time-independent "envelope function" $A/(\omega-\omega_0)^2$. The transition probability vanishes when

$$\omega-\omega_0 = N\pi/t, \qquad N = \pm 2, \pm 4, \pm 6, \ldots$$

while it touches the envelope when

$$\omega-\omega_0 = N\pi/t, \qquad N = \pm 1, \pm 3, \pm 5, \ldots.$$

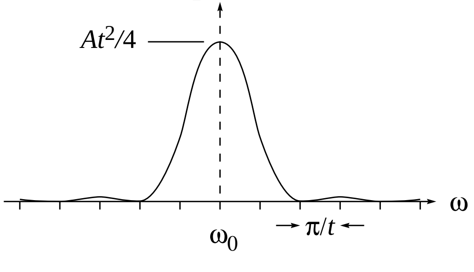
What about when $\omega = \omega_0$? Here you may use l'Hôpital's rule, or the approximation

$$\sin\theta \approx \theta \qquad \text{for } \theta \ll 1,$$

but either way you'll find that

$$\text{when } \omega = \omega_0, \quad P_{a\to b} = At^2/4. \tag{23.34}$$

In short, the transition probability as a function of ω looks like this graph:



> **Problem:** Show that if the central maximum has value P_{max}, then the first touching of the envelope (at $\omega - \omega_0 = \pi/t$) has value $(4/\pi^2)P_{\text{max}} = 0.405\,P_{\text{max}}$, the second touching (at $\omega - \omega_0 = 3\pi/t$) has value $(4/9\pi^2)P_{\text{max}} = 0.045\,P_{\text{max}}$, and the third (at $\omega - \omega_0 = 5\pi/t$) has value $(4/25\pi^2)P_{\text{max}} = 0.016\,P_{\text{max}}$. Notice that these ratios are independent of time.

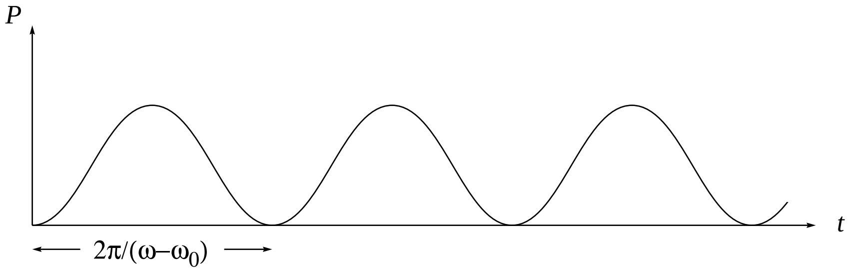
There are several unphysical aspects of this graph it gives a result even at $\omega = 0$ ... indeed, even when ω is negative! But the formula was derived assuming $\omega \approx \omega_0$, so we don't expect it to give physically reasonable results in this regime. In time, the maximum transition probability $At^2/4$ will grow to be very large, in fact even larger than one! But the formula was derived assuming a small transition probability, and becomes invalid long before such an absurdity happens.

This result may help you with a conundrum. You have perhaps been told something like: "To excite hydrogen from the ground state to the first excited state, a transition with $\Delta E = \frac{1}{4}\text{Ry}$, you must supply a photon with energy exactly equal to $\frac{1}{4}\text{Ry}$, what is with frequency $\omega_0 = \frac{1}{4}\text{Ry}/\hbar$, or in other words with wavelength 364.506 820 nm." You know that no laser produces light with the *exact* wavelength of 364.506 820 nm. If the photon had to have *exactly* that wavelength, there would almost never be a transition. But the laser doesn't need to have *exactly* that wavelength: as you can see, there's some probability of absorbing light that differs a bit from the natural frequency ω_0.

> **Problem:** Show that the width of the central peak, from zero to zero, is $4\pi/t$.

One aspect of the transition probability expression is quite natural: The light most effective at promoting a transition is light with frequency ω equal to the transition's natural frequency ω_0. Also natural is that the effectiveness decreases as ω moves away from ω_0, until the transition probability vanishes entirely at $\omega = \omega_0 \pm 2\pi/t$. But then a puzzling phenomenon sets in: as ω moves still further away from ω_0, the transition probability *increases*. This increase is admittedly slight, but nonetheless it exists, and I know of no way to explain it in physical terms. I do point out, however, that this puzzling phenomenon does *not* exist for light pulses of Gaussian form: see problem 23.5, "Gaussian light pulse".

Now, investigate the formula (23.33) as a function of time t at fixed light frequency ω. This seems at first to be a much simpler task, because the graph is trivial:



But now reflect upon the graph. We have a laser set to make transitions from $|a\rangle$ to $|b\rangle$. We turn on the laser, and the probability of that transition increases. So far, so good. Now we keep the laser on, but the probability decreases! And if we keep it on for exactly the right amount of time, there is zero probability for a transition. It's as if we were driving a nail into a board with a hammer. The first few strikes push the nail into the board, but with continued strikes the nail backs out of the board, and it eventually pops out altogether!

How can this be? Certainly, no nail that I've hammered has ever behaved this way! The point is that there are two routes to get from $|a\rangle$ to $|a\rangle$: You can go from $|a\rangle$ to $|b\rangle$ and then back to $|a\rangle$, or you can stay always in $|a\rangle$, that is go from $|a\rangle$ to $|a\rangle$ to $|a\rangle$. There is an amplitude associated with each route. If these two amplitudes interfere constructively, there is a high probability of remaining in $|a\rangle$ (a low probability of transitioning to $|b\rangle$). If these two amplitudes interfere destructively, there is a low probability of remaining in $|a\rangle$ (a high probability of transitioning to $|b\rangle$). This wavy graph is a result of interference of two routes that are, not paths in position space, but routes through energy eigenstates.[3]

This phenomenon is called "Rabi oscillation", and it's the pulse at the heart of an atomic clock.

---

[3] This point of view is developed extensively in R.P. Feynman and A.R. Hibbbs, *Quantum Mechanics and Path Integrals* (D.F. Styer, emending editor, Dover Publications, Mineola, New York, 2010) pages 116–117, 144–147.

23.4 Absorbing incoherent light

For coherent, z-polarized, x-directed, long-wavelength, non-magnetic, classical, non-diminishing light, in the approximation of first-order time-dependent perturbation theory, and with $\omega \approx \omega_0$, the transition probability is

$$P_{a\to b} = \frac{e^2 E_0^2}{\hbar^2} |\langle b|\hat{z}|a\rangle|^2 \frac{\sin^2((\omega-\omega_0)t/2)}{(\omega-\omega_0)^2}. \tag{23.35}$$

The classical energy density (average energy per volume) of an electromagnetic wave is $u = \epsilon_0 E_0^2/2$, where ϵ_0 is the famous vacuum permittivity that appears as $1/(4\pi\epsilon_0)$ in Coulomb's law, so this result is often written

$$P_{a\to b} = \frac{2e^2 u}{\epsilon_0 \hbar^2} |\langle b|\hat{z}|a\rangle|^2 \frac{\sin^2((\omega-\omega_0)t/2)}{(\omega-\omega_0)^2}. \tag{23.36}$$

What if the light is polarized but not coherent? In this case light comes at varying frequencies. Writing the energy density per frequency as $\rho(\omega)$, the transition probability due to light of frequency ω to $\omega + d\omega$ is

$$P'_{a\to b} = \frac{2e^2 \rho(\omega)\, d\omega}{\epsilon_0 \hbar^2} |\langle b|\hat{z}|a\rangle|^2 \frac{\sin^2((\omega-\omega_0)t/2)}{(\omega-\omega_0)^2}, \tag{23.37}$$

whence the total transition probability is

$$P_{a\to b} = \frac{2e^2}{\epsilon_0 \hbar^2} |\langle b|\hat{z}|a\rangle|^2 \int_0^\infty \frac{\sin^2((\omega-\omega_0)t/2)}{(\omega-\omega_0)^2} \rho(\omega)\, d\omega. \tag{23.38}$$

⟦We have assumed that the light components at various frequencies is independent, so that the total transition probability is the sum of the individual transition probabilities. If instead the light components were completely correlated, then the total transition *amplitude* would be the sum of the individual transition *amplitudes*. This is the case in problem 23.5, "Gaussian light pulse". If the light components were incompletely correlated but not completely independent, then a hybrid approach would be needed.⟧ If $\rho(\omega)$ is slowly varying relative to the absorption profile (23.33) — which it almost always is — then it is accurate to approximate

$$P_{a\to b} = \frac{2e^2}{\epsilon_0 \hbar^2} |\langle b|\hat{z}|a\rangle|^2 \rho(\omega_0) \int_{-\infty}^{+\infty} \frac{\sin^2((\omega-\omega_0)t/2)}{(\omega-\omega_0)^2}\, d\omega, \tag{23.39}$$

where I have changed the lower integration limit from 0 to $-\infty$, with negligible change in $P_{a\to b}$, because the integrand nearly vanishes whenever $\omega < 0$. Finally, the definite integral

$$\int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2}\, dx = \pi$$

gives, for polarized incoherent light,

$$P_{a\to b} = \frac{\pi e^2}{\epsilon_0 \hbar^2} |\langle b|\hat{z}|a\rangle|^2 \rho(\omega_0) t. \tag{23.40}$$

The primary thing to note about this formula is the absence of Rabi oscillations: it gives a far more familiar *rate* of transition. The second thing is that the rate from $|b\rangle$ to $|a\rangle$ is equal to the rate from $|a\rangle$ to $|b\rangle$, which is somewhat unusual: you might think that the rate to lose energy ($|b\rangle$ to $|a\rangle$) should be greater than the rate to gain energy ($|a\rangle$ to $|b\rangle$). [Just as it's easier to walk down a staircase than up the same staircase.]

Finally, what if the light is not coherent, not polarized, and not directed? (Such as the light in a room, that comes from all directions.) In this case

$$P_{a\to b} = \frac{\pi e^2}{3\epsilon_0 \hbar^2} \left[|\langle b|\hat{x}|a\rangle|^2 + |\langle b|\hat{y}|a\rangle|^2 + |\langle b|\hat{z}|a\rangle|^2 \right] \rho(\omega_0) t. \tag{23.41}$$

23.5 Absorbing and emitting light

Qualitative quantum electrodynamics

Of course we want to do better than the treatment above: Instead of treating a quantum mechanical atom immersed in a classical electromagnetic field, we want a full quantum-mechanical treatment of the atom *and* the light. Such a theory — quantum electrodynamics — has been developed and it is a beautiful thing. Because light must travel at speed c this theory is intrinsically relativistic and, while beautiful, also a very difficult thing. We will not give it a rigorous treatment in this book. But this section motivates the theory and discusses its qualitative character.

Most of this book discusses the quantum mechanics of atoms: The Hamiltonian operator $\hat{H}_{\text{atom}}$ has energy eigenstates like the ground state $|a\rangle$ and the excited state $|b\rangle$. The system can exist in any linear combination of these states, such as $(|a\rangle - |b\rangle)/\sqrt{2}$. If the system starts off in one of the energy states, including the excited state $|b\rangle$, it stays there forever.

You can also write down a Hamiltonian operator $\hat{H}_{\text{EM}}$ for the electromagnetic field. This operator has energy eigenstates. By convention, the ground state is called $|\text{vacuum}\rangle$, one excited state is called $|1 \text{ photon}\rangle$, an even more excited state is called $|2 \text{ photons}\rangle$. The field can also exist in linear combinations such as $(|\text{vacuum}\rangle - |2\text{ photons}\rangle)/\sqrt{2}$, but this state is *not* a stationary state, and it does *not* have an energy.

You can do the classic things with field energy states: There's an operator for energy and an operator for photon position, but they don't commute. So in the state $|1\text{ photon}\rangle$ the photon has an energy but no position. There's a linear combinations of energy states in which th photon *does* have a position, but in these position states the electromagnetic field has no energy.

But there's even more: There is an operator for electric field at a given location. And this operator doesn't commute with either the Hamiltonian or with the photon position operator.[4] So in a state of electric field at some given point, the photon does not have a position, and does not have an energy. Anyone thinking of the photon as a "ball of light" — a wavepacket of electric and magnetic fields — is thinking of a misconception. A photon might have a "pretty well defined" position and a "pretty well defined" energy and a "pretty well defined" field, but it can't have an exact position and an exact energy and an exact field at the same time.

If the entire Hamiltonian were $\hat{H}_{\text{atom}} + \hat{H}_{\text{EM}}$, then energy eigenstates of the atom plus field would have the character of $|a\rangle|2\text{ photons}\rangle$, or $|b\rangle|\text{vacuum}\rangle$ and if you started off in such a state you would stay in it forever. Note particularly the second example: if the atom started in an excited state, it would never decay to the ground state, emitting light.

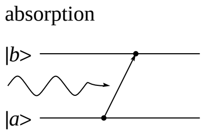
But since that process (called "spontaneous emission") *does* happen, the Hamiltonian $\hat{H}_{\text{atom}} + \hat{H}_{\text{EM}}$ *must not* be the whole story. There must be some additional term in the Hamiltonian that involves both the atom and the field: This term is called the "interaction Hamiltonian" $\hat{H}_{\text{int}}$. (Sometimes called the "coupling Hamiltonian", because it couples — connects — the atom and the field.) The full Hamiltonian is $\hat{H}_{\text{atom}} + \hat{H}_{\text{EM}} + \hat{H}_{\text{int}}$. The state $|b\rangle|\text{vacuum}\rangle$ is *not* an eigenstate of this full Hamiltonian: If you start off in $|b\rangle|\text{vacuum}\rangle$, then at a later time there will be some amplitude to remain in $|b\rangle|\text{vacuum}\rangle$, but also some amplitude to be in $|a\rangle|1\text{ photon}\rangle$.

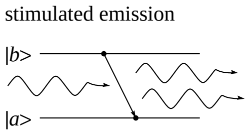
---

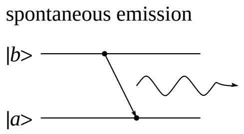
[4] It's clear, even without writing down the "EM field Hamiltonian" and the "electric field at a given point" operators, that they do not commute: any operator that commutes with the Hamiltonian is conserved, so if these two operators commuted then the electric field at a given point would never change with time!

Einstein A and B argument

Back in 1916, Einstein wanted to know about both absorption *and* emission of light by atoms, and — impatient as always — he didn't want to wait until a full theory of quantum electrodynamics was developed. So he came up with the following argument — one of the cleverest in all of physics.







Einstein said that there were three processes going on, represented schematically in the figure above. In *absorption* of radiation the atom starts in its ground state $|a\rangle$ and ends in excited state $|b\rangle$, while the light intensity at frequency ω_0 is reduced. Although the reasoning leading to equation (23.41) hadn't yet been performed in 1916, Einstein thought it reasonable that the probability of absorption would be given by some rate coefficient B_{ab}, times the energy density of radiation with the proper frequency for exciting the atom, times the time:

$$P_{a\to b} = B_{ab}\,\rho(\omega_0)\,t. \tag{23.42}$$

In *stimulated emission* the atom starts in excited state $|b\rangle$ and, under the influence of light, ends in ground state $|a\rangle$. After this happens the light intensity at frequency ω_0 increases due to the emitted light. In this process the incoming light of frequency ω_0 "shakes" the atom out of its excited state. Einstein thought the probability for this process would be

$$P_{b\to a} = B_{ba}\,\rho(\omega_0)\,t. \tag{23.43}$$

We know, from equation (23.41), that in fact $B_{ba} = B_{ab}$, but Einstein didn't know this so his argument doesn't use this fact.

Finally, in *spontaneous emission* the atom starts in excited state $|b\rangle$ and ends in ground state $|a\rangle$, but it does so without any incoming light to "shake" it. After spontaneous emission the light intensity at frequency ω_0 increases due to the emitted light. Because this process doesn't rely on incoming light, the probability of it happening doesn't depend on $\rho(\omega_0)$. Instead, Einstein thought, the probability would be simply

$$P'_{b\to a} = At. \tag{23.44}$$

Einstein knew that this process had to happen, because excited atoms in the dark can give off light and go to their ground state, but he didn't have a theory of quantum electrodynamics that would enable him to calculate the rate coefficient A.

The coefficients B_{ab}, B_{ba}, and A are independent of the properties of the light, the number of atoms in state $|a\rangle$, the number of atoms in state $|b\rangle$, etc. — they depend only upon the characteristics of the atom.

Now if you have a bunch of atoms, with N_a of them in the ground state and N_b in the excited state, the rate of change of N_a through these three processes is

$$\frac{dN_a}{dt} = -B_{ab}\,\rho(\omega_0)\,N_a + B_{ba}\,\rho(\omega_0)\,N_b + AN_b. \tag{23.45}$$

In equilibrium, by definition,

$$\frac{dN_a}{dt} = 0. \tag{23.46}$$

In addition, in thermal equilibrium at temperature T, the following two facts are true: The first is called "Boltzmann distribution"

$$\frac{N_b}{N_a} = e^{-(E_b-E_a)/k_BT} = e^{-\hbar\omega_0/k_BT}, \tag{23.47}$$

where k_B is the so-called "Boltzmann constant" that arises frequently in thermal physics. The second is called "energy density for light in thermal equilibrium (backbody radiation)"

$$\rho(\omega) = \frac{\hbar}{\pi^2c^3}\frac{\omega^3}{e^{\hbar\omega/k_BT}-1}, \tag{23.48}$$

where c is the speed of light. [If you have taken a course in statistical mechanics, you have certainly seen the first result. You might think you haven't seen the second result, but in fact it is a property of the ideal Bose gas when the chemical potential μ vanishes.]

You might not yet know these two facts, but Einstein did. He combined equation (23.46) and equation (23.45) finding

$$\rho(\omega_0) = \frac{AN_b}{B_{ab}N_a - B_{ba}N_b}.$$

Then he used the Boltzmann distribution (23.47) to produce

$$\rho(\omega_0) = \frac{A}{B_{ab}e^{\hbar\omega_0/k_BT} - B_{ba}} \tag{23.49}$$

and compared that to the blackbody result (23.48) producing

$$\frac{A}{B_{ab}e^{\hbar\omega_0/k_BT} - B_{ba}} = \frac{\hbar}{\pi^2 c^3}\frac{\omega_0^3}{e^{\hbar\omega_0/k_BT} - 1}.$$

This result must hold for all temperatures T, and the coefficients B_{ab}, B_{ba}, and A are independent of T. Thus, Einstein reasoned, we must have

$$B_{ab} = B_{ba} \equiv B \tag{23.50}$$

(which we already knew, but which was a discovery to Einstein) and hence

$$\frac{A}{B(e^{\hbar\omega_0/k_BT} - 1)} = \frac{\hbar}{\pi^2 c^3}\frac{\omega_0^3}{e^{\hbar\omega_0/k_BT} - 1}$$

or, with temperature-dependent parts canceling on both sides,

$$\frac{A}{B} = \frac{\hbar\omega_0^3}{\pi^2 c^3}. \tag{23.51}$$

The result is, of necessity, independent of temperature T. Einstein's argument uses thermal equilibrium not to discover the macroscopic properties of matter, but as a vehicle to uncover microscopic details about the relation between matter and radiation. We have no way to find A from first principles, but from the fact that thermal equilibrium exits we can find A through

$$A = \frac{\hbar\omega_0^3}{\pi^2 c^3}B = \frac{4h}{\lambda_0^3}B. \tag{23.52}$$

I hope you find this argument as astounding, and as beautiful, as I do. It has the character of Einstein: First, it is not technically difficult, but it combines the various features in a way that I never would have thought of, to produce a result that I thought would require working out full theory of quantum electrodynamics. Second, it turns the problem on its head: The fundamental question is "Will microscopic actions always result in macroscopic thermal equilibrium? If so, how fast will that equilibrium be approached?" Einstein skips over the fundamental question and asks "We know from observation that macroscopic thermal equilibrium does in fact exist. How can we exploit this fact to find out about microscopic actions?"

Numerical example: I would expect the stimulated decay rate $B\rho(\omega_0)$ to exceed the spontaneous emission rate A (just as a jar on a shelf is more likely to fall off when shaken than when left alone). On the other hand I've found my expectations violated by quantum mechanics so frequently that I can't be sure. What is the ratio of A to $B\rho(\omega_0)$ at room temperature for the transition associated with the red light of a Helium-Neon laser ($\lambda_0 =$ 633 nm)?

Use equation (23.49) to write

$$\frac{B\rho(\omega_0)}{A} = \frac{1}{e^{\hbar\omega_0/k_BT} - 1}. \tag{23.53}$$

Now at room temperature, $k_BT = \frac{1}{40}$ eV, so

$$\frac{\hbar\omega_0}{k_BT} = \frac{hc}{\lambda_0 k_BT} = \frac{1240 \text{ eV}\cdot\text{nm}}{(633 \text{ nm})(\frac{1}{40} \text{ eV})} = 78$$

resulting in

$$\frac{B\rho(\omega_0)}{A} = \frac{1}{e^{78} - 1} = e^{-78} = 10^{-34}.$$

My intuition about shaking has been vindicated! At what temperature will the stimulated and spontaneous rates be equal?

23.6 Problems

23.1 **On being kicked upstairs**

A particle in the ground state of an infinite square well is perturbed by a transient effect described by the Hamiltonian (in coordinate representation)

$$H'(x,t) = A_0 \sin\left(\frac{2\pi x}{L}\right)\delta(t), \tag{23.54}$$

where A_0 is a constant with the dimensions of action. What is the probability that after this jolt an energy measurement will find the system in the first excited state?

23.2 **Second-order time-dependent perturbation theory**

At equation (23.16) we treated, to first order in perturbation theory, the problem of a simple harmonic oscillator in its ground state exposed to a sinusoidal external force (with frequency ω and amplitude eE_0). We concluded that the only non-vanishing first-order transition amplitudes were $c_0^{(1)}(t) = 1$ and $c_1^{(1)}(t)$. (Here the superscript (1) denotes "first-order".) Show that to second order the non-vanishing transition amplitudes are:

$$c_0^{(2)}(t) = 1 - \frac{i}{\hbar}\int_0^t H'_{01}(t')e^{-i\omega_0 t'}c_1^{(1)}(t')\,dt', \tag{23.55}$$

$$c_1^{(2)}(t) = -\frac{i}{\hbar}\int_0^t H'_{10}(t')e^{+i\omega_0 t'}c_0^{(1)}(t')\,dt', \tag{23.56}$$

$$c_2^{(2)}(t) = -\frac{i}{\hbar}\int_0^t H'_{21}(t')e^{+i\omega_0 t'}c_1^{(1)}(t')\,dt', \tag{23.57}$$

where

$$H'_{01}(t) = H'_{10}(t) = eE_0\sqrt{\frac{\hbar}{2m\omega_0}}\,\sin(\omega t), \tag{23.58}$$

and

$$H'_{21}(t) = eE_0\sqrt{\frac{2\hbar}{2m\omega_0}}\,\sin(\omega t). \tag{23.59}$$

The integrals for $c_0^{(2)}(t)$ and $c_2^{(2)}(t)$ are not worth working out, but it is worth noticing that $c_2^{(2)}(t)$ involves a factor of $(eE_0)^2$ (where eE_0 is in some sense "small"), and that $c_1^{(2)}(t) = c_1^{(1)}(t)$.

23.3 **Is light a perturbation?**
Is it legitimate to use perturbation theory in the case of light absorbed by an atom? After all, we're used to thinking of the light from a powerful laser as a big effect, not a tiny perturbation. However, whether an effect is big or small depends on context. Estimate the maximum electric field due to a laser of XX watts, and the electric field at an electron due to its nearby nucleus. Conclude that while the laser is very powerful on a human scale (and you should *not* stick your eye into a laser beam), it is nevertheless very weak on an atomic scale.

23.4 **Magnitude of transitions**
At equation (23.33) we defined

$$A \equiv \frac{e^2E_0^2|\langle b|\hat{z}|a\rangle|^2}{\hbar^2}$$

and then noted that it was independent of ω and t, but otherwise ignored it. (Although we used it when we said that the maximum transition probability was $At^2/4$.) This problem investigates the character of A.

The maximum classical force on the electron due to light is eE_0. A typical force is less, so define the characteristic force due to light as

$$F_{c,L} \equiv \tfrac{1}{2}eE_0.$$

A typical classical force on the electron due to the nucleus is

$$F_{c,N} \equiv \left(\frac{e^2}{4\pi\epsilon_0}\right)\frac{1}{a_0^2}.$$

Using these two definitions, and taking a typical matrix element $|\langle b|\hat{z}|a\rangle|$ to be a_0, show that a typical value of A is

$$4\left(\frac{F_{c,L}}{F_{c,N}}\right)^2\frac{1}{\tau_0^2}.$$

If this excites you, you may also show that the exact value is

$$A = 4\left(\frac{F_{c,L}}{F_{c,N}}\right)^2\frac{1}{\tau_0^2}\left|\frac{\langle b|\hat{z}|a\rangle}{a_0}\right|^2.$$

23.5 **Gaussian light pulse**

An atom is exposed to a Gaussian packet of light

$$E(t) = E_0 e^{-t^2/\tau^2}\sin(\omega t). \tag{23.60}$$

At time $t = -\infty$, the atom was in state $|a\rangle$. Find the amplitude, to first order in perturbation theory, that at time $t = \infty$ the atom is in state $|b\rangle$. *Clue:* Use the Gaussian integral (G.8). *Answer:*

$$c_b = \frac{eE_0\langle b|\hat{z}|a\rangle}{\hbar}\left[-\frac{\sqrt{\pi}}{2\tau}\right]\left[e^{-\tau^2(\omega+\omega_0)/4} + e^{-\tau^2(\omega-\omega_0)/4}\right].$$

Chapter 24

The Territory Ahead

> I reckon I got to light out for the territory ahead. . .
> — Mark Twain (last sentence of *Huckleberry Finn*)

This is the last chapter of the book, but not the last chapter of quantum mechanics. There are many fascinating topics that this book hasn't even touched on. Quantum mechanics will — if you allow it — surprise and delight and mystify you for the rest of your life.

This book started by considering **qubits**, also called spin-$\frac{1}{2}$ systems. Plenty remains to investigate: "which path" interference experiments, delayed-choice interference experiments, many different entanglement situations. For example, we developed entanglement through a situation where the quantal probability was $\frac{1}{2}$ while the local deterministic probability was $\frac{5}{9}$ or more (page 47). Different, to be sure, but not dramatically different. In the Greenberger–Horne–Zeilinger entanglement situation the quantal probability is 1 and the local deterministic probability is 0. You can't find probabilities more different than that! If you find these situations as fascinating as I do, then I recommend George Greenstein and Arthur G. Zajonc, *The Quantum Challenge: Modern Research on the Foundations of Quantum Mechanics*.

For many decades, research into qubits yielded insight and understanding, but no practical applications. All that changed with the advent of quantum computing. This is a rapidly changing field, but the essay "Quantum Entanglement: A Modern Perspective" by Barbara M. Terhal, Michael M. Wolf, and Andrew C. Doherty (*Physics Today*, April 2003) contains core insights that will outlive any transient. From the abstract: "It's not your grandfather's quantum mechanics. Today, researchers treat entanglement as a physical resource: Quantum information can now be measured, mixed, distilled, concentrated, and diluted."

Because quantum mechanics is both intricate and unfamiliar, a formidable yet beautiful mathematical **formalism** has developed around it: position wavefunctions, momentum wavefunctions, Fourier transforms, operators, Wigner functions. These are powerful precision tools, so magnificent that some confuse the tools with nature itself. This textbook has started but not finished that development. I also recommend the cute book by Leonard Susskind and Art Friedman, *Quantum Mechanics: The Theoretical Minimum.*

We have applied quantum mechanics to cryptography, to model systems, to atoms, and to molecules. **Applications** continue to solids, to nuclei and to elementary particles, to superfluids, superconductors, and lasers, to liquid crystals, polymers, and membranes; the list is endless. Indeed, sunlight itself is generated through a quantal tunneling process! White dwarf stars work because of quantum mechanics, so do transistors and light-emitting diodes. In 1995 a new state of matter, the Bose-Einstein condensate, came into existence in a laboratory in Boulder, Colorado. In 2003 an even more delicate state, the fermionic condensate, was produced, again in Boulder. Both of these states of matter exist because of the Pauli principle, applied over and over again to millions of atoms.

Way back on page 3 we mentioned the need for a **relativistic quantum mechanics** and its associate, quantum field theory. The big surprise is that these theories don't just treat particles moving from place to place. They predict that particles can be created and destroyed, and sure enough that happens in nature under appropriate conditions.

There's **plenty more to investigate**: quantal chaos and the classical limit of quantum mechanics, friction and the transition to ground state, applications to astrophysics and cosmology and elementary particles.

But I want to close with one important yet rarely mentioned item: it's valuable to **develop your intuition** concerning quantum mechanics. Hilbert said[1] that "clearness and ease of comprehension" were required before a mathematical theory could be considered complete. Quantum theory has not yet reached this standard. On page 49 we found that *no* picture drawn with classical ink could successfully capture all aspects of

---

[1] David Hilbert, "Mathematical Problems" translation by Maby Winton Newson appearing in *Bulletin of the American Mathematical Society* **8** (1902), 437–479.

quantum mechanics. How, then, can one develop a visualization or intuition for quantum mechanics? This is a lifelong journey which you have already begun. A good next step is to read the slim but profound book by Richard Feynman titled *QED: The Strange Theory of Light and Matter*.

None of this is to denigrate what you already know, because all of these extensions and elaborations fall solidly within the amplitude framework developed in this book. Much remains to be discovered, and I hope that you will do some of that discovery yourself.

Problem

24.1 **Questions** (recommended problem)
This is the end of the book, not the end of quantum mechanics. Write down any questions you have concerning quantum mechanics. Perhaps you will answer some of these through future study. Others might suggest future research directions for you.

Appendix A

Tutorial on Matrix Diagonalization

You know from as far back as your introductory mechanics course that some problems are difficult given one choice of coordinate axes and easy or even trivial given another. (For example, the famous "monkey and hunter" problem is difficult using a horizontal axis, but easy using an axis stretching from the hunter to the monkey.) The mathematical field of linear algebra is devoted, in large part, to systematic techniques for finding coordinate systems that make problems easy. This tutorial introduces the most valuable of these techniques. It assumes that you are familiar with matrix multiplication and with the ideas of the inverse, the transpose, and the determinant of a square matrix. It is also useful to have a nodding acquaintance with the inertia tensor.

This presentation is intentionally non-rigorous. A rigorous, formal treatment of matrix diagonalization can be found in any linear algebra textbook,[1] and there is no need to duplicate that function here. What is provided here instead is a heuristic picture of what's going on in matrix diagonalization, how it works, and why anyone would want to do such a thing anyway. Thus this presentation complements, rather than replaces, the logically impeccable ("bulletproof") arguments of the mathematics texts.

Essential problems in this tutorial are marked by asterisks (*).

A.1 What's in a name?

There is a difference between an entity and its name. For example, a tree is made of wood, whereas its name "tree" made of ink. One way to see this is to note that in German, the name for a tree is "Baum", so the name changes upon translation, but the tree itself does not change. (Throughout this tutorial, the term "translate" is used as in "translate from one language to another" rather than as in "translate by moving in a straight line".)

The same holds for mathematical entities. Suppose a length is represented by the number "2" because it is two feet long. Then the same length is represented by the number "24" because it is twenty-four inches long. The same length is represented by two different numbers, just as the same tree has two different names. The representation of a length as a number depends not only upon the length, but also upon the coordinate system used to measure the length.

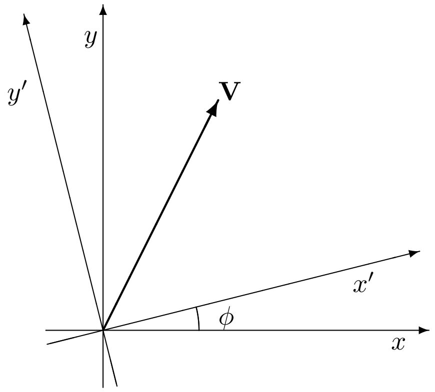
A.2 Vectors in two dimensions

One way of describing a two-dimensional vector **V** is by giving its x and y components in the form of a 2×1 column matrix

$$\begin{pmatrix} V_x \\ V_y \end{pmatrix}. \tag{A.1}$$

Indeed it is sometimes said that the vector **V** is *equal* to the column matrix (A.1). This is not precisely correct—it is better to say that the vector is *described by* the column matrix or *represented by* the column matrix or that its *name* is the column matrix. This is because if you describe the vector using a different set of coordinate axes you will come up with a different column matrix to describe the same vector. For example, in the situation shown below the descriptions in terms of the two different coordinate systems are related through the matrix equation

$$\begin{pmatrix} V_{x'} \\ V_{y'} \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} V_x \\ V_y \end{pmatrix}. \tag{A.2}$$



The 2×2 matrix above is called the "rotation matrix" and is usually denoted by $\mathsf{R}(\phi)$:

$$\mathsf{R}(\phi) \equiv \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}. \tag{A.3}$$

One interesting property of the rotation matrix is that it is always invertible, and that its inverse is equal to its transpose. Such matrices are called *orthogonal.*[2] You could prove this by working a matrix multiplication, but it is easier to simply realize that the inverse of a rotation by ϕ is simply a rotation by $-\phi$, and noting that

$$\mathsf{R}^{-1}(\phi) = \mathsf{R}(-\phi) = \mathsf{R}^{\dagger}(\phi). \tag{A.4}$$

(The dagger represents matrix transposition.)

There are, of course, an infinite number of column matrix representations for any vector, corresponding to the infinite number of coordinate axis rotations with ϕ from 0 to 2π. But one of these representations is special: It is the one in which the x'-axis lines up with the vector, so the column matrix representation is just

$$\begin{pmatrix} V \\ 0 \end{pmatrix}, \tag{A.5}$$

---

[2] Although all rotation matrices are orthogonal, there are orthogonal matrices that are not rotation matrices: see problem A.4.

where $V = |\mathbf{V}| = \sqrt{V_x^2 + V_y^2}$ is the magnitude of the vector. This set of coordinates is the preferred (or "canonical") set for dealing with this vector: one of the two components is zero, the easiest number to deal with, and the other component is a physically important number. You might wonder how I can claim that this representation has full information about the vector: The initial representation (A.1) contains two independent numbers, whereas the preferred representation (A.5) contains only one. The answer is that the preferred representation contains one number (the magnitude of the vector) explicitly while another number (the polar angle of the vector relative to the initial x-axis) is contained implicitly in the rotation needed to produce the preferred coordinate system.

A.1 **Problem: Right angle rotations**

Verify equation (A.2) in the special cases $\phi = 90°$, $\phi = 180°$, $\phi = 270°$, and $\phi = 360°$.

A.2 **Problem: The rotation matrix**

a. Derive equation (A.2) through purely geometrical arguments.

b. Express $\hat{\mathbf{i}}'$ and $\hat{\mathbf{j}}'$, the unit vectors of the (x', y') coordinate system, as linear combinations of $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$. Then use

$$V_{x'} = \mathbf{V}\cdot\hat{\mathbf{i}}' \quad \text{and} \quad V_{y'} = \mathbf{V}\cdot\hat{\mathbf{j}}' \tag{A.6}$$

to derive equation (A.2).

c. Which derivation do you find easier?

A.3 **Problem: Rotation to the preferred coordinate system***

In the preferred coordinate system, $V_{y'} = 0$. Use this requirement to show that the preferred system is rotated from the initial system by an angle ϕ with

$$\tan\phi = \frac{V_y}{V_x}. \tag{A.7}$$

For any value of V_y/V_x, there are *two* angles that satisfy this equation. What is the representation of $\mathbf{V}$ in each of these two coordinate systems?

A.4 **Problem: A non-rotation orthogonal transformation**

In one coordinate system the y-axis is vertical and the x-axis points to the right. In another the y'-axis is vertical and the x'-axis points to the left. Find the matrix that translates vector coordinates from one system to the other. Show that this matrix is orthogonal but not a rotation matrix.

A.5 **Problem: Other changes of coordinate***

Suppose vertical distances (distances in the y direction) are measured in feet while horizontal distances (distances in the x direction) are measured in miles. (This system is not perverse. It is used in nearly all American road maps.) Find the matrix that changes the representation of a vector in this coordinate system to the representation of a vector in a system where all distances are measured in feet. Find the matrix that translates back. Are these matrices orthogonal?

A.6 **Problem: Other special representations**

At equation (A.5) we mentioned one "special" (or "canonical") representation of a vector. There are three others, namely

$$\begin{pmatrix} 0 \\ -V \end{pmatrix}, \qquad \begin{pmatrix} -V \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ V \end{pmatrix}. \tag{A.8}$$

If coordinate-system rotation angle ϕ brings the vector representation into the form (A.5), then what rotation angle will result in these three representations?

A.3 Tensors in two dimensions

A tensor, like a vector, is a geometrical entity that may be described ("named") through components, but a d-dimensional tensor requires d^2 rather than d components. Tensors are less familiar and more difficult to visualize than vectors, but they are neither less important nor "less physical". We will introduce tensors through the concrete example of the inertia tensor of classical mechanics (see, for example, reference [2]), but the results we present will be perfectly general.

Just as the two components of a two-dimensional vector are most easily kept track of through a 2×1 matrix, so the four components of two-dimensional tensor are most conveniently written in the form of a 2×2 matrix. For example, the inertia tensor $\mathbf{T}$ of a point particle with mass m located[3] at (x, y) has components

$$\mathsf{T} = \begin{pmatrix} my^2 & -mxy \\ -mxy & mx^2 \end{pmatrix}. \tag{A.9}$$

---

[3] Or, to be absolutely precise, the particle located at the point represented by the vector with components (x, y).

(Note the distinction between the tensor **T** and its matrix of components, its "name", T.) As with vector components, the tensor components are different in different coordinate systems, although the tensor itself does not change. For example, in the primed coordinate system of the figure on page 507, the tensor components are of course

$$\mathsf{T}' = \begin{pmatrix} my'^2 & -mx'y' \\ -mx'y' & mx'^2 \end{pmatrix}. \tag{A.10}$$

A little calculation shows that the components of the inertia tensor in two different coordinate systems are related through

$$\mathsf{T}' = \mathsf{R}(\phi)\mathsf{T}\mathsf{R}^{-1}(\phi). \tag{A.11}$$

This relation holds for *any* tensor, not just the inertia tensor. (In fact, one way to define "tensor" is as an entity with four components that satisfy the above relation under rotation.) If the matrix representing a tensor is symmetric (i.e. the matrix is equal to its transpose) in one coordinate system, then it is symmetric in all coordinate systems (see problem A.7). Therefore the symmetry is a property of the tensor, not of its matrix representation, and we may speak of "a symmetric tensor" rather than just "a tensor represented by a symmetric matrix".

As with vectors, one of the many matrix representations of a given tensor is considered special (or "canonical"): It is the one in which the lower left component is zero. Furthermore if the tensor is symmetric (as the inertia tensor is) then in this preferred coordinate system the upper right component will be zero also, so the matrix will be all zeros except for the diagonal elements. Such a matrix is called a "diagonal matrix" and the process of finding the rotation that renders the matrix representation of a symmetric tensor diagonal is called "diagonalization".[4] We may do an "accounting of information" for this preferred coordinate system just as we did with vectors. In the initial coordinate system, the symmetric tensor had three independent components. In the preferred system, it has two independent components manifestly visible in the diagonal matrix representation, and one number hidden through the specification of the rotation.

A.7 Problem: Representations of symmetric tensors*

Show that if the matrix S representing a tensor is symmetric, and if B

---

[4] An efficient algorithm for diagonalization is discussed in section A.8. For the moment, we are more interested in knowing that a diagonal matrix representation must exist than in knowing how to most easily find that preferred coordinate system.

is any orthogonal matrix, then all of the representations

$$\mathsf{BSB}^\dagger \tag{A.12}$$

are symmetric. (*Clue:* If you try to solve this problem for rotations in two dimensions using the explicit rotation matrix (A.3), you will find it solvable but messy. The clue is that this problem asks you do prove the result in any number of dimensions, and for any orthogonal matrix B, not just rotation matrices. This more general problem is considerably easier to solve.)

A.8 **Problem: Diagonal inertia tensor**

The matrix (A.9) represents the inertia tensor of a point particle with mass m located a distance r from the origin. Show that the matrix is diagonal in four different coordinate systems: one in which the x'-axis points directly toward the particle, one in which the y'-axis points directly away from the particle, one in which the x'-axis points directly away from the particle, and one in which the y'-axis points directly toward the particle. Find the matrix representation in each of these four coordinate systems.

A.9 **Problem: Representations of a certain tensor**

Show that a tensor represented in one coordinate system by a diagonal matrix with equal elements, namely

$$\begin{pmatrix} d_0 & 0 \\ 0 & d_0 \end{pmatrix}, \tag{A.13}$$

has the same representation in *all* orthogonal coordinate systems.

A.10 **Problem: Rotation to the preferred coordinate system***

A tensor is represented in the initial coordinate system by

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}. \tag{A.14}$$

Show that the tensor is diagonal in a preferred coordinate system which is rotated from the initial system by an angle ϕ with

$$\tan(2\phi) = \frac{2b}{a-c}. \tag{A.15}$$

This equation has four solutions. Find the rotation matrix for $\phi = 90°$, then show how the four different diagonal representations are related. You do not need to find any of the diagonal representations in terms of

a, b and c. . . just show what the other three are given that one of them is

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}. \tag{A.16}$$

A.11 Problem: Inertia tensor in outer product notation

The discussion in this section has emphasized the tensor's matrix representation ("name") T rather than the tensor $\mathbf{T}$ itself.

a. Define the "identity tensor" **1** as the tensor represented in some coordinate system by

$$\mathsf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{A.17}$$

Show that this tensor has the same representation in any coordinate system.

b. Show that the inner product between two vectors results in a scalar: Namely

if vector bfa is represented by $\begin{pmatrix} a_x \\ a_y \end{pmatrix}$ and vector bfb is represented by $\begin{pmatrix} b_x \\ b_y \end{pmatrix}$

then the inner product $\mathbf{a} \cdot \mathbf{b}$ is given through

$$\begin{pmatrix} a_x & a_y \end{pmatrix} \begin{pmatrix} b_x \\ b_y \end{pmatrix} = a_x b_x + a_y b_y,$$

and this inner product is a scalar. (A 1×2 matrix times a 2×1 matrix is a 1×1 matrix.) That is, the vector **a** is represented by different coordinates in different coordinate systems, and the vector **b** is represented by different coordinates in different coordinate systems, but the inner product $\mathbf{a} \cdot \mathbf{b}$ is the same in all coordinate systems.

c. In contrast, show that the outer product of two vectors is a tensor: Namely

$$\mathbf{ab} \doteq \begin{pmatrix} a_x \\ a_y \end{pmatrix} \begin{pmatrix} b_x & b_y \end{pmatrix} = \begin{pmatrix} a_x b_x & a_x b_y \\ a_y b_x & a_y b_y \end{pmatrix}.$$

(A 2×1 matrix times a 1×2 matrix is a 2×2 matrix.) That is, show that the representation of **ab** transforms from one coordinate system to another as specified through (A.11).

d. Show that the inertia tensor for a single particle of mass m located at position $\mathbf{r}$ can be written in coordinate-independent fashion as

$$\mathbf{T} = m\mathbf{1}r^2 - m\mathbf{rr}. \tag{A.18}$$

A.4 Tensors in three dimensions

A three-dimensional tensor is represented in component form by a 3×3 matrix with nine entries. If the tensor is symmetric, there are six independent elements. . . three on the diagonal and three off-diagonal. The components of a tensor in three dimensions change with coordinate system according to

$$\mathsf{T}' = \mathsf{R}\mathsf{T}\mathsf{R}^{\dagger}, \tag{A.19}$$

where R is the 3×3 rotation matrix.

A rotation in two dimension is described completely by giving a single angle. In three dimensions more information is required. Specifically, we need not only the amount of the rotation, but we must also know the plane in which the rotation takes place. We can specify the plane by giving the unit vector perpendicular to that plane. Specifying an arbitrary vector in three dimensions requires three numbers, but specifying a unit vector in three dimensions requires only two numbers because the magnitude is already fixed at unity. Thus three numbers are required to specify a rotation in three dimensions: two to specify the rotation's plane, one to specify the rotation's size. (One particularly convenient way to specify a three-dimensional rotation is through the three *Euler angles.* Reference [3] defines these angles and shows how to write the 3×3 rotation matrix in terms of these variables. For the purposes of this tutorial, however, we will not need an explicit rotation matrix. . . all we need is to know is the number of angles required to specify a rotation.)

In two dimensions, any symmetric tensor (which has three independent elements), could be represented by a diagonal tensor (with two independent elements) plus a rotation (one angle). We were able to back up this claim with an explicit expression for the angle.

In three dimensions it seems reasonable that any symmetric tensor (six independent elements) can be represented by a diagonal tensor (three independent elements) plus a rotation (three angles). The three angles just have to be selected carefully enough to make sure that they cause the off-diagonal elements to vanish. This supposition is indeed correct, although we will not pause for long enough to prove it by producing explicit formulas for the three angles.

A.5 Tensors in d dimensions

A d-dimensional tensor is represented by a $d \times d$ matrix with d^2 entries. If the tensor is symmetric, there are d independent on-diagonal elements and $d(d-1)/2$ independent off-diagonal elements. The tensor components will change with coordinate system in the now-familiar form

$$\mathsf{T}' = \mathsf{R}\mathsf{T}\mathsf{R}^\dagger, \tag{A.20}$$

where R is the $d \times d$ rotation matrix.

How many angles does it take to specify a rotation in d dimensions? Remember how we went from two dimensions to three: The three dimensional rotation took place "in a plane", i.e. in a two-dimensional subspace. It required two (i.e. $d-1$) angles to specify the orientation of the plane plus one to specify the rotation within the plane... a total of three angles.

A rotation in four dimensions takes place within a three-dimensional subspace. It requires $3 = d-1$ angles to specify the orientation of the three-dimensional subspace, plus, as we found above, three angles to specify the rotation within the three-dimensional subspace... a total of six angles.

A rotation in five dimensions requires $4 = d-1$ angles to specify the four-dimensional subspace in which the rotation occurs, plus the six angles that we have just found specify a rotation within that subspace... a total of ten angles.

In general, the number of angles needed to specify a rotation in d dimensions is

$$A_d = d - 1 + A_{d-1} = d(d-1)/2. \tag{A.21}$$

This is exactly the number of independent off-diagonal elements in a symmetric tensor. It seems reasonable that we can choose the angles to ensure that, in the resulting coordinate system, all the off-diagonal elements vanish. The proof of this result is difficult and proceeds in a very different manner from the plausibility argument sketched here. (The proof involves concepts like eigenvectors and eigenvalues, and it gives an explicit recipe for constructing the rotation matrix. It has the advantage of rigor and the disadvantage of being so technical that it's easy to lose track of the fact that that all you're doing is choosing a coordinate system.)

A.12 **Problem: Non-symmetric tensors***

Argue that a non-symmetric tensor can be brought into a "triangular" representation in which all the elements below the diagonal are equal to zero and all the elements on and above the diagonal are independent. (This is indeed the case, although in general some of the non-zero elements remaining will be complex-valued, and some of the angles will involve rotations into complex-valued vectors.)

A.6 Linear transformations in two dimensions

Section A.3 considered 2×2 matrices as representations of tensors. This section gains additional insight by considering 2×2 matrices as representations of linear transformations. It demonstrates how diagonalization can be useful and gives a clue to an efficient algorithm for diagonalization.

A linear transformation is a function from vectors to vectors that can be represented in any given coordinate system as

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{A.22}$$

If the equation above represents ("names") the transformation in one coordinate system, what is its representation in some other coordinate system? We assume that the two coordinate systems are related through an orthogonal matrix B such that

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \mathsf{B} \begin{pmatrix} u \\ v \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \mathsf{B} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{A.23}$$

(For example, if the new coordinate system is the primed coordinate system of the figure on page 507, then the matrix B that translates from the original to the new coordinates is the rotation matrix $\mathsf{R}(\phi)$.) Given this "translation dictionary", we have

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \mathsf{B} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{A.24}$$

But B is invertible, so

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathsf{B}^{-1} \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{A.25}$$

whence

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \mathsf{B} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mathsf{B}^{-1} \begin{pmatrix} x' \\ y' \end{pmatrix}. \tag{A.26}$$

Thus the representation of the transformation in the primed coordinate system is

$$\mathsf{B}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\mathsf{B}^{-1} \tag{A.27}$$

(compare equation A.11). This equation has a very direct physical meaning. Remember that the matrix B translates from the old (x, y) coordinates to the new (x', y') coordinates, while the matrix B^{-1} translates in the opposite direction. Thus the equation above says that the representation of a transformation in the new coordinates is given by translating from new to old coordinates (through the matrix B^{-1}), then applying the old representation (the "a matrix") to those old coordinates, and finally translating back from old to new coordinates (through the matrix B).

The rest of this section considers only transformations represented by symmetric matrices, which we will denote by

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{A.28}$$

Let's try to understand this transformation as something more than a jumble of symbols awaiting a plunge into the calculator. First of all, suppose the vector $\mathbf{V}$ maps to the vector $\mathbf{W}$. Then the vector $5\mathbf{V}$ will be mapped to vector $5\mathbf{W}$. In short, if we know how the transformation acts on vectors with magnitude unity, we will be able to see immediately how it acts on vectors with other magnitudes. Thus we focus our attention on vectors on the unit circle:

$$x^2 + y^2 = 1. \tag{A.29}$$

A brief calculation shows that the length of the output vector is then

$$\sqrt{u^2+v^2} = \sqrt{a^2x^2 + b^2 + c^2y^2 + 2b(a+c)xy}, \tag{A.30}$$

which isn't very helpful. Another brief calculation shows that if the input vector has polar angle θ, then the output vector has polar angle φ with

$$\tan\varphi = \frac{b + c\tan\theta}{a + b\tan\theta}, \tag{A.31}$$

which is similarly opaque and messy.

Instead of trying to understand the transformation in its initial coordinate system, let's instead convert (rotate) to the special coordinate system in which the transformation is represented by a diagonal matrix. In this system,

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} d_1 x' \\ d_2 y' \end{pmatrix}. \tag{A.32}$$

The unit circle is still

$$x'^2 + y'^2 = 1, \tag{A.33}$$

so the image of the unit circle is

$$\left(\frac{u'}{d_1}\right)^2 + \left(\frac{v'}{d_2}\right)^2 = 1, \tag{A.34}$$

namely an ellipse! This result is transparent in the special coordinate system, but almost impossible to see in the original one.

Note particularly what happens to a vector pointing along the x' coordinate axis. For example, the unit vector in this direction transforms to

$$\begin{pmatrix} d_1 \\ 0 \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{A.35}$$

In other words, the when the vector is transformed it changes in magnitude, but *not* in direction. Vectors with this property are called *eigenvectors*. It is easy to see that any vector on either the x' or y' coordinate axes are eigenvectors.

A.7 What does "eigen" mean?

If a vector $\mathbf{x}$ is acted upon by a linear transformation B, then the output vector

$$\mathbf{x}' = \mathsf{B}\mathbf{x} \tag{A.36}$$

will usually be skew to the original vector $\mathbf{x}$. However, for some very special vectors it might just happen that $\mathbf{x}'$ is parallel to $\mathbf{x}$. Such vectors are called "eigenvectors". (This is a terrible name because (1) it gives no idea of what eigenvectors are or why they're so important and (2) it sounds gross. However, that's what they're called.) We have already seen, in the previous section, that eigenvectors are related to coordinate systems in which the transformation is particularly easy to understand.

If $\mathbf{x}$ is an eigenvector, then

$$\mathsf{B}\mathbf{x} = \lambda\mathbf{x}, \tag{A.37}$$

where λ is a scalar called "the eigenvalue associated with eigenvector $\mathbf{x}$". If $\mathbf{x}$ is an eigenvector, then any vector parallel to $\mathbf{x}$ is also an eigenvector with the same eigenvalue. (That is, any vector of the form $c\mathbf{x}$, where c is any scalar, is also an eigenvector with the same eigenvalue.) Sometimes we speak of a "line of eigenvectors".

The vector $\mathbf{x} = \mathbf{0}$ is never considered an eigenvector, because

$$\mathsf{B}\mathbf{0} = \lambda\mathbf{0}, \tag{A.38}$$

for any value of λ for any linear transformation. On the other hand, if

$$\mathsf{B}\mathbf{x} = 0\mathbf{x} = \mathbf{0} \tag{A.39}$$

for some non-zero vector $\mathbf{x}$, then $\mathbf{x}$ is an eigenvector with eigenvalue $\lambda = 0$.

A.13 **Problem: Plane of eigenvectors**
Suppose $\mathbf{x}$ and $\mathbf{y}$ are two non-parallel vectors with the same eigenvalue. (In this case the eigenvalue is said to be "degenerate", which sounds like an aspersion cast upon the morals of the eigenvalue but which is really just poor choice of terminology again.) Show that any vector of the form $c_1\mathbf{x} + c_2\mathbf{y}$ is an eigenvector with the same eigenvalue.

A.8 How to diagonalize a symmetric matrix

We saw in section A.3 that for any 2×2 symmetric matrix, represented in its initial basis by, say,

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}, \tag{A.40}$$

a simple rotation of axes would produce a new coordinate system in which the matrix representation is diagonal:

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}. \tag{A.41}$$

These two matrices are related through

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} = \mathsf{R}(\phi) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \mathsf{R}^{-1}(\phi), \tag{A.42}$$

where $\mathsf{R}(\phi)$ is the rotation matrix (A.3). Problem A.10 gave a direct way to find the desired rotation. However this direct technique is cumbersome and doesn't generalize readily to higher dimensions. This section presents a different technique, which relies on eigenvalues and eigenvectors, that is more efficient and that generalizes readily to complex-valued matrices and to matrices in any dimension, but that is somewhat sneaky and conceptually roundabout.

We begin by noting that any vector lying along the x'-axis (of the preferred coordinate system) is an eigenvector. For example, the vector $5\hat{\mathbf{i}}'$ is represented (in the preferred coordinate system) by

$$\begin{pmatrix} 5 \\ 0 \end{pmatrix}. \tag{A.43}$$

Multiplying this vector by the matrix in question gives

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} 5 \\ 0 \end{pmatrix} = d_1 \begin{pmatrix} 5 \\ 0 \end{pmatrix}, \tag{A.44}$$

so $5\hat{\mathbf{i}}'$ is an eigenvector with eigenvalue d_1. The same holds for any scalar multiple of $\hat{\mathbf{i}}'$, whether positive or negative. Similarly, any scalar multiple of $\hat{\mathbf{j}}'$ is an eigenvector with eigenvalue d_2. In short, the two elements on the diagonal in the preferred (diagonal) representation are the two eigenvalues, and the two unit vectors $\hat{\mathbf{i}}'$ and $\hat{\mathbf{j}}'$ of the preferred coordinate system are two of the eigenvectors.

Thus finding the eigenvectors and eigenvalues of a matrix gives you the information needed to diagonalize that matrix. The unit vectors $\hat{\mathbf{i}}'$ and $\hat{\mathbf{j}}'$ constitute an "orthonormal basis of eigenvectors". The eigenvectors even give the rotation matrix directly, as described in the next paragraph.

Let's call the rotation matrix

$$\mathsf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \tag{A.45}$$

so that the inverse (transpose) matrix is

$$\mathsf{B}^{-1} = \mathsf{B}^{\dagger} = \begin{pmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{pmatrix}. \tag{A.46}$$

The representation of $\hat{\mathbf{i}}'$ in the preferred basis is

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{A.47}$$

so its representation in the initial basis is (see equation A.2)

$$\mathsf{B}^\dagger \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} b_{11} \\ b_{12} \end{pmatrix}. \tag{A.48}$$

Similarly, the representation of $\hat{\mathbf{j}}'$ in the initial basis is

$$\mathsf{B}^\dagger \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b_{21} \\ b_{22} \end{pmatrix}. \tag{A.49}$$

Thus the rotation matrix is

$$\mathsf{B} = \begin{pmatrix} \text{initial rep. of } \hat{\mathbf{i}}', \text{ on its side} \\ \text{initial rep. of } \hat{\mathbf{j}}', \text{ on its side} \end{pmatrix}. \tag{A.50}$$

Example

Suppose we need to find a diagonal representation for the matrix

$$\mathsf{T} = \begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix}. \tag{A.51}$$

First we search for the special vectors—the eigenvectors—such that

$$\begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}. \tag{A.52}$$

At the moment, we don't know either the eigenvalue λ or the associated eigenvector (x, y). Thus it seems that (bad news) we are trying to solve two equations for three unknowns:

$$\begin{aligned} 7x + 3y &= \lambda x \\ 3x + 7y &= \lambda y \end{aligned} \tag{A.53}$$

Remember, however, that there is not one *single* eigenvector: any multiple of an eigenvector is also an eigenvector. (Alternatively, any vector on the line that extends the eigenvector is another eigenvector.) We only need one of these eigenvectors, so let's take the one that has $x = 1$ (i.e. the vector on the extension line where it intersects the vertical line $x = 1$). (This technique will fail if we have the bad luck that our actual eigenvector is vertical and hence never passes through the line $x = 1$.) So we really have two equations in two unknowns:

$$\begin{aligned} 7 + 3y &= \lambda \\ 3 + 7y &= \lambda y \end{aligned}$$

but note that they are not *linear* equations…the damnable product λy in the lower right corner means that all our techniques for solving linear equations go right out the window. We *can* solve these two equations for λ and y, but there's an easier, if somewhat roundabout, approach.

Finding eigenvalues

Let's go back to equation (A.52) and write it as

$$\begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \lambda \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A.54}$$

Then

$$\begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{A.55}$$

or

$$\begin{pmatrix} 7-\lambda & 3 \\ 3 & 7-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A.56}$$

Let's think about this. It says that for some matrix $\mathsf{M} = \mathsf{T} - \lambda \mathsf{1}$, we have

$$\mathsf{M} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A.57}$$

You know right away one vector (x, y) that satisfies this equation, namely $(x, y) = (0, 0)$. And most of the time, this is the *only* vector that satisfies the equation, because

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathsf{M}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A.58}$$

We appear to have reached a dead end. The solution is $(x, y) = (0, 0)$, but the zero vector is not, by definition, considered an eigenvector of any transformation. (Because it always gives eigenvalue zero for any transformation.)

However, if the matrix M is *not invertible*, then there will be other solutions to

$$\mathsf{M} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A.59}$$

in addition to the trivial solution $(x, y) = (0, 0)$. Thus we must look for those special values of λ such that the so-called *characteristic matrix* M is not invertible. These values come if and only if the determinant of M vanishes. For this example, we have to find values of λ such that

$$\det \begin{pmatrix} 7-\lambda & 3 \\ 3 & 7-\lambda \end{pmatrix} = 0. \tag{A.60}$$

This is a quadratic equation in λ

$$(7-\lambda)^2 - 3^2 = 0 \tag{A.61}$$

called the *characteristic equation*. Its two solutions are

$$7 - \lambda = \pm 3 \tag{A.62}$$

or

$$\lambda = 7 \pm 3 = 10 \text{ or } 4. \tag{A.63}$$

We have found the two eigenvalues of our matrix!

Finding eigenvectors

Let's look now for the eigenvector associated with $\lambda = 4$. Equation (A.53)

$$\begin{aligned} 7x + 3y &= \lambda x \\ 3x + 7y &= \lambda y \end{aligned}$$

still holds, but no longer does it look like two equations in three unknowns, because we are now interested in the case $\lambda = 4$:

$$\begin{aligned} 7x + 3y &= 4x \\ 3x + 7y &= 4y \end{aligned}$$

Following our nose gives

$$\begin{aligned} 3x + 3y &= 0 \\ 3x + 3y &= 0 \end{aligned}$$

and when we see this our heart skips a beat or two... a degenerate system of equations! Relax and rest your heart. This system has an infinite number of solutions and it's supposed to have an infinite number of solutions, because any multiple of an eigenvector is also an eigenvector. The eigenvectors associated with $\lambda = 4$ are any multiple of

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{A.64}$$

An entirely analogous search for the eigenvectors associated with $\lambda = 10$ finds any multiple of

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{A.65}$$

Tidying up

We have the two sets of eigenvectors, but which shall we call $\hat{\mathbf{i}}'$ and which $\hat{\mathbf{j}}'$? This is a matter of individual choice, but my choice is usually to make the transformation be a rotation (without reflection) through a small positive angle. Our new, preferred coordinate system is related to the original coordinates by a simple rotation of 45° if we choose

$$\hat{\mathbf{i}}' = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \hat{\mathbf{j}}' = \tfrac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{A.66}$$

(Note that we have also “normalized the basis”, i.e. selected the basis vectors to have magnitude unity.) Given this choice, the orthogonal rotation matrix that changes coordinates from the original to the preferred system is (see equation A.50)

$$\mathsf{B} = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \tag{A.67}$$

and the diagonalized matrix (or, more properly, the representation of the matrix in the preferred coordinate system) is

$$\begin{pmatrix} 10 & 0 \\ 0 & 4 \end{pmatrix}. \tag{A.68}$$

You don’t believe me? Then multiply out

$$\mathsf{B} \begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix} \mathsf{B}^\dagger \tag{A.69}$$

and see for yourself.

Problems

A.14 **Problem: Diagonalize a 2×2 matrix***

Diagonalize the matrix

$$\begin{pmatrix} 26 & 12 \\ 12 & 19 \end{pmatrix}. \tag{A.70}$$

a. Find its eigenvalues.
b. Find its eigenvectors, and verify that they are orthogonal.
c. Sketch the eigenvectors, and determine the signs and sequence most convenient for assigning axes. (That is, should the first eigenvector you found be called $\hat{\mathbf{i}}'$, $-\hat{\mathbf{i}}'$, or $\hat{\mathbf{j}}'$?)
d. Find the matrix that translates from the initial basis to the basis of eigenvectors produced in part (c.).
e. Verify that the matrix produced in part (d.) is orthogonal.
f. Verify that the representation of the matrix above in the basis of eigenvectors is diagonal.
g. (Optional.) What is the rotation angle?

A.15 **Problem: Eigenvalues of a 2×2 matrix**

Show that the eigenvalues of

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \tag{A.71}$$

are

$$\lambda = \tfrac{1}{2}\left[(a+c) \pm \sqrt{(a-c)^2 + 4b^2}\right]. \qquad \text{(A.72)}$$

Under what circumstances is an eigenvalue complex valued? Under what circumstances are the two eigenvalues the same?

A.16 **Problem: Diagonalize a 3×3 matrix**

Diagonalize the matrix

$$\frac{1}{625}\begin{pmatrix} 1182 & -924 & 540 \\ -924 & 643 & 720 \\ 540 & 720 & -575 \end{pmatrix}. \qquad \text{(A.73)}$$

a. Find its eigenvalues by showing that the characteristic equation is

$$\lambda^3 - 2\lambda^2 - 5\lambda + 6 = (\lambda - 3)(\lambda + 2)(\lambda - 1) = 0. \qquad \text{(A.74)}$$

b. Find its eigenvectors, and verify that they are orthogonal.

c. Show that the translation matrix can be chosen to be

$$\mathsf{B} = \frac{1}{25}\begin{pmatrix} 20 & -15 & 0 \\ 9 & 12 & -20 \\ 12 & 16 & 15 \end{pmatrix}. \qquad \text{(A.75)}$$

Why did I use the phrase “the translation matrix can be chosen to be” rather then “the translation matrix is”?

A.17 **Problem: A 3×3 matrix eigenproblem**

Find the eigenvalues and associated eigenvectors for the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}. \qquad \text{(A.76)}$$

A.9 A glance at computer algorithms

Anyone who has worked even one of the problems in section A.8 knows that diagonalizing a matrix is no picnic: there’s a lot of mundane arithmetic involved and it’s very easy to make mistakes. This is a problem ripe for computer solution. One’s first thought is to program a computer to solve the problem using the same technique that we used to solve it on paper: first find the eigenvalues through the characteristic equation, then find the eigenvectors through a degenerate set of linear equations.

This turns out to be a very poor algorithm for automatic computation. The effective algorithm is to choose a matrix B such that the off-diagonal elements of

$$\mathsf{BAB}^{-1} \tag{A.77}$$

are smaller than the off-diagonal elements of A. Then choose another, and another. Go through this process again and again until the off-diagonal elements have been ground down to machine zero. There are many strategies for choosing the series of B matrices. These are well-described in any edition of *Numerical Recipes.*[4]

When you need to diagonalize matrices numerically, I urge you to look at *Numerical Recipes* to see what's going on, but I urge you *not* to code these algorithms yourself. These algorithms rely in an essential way on the fact that computer arithmetic is approximate rather than exact, and hence they are quite tricky to implement. Instead of coding the algorithms yourself, I recommend that you use the implementations in either LAPACK[5] (the Linear Algebra PACKage) or EISPACK.[6] These packages are probably the finest computer software ever written, and they are free. They can be obtained through the "Guide to Available Mathematical Software" (GAMS) at http://gams.nist.gov.

A.10 A glance at non-symmetric matrices and the Jordan form

Many of the matrices that arise in applications are symmetric and hence the results of the previous sections are the only ones needed. But every once in a while you do encounter a non-symmetric matrix and this section gives you a guide to treating them. It is just an introduction and treats only 2 × 2 matrices.

Given a non-symmetric matrix, the first thing to do is rotate the axes to make the matrix representation triangular, as discussed in problem A.12:

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}. \tag{A.78}$$

Note that $b \neq 0$ because otherwise the matrix would be symmetric and we would already be done. In this case vectors on the x-axis are eigenvectors because

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{A.79}$$

Are there any other eigenvectors? The equation

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \tag{A.80}$$

tells us that

$$\begin{aligned} ax + by &= \lambda x \\ cy &= \lambda y \end{aligned}$$

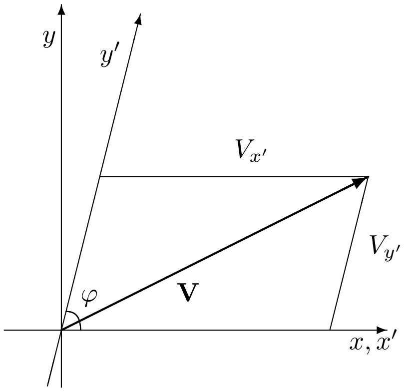
whence $\lambda = c$ and the eigenvector has polar angle θ where

$$\tan\theta = \frac{c-a}{b}. \tag{A.81}$$

Note that if $c = a$ (the "degenerate" case: both eigenvalues are the same) then $\theta = 0$ or $\theta = \pi$. In this case *all* of the eigenvectors are on the x-axis.

Diagonal form

We already know that that a rotation of orthogonal (Cartesian) coordinates will not diagonalize this matrix. We must instead transform to a skew coordinate system in which the axes are not perpendicular.



Note that in with oblique axes, the coordinates are given by

$$\mathbf{V} = V_{x'}\hat{\mathbf{i}}' + V_{y'}\hat{\mathbf{j}}' \tag{A.82}$$

but, because $\hat{\mathbf{i}}'$ and $\hat{\mathbf{j}}'$ are not perpendicular, it is *not true* that

$$V_{x'} = \mathbf{V}\cdot\hat{\mathbf{i}}'. \qquad \textbf{NO!} \tag{A.83}$$

A little bit of geometry will convince you that the name of the vector $\mathbf{V}$ changes according to

$$\begin{pmatrix} V_{x'} \\ V_{y'} \end{pmatrix} = \mathsf{B}\begin{pmatrix} V_x \\ V_y \end{pmatrix}, \tag{A.84}$$

where

$$\mathsf{B} = \frac{1}{\sin\varphi}\begin{pmatrix} \sin\varphi & -\cos\varphi \\ 0 & 1 \end{pmatrix}. \tag{A.85}$$

This matrix is *not* orthogonal. In fact its inverse is

$$\mathsf{B}^{-1} = \begin{pmatrix} 1 & \cos\varphi \\ 0 & \sin\varphi \end{pmatrix}. \tag{A.86}$$

Finally, note that we cannot have $\varphi = 0$ or $\varphi = \pi$, because then both $V_{x'}$ and $V_{y'}$ would give information about the horizontal component of the vector, and there would be no information about the vertical component of the vector.

What does this say about the representations of tensors (or, equivalently, of linear transformations)? The "name translation" argument of equation (A.27) still applies, so

$$\mathsf{T}' = \mathsf{B}\mathsf{T}\mathsf{B}^{-1}. \tag{A.87}$$

Using the explicit matrices already given, this says

$$\mathsf{T}' = \frac{1}{\sin\varphi}\begin{pmatrix} \sin\varphi & -\cos\varphi \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}\begin{pmatrix} 1 & \cos\varphi \\ 0 & \sin\varphi \end{pmatrix} = \begin{pmatrix} a & (a-c)\cos\varphi + b\sin\varphi \\ 0 & c \end{pmatrix}. \tag{A.88}$$

To make this diagonal, we need only choose a skew coordinate system where the angle φ gives

$$(a-c)\cos\varphi + b\sin\varphi = 0, \tag{A.89}$$

that is, one with

$$\tan\varphi = \frac{c-a}{b}. \tag{A.90}$$

Comparison with equation (A.81) shows that this simply means that the skew coordinate system should have its axes pointing along two eigenvectors. We have once again found an intimate connection between diagonal representations and eigenvectors, a connection which is exploited fully in abstract mathematical treatments of matrix diagonalization.

Once again we can do an accounting of information. In the initial coordinate system, the four elements of the matrix contain four independent pieces of information. In the diagonalizing coordinate system, two of those pieces are explicit in the matrix, and two are implicit in the two axis rotation angles needed to implement the diagonalization.

This procedure works almost all the time. But, if $a = c$, then it would involve $\varphi = 0$ or $\varphi = \pi$, and we have already seen that this is not an acceptable change of coordinates.

Degenerate case

Suppose our matrix has equal eigenvalues, $a = c$, so that it reads

$$\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}. \tag{A.91}$$

If $b = 0$, then the matrix is already diagonal. (Indeed, in this case *all* vectors are eigenvectors with eigenvalue a, and the linear transformation is simply multiplication of each vector by a).

But if $b \neq 0$, then, as we have seen, the only eigenvectors are on the x-axis, and it is impossible to make a basis of eigenvectors. Only one thing can be done to make the matrix representation simpler than it stands in equation (A.91), and that is a shift in the scale used to measure the y-axis.

For example, suppose that in the (x, y) coordinate system, the y-axis is calibrated in inches. We wish to switch to the (x', y') system in which the y'-axis is calibrated in feet. There is no change in axis orientation or in the x-axis. It is easy to see that the two sets of coordinates are related through

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1/12 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 12 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{A.92}$$

This process is sometimes called a "stretching" or a "scaling" of the y-axis.

The transformation represented by matrix (A.91) in the initial coordinate system is represented in the new coordinate system by

$$\begin{pmatrix} 1 & 0 \\ 0 & 1/12 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 12 \end{pmatrix} = \begin{pmatrix} a & 12b \\ 0 & a \end{pmatrix}. \tag{A.93}$$

The choice of what to do now is clear. Instead of scaling the y-axis by a factor of 12, we can scale it by a factor of $1/b$, and produce a new matrix representation of the form

$$\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}. \tag{A.94}$$

Where is the information in this case? In the initial coordinate system, the four elements of the matrix contain four independent pieces of information. In the new coordinate system, two of those pieces are explicit in the matrix, one is implicit in the rotation angle needed to implement the initial triangularization, and one is implicit in the y-axis scale transformation.

The Jordan form

Remarkably, the situation discussed above for 2×2 matrices covers all the possible cases for $n \times n$ matrices. That is, in n dimensional space, the proper combination of rotations, skews, and stretches of coordinate axes will bring the matrix representation (the "name") of any tensor or linear transformation into a form where every element is zero except on the diagonal and on the superdiagonal. The elements on the diagonal are eigenvalues, and each element on the superdiagonal is either zero or one: zero if the two adjacent eigenvalues differ, either zero or one if they are the same. The warning of problem A.12 applies here as well: The eigenvalues on the diagonal may well be complex valued, and the same applies for the elements of the new basis vectors.

References

[1] For example, Kenneth Hoffman and Ray Kunze, *Linear Algebra*, second edition (Prentice-Hall, Englewood Cliffs, New Jersey, 1971).

[2] For example, Jerry Marion and Stephen Thorton, *Classical Dynamics of Particles and Systems*, fourth edition (Saunders College Publishing, Fort Worth, Texas, 1995) section 11.2.

[3] For example, Jerry Marion and Stephen Thorton, *Classical Dynamics of Particles and Systems*, fourth edition (Saunders College Publishing, Fort Worth, Texas, 1995) section 11.7.

[4] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes* (Cambridge University Press, Cambridge, U.K., 1992).

[5] E. Anderson, *et al.*, *LAPACK Users' Guide* (SIAM, Philadelphia, 1992).

[6] B.T. Smith, *et al.*, *Matrix Eigensystem Routines—EISPACK Guide* (Springer-Verlag, Berlin, 1976).

Appendix B

The Dirac Delta Function

In classical mechanics a central idealization is the "point particle": it has a mass, it has a position, it has a velocity, but it has zero volume. You know that no planet, no football, no ball bearing, no atom actually *is* a point particle. It can nevertheless be a useful idealization.[5]

The Dirac delta function $\delta(x)$ is a useful idealization quite analogous to the classical point particle. It is not really a function: mathematicians call it a "generalized function" or a "Schwartz distribution". Whatever name you give it, it has the property that

$$\int_a^b f(x)\delta(x-x_0)\,dx = \begin{cases} 0 & \text{for } x_0 < a \\ f(x_0) & \text{for } a < x_0 < b \\ 0 & \text{for } b < x_0 \end{cases}. \tag{B.1}$$

You can see that $\delta(x)$ must have two properties: First, $\delta(x) = 0$ for $x \neq 0$. Second,

$$\int_{-\infty}^{+\infty} \delta(x)\,dx = 1. \tag{B.2}$$

There are several analytic expressions for the Dirac delta function. First, as a limit of box functions, each of unit area: The box function is defined through

$$b_a(x) = \begin{cases} 0 & \text{for } x < -a/2 \\ 1/a & \text{for } -a/2 < x < a/2 \\ 0 & \text{for } a/2 < x \end{cases}. \tag{B.3}$$

---

[5] For example, when investigating the orbit of the Earth around the Sun, it is useful to approximate the Earth as a point particle. In contrast, when constructing a house it is useful to approximate the Earth's surface as an infinite plane. The Earth is in fact neither a point particle nor an infinite plane, but in different situations these two very different approximations can be useful.

And the Dirac delta function is then

$$\delta(x) = \lim_{a \to 0} [b_a(x)]. \tag{B.4}$$

(This expression for the Dirac delta function arises implicitly in equation 6.13, which uses Δx instead of a.)

Second, as a limit of Gaussian functions, each of unit area:

$$\delta(x) = \lim_{a \to 0} \left[\frac{1}{\sqrt{\pi a^2}} e^{-x^2/a^2} \right]. \tag{B.5}$$

Third, through the Dirichlet form:

$$\delta(x) = \lim_{a \to 0} \left[\frac{\sin(x/a)}{\pi x} \right]. \tag{B.6}$$

Exercise B.A. Show that the functions within square brackets in equations (B.4), (B.5), and (B.6) all have unit area under the curve, regardless of the value of a. You may use the result

$$\int_{-\infty}^{+\infty} \frac{\sin u}{u}\, du = \pi.$$

Exercise B.B. Show that the functions within square brackets in equations (B.4) and (B.5) all approach zero when $a \to 0$ with $x \neq 0$.

Exercise B.C. Argue that, for the function within square brackets in equation (B.6), the mean value over a tiny window centered on $x \neq 0$ approaches zero when $a \to 0$.

Exercise B.D. The "Lorentzian form" of the Dirac delta function is

$$\lim_{a \to 0} \left[\frac{A}{x^2 + a^2} \right]. \tag{B.7}$$

a. How should A be chosen so that there is unit area under the curve, regardless of the value of a? You may use the result

$$\int_{-\infty}^{+\infty} \frac{du}{u^2 + 1} = \pi.$$

b. Show that with this expression for A, the function within square brackets in equation (B.7) approaches zero when when $a \to 0$ with $x \neq 0$.

The most useful analytic expression for the Dirac delta function derives from the Dirichlet form:

$$\begin{aligned}\delta(x) &= \lim_{K\to\infty} \frac{\sin(Kx)}{\pi x} \\ &= \lim_{K\to\infty} \frac{1}{2\pi} \int_{-K}^{+K} e^{ikx}\, dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ikx}\, dk. \qquad \text{(B.8)}\end{aligned}$$

This result is so useful that it is the very first expression (equation G.1) in the "Quantum Mechanics Cheat Sheet".

Appendix C

Problem-Solving Tips

A physicist can wax eloquent about concepts like interference and entanglement, but can also use those concepts to solve problems about the behavior of nature and the results of experiments. This appendix serves as a guide to the tips on problem solving scattered throughout this book.

You have heard that "practice makes perfect", but in fact practice makes permanent. If you practice slouchy posture, sloppy reasoning, or inefficient problem-solving technique, these bad habits will become second nature to you. For proof of this, just consider the career of ⟦insert here the name of your least favorite public figure, current or historical, foreign or domestic⟧. So I urge you to start now with straight posture, dexterous reasoning, and facile problem-solving technique, lest you end up like ⟦insert same name here⟧.

List of problem-solving tools

check your result, 129
dimensional analysis, 259
easy part first, 220
everyone makes errors, 210
ODE, informal solution of, 242–261
scaled quantities, 262–265
scaling, 261
test and reflect on your solution, 32–33, 151–152, 159–160, 221–224

Appendix D

Catalog of Misconceptions

Effective teaching does not merely instruct on what is correct — it also guards against beliefs that are not correct. There are a number of prevalent misconceptions concerning quantum mechanics. This catalog presents misconceptions mentioned in this book, together with the page number where that misconception is pointed out and corrected.

Appendix E

The Spherical Harmonics

A "function on the unit sphere" is a function $f(\theta, \phi)$. Another convenient variable is $\zeta = \cos\theta = z/r$. "Integration over the unit sphere" means

$$\int d\Omega\, f(\theta, \phi) = \int_0^\pi \sin\theta\, d\theta \int_0^{2\pi} d\phi\, f(\theta, \phi) = \int_{-1}^{+1} d\zeta \int_0^{2\pi} d\phi\, f(\theta, \phi).$$

$$\nabla^2 Y_\ell^m(\theta, \phi) = -\frac{1}{r^2}\ell(\ell+1) Y_\ell^m(\theta, \phi) \tag{E.1}$$

$$\int Y_{\ell'}^{m'*}(\theta, \phi) Y_\ell^m(\theta, \phi)\, d\Omega = \delta_{\ell',\ell}\delta_{m',m} \tag{E.2}$$

$$f(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell,m} Y_\ell^m(\theta, \phi) \quad \text{where} \tag{E.3}$$

$$f_{\ell,m} = \int Y_\ell^{m*}(\theta, \phi) f(\theta, \phi) d\Omega \tag{E.4}$$

In the table, square roots are always taken to be positive.

$$Y_0^0(\zeta,\phi) = \left(\frac{1}{2^2\pi}\right)^{1/2}$$

$$\begin{aligned}
Y_1^0(\zeta,\phi) &= \left(\frac{3}{2^2\pi}\right)^{1/2} \zeta && = \left(\frac{3}{2^2\pi}\right)^{1/2} \frac{z}{r} \\
Y_1^{\pm1}(\zeta,\phi) &= \mp\left(\frac{3}{2^3\pi}\right)^{1/2} \sqrt{1-\zeta^2}\, e^{\pm i\phi} && = \mp\left(\frac{3}{2^3\pi}\right)^{1/2} \frac{1}{r}(x \pm iy)
\end{aligned}$$

$$\begin{aligned}
Y_2^0(\zeta,\phi) &= \left(\frac{5}{2^4\pi}\right)^{1/2} (3\zeta^2 - 1) && = \left(\frac{5}{2^4\pi}\right)^{1/2} \left(3\frac{z^2}{r^2} - 1\right) \\
Y_2^{\pm1}(\zeta,\phi) &= \mp\left(\frac{3\cdot 5}{2^3\pi}\right)^{1/2} \zeta\sqrt{1-\zeta^2}\, e^{\pm i\phi} && = \mp\left(\frac{3\cdot 5}{2^3\pi}\right)^{1/2} \frac{z}{r^2}(x \pm iy) \\
Y_2^{\pm2}(\zeta,\phi) &= \left(\frac{3\cdot 5}{2^5\pi}\right)^{1/2} (1-\zeta^2)e^{\pm 2i\phi} && = \left(\frac{3\cdot 5}{2^5\pi}\right)^{1/2} \frac{1}{r^2}(x \pm iy)^2
\end{aligned}$$

$$\begin{aligned}
Y_3^0(\zeta,\phi) &= \left(\frac{7}{2^4\pi}\right)^{1/2} (5\zeta^3 - 3\zeta) && = \left(\frac{7}{2^4\pi}\right)^{1/2} \left(5\frac{z^3}{r^3} - 3\frac{z}{r}\right) \\
Y_3^{\pm1}(\zeta,\phi) &= \mp\left(\frac{3\cdot 7}{2^6\pi}\right)^{1/2} (5\zeta^2 - 1)\sqrt{1-\zeta^2}\, e^{\pm i\phi} && = \mp\left(\frac{3\cdot 7}{2^6\pi}\right)^{1/2} \left(5\frac{z^2}{r^2} - 1\right) \frac{1}{r}(x \pm iy) \\
Y_3^{\pm2}(\zeta,\phi) &= \left(\frac{3\cdot 5\cdot 7}{2^5\pi}\right)^{1/2} \zeta(1-\zeta^2)e^{\pm 2i\phi} && = \left(\frac{3\cdot 5\cdot 7}{2^5\pi}\right)^{1/2} \frac{z}{r^3}(x \pm iy)^2 \\
Y_3^{\pm3}(\zeta,\phi) &= \mp\left(\frac{5\cdot 7}{2^6\pi}\right)^{1/2} (1-\zeta^2)\sqrt{1-\zeta^2}\, e^{\pm 3i\phi} && = \mp\left(\frac{5\cdot 7}{2^6\pi}\right)^{1/2} \frac{1}{r^3}(x \pm iy)^3
\end{aligned}$$

Appendix F

Radial Wavefunctions for the Coulomb Problem

Based on Griffiths, page 154, but with scaled variables and with integers factorized.

$$R_{10}(r) = 2e^{-r}$$

$$\begin{aligned}
R_{20}(r) &= \frac{1}{\sqrt{2}}\left(1 - \frac{1}{2}r\right)e^{-r/2} \\
R_{21}(r) &= \frac{1}{\sqrt{2^3 \cdot 3}}\, r\, e^{-r/2}
\end{aligned}$$

$$\begin{aligned}
R_{30}(r) &= \frac{2}{\sqrt{3^3}}\left(1 - \frac{2}{3}r + \frac{2}{3^3}r^2\right)e^{-r/3} \\
R_{31}(r) &= \frac{2^3}{3^3\sqrt{2 \cdot 3}}\left(1 - \frac{1}{2 \cdot 3}r\right) r\, e^{-r/3} \\
R_{32}(r) &= \frac{2^2}{3^4\sqrt{2 \cdot 3 \cdot 5}}\, r^2\, e^{-r/3}
\end{aligned}$$

$$\begin{aligned}
R_{40}(r) &= \frac{1}{2^2}\left(1 - \frac{3}{2^2}r + \frac{1}{2^3}r^2 - \frac{1}{2^6 \cdot 3}r^3\right)e^{-r/4} \\
R_{41}(r) &= \frac{\sqrt{5}}{2^4\sqrt{3}}\left(1 - \frac{1}{2^2}r + \frac{1}{2^4 \cdot 5}r^2\right) r\, e^{-r/4} \\
R_{42}(r) &= \frac{1}{2^6\sqrt{5}}\left(1 - \frac{1}{2^2 \cdot 3}r\right) r^2\, e^{-r/4} \\
R_{43}(r) &= \frac{1}{2^8 \cdot 3\sqrt{5 \cdot 7}}\, r^3\, e^{-r/4}
\end{aligned}$$

Appendix G

Quantum Mechanics Cheat Sheet

Delta functions:

$$\int_{-\infty}^{+\infty} e^{ikx}\, dk = 2\pi\delta(x) \tag{G.1}$$

$$\int_{-\infty}^{+\infty} e^{i(p/\hbar)x}\, dp = 2\pi\hbar\delta(x) \tag{G.2}$$

$$\int_{-\infty}^{+\infty} e^{i\omega t}\, d\omega = 2\pi\delta(t) \tag{G.3}$$

Fourier transforms:

$$\widetilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi(x) e^{-i(p/\hbar)x}\, dx \tag{G.4}$$

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \widetilde{\psi}(p) e^{+i(p/\hbar)x}\, dp \tag{G.5}$$

$$\widetilde{f}(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t}\, dt \tag{G.6}$$

$$f(t) = \int_{-\infty}^{+\infty} \widetilde{f}(\omega) e^{+i\omega t}\, \frac{d\omega}{2\pi} \tag{G.7}$$

Gaussian integrals:

$$\int_{-\infty}^{+\infty} e^{-ax^2+bx}\, dx = \sqrt{\frac{\pi}{a}}\, e^{b^2/4a} \qquad \text{for } \Re e\{a\} \geq 0 \text{ and } a \neq 0 \tag{G.8}$$

$$\frac{\displaystyle\int_{-\infty}^{+\infty} x^2 e^{-x^2/2\sigma^2}\, dx}{\displaystyle\int_{-\infty}^{+\infty} e^{-x^2/2\sigma^2}\, dx} = \sigma^2 \tag{G.9}$$

Time evolution:

$$\frac{d|\psi(t)\rangle}{dt} = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle \tag{G.10}$$

$$\frac{\partial\psi(\mathbf{x},t)}{\partial t} = -\frac{i}{\hbar}\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})\right]\psi(\mathbf{x},t) \tag{G.11}$$

$$|\psi(t)\rangle = \sum_n e^{-(i/\hbar)E_n t} c_n|\eta_n\rangle \tag{G.12}$$

$$\frac{d\langle\hat{A}\rangle}{dt} = -\frac{i}{\hbar}\langle[\hat{A},\hat{H}]\rangle \tag{G.13}$$

Momentum:

$$\hat{p} \Longleftrightarrow -i\hbar\frac{\partial}{\partial x} \tag{G.14}$$

$$[\hat{x},\hat{p}] \quad = \quad i\hbar \tag{G.15}$$

$$\langle x|p\rangle \quad = \quad \frac{1}{\sqrt{2\pi\hbar}}e^{i(p/\hbar)x} \tag{G.16}$$

Dimensions:

$$\psi(x) \text{ has dimensions [length]}^{-1/2} \tag{G.17}$$

$$\psi(\vec{r}_1,\vec{r}_2) \text{ has dimensions [length]}^{-6/2} \tag{G.18}$$

$$\widetilde{\psi}(p) \text{ has dimensions [momentum]}^{-1/2} \tag{G.19}$$

$$\begin{aligned}\hbar \text{ has dimensions [length} \times \text{momentum]}\\ \text{or [energy} \times \text{time]}\end{aligned} \tag{G.20}$$

Energy eigenfunction sketching: (one dimension)

nth excited state has n nodes (G.21)

if classically allowed: regions of high $V(x)$ have large amplitude and long wavelength (G.22)

if classically forbidden: regions of high $V(x)$ have faster cutoff (G.23)

Infinite square well: (width L)

$$\eta_n(x) = \sqrt{2/L}\sin k_n x \quad k_n = n\pi/L \quad n = 1,2,3,\ldots \tag{G.24}$$

$$E_n = \frac{\hbar^2 k_n^2}{2m} = n^2\frac{\pi^2\hbar^2}{2mL^2} \tag{G.25}$$

Simple harmonic oscillator: ($V(x) = \frac{1}{2}Kx^2$, $\omega = \sqrt{K/m}$)

$$E_n = (n + \tfrac{1}{2})\hbar\omega \quad n = 0, 1, 2, \ldots \tag{G.26}$$

$$[\hat{a}, \hat{a}^\dagger] = \hat{1} \tag{G.27}$$

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \tfrac{1}{2}) \tag{G.28}$$

$$\hat{a}|n\rangle = \sqrt{n}\,|n-1\rangle \tag{G.29}$$

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle \tag{G.30}$$

$$\hat{x} = \sqrt{\hbar/2m\omega}\,(\hat{a} + \hat{a}^\dagger) \tag{G.31}$$

$$\hat{p} = -i\sqrt{\hbar m\omega/2}\,(\hat{a} - \hat{a}^\dagger) \tag{G.32}$$

Coulomb problem:

$$E_n = -\frac{\mathrm{Ry}}{n^2} \qquad \mathrm{Ry} = \frac{m_e(e^2/4\pi\epsilon_0)^2}{2\hbar^2} = 13.6 \text{ eV} \tag{G.33}$$

$$a_0 = \frac{(e^2/4\pi\epsilon_0)}{2\,\mathrm{Ry}} = 0.0529 \text{ nm} \qquad \text{(Bohr radius)} \tag{G.34}$$

$$\tau_0 = \frac{\hbar}{2\,\mathrm{Ry}} = 0.0242 \text{ fsec} \qquad \text{(characteristic time)} \tag{G.35}$$

Angular momentum:

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z, \text{ and cyclic permutations} \tag{G.36}$$

The eigenvalues of $\hat{J}^2$ are

$$\hbar^2 j(j+1) \qquad j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \ldots. \tag{G.37}$$

For a given j, the eigenvalues of $\hat{J}_z$ are

$$\hbar m \qquad m = -j, -j+1, \ldots, j-1, j. \tag{G.38}$$

The eigenstates $|j, m\rangle$ are related through the operators

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \qquad \hat{J}_- = \hat{J}_x - i\hat{J}_y \tag{G.39}$$

by

$$\hat{J}_+|j,m\rangle = \hbar\sqrt{j(j+1) - m(m+1)}\,|j, m+1\rangle \tag{G.40}$$

$$\hat{J}_-|j,m\rangle = \hbar\sqrt{j(j+1) - m(m-1)}\,|j, m-1\rangle. \tag{G.41}$$

Spherical harmonics:
A "function on the unit sphere" is a function $f(\theta, \phi)$. Another convenient variable is $\zeta = \cos\theta = z/r$. "Integration over the unit sphere" means

$$\int d\Omega\, f(\theta,\phi) = \int_0^\pi \sin\theta\, d\theta \int_0^{2\pi} d\phi\, f(\theta,\phi) = \int_{-1}^{+1} d\zeta \int_0^{2\pi} d\phi\, f(\theta,\phi).$$

$$\nabla^2 Y_\ell^m(\theta,\phi) = -\frac{1}{r^2}\ell(\ell+1)Y_\ell^m(\theta,\phi) \tag{G.42}$$

$$\int Y_{\ell'}^{m'*}(\theta,\phi)Y_\ell^m(\theta,\phi)\, d\Omega = \delta_{\ell',\ell}\delta_{m',m} \tag{G.43}$$

$$f(\theta,\phi) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} f_{\ell,m}Y_\ell^m(\theta,\phi) \qquad \text{where} \tag{G.44}$$

$$f_{\ell,m} = \int Y_\ell^{m*}(\theta,\phi)f(\theta,\phi)d\Omega \tag{G.45}$$

Index