

Notes on

The Physics of Quantum Mechanics

Daniel F. Styer

Professor of Physics, Oberlin College

Dan.Styer@oberlin.edu

Copyright © 15 August 2011

Everything in this book is a sketch — it is not polished or complete.

Contents

Preface	7
1 What is Quantum Mechanics About?	9
1.1 Quantization	9
1.2 Probability	9
1.3 Interference	10
1.4 Entanglement	11
2 Forging Mathematical Tools	12
2.1 Introduction	12
2.2 Amplitude	12
2.3 Reversal-conjugation relation	15
2.4 Phase conventions for quantal spin- $\frac{1}{2}$ systems	17
2.5 Terms concerning quantum states	17
2.6 How to specify a quantal state	18
2.7 Outer products, operators, measurement	23
2.8 Photon polarization	29
2.9 Lightning linear algebra	31
2.10 Problems	39
3 Formalism	43
3.1 The density matrix	43

<i>CONTENTS</i>	3
4 Time Evolution	45
4.1 Operator for time evolution	45
4.2 Working with the Schrödinger equation	48
4.3 Formal properties of time evolution; Conservation laws	59
4.4 The neutral K meson	62
5 Continuum Systems	65
5.1 Describing states in continuum systems	65
5.2 How does position amplitude change with time?	71
5.3 Operators and their representations; The momentum basis	77
5.4 Time evolution of average quantities	85
6 The Free Particle	88
6.1 Problems	88
7 Square Wells	91
7.1 What does an electron look like?	92
8 The Simple Harmonic Oscillator	94
8.1 Resume of energy eigenproblem	94
8.2 Solution of the energy eigenproblem: Differential equation approach	94
8.3 Solution of the energy eigenproblem: Operator factorization approach	98
8.4 Problems	102
9 Qualitative Solution of Energy Eigenproblems	104
10 Perturbation Theory	105
10.1 The \mathcal{O} notation	105
10.2 Perturbation theory for cubic equations	107
10.3 Derivation of perturbation theory for the energy eigenproblem	110
10.4 Perturbation theory for the energy eigenproblem: Summary of results	113
10.5 Problems	114

11 Quantum Mechanics in Two and Three Dimensions	118
11.1 More degrees of freedom	118
11.2 Vector operators	121
11.3 Multiple particles	122
11.4 The phenomena of quantum mechanics	122
12 Angular Momentum	125
12.1 Solution of the angular momentum eigenproblem	125
12.2 Summary of the angular momentum eigenproblem	128
12.3 Ordinary differential equations for the $d_{m,m'}^{(j)}(\theta)$	129
12.4 Problems	129
13 Central Force Motion	131
13.1 Energy eigenproblem in two dimensions	131
13.2 Energy eigenproblem in three dimensions	137
13.3 Bound state energy eigenproblem for Coulombic potentials	140
13.4 Summary of the bound state energy eigenproblem for a Coulombic potential	144
13.5 Problems	145
14 Identical Particles	148
14.1 Many-particle systems in quantum mechanics	148
14.2 An antisymmetric basis for the helium problem	159
15 Breather	165
15.1 Problems	165
16 Hydrogen	169
17 Helium	171
17.1 Ground state energy of helium	171
18 Atoms	175

19 Molecules	176
19.1 The hydrogen molecule ion	176
19.2 Problems	179
19.3 The hydrogen molecule	180
20 WKB: The Quasiclassical Approximation	181
21 Perturbation Theory for the Time Development Problem	182
21.1 Setup	182
21.2 Fermi's golden rule	186
21.3 Problems	188
22 The Interaction of Matter and Radiation	189
23 Quantization of the Electromagnetic Field	190
23.1 Overview	190
23.2 Classical electromagnetism	191
23.3 Quantal electromagnetism — What is a photon?	196
24 The Vistas Open to Us	198
A Tutorial on Matrix Diagonalization	199
A.1 What's in a name?	199
A.2 Vectors in two dimensions	200
A.3 Tensors in two dimensions	202
A.4 Tensors in three dimensions	204
A.5 Tensors in d dimensions	205
A.6 Linear transformations in two dimensions	206
A.7 What does "eigen" mean?	208
A.8 How to diagonalize a symmetric matrix	208
A.9 A glance at computer algorithms	214
A.10 A glance at non-symmetric matrices and the Jordan form	215

B The Spherical Harmonics	219
C Radial Wavefunctions for the Coulomb Problem	221
D Quantum Mechanics Cheat Sheet	222

Preface

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

Just over 100 years ago, 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 in the last 10 or 20 years that we 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 quantum 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.

What you can expect from this book

This is a book about physics, not about mathematics. The emphasis is on nature, not on 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.

However, the never abandons its focus on nature. It provides a balanced, interwoven treatment of concepts, formalism, and applications so that each thread reinforces the other. 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”. Example problems build both mathematical technique and physical insight.

The book does not merely convey correct ideas, but also refutes misconceptions. Just to get started on the right foot, 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 eigenstates.
- (c) The wavefunction $\psi(\mathbf{x}, t)$ is “out there” in space and you could reach out and touch it if only your fingers were sufficiently sensitive.

Chapter 1

What is Quantum Mechanics About?

We will, to the extent possible, do a quantum mechanical treatment of an atom's spin while keeping a classical treatment of all other aspects — namely its position and momentum.

1.1 Quantization

Look at *The Strange World of Quantum Mechanics*, chapter 3.

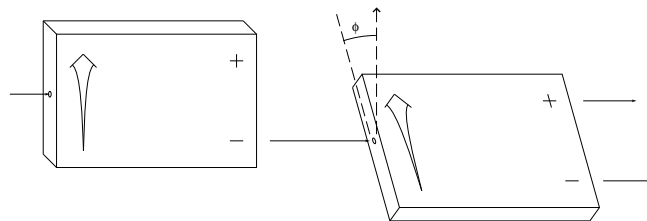
1.2 Probability

Look at *The Strange World of Quantum Mechanics*, chapter 4.

Problems

1.1 Exit probabilities

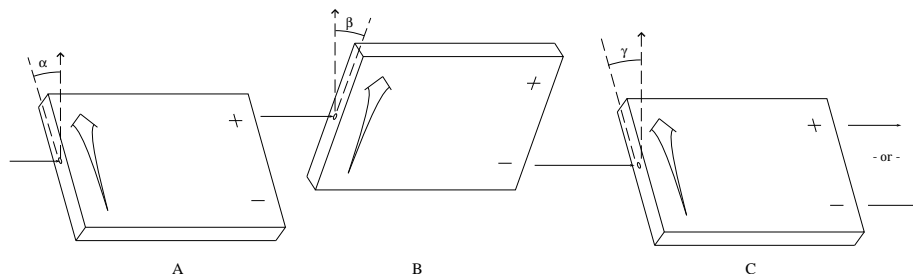
Consider the following combination of Stern-Gerlach analyzers.



(The dashed line with an arrowhead points in the positive z direction.) Given the fact that an atom with $\mu_z = +\mu_B$ exits the + port of a θ -analyzer with probability $\cos^2(\theta/2)$, show that an atom entering the right-most analyzer exits the + port with probability $\sin^2(\phi/2)$.

1.2 Multiple analyzers

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



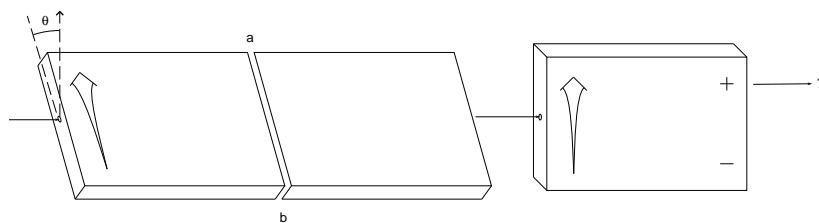
What is the probability that it emerges from the + output of analyzer C? From the - output? Why don't these probabilities sum to one?

1.3 Interference

Look at *The Strange World of Quantum Mechanics*, chapters 9 and 12.

Problems

1.3 Analyzer loop

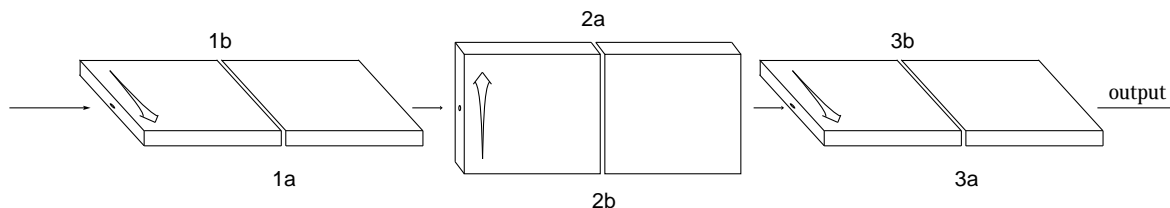


An atom with $\mu_z = +\mu_B$ is injected into an analyzer loop tilted an angle θ to the z direction. The output atom is then fed into a z Stern-Gerlach analyzer. What is the probability of the atom leaving the + channel of the last analyzer when:

- Branches a and b are both open?
- Branch b is closed?
- Branch a is closed?

1.4 Three analyzer loops

A number of atoms with $\mu_z = +\mu_B$ are shot into the following line of three analyzer loops.



If all branches are open, 100% of the incoming atoms exit from the output. What percent of the incoming atoms leave from the output if the following branches are closed? (The atoms are not observed as they pass through the analyzer loops.)

- | | | |
|-------|--------------|---------------------|
| a. 1b | d. 2a | g. 1b and 3b |
| b. 3a | e. 2b | h. 1b and 3a |
| c. 3b | f. 2a and 3b | i. 1b and 3a and 2a |

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

1.5 Tilted analyzer loop

In class we discussed an interference experiment in which atoms with $\mu_z = +\mu_B$ were fed first through a horizontal analyzer loop (interferometer), then through a z -analyzer, and the output was whatever came out the $-$ port of the z -analyzer. Consider this experiment if the analyzer loop is tilted by angle θ instead of by angle 90° . What is the probability for passage from input to output if both branches are unblocked? If branch a is blocked? If branch b is blocked? (Be sure that your answers give the correct results when $\theta = 90^\circ$.)

1.4 Entanglement

Look at *The Strange World of Quantum Mechanics*, chapters 6 and 7.

Chapter 2

Forging Mathematical Tools

2.1 Introduction

Look at *The Strange World of Quantum Mechanics*, chapters 10 and 11.

What is an amplitude?

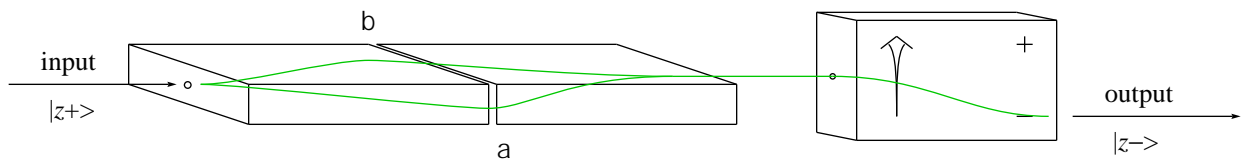
Amplitudes for entangled and measurement states.

General rules for amplitudes, including series and parallel.

Any working set of amplitudes, multiplied by any complex number with square modulus unity, will give a set of amplitudes just as good as the original.

Apply general rules to analysis experiment (find magnitudes) and to interference experiment (find a plausible set of phases).

2.2 Amplitude



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

probability to go from input to output \neq
 probability to go from input to output via a
 + probability to go from input to output via b.

On the other hand, it makes sense to associate some sort of “influence to go from input to output via a” with the path via a. This postulated influence is called “probability amplitude” or just “amplitude”. Whatever amplitude is, its desired property is that

amplitude to go from input to output =
 amplitude to go from input to output via a
 + amplitude to go from input to output via b.

For the moment, the very existence of amplitude is nothing but surmise. Indeed we cannot now and never will be able to prove that “the amplitude framework” applies to all situations. That’s because new situations are being investigated every day, and perhaps tomorrow a new situation will be discovered that cannot fit into the amplitude framework. But up until today, that hasn’t happened.

The role of amplitude, whatever it may prove to be, is to calculate probabilities. We establish the three desirable rules:

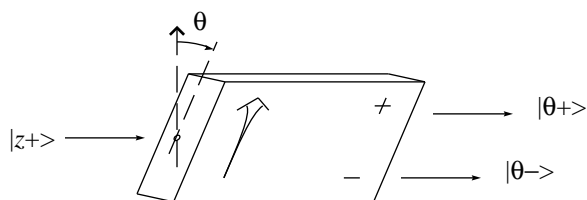
1. *From amplitude to probability.* For every possible action there is an associated amplitude, such that

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

2. *Actions in series.* If an action takes place through two stages, the amplitude for that action is the product of the amplitudes for each stage.
3. *Actions in parallel.* If an action can be performed in two ways, the amplitude for that action is the sum of the amplitudes for each way.

We apply these rules to various situations that we’ve already encountered, beginning with the situation sketched above. Remember that the probability to go from input to output is 0, whereas the probability to go from input to output via a is $\frac{1}{2}$ and the probability to go from input to output via b is also $\frac{1}{2}$. If rule 1 is to hold, then the amplitude to go from input to output must also be 0, and the amplitude to go via a has magnitude $\frac{1}{\sqrt{2}}$ and the amplitude to go via b also has magnitude $\frac{1}{\sqrt{2}}$. So according to rule 3, the two amplitudes to go via a and via b must sum to zero, so they cannot be represented by positive numbers. Whatever mathematical entity is used to represent amplitude, it must be such that two of these entities, each with non-zero magnitude, can sum to zero. There are many such entities: real numbers, complex numbers, hypercomplex numbers, and vectors in three dimensions are all possibilities. It turns out that, for all situations yet encountered, it is adequate to represent amplitude mathematically through complex numbers.

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



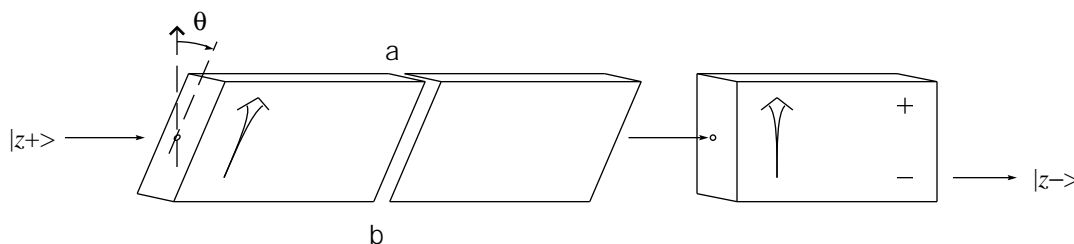
The amplitude for projection from $|z+\rangle$ to $|\theta+\rangle$ through a θ -analyzer is called $\langle\theta+|z+\rangle$. This is read “the amplitude for projection from $|z+\rangle$ to $|\theta+\rangle$ ” or “the amplitude to go from $|z+\rangle$ to $|\theta+\rangle$ ”. From rule 1, we know that

$$|\langle\theta+|z+\rangle|^2 = \cos^2(\theta/2) \quad (2.1)$$

$$|\langle\theta-|z+\rangle|^2 = \sin^2(\theta/2) \quad (2.2)$$

These projection experiments, in other words, will determine the magnitudes of the amplitudes. But no projection 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 branch **a** is the product of the amplitude to go from $|z+\rangle$ to $|\theta+\rangle$ times the amplitude to go from $|\theta+\rangle$ to $|z-\rangle$:

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

Similarly

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

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 branch **a** and the amplitude to go via branch **b**. In other words

$$\langle z-|z+\rangle = \langle z-|\theta+\rangle\langle\theta+|z+\rangle + \langle z-|\theta-\rangle\langle\theta-|z+\rangle. \quad (2.3)$$

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

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

The task now is to assign phases to these magnitudes in such a way that equation (2.3) is satisfied. In doing so we are faced with an embarrassment of riches: there are *many* consistent ways that this assignment can be made. Here are two commonly-used conventions:

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

There are a few 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, all of the conventions show some unexpected behavior: Since the angle θ is the same as the angle $2\pi + \theta$, one would expect that $\langle \theta + |z+\rangle$ would equal $\langle (2\pi + \theta) + |z+\rangle$ whereas in fact $\langle \theta + |z+\rangle = -\langle (2\pi + \theta) + |z+\rangle$. Because the state $|\pi-\rangle$ is the same as the state $|z+\rangle$, one would expect that $\langle \pi - |z+\rangle = 1$, whereas in fact in $\langle \pi - |z+\rangle$ is either $-i$ or -1 , 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. Amplitude cannot be measured, it is not “out there, physically present in space” in the way that, say, a nitrogen molecule is.

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:

$$\langle \psi|\phi\rangle = \langle \phi|\psi\rangle^*. \quad (2.4)$$

The proof below works for single-particle spin states, the kind of states we’ve been working with so far. But in fact the result holds true 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 spin and the analyzer, and not on the sense of the

angle. (An atom in state $|z+\rangle$ has the same probability of leaving the + port of an analyzer whether it is rotated 17° to the right or 17° to the left.) Thus

$$|\langle\phi|\psi\rangle|^2 = |\langle\psi|\phi\rangle|^2. \quad (2.5)$$

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

$$\langle\phi|\psi\rangle = \langle\phi|\theta+\rangle\langle\theta+|\psi\rangle + \langle\phi|\theta-\rangle\langle\theta-|\psi\rangle. \quad (2.6)$$

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

$$1 = |\langle\theta+|\psi\rangle|^2 + |\langle\theta-|\psi\rangle|^2. \quad (2.7)$$

From the first fact, the amplitude $\langle\phi|\psi\rangle$ differs from the amplitude $\langle\psi|\phi\rangle$ only by a phase, so

$$\langle\phi|\psi\rangle = e^{i\delta}\langle\psi|\phi\rangle^* \quad (2.8)$$

where the phase δ might depend on the states $|\phi\rangle$ and $|\psi\rangle$. Apply the second fact with $|\phi\rangle = |\psi\rangle$, giving

$$\begin{aligned} 1 &= \langle\psi|\theta+\rangle\langle\theta+|\psi\rangle + \langle\psi|\theta-\rangle\langle\theta-|\psi\rangle \\ &= e^{i\delta_+}\langle\theta+|\psi\rangle^*\langle\theta+|\psi\rangle + e^{i\delta_-}\langle\theta-|\psi\rangle^*\langle\theta-|\psi\rangle \\ &= e^{i\delta_+}|\langle\theta+|\psi\rangle|^2 + e^{i\delta_-}|\langle\theta-|\psi\rangle|^2 \end{aligned} \quad (2.9)$$

where the phase δ_+ might depend upon the states $|\theta_+\rangle$ and $|\psi\rangle$ while the phase δ_- might depend upon the states $|\theta_-\rangle$ and $|\psi\rangle$. Compare this result to the third fact

$$1 = |\langle\theta+|\psi\rangle|^2 + |\langle\theta-|\psi\rangle|^2 \quad (2.10)$$

and you will see that the only way the two positive numbers $|\langle\theta+|\psi\rangle|^2$ and $|\langle\theta-|\psi\rangle|^2$ can sum to 1 is for the two the phases δ_+ and δ_- in equation (2.9) to vanish. (This is sometimes called the “triangle inequality”.)

2.4 Phase conventions for quantal spin- $\frac{1}{2}$ systems

We use the following phase conventions for amplitudes:

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

In particular, for $\theta = 90^\circ$ we have

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

These conventions have a desirable special case for $\theta = 0^\circ$, namely

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

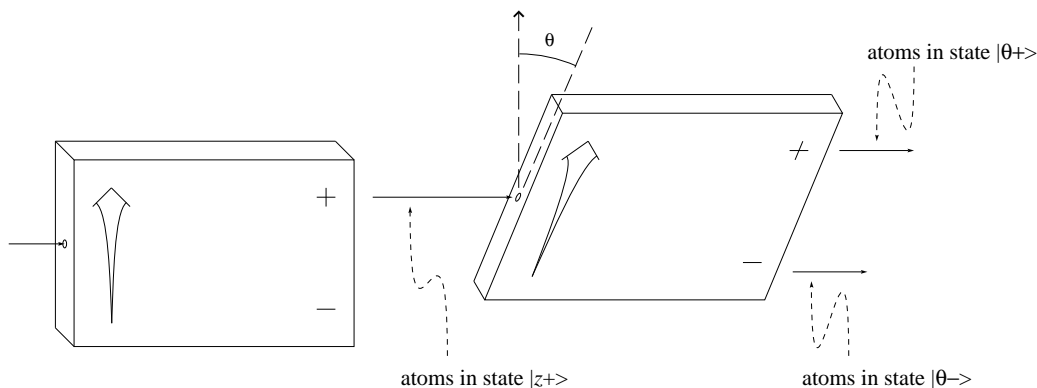
but an unexpected special case for $\theta = 360^\circ$, namely

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

This is perplexing, given that $\theta = 0^\circ$ is the same as $\theta = 360^\circ$! Any convention will have similar perplexing cases. To me, this underscores the fact that amplitudes are mathematical tools used to calculate probabilities. They are not “physically real”.

2.5 Terms concerning quantum states

Just as the symbol for a vector is identified as such by a decoration, namely \vec{r} or \mathbf{r} , so the symbol for a quantum state is identified by a decoration, namely $|A\rangle$. For example $|z+\rangle$, $|z-\rangle$, $|x+\rangle$, $|x-\rangle$, $|\theta+\rangle$, $|\theta-\rangle$. In terms of our diagrams we have



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 $\{|A\rangle, |B\rangle, \dots, |N\rangle\}$, 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.)

2.6 How to specify a quantal state

I’ll begin with an analogy:

How to specify a position vector

To specify a position vector \mathbf{r} , we use the components of \mathbf{r} in a particular basis, usually denoted (x, y, z) . We often write “ $\mathbf{r} = (x, y, z)$ ” but in fact that’s exactly correct. The vector \mathbf{r} represents a position — it is independent of basis and in fact the concept of “position” was known to cavemen who did not have any concept of “basis”. The row matrix (x, y, z) represents the components of that position vector in a particular basis — it is the “name” of the position in a particular basis.

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.

Back to quantal states.

We've been specifying a state like $|\psi\rangle = |17^\circ\rangle$ by listing the axis upon which the projection of μ is definite and equal to $+\mu_B$ — in this case, it's the axis tilted 17° from the vertical.

Another way to specify a state $|\psi\rangle$ would be to give the projection amplitude of $|\psi\rangle$ onto all possible states: that is, to list $\langle\theta + |\psi\rangle$ 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 magnitude 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.

There's a third way to specify a state that has combines the brevity of the first way with the convenience of the second way. As we'll see in a moment, we don't have to list the projection amplitude of $|\psi\rangle$ onto *every* state. If we list the projection amplitude of $|\psi\rangle$ onto the two elements of a basis, then it's easy to find the projection amplitude onto any state. How?

Suppose the two amplitudes $\langle z + |\psi\rangle$ and $\langle z - |\psi\rangle$ are known. Then we can easily find the projection 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 an 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 $\{|\theta+\rangle, |\theta-\rangle\}$:

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

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

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

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

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

Examples:

$$\begin{aligned} |z+\rangle &\doteq \begin{pmatrix} \langle z+|z+\rangle \\ \langle z-|z+\rangle \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |z-\rangle &\doteq \begin{pmatrix} \langle z+|z-\rangle \\ \langle z-|z-\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ |\theta+\rangle &\doteq \begin{pmatrix} \langle z+|\theta+\rangle \\ \langle z-|\theta+\rangle \end{pmatrix} = \begin{pmatrix} \cos \theta/2 \\ \sin \theta/2 \end{pmatrix} \\ |\theta-\rangle &\doteq \begin{pmatrix} \langle z+|\theta-\rangle \\ \langle z-|\theta-\rangle \end{pmatrix} = \begin{pmatrix} -\sin \theta/2 \\ \cos \theta/2 \end{pmatrix} \end{aligned}$$

Analysis of states into basis states

For all states $|\psi\rangle, |\phi\rangle$, and any basis $|a\rangle, |b\rangle$:

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

This equation is exactly the interference experiment: If the atom goes through both branch a and through branch b , then it emerges unchanged.

In particular, we write this equation for $|\phi\rangle = |z+\rangle$ and then for $|\phi\rangle = |z-\rangle$:

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

This equation is the representation, in the $\{|z+\rangle, |z-\rangle\}$ basis, of

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

What is this equation supposed to mean? The quantities $|\psi\rangle, |a\rangle$, and $|b\rangle$ represent states, the quantities $\langle a|\psi\rangle$ and $\langle b|\psi\rangle$ represent complex numbers. This is the first time we have ever added states. What does it mean? It means nothing more nor less than the interferometer experiment. (Just as cavemen were able to spear mammoths through vector addition, even though they didn’t know about basis states or the coordinate representations of vectors in a particular basis, so atoms are able to interfere through the above equation, even though they don’t know about bases or representations. The representations make it easier to work with position vectors or with quantal states, but they aren’t required.)

Reinforcing the meaning of this equation by writing it down for the basis $\{|\theta+\rangle, |\theta-\rangle\}$: When I write

$$|\psi\rangle = |\theta+\rangle\langle \theta+|\psi\rangle + |\theta-\rangle\langle \theta-|\psi\rangle,$$

I mean that if an atom passes through a θ -analyzer, it has amplitude $\langle \theta + | \psi \rangle$ to behave like an atom in state $|\theta + \rangle$, and it has amplitude $\langle \theta - | \psi \rangle$ to behave like an atom in state $|\theta - \rangle$.

There's an easier way to derive the equation (2.11). If two states $|\alpha \rangle$ and $|\beta \rangle$ have the same projection amplitudes

$$\langle \phi | \alpha \rangle = \langle \phi | \beta \rangle$$

for all states $|\phi \rangle$, then they are the same state:

$$|\alpha \rangle = |\beta \rangle.$$

(Two atoms that behave identically under *all* circumstances are in the same state.) In other words, because the state $|\phi \rangle$ is arbitrary, we can erase the symbol $\langle \phi |$ from both sides of the equation. Applying this principle to

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

where the state $|\phi \rangle$ is arbitrary, results in

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

This equation is analogous to the relation for position vectors that

$$\mathbf{r} = \hat{\mathbf{i}}x + \hat{\mathbf{j}}y + \hat{\mathbf{k}}z.$$

Example: If the basis $\{|a \rangle, |b \rangle\}$ is $\{|z + \rangle, |z - \rangle\}$, then

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

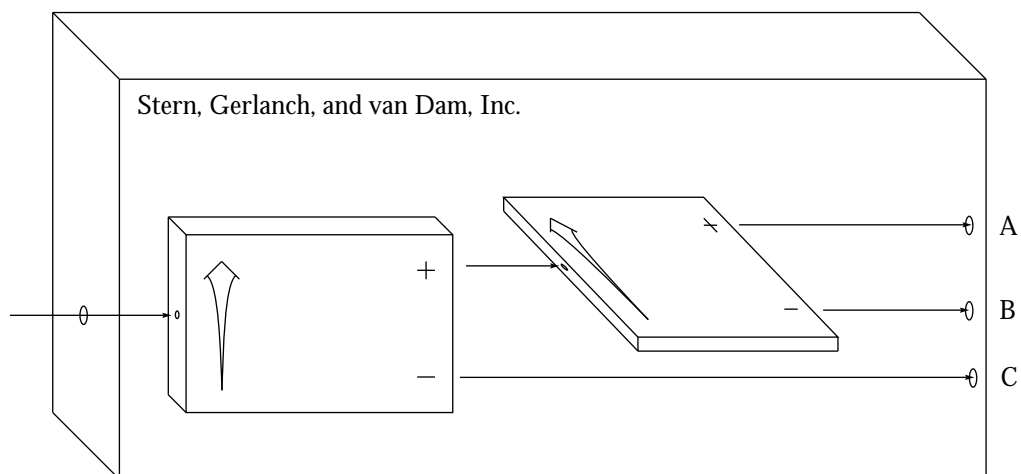
Using the representations for $|z + \rangle$ and $|z - \rangle$ in the $\{|z + \rangle, |z - \rangle\}$ basis, this equation becomes

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

Problems

2.1 What is a basis state?

Mr. van Dam claims that a silver atom has three, not two, basic magnetic dipole states. To back up his claim, he has constructed the following ‘‘Stern-Gerlach-van Dam’’ analyzer out of a z Stern-Gerlach analyzer and an x Stern-Gerlach analyzer.



(The output of the x analyzer is piped to output holes A and B using atomic pipes that do not affect the magnetic dipole state.) Show that the set of exit states $\{|A\rangle, |B\rangle, |C\rangle\}$ is complete, but that $|B\rangle$ is not orthogonal to $|C\rangle$.

2.2 Representations

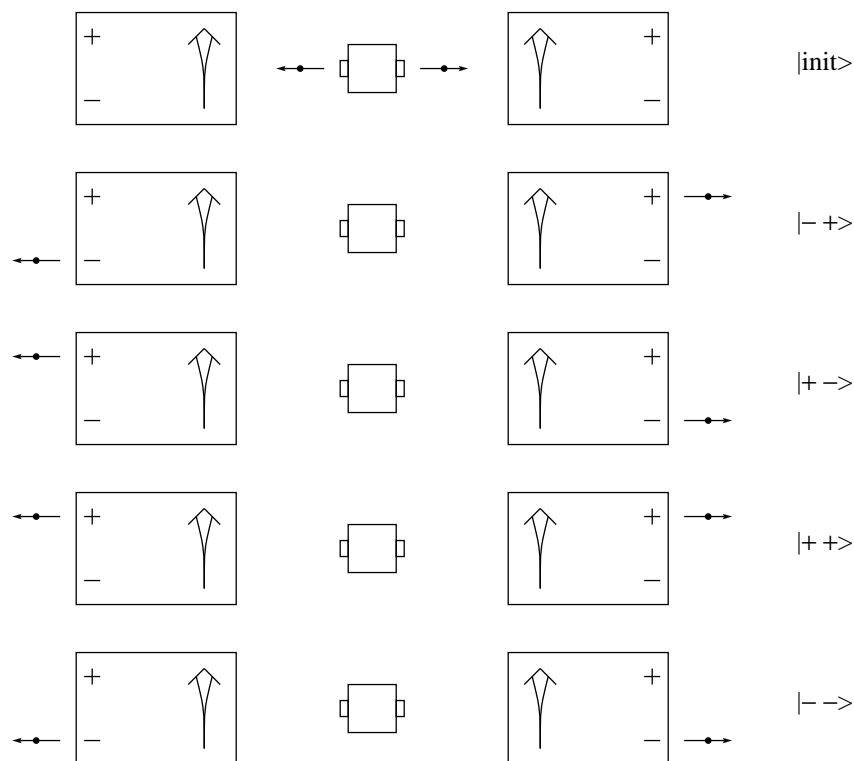
Suppose that the representation of $|\psi\rangle$ in the $\{|z+\rangle, |z-\rangle\}$ basis is

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

(i.e., $\psi_+ = \langle z+|\psi\rangle$, $\psi_- = \langle z-|\psi\rangle$). If ψ_+ and ψ_- are both real, show that there is an axis upon which the projection of $\boldsymbol{\mu}$ has a definite, positive value, and find the angle between that axis and the z axis in terms of ψ_+ and ψ_- .

2.3 Amplitudes for entangled states

In the Einstein-Podolsky-Rosen-Bohm experiment, the initial state is represented by $|\text{init}\rangle$, and various hypothetical final states are represented by $|-\rangle|+\rangle$ and so forth, as indicated below.



What are the magnitudes of the amplitudes $\langle - + | \text{init} \rangle$, $\langle + - | \text{init} \rangle$, $\langle + + | \text{init} \rangle$, and $\langle - - | \text{init} \rangle$?

(Notice that in this 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 *doesn't have a state*, in the same way that an atom passing through an interferometer *doesn't have a position* and love *doesn't have a color*.)

2.7 Outer products, operators, measurement

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 amplitudes into inner product of states:

$$\langle \phi | \psi \rangle = \langle \phi | \left\{ |a\rangle \langle a| + |b\rangle \langle b| \right\} | \psi \rangle. \quad (2.12)$$

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

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

then

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

Thus 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: $|\alpha\rangle\langle\beta|$ 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} .

With these ideas in place, we see what's inside the curly brackets of expression (2.12) — 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

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

so

$$|z+\rangle\langle z+| \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (2.13)$$

Meanwhile

$$|z-\rangle\langle z-| \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \quad (2.14)$$

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 that in this basis

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

so

$$|\theta+\rangle\langle\theta+| \doteq \begin{pmatrix} \cos \theta/2 \\ \sin \theta/2 \end{pmatrix} \begin{pmatrix} \cos \theta/2 & \sin \theta/2 \end{pmatrix} = \begin{pmatrix} \cos^2 \theta/2 & \cos \theta/2 \sin \theta/2 \\ \sin \theta/2 \cos \theta/2 & \sin^2 \theta/2 \end{pmatrix}. \quad (2.15)$$

Meanwhile

$$|\theta-\rangle\langle\theta-| \doteq \begin{pmatrix} -\sin \theta/2 \\ \cos \theta/2 \end{pmatrix} \begin{pmatrix} -\sin \theta/2 & \cos \theta/2 \end{pmatrix} = \begin{pmatrix} \sin^2 \theta/2 & -\sin \theta/2 \cos \theta/2 \\ -\cos \theta/2 \sin \theta/2 & \cos^2 \theta/2 \end{pmatrix}. \quad (2.16)$$

(As a check, notice that when $\theta = 0$, equation (2.15) reduces to equation (2.13), and equation (2.16) reduces to equation (2.14).) Thus

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

Yes! Once again this combination is the identity matrix.

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 $\boldsymbol{\mu}$ on the θ axis? (We call the projection of $\boldsymbol{\mu}$ 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 average value of μ_θ ?

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

In the last line we have effected the divorce — writing amplitudes in terms of inner products between states.

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

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

so that

$$\langle \mu_\theta \rangle = \langle \phi | \hat{\mu}_\theta | \phi \rangle.$$

Notice what we've done here: To find the average 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.

Example: What is the matrix representation of $\hat{\mu}_\theta$ in the basis $\{|z+\rangle, |z-\rangle\}$? We have already found representations for the outer product $|\theta+\rangle\langle\theta+|$ in equation (2.15) and for the outer product $|\theta-\rangle\langle\theta-|$ in equation (2.16). Using these expressions

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

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

In particular, using the values $\theta = 0$ and $\theta = 90^\circ$,

$$\hat{\mu}_z = \mu_B \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \hat{\mu}_x = \mu_B \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and furthermore

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

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

$$\hat{\mathbf{r}} = \cos \theta \hat{\mathbf{k}} + \sin \theta \hat{\mathbf{i}}.$$

So, knowing the operator associated with a measurement, we can easily find the resulting average value for any given state when measured. But we often want to know more than the average. 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 $\hat{\mu}_\theta$?

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

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

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

But

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

so

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

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 $\theta = 180^\circ$.

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 $-\mu_B$.

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

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

Summary: The quantum theory of measurement

This summarizes the quantum theory of measurement as applied to the measurement of μ projected onto the unit vector in the direction of θ :

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

$$\begin{aligned}\text{state } |\theta+\rangle &\text{ with eigenvalue } +\mu_B \\ \text{state } |\theta-\rangle &\text{ with eigenvalue } -\mu_B\end{aligned}$$

(a) If you measure μ_θ of an atom in an eigenstate of $\hat{\mu}_\theta$, then the number you measure 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 you measure 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: Show that (a) follows from (b).

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 we find through repeated measurement experiments, interference experiments, and indeed all experiments. 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 and Albert Hibbs do just this in their book *Quantum Mechanics and Path Integrals*. States and operators do not make an appearance until deep into this book, and even when they do appear it is only to make a connection between Feynman's formulation and more traditional formulations of quantum mechanics — states and operators 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.

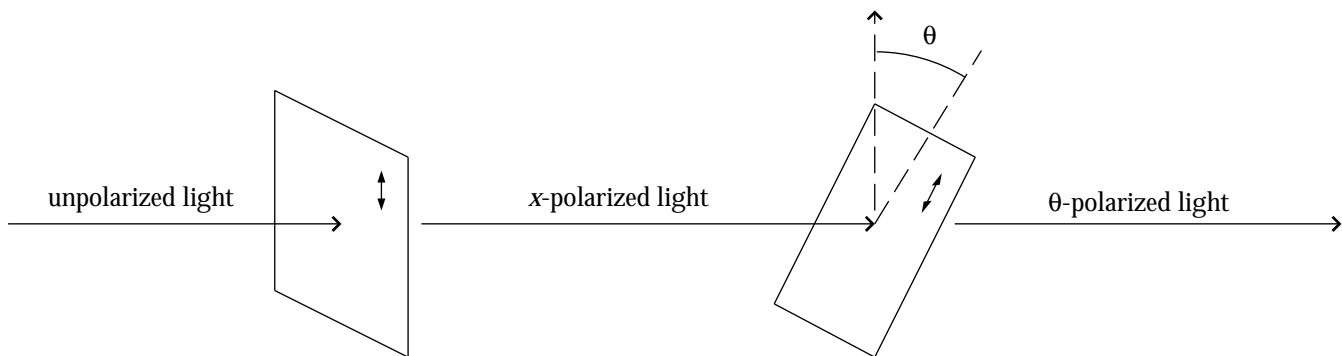
For more extensive discussion, see N. David Mermin, “What's bad about this habit?” *Physics Today*, **62** (5), May 2009, pages 8–9, and discussion about this essay in *Physics Today*, **62** (9), September 2009, pages 10–15.

2.8 Photon polarization

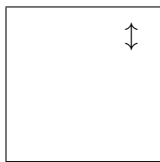
So far we have developed the principles of quantum mechanics using a particular system, the magnetic moment of a silver atom (a so-called “spin- $\frac{1}{2}$ ” system), which has two basis states. Another system with two basis states is polarized light. I did not use this system mainly because photons are less familiar than atoms. These problems develop the quantum mechanics of photon polarization much as the text developed the quantum mechanics of spin- $\frac{1}{2}$.

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.

2.4 Classical description of polarized light



When a beam of unpolarized light passes through a sheet of perfect polarizing material (called “Polaroid” and represented by the symbol

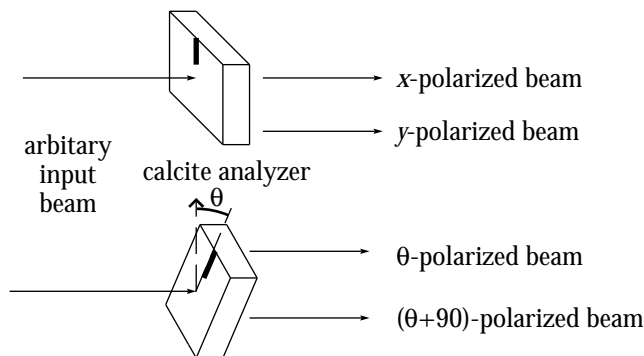


where the arrow shows the polarizing axis), the emerging beam is of lower intensity and it is “polarized”, i.e. the electric field vector undulates but points only parallel or antiparallel to the polarizing axis. When a beam of vertically polarized light (an “ x -polarized beam”) is passed through a sheet of ideal Polaroid with polarizing axis oriented at an angle θ to the vertical, the beam is reduced in intensity and emerges with an electric field undulating parallel to the sheet’s polarizing axis (a “ θ -polarized beam”).

Polaroid sheet performs these feats by absorbing any component of electric field perpendicular to its polarizing axis. Show that if the incoming x -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 proper results when θ is 0° , 90° , 180° or 270° .

2.5 Quantal description of polarized light: Analyzers

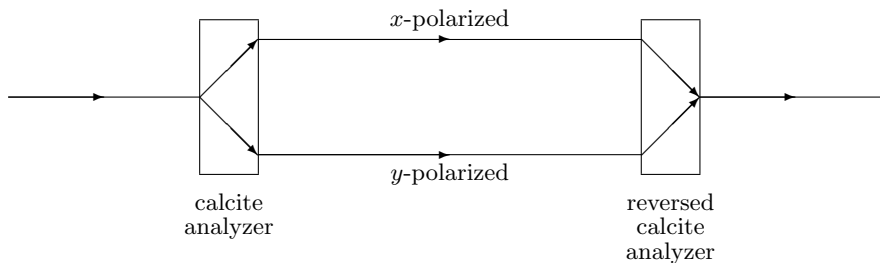
In quantum mechanics, a photon state is described by three quantities: 1) energy, 2) direction of motion, 3) polarization. We ignore the first two quantities. There is an infinite number of possible polarization states: the photons in an x -polarized beam are all in the $|x\rangle$ state, the photons in a θ -polarized beam ($0^\circ \leq \theta < 180^\circ$) are all in the $|\theta\rangle$ state, etc. In the quantum description, when an $|x\rangle$ photon encounters a polarizing sheet oriented at an angle θ to the vertical, then either it is absorbed (with probability $\sin^2 \theta$) or else it is changed into a $|\theta\rangle$ photon (with probability $\cos^2 \theta$). A polarizing sheet is thus not an analyzer: instead of splitting the incident beam into two (or more) beams, it absorbs one of the beams that we would like an analyzer to emit. An analyzer can be constructed out of any material that exhibits double refraction. It is conventional to use a simple calcite crystal:



What are the projection probabilities $|\langle x|\theta\rangle|^2$, $|\langle x|\theta + 90^\circ\rangle|^2$? Show that the states $\{|\theta\rangle, |\theta + 90^\circ\rangle\}$ constitute a basis.

2.6 Interference

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

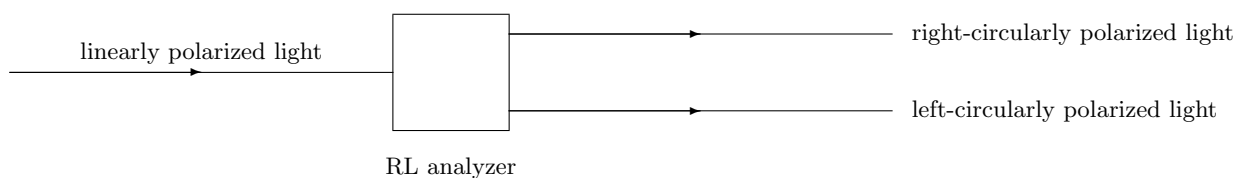


Invent a series of experiments that demonstrates quantum interference. Show that the results of these experiments, and the results of problem 2.5, are consistent with the amplitudes

$$\begin{aligned}\langle x|\theta\rangle &= \cos\theta & \langle x|\theta+90^\circ\rangle &= -\sin\theta \\ \langle y|\theta\rangle &= \sin\theta & \langle y|\theta+90^\circ\rangle &= \cos\theta.\end{aligned}\tag{2.17}$$

2.7 Circular polarization

Just as it is possible to analyze any light beam into x - and y -polarized beams, or θ - and $\theta+90^\circ$ -polarized beams, so it is possible to analyze any beam into right- and left-circularly polarized beams. You might remember from classical optics 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 projection amplitudes thus have magnitudes

$$\begin{aligned}|\langle R|\ell p\rangle| &= 1/\sqrt{2} \\ |\langle L|\ell p\rangle| &= 1/\sqrt{2}\end{aligned}\tag{2.18}$$

where $|\ell p\rangle$ is any linearly polarized state. By building an RL analyzer loop you can convince yourself that

$$\langle\theta|R\rangle\langle R|x\rangle + \langle\theta|L\rangle\langle L|x\rangle = \langle\theta|x\rangle = \cos\theta.\tag{2.19}$$

Show that no real valued projection amplitudes can satisfy both relations (2.18) and (2.19), but that the complex values

$$\begin{aligned}\langle L|\theta\rangle &= e^{i\theta}/\sqrt{2} & \langle L|x\rangle &= 1/\sqrt{2} \\ \langle R|\theta\rangle &= e^{-i\theta}/\sqrt{2} & \langle R|x\rangle &= 1/\sqrt{2}\end{aligned}\tag{2.20}$$

are satisfactory!

2.9 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.

Scalars: either real numbers (x) or complex numbers (z)

Vectors: notation \mathbf{a} , \mathbf{b} , \mathbf{c} , or $|\psi\rangle$, $|\phi\rangle$, $|\chi\rangle$

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

I won't define "vector" any more than I defined "number". But I will give some examples:

arrows in N -dimensional space

n -tuples, with real entries or with complex entries

polynomials

functions

$n \times m$ matrices

Inner product

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

$$\text{I.P.}(\mathbf{a}, \mathbf{b}) = \text{a real or complex number,} \quad (2.21)$$

that satisfies

$$\text{I.P.}(\mathbf{a}, \mathbf{b} + \mathbf{c}) = \text{I.P.}(\mathbf{a}, \mathbf{b}) + \text{I.P.}(\mathbf{a}, \mathbf{c}) \quad (2.22)$$

$$\text{I.P.}(\mathbf{a}, z\mathbf{b}) = z\text{I.P.}(\mathbf{a}, \mathbf{b}) \quad (2.23)$$

$$\text{I.P.}(\mathbf{a}, \mathbf{b}) = [\text{I.P.}(\mathbf{b}, \mathbf{a})]^* \quad (2.24)$$

$$\text{I.P.}(\mathbf{a}, \mathbf{a}) > 0 \quad \text{unless } \mathbf{a} = \mathbf{0} \quad (2.25)$$

It follows from equation (2.24) that $\text{I.P.}(\mathbf{a}, \mathbf{a})$ is real. Equation (2.25) demands also that it's positive.

Why is there a complex conjugation in equation (2.24)? Why not just demand that $\text{I.P.}(\mathbf{a}, \mathbf{b}) = \text{I.P.}(\mathbf{b}, \mathbf{a})$? The complex conjugation is needed for consistency with (2.25). If it weren't there, then

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

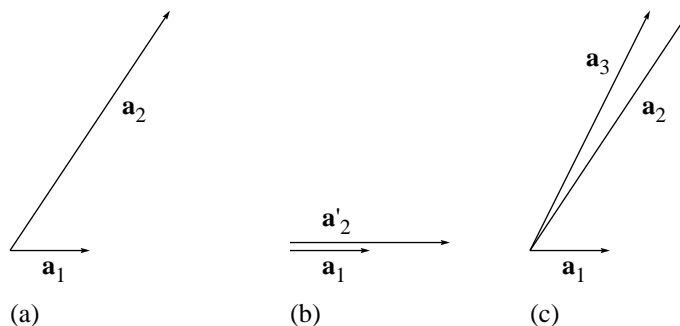
Notation: $\text{I.P.}(\mathbf{a}, \mathbf{b}) = (\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$, $\text{I.P.}(|\phi\rangle, |\psi\rangle) = \langle \phi | \psi \rangle$.

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

Making 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.



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 $\{\mathbf{a}_1, \mathbf{a}_2\}$.

The set $\{\mathbf{a}_1, \mathbf{a}_2\}$ is “linearly independent”, the set $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ is not. *Linearly independent*: You can’t build any member of the set out of the other members.

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,

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

Basis: A spanning set of linearly independent vectors. (That 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 elements, and this number is called the dimensionality, N .

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

$$\langle n|m\rangle = \delta_{n,m}. \quad (2.26)$$

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

$$|\psi\rangle = \psi_1|1\rangle + \psi_2|2\rangle + \dots + \psi_N|N\rangle = \sum_{n=1}^N \psi_n|n\rangle, \quad (2.27)$$

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 element $|m\rangle$ with $|\psi\rangle$, giving

$$\langle m|\psi\rangle = \sum_{n=1}^N \psi_n \langle m|n\rangle = \sum_{n=1}^N \psi_n \delta_{m,n} = \psi_m. \quad (2.28)$$

Thus the expansion (2.27) is

$$|\psi\rangle = \sum_{n=1}^N |n\rangle \langle n|\psi\rangle. \quad (2.29)$$

You have seen this formula in the context of arrows. For example, using two-dimensional arrows with the orthonormal basis $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}\}$, also called $\{\mathbf{e}_x, \mathbf{e}_y\}$, you know that

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y,$$

where

$$x = \mathbf{r} \cdot \mathbf{e}_x \quad \text{and} \quad y = \mathbf{r} \cdot \mathbf{e}_y.$$

Thus

$$\mathbf{r} = \mathbf{e}_x(\mathbf{e}_x \cdot \mathbf{r}) + \mathbf{e}_y(\mathbf{e}_y \cdot \mathbf{r}),$$

which is just an instance of the more general expression (2.29).

Representations

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

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} = \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \\ \vdots \\ \langle N|\psi\rangle \end{pmatrix}. \quad (2.30)$$

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

$$|\psi\rangle \doteq \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} = \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \\ \vdots \\ \langle N|\psi\rangle \end{pmatrix}. \quad (2.31)$$

What can we do with representations? Here's a way to connect an inner product, which is defined solely through the list of properties (2.22)–(2.25), to a formula in terms of representations.

$$\begin{aligned}
 & \langle \phi | \psi \rangle && \text{[[using (2.29) ...]]} \\
 = & \langle \phi | \left\{ \sum_n |n\rangle \langle n| \psi \right\} && \text{[[using (2.22) ...]]} \\
 = & \sum_n \langle \phi | n \rangle \langle n | \psi \rangle && \text{[[using (2.24) ...]]} \\
 = & \sum_n \phi_n^* \psi_n \\
 = & (\phi_1^* \quad \phi_2^* \quad \cdots \quad \phi_N^*) \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}
 \end{aligned}$$

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

$$(\phi_1^* \quad \phi_2^* \quad \cdots \quad \phi_N^*). \quad (2.32)$$

Transformation of representations

In the orthonormal basis $\{|1\rangle, |2\rangle, \dots, |N\rangle\}$, the vector $|\psi\rangle$ is represented by

$$\begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}. \quad (2.33)$$

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

$$\begin{pmatrix} \psi'_1 \\ \psi'_2 \\ \vdots \\ \psi'_N \end{pmatrix}. \quad (2.34)$$

How are these two representations related?

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

so

$$\begin{pmatrix} \psi'_1 \\ \psi'_2 \\ \vdots \\ \psi'_N \end{pmatrix} = \begin{pmatrix} \langle 1'|1 \rangle & \langle 1'|2 \rangle & \cdots & \langle 1'|N \rangle \\ \langle 2'|1 \rangle & \langle 2'|2 \rangle & \cdots & \langle 2'|N \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle N'|1 \rangle & \langle N'|2 \rangle & \cdots & \langle N'|N \rangle \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}. \quad (2.35)$$

Operators

A *linear operator* \hat{A} is a function from vectors to vectors

$$\hat{A} : |\psi\rangle \mapsto |\phi\rangle \quad \text{or in other words} \quad |\phi\rangle = \hat{A}|\psi\rangle, \quad (2.36)$$

with the property that

$$\hat{A}(z_1|\psi_1\rangle + z_2|\psi_2\rangle) = z_1\hat{A}|\psi_1\rangle + z_2\hat{A}|\psi_2\rangle. \quad (2.37)$$

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

$$\hat{A}|\psi\rangle = \hat{A} \left\{ \sum_n \psi_n |n\rangle \right\} = \sum_n \psi_n \hat{A}|n\rangle, \quad (2.38)$$

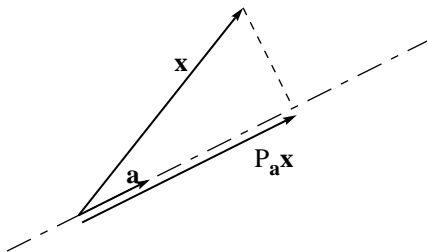
and the vectors $\hat{A}|n\rangle$ are known.

Examples of linear operators:

- The identity operator: $\hat{1}|\psi\rangle = |\psi\rangle$.
- 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” $\hat{P}_{|a\rangle}$, defined in terms of some fixed vector $|a\rangle$ as

$$\hat{P}_{|a\rangle} : |\psi\rangle \mapsto (\langle a|\psi\rangle) |a\rangle \quad (2.39)$$

This is often used for vectors $|a\rangle$ of norm 1, in which case, for arrows in space, it looks like:



- More generally, for a given $|a\rangle$ and $|b\rangle$ the operator

$$\hat{S} : |\psi\rangle \mapsto (\langle b|\psi\rangle) |a\rangle \quad (2.40)$$

is linear.

Operators may not commute. That is, we might well have

$$\hat{A}_1 \hat{A}_2 |\psi\rangle \neq \hat{A}_2 \hat{A}_1 |\psi\rangle. \quad (2.41)$$

Non-linear operators also exist, such as $\hat{N} : |\psi\rangle \mapsto (\langle \psi|\psi\rangle) |\psi\rangle$, but are not so important in applications to quantum mechanics. [[This is the source of the name *linear* algebra. For non-linear operators, knowledge of the action of \hat{N} on the basis vectors is *not* sufficient to define the operator. It is a mystery why all the exact operators in quantum mechanics are linear.]]

Outer products

Recall the operator

$$\begin{aligned} \hat{S} : |\psi\rangle &\mapsto (\langle b|\psi\rangle) |a\rangle = |a\rangle \langle b|\psi\rangle \\ \hat{S}|\psi\rangle &= |a\rangle \langle b|\psi\rangle \end{aligned} \quad (2.42)$$

We will write the operator \hat{S} as $|a\rangle \langle b|$ and call it “the outer product of $|a\rangle$ and $|b\rangle$ ”. This means neither more nor less than the defining equation (2.42).

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

$$\hat{T} \equiv |1\rangle \langle 1| + |2\rangle \langle 2| + \dots + |N\rangle \langle N|. \quad (2.43)$$

The effect of this operator on an arbitrary vector $|\psi\rangle$ is given in equation (2.29), which shows that $\hat{T}|\psi\rangle = |\psi\rangle$ for any $|\psi\rangle$. Hence my favorite equation

$$\hat{1} = \sum_n |n\rangle \langle n|. \quad (2.44)$$

This might look like magic, but it means nothing more than equation (2.29): that a vector may be resolved into its components. The operator of equation (2.44) 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 operators

Operators are represented by $N \times N$ matrices. If

$$|\phi\rangle = \hat{A}|\psi\rangle, \quad (2.45)$$

then

$$\begin{aligned}
 \langle n|\phi\rangle &= \langle n|\hat{A}|\phi\rangle \\
 &= \langle n|\hat{A}\hat{1}|\phi\rangle \\
 &= \langle n|\hat{A}\left\{\sum_m|m\rangle\langle m|\right\}|\phi\rangle \\
 &= \sum_m \langle n|\hat{A}|m\rangle\langle m|\psi\rangle,
 \end{aligned} \tag{2.46}$$

or, in matrix form,

$$\begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{pmatrix} = \begin{pmatrix} \langle 1|\hat{A}|1\rangle & \langle 1|\hat{A}|2\rangle & \cdots & \langle 1|\hat{A}|N\rangle \\ \langle 2|\hat{A}|1\rangle & \langle 2|\hat{A}|2\rangle & \cdots & \langle 2|\hat{A}|N\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle N|\hat{A}|1\rangle & \langle N|\hat{A}|2\rangle & \cdots & \langle N|\hat{A}|N\rangle \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}. \tag{2.47}$$

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

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 “Hermitian¹ conjugate” (or “Hermitian adjoint”) of \hat{A} such that

$$\langle \psi|\hat{A}^\dagger|\phi\rangle = \langle \phi|\hat{A}|\psi\rangle^* \tag{2.48}$$

for all vectors $|\psi\rangle$ and $|\phi\rangle$.

If the matrix elements for \hat{A} are $M_{n,m}$, then the matrix elements for \hat{A}^\dagger are $K_{n,m} = M_{m,n}^*$.

¹Charles 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.

Hermitian operators

If $\hat{H}^\dagger = \hat{H}$, then \hat{H} is said to be *Hermitian*. Matrix representations of Hermitian operators have $M_{n,m} = M_{m,n}^*$. Hermitian operators are important in quantum mechanics because if an operator is to correspond to an observable, then that operator must be Hermitian.

Theorem: 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 $\{|1\rangle, |2\rangle, \dots, |N\rangle\}$, and $\hat{H}|n\rangle = \lambda_n|n\rangle$, then

$$\hat{H} = \lambda_1|1\rangle\langle 1| + \lambda_2|2\rangle\langle 2| + \dots + \lambda_N|N\rangle\langle N|. \quad (2.49)$$

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

$$\hat{H} \doteq \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \dots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{pmatrix}. \quad (2.50)$$

2.10 Problems

2.8 Change of basis

The set $\{|a\rangle, |b\rangle\}$ is an orthonormal basis.

- a. Show that the set $\{|a'\rangle, |b'\rangle\}$, where

$$\begin{aligned} |a'\rangle &= +\cos\phi|a\rangle + \sin\phi|b\rangle \\ |b'\rangle &= -\sin\phi|a\rangle + \cos\phi|b\rangle \end{aligned}$$

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 $\phi = 45^\circ$.)

2.9 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?

2.10 **Change of representation, II**

Same as the previous problem, but use $|a'\rangle = i|a\rangle$ and $|b'\rangle = -i|b\rangle$.

2.11 **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?

2.12 **Outer product**

Using the $\{|z+\rangle, |z-\rangle\}$ basis representations

$$|\psi\rangle \doteq \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} \quad |\phi\rangle \doteq \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix} \quad |\theta+\rangle \doteq \begin{pmatrix} \cos \theta/2 \\ \sin \theta/2 \end{pmatrix} \quad |\theta-\rangle \doteq \begin{pmatrix} -\sin \theta/2 \\ \cos \theta/2 \end{pmatrix},$$

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

$$\langle\phi|\psi\rangle = \langle\phi|\theta+\rangle\langle\theta+|\psi\rangle + \langle\phi|\theta-\rangle\langle\theta-|\psi\rangle.$$

2.13 **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.

2.14 **The trace**

For any $N \times N$ matrix A (with components a_{ij}) the *trace* of A is defined by

$$\text{tr}(A) = \sum_{i=1}^N a_{ii}$$

Show that $\text{tr}(AB) = \text{tr}(BA)$, and hence that $\text{tr}(ABCD) = \text{tr}(DABC) = \text{tr}(CDAB)$, etc. (the so-called “cyclic invariance” of the trace). However, show that $\text{tr}(ABC)$ does not generally equal $\text{tr}(CBA)$ by constructing a counterexample. (Assume all matrices to be square.)

2.15 **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.)

$$\begin{aligned} x &= (x_1 \ x_2 \ \dots \ x_N) \\ y &= (y_1 \ y_2 \ \dots \ y_N) \\ x \otimes y &= \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} (y_1^* \ y_2^* \ \dots \ y_N^*) = \begin{pmatrix} x_1 y_1^* & x_1 y_2^* & \dots & x_1 y_N^* \\ x_2 y_1^* & x_2 y_2^* & \dots & x_2 y_N^* \\ \vdots & \vdots & \ddots & \vdots \\ x_N y_1^* & x_N y_2^* & \dots & x_N y_N^* \end{pmatrix}. \end{aligned}$$

This so-called “outer product” is quite different from the familiar “dot product” or “inner product”

$$x \cdot y = (x_1^* \ x_2^* \ \dots \ x_N^*) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = x_1^* y_1 + x_2^* y_2 + \dots + x_N^* y_N.$$

Write a formula for the i, j component of $x \otimes y$ and use it to show that $\text{tr}(y \otimes x) = x \cdot y$.

2.16 Pauli matrix algebra

Three important matrices are the *Pauli matrices*:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(Sometimes they are called $\sigma_1, \sigma_2, \sigma_3$ and other times they are called $\sigma_x, \sigma_y, \sigma_z$.)

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 in the “merry-go-round.”

c. Show that for any complex numbers c_1, c_2, c_3 ,

$$(c_1\sigma_1 + c_2\sigma_2 + c_3\sigma_3)^2 = (c_1^2 + c_2^2 + c_3^2)I.$$

2.17 Diagonalizing the Pauli matrices

Find the eigenvalues and corresponding (normalized) eigenvectors for all three Pauli matrices.

2.18 Exponentiation of Pauli matrices

Define exponentiation of matrices through

$$e^M = \sum_{n=0}^{\infty} \frac{M^n}{n!}.$$

a. Show that

$$e^{z\sigma_i} = \cosh(z)I + \sinh(z)\sigma_i \quad \text{for } i = 1, 2, 3.$$

(Hint: 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 $e^{\sigma_1}e^{\sigma_3} \neq e^{(\sigma_1 + \sigma_3)}$.

2.19 Hermitian operators

a. Show that if \hat{A} is a linear operator and $(a, \hat{A}a)$ is real for all vectors a , then \hat{A} is Hermitian. (Hint: Employ the hypothesis with $a = b + c$ and $a = b + ic$.)

b. 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.

c. You know that if an operator is Hermitian then all of its eigenvalues are real. Show that the converse is false by producing a counterexample. (Hint: Try a 2×2 upper triangular matrix.)

2.20 Unitary operators

Show that all the eigenvalues of a unitary operator have square modulus unity.

2.21 Commutator algebra

Prove that

$$\begin{aligned} [\hat{A}, b\hat{B} + c\hat{C}] &= b[\hat{A}, \hat{B}] + c[\hat{A}, \hat{C}] \\ [a\hat{A} + b\hat{B}, \hat{C}] &= a[\hat{A}, \hat{C}] + b[\hat{B}, \hat{C}] \\ [\hat{A}, \hat{B}\hat{C}] &= \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C} \\ [\hat{A}\hat{B}, \hat{C}] &= \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B} \\ 0 &= [\hat{A}, [\hat{B}, \hat{C}]] + [\hat{B}, [\hat{C}, \hat{A}]] + [\hat{C}, [\hat{A}, \hat{B}]] \quad (\text{the "Jacobi identity"}). \end{aligned}$$

Chapter 3

Formalism

Look at *A Quantum Mechanics Primer* by Daniel T. Gillespie, pages 1 through 70.

3.1 The density matrix

3.1 Definition

Consider a system in quantum state $|\psi\rangle$. Define the operator

$$\hat{\rho} = |\psi\rangle\langle\psi|,$$

called the *density matrix*, and show that the expectation value of the observable associated with operator \hat{A} in $|\psi\rangle$ is

$$\text{tr}\{\hat{\rho}\hat{A}\}.$$

3.2 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 μ 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 the 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 uncertainty, due to the way the world works). The density matrix formalism neatly handles both kinds of uncertainty at once.

If the system could be in any of the states $|a\rangle, |b\rangle, \dots, |i\rangle, \dots$ (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 expectation value of the observable associated with \hat{A} is still given by

$$\text{tr}\{\hat{\rho}\hat{A}\}.$$

3.3 Trace of the density matrix

Show that $\text{tr}\{\hat{\rho}\} = 1$. (This can be either a long and tedious proof, or a short and insightful one.)

Chapter 4

Time Evolution

4.1 Operator for time evolution

You now are at the point in quantum mechanics where you were when you first stepped into the door of your classical mechanics classroom: you know what you're trying to calculate.

But! How to calculate it? If quantum mechanics is to have a classical limit, then quantal states have to change with time. We write this time dependence explicitly as

$$|\psi(t)\rangle. \tag{4.1}$$

We seek the equations that govern this time evolution, the ones parallel to the classical time development equations, be they the Newtonian equations

$$\sum \vec{F} = m\vec{a} \tag{4.2}$$

or the Lagrange equations

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \tag{4.3}$$

or the Hamilton equations

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i, \quad \frac{\partial H}{\partial p_i} = \dot{q}_i. \tag{4.4}$$

Assume the existence of some “time development operator” $\hat{U}(\Delta t)$ such that

$$|\psi(t + \Delta t)\rangle = \hat{U}(\Delta t)|\psi(t)\rangle. \tag{4.5}$$

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-development 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 to assume that the time-development operator can be expanded in a Taylor series:

$$\hat{U}(\Delta t) = \hat{U}(0) + \hat{A}\Delta t + \hat{B}(\Delta t)^2 + \dots \quad (4.6)$$

We know that $\hat{U}(0) = \hat{1}$, and we'll write the quadratic and higher-order terms as $\hat{B}(\Delta t)^2 + \dots = \mathcal{O}(\Delta t^2)$... 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

$$\hat{U}(\Delta t) = \hat{1} - \frac{i}{\hbar}\hat{H}\Delta t + \mathcal{O}(\Delta t^2). \quad (4.7)$$

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, we will see that it corresponds to the classical energy.¹

[[The energy operator is called "the Hamiltonian" and represented by the letter \hat{H} in honor of William Rowan Hamilton, who first pointed out the central role that energy can play in time development in the formal theory of classical mechanics. Hamilton (1805–1865) made important contributions to mathematics, optics, classical mechanics, and astronomy. At the age of 22 years, while still an undergraduate, he was appointed professor of astronomy at his university and the Royal Astronomer of Ireland. He was not related to the American founding father Alexander Hamilton.]]

Theorem: The operator \hat{H} defined above is Hermitian.

Proof: The proof uses the fact that the norm of $|\psi(t + \Delta t)\rangle$ equals the norm of $|\psi(t)\rangle$:

$$|\psi(t + \Delta t)\rangle = |\psi(t)\rangle - \frac{i}{\hbar}\Delta t \underbrace{\hat{H}|\psi(t)\rangle}_{\equiv |\psi_H(t)\rangle} + \mathcal{O}(\Delta t^2). \quad (4.8)$$

¹For now, you can just use dimensional analysis to see that it has the correct *dimensions* for energy.

Thus

$$\langle \psi(t + \Delta t) | \psi(t + \Delta t) \rangle = \left[\langle \psi(t) | + \frac{i}{\hbar} \Delta t \langle \psi_H(t) | + \mathcal{O}(\Delta t^2) \right] \left[|\psi(t)\rangle - \frac{i}{\hbar} \Delta t |\psi_H(t)\rangle + \mathcal{O}(\Delta t^2) \right] \quad (4.9)$$

$$= \langle \psi(t) | \psi(t) \rangle + \frac{i}{\hbar} \Delta t \left[\langle \psi_H(t) | \psi(t) \rangle - \langle \psi(t) | \psi_H(t) \rangle \right] + \mathcal{O}(\Delta t^2) \quad (4.10)$$

$$1 = 1 + \frac{i}{\hbar} \Delta t \left[\langle \psi(t) | \psi_H(t) \rangle^* - \langle \psi(t) | \psi_H(t) \rangle \right] + \mathcal{O}(\Delta t^2) \quad (4.11)$$

$$0 = \frac{i}{\hbar} \Delta t \left[\langle \psi(t) | \hat{H} | \psi(t) \rangle^* - \langle \psi(t) | \hat{H} | \psi(t) \rangle \right] + \mathcal{O}(\Delta t^2). \quad (4.12)$$

This equation has to hold *for all values of Δt* , so the quantity in square brackets must vanish! That is,

$$\langle \psi(t) | \hat{H} | \psi(t) \rangle^* = \langle \psi(t) | \hat{H} | \psi(t) \rangle \quad (4.13)$$

for *all* vectors $|\psi(t)\rangle$. A simple exercise (problem 2.19, part a) then shows that operator \hat{H} is Hermitian. ■

We have written the time-development equation as

$$|\psi(t + \Delta t)\rangle = |\psi(t)\rangle - \frac{i}{\hbar} \Delta t \hat{H} |\psi(t)\rangle + \mathcal{O}(\Delta t^2). \quad (4.14)$$

Rearrangement gives

$$\frac{|\psi(t + \Delta t)\rangle - |\psi(t)\rangle}{\Delta t} = -\frac{i}{\hbar} \hat{H} |\psi(t)\rangle + \mathcal{O}(\Delta t). \quad (4.15)$$

In the limit $\Delta t \rightarrow 0$, this gives

$$\frac{d|\psi(t)\rangle}{dt} = -\frac{i}{\hbar} \hat{H} |\psi(t)\rangle \quad (4.16)$$

which is an important result known as the Schrödinger² equation!

Time evolution of projection probabilities

Theorem: If $|\phi\rangle$ is a time-independent state and $\hat{P}_\phi = |\phi\rangle\langle\phi|$ is its associated projection operator, then

$$\frac{d}{dt} |\langle\phi|\psi(t)\rangle|^2 = -\frac{i}{\hbar} \langle[\hat{P}_\phi, \hat{H}]\rangle. \quad (4.17)$$

Proof:

$$\begin{aligned} \frac{d}{dt} |\langle\phi|\psi(t)\rangle|^2 &= \frac{d}{dt} [\langle\phi|\psi(t)\rangle \langle\phi|\psi(t)\rangle^*] \\ &= \left[\langle\phi| \frac{d}{dt} |\psi(t)\rangle \right] \langle\phi|\psi(t)\rangle^* + \langle\phi|\psi(t)\rangle \left[\langle\phi| \frac{d}{dt} |\psi(t)\rangle \right]^* \end{aligned}$$

²Erwin 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-development equation that now bears his name. This led, in 1927, to a prestigious appointment in Berlin. In 1933, disgusted with the Nazi regime, he left Berlin for Oxford, England. He held several positions in various cities before ending up in Dublin. There he wrote a book titled *What is Life?* which is widely credited for stimulating interest in what had been a backwater of science: biochemistry.

But $\langle \phi | \frac{d}{dt} |\psi(t)\rangle = -\frac{i}{\hbar} \langle \phi | \hat{H} |\psi(t)\rangle$, so

$$\begin{aligned} \frac{d}{dt} |\langle \phi | \psi(t)\rangle|^2 &= -\frac{i}{\hbar} \left[\langle \phi | \hat{H} |\psi(t)\rangle \langle \phi | \psi(t)\rangle^* - \langle \phi | \psi(t)\rangle \langle \phi | \hat{H} |\psi(t)\rangle^* \right] \\ &= -\frac{i}{\hbar} \left[\langle \psi(t) | \phi \rangle \langle \phi | \hat{H} |\psi(t)\rangle - \langle \psi(t) | \hat{H} | \phi \rangle \langle \phi | \psi(t)\rangle \right] \\ &= -\frac{i}{\hbar} \left[\langle \psi(t) | \left\{ | \phi \rangle \langle \phi | \hat{H} - \hat{H} | \phi \rangle \langle \phi | \right\} | \psi(t)\rangle \right] \\ &= -\frac{i}{\hbar} \langle \psi(t) | [\hat{P}_\phi, \hat{H}] | \psi(t)\rangle \quad \blacksquare \end{aligned}$$

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: From the compatibility theorem, 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

$$\hat{B} |b_1\rangle\langle b_1| = \sum_n |b_n\rangle\langle b_n |b_1\rangle\langle b_1| = \sum_n |b_n\rangle\delta_{n,1}\langle b_1| = b_1 |b_1\rangle\langle b_1|$$

while

$$|b_1\rangle\langle b_1| \hat{B} = \sum_n |b_1\rangle\langle b_1 |b_n\rangle\langle b_n| = \sum_n |b_1\rangle\delta_{1,n}\langle b_n| = b_1 |b_1\rangle\langle b_1|.$$

4.2 Working with the Schrödinger equation

Quantal states evolve according to the Schrödinger time-development equation

$$\frac{d}{dt} |\psi(t)\rangle = -\frac{i}{\hbar} \hat{H} |\psi(t)\rangle. \quad (4.18)$$

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 development is governed by

$$\sum \mathbf{F} = m\mathbf{a}, \quad (4.19)$$

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.

Representations of the Schrödinger equation

As usual, we become familiar with states through their components, that is through their representations in a particular basis:

$$|\psi(t)\rangle = \sum_n \psi_n |n\rangle. \quad (4.20)$$

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 development 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
$\{ n\rangle\}$	$\psi_n(t)$	Schrödinger picture
ψ_n	$\{ n(t)\rangle\}$	Heisenberg picture
nothing	$\psi_n(t), \{ n(t)\rangle\}$	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

$$\frac{d}{dt} \langle n|\psi(t)\rangle = -\frac{i}{\hbar} \langle n|\hat{H}|\psi(t)\rangle = -\frac{i}{\hbar} \sum_m \langle n|\hat{H}|m\rangle \langle m|\psi(t)\rangle, \quad (4.21)$$

or, in other words, according to

$$\frac{d\psi_n(t)}{dt} = -\frac{i}{\hbar} \sum_m H_{n,m} \psi_m(t) \quad \text{where, recall} \quad H_{n,m} = H_{m,n}^*. \quad (4.22)$$

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 development problem is

$$\begin{aligned} \text{Initial condition:} & \quad \psi_1(0) = 1 \\ \text{Differential equation:} & \quad \frac{d\psi_1(t)}{dt} = -\frac{i}{\hbar} E_g \psi_1(t), \end{aligned}$$

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,

$$|\psi(t)\rangle = e^{-(i/\hbar)E_g t} |1\rangle. \quad (4.23)$$

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

$$e^{-(i/\hbar)E_g t} = \cos((E_g/\hbar)t) - i \sin((E_g/\hbar)t). \quad (4.24)$$

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: 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.

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

$$|1\rangle = |z+\rangle \quad \text{and} \quad |2\rangle = |z-\rangle. \quad (4.25)$$

It's very easy to write down the differential equation

$$\frac{d}{dt} \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix} = -\frac{i}{\hbar} \begin{pmatrix} H_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{pmatrix} \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix} \quad (4.26)$$

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

$$U = -\boldsymbol{\mu} \cdot \mathbf{B} = -\mu_z B. \quad (4.27)$$

Our guess for the quantum Hamiltonian is simply to change quantities into operators

$$\hat{H} = -\hat{\mu}_z B \quad (4.28)$$

where

$$\hat{\mu}_z = (+\mu_B)|z+\rangle\langle z+| + (-\mu_B)|z-\rangle\langle z-| \quad (4.29)$$

is the quantum mechanical operator corresponding to the observable μ_z . (See equation 2.7.) 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 (4.28) as a reasonable starting point, and indeed it turns out to describe this system to high accuracy, although not perfectly.³

It is an easy exercise to show that in the basis

$$\{|z+\rangle, |z-\rangle\} = \{|1\rangle, |2\rangle\},$$

the Hamiltonian operator (4.28) is represented by the matrix

$$\begin{pmatrix} H_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{pmatrix} = \begin{pmatrix} -\mu_B B & 0 \\ 0 & +\mu_B B \end{pmatrix}. \quad (4.30)$$

Thus the differential equations (4.26) become

$$\begin{aligned} \frac{d\psi_1(t)}{dt} &= -\frac{i}{\hbar}(-\mu_B B)\psi_1(t) \\ \frac{d\psi_2(t)}{dt} &= -\frac{i}{\hbar}(+\mu_B B)\psi_2(t). \end{aligned}$$

The solutions are straightforward:

$$\begin{aligned} \psi_1(t) &= \psi_1(0)e^{-(i/\hbar)(-\mu_B B)t} \\ \psi_2(t) &= \psi_2(0)e^{-(i/\hbar)(+\mu_B B)t}. \end{aligned}$$

Stuff about initial state $|z+\rangle$.

Suppose the initial state is

$$|x+\rangle = |z+\rangle\langle z+|x+\rangle + |z-\rangle\langle z-|x+\rangle = |z+\rangle\frac{1}{\sqrt{2}} + |z-\rangle\frac{1}{\sqrt{2}}.$$

Then

$$\psi_1(0) = \frac{1}{\sqrt{2}} \quad \psi_2(0) = \frac{1}{\sqrt{2}}$$

so

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}e^{-(i/\hbar)(-\mu_B B)t}|z+\rangle + \frac{1}{\sqrt{2}}e^{-(i/\hbar)(+\mu_B B)t}|z-\rangle.$$

³If you want perfection, you'll need to go into some discipline other than science.

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's $+\mu_B$ again? That probability is the square of the amplitude

$$\begin{aligned} \langle x+ | \psi(t) \rangle &= \frac{1}{\sqrt{2}} e^{-(i/\hbar)(-\mu_B B)t} \langle x+ | z+\rangle + \frac{1}{\sqrt{2}} e^{-(i/\hbar)(+\mu_B B)t} \langle x+ | z-\rangle \\ &= \frac{1}{\sqrt{2}} e^{-(i/\hbar)(-\mu_B B)t} \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} e^{-(i/\hbar)(+\mu_B B)t} \frac{1}{\sqrt{2}} \\ &= \frac{1}{2} \left[e^{-(i/\hbar)(-\mu_B B)t} + e^{-(i/\hbar)(+\mu_B B)t} \right] \\ &= \frac{1}{2} \left[2 \cos((1/\hbar)(\mu_B B)t) \right] \\ &= \cos\left(\frac{\mu_B B}{\hbar} t\right) \end{aligned}$$

The probability is

$$|\langle x+ | \psi(t) \rangle|^2 = \cos^2\left(\frac{\mu_B B}{\hbar} t\right) \quad (4.31)$$

which starts at one when $t = 0$, then goes down to zero, then goes back up to one, with an oscillation period of

$$\frac{\pi\hbar}{\mu_B B}.$$

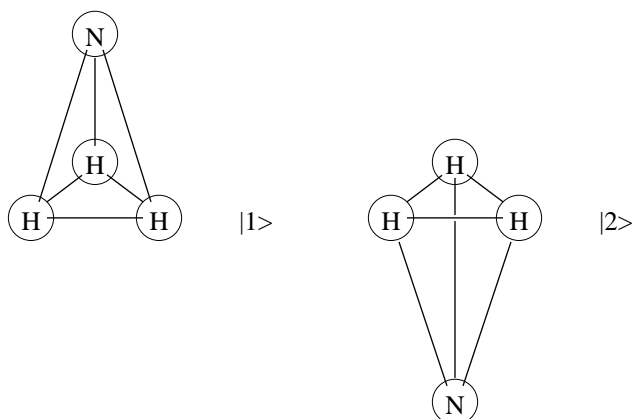
This phenomena is called “Rabi oscillation” — it is responsible for the workings of atomic clocks.⁴

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,⁵ then there are still two possible states for the molecule: state $|1\rangle$ with the nitrogen atom pointing up, and state $|2\rangle$ 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 $|1\rangle$, then some time later it might be in state $|2\rangle$, because the nitrogen atom tunneled through the plane of hydrogen atoms.

⁴Isidor Isaac Rabi won the Nobel Prize for nuclear magnetic resonance, but he also contributed to the invention of the laser and the atomic clock. No capsule biography suffices because he did so much. If you read Jeremy Bernstein’s profile of Rabi in the *New Yorker* (13 and 20 October 1975) you will see that he was a *very* clever man.

⁵That is, ignore vibration. These approximations seem, at first glance, to be absurd. They are in fact excellent approximations, because the tunneling happens so fast that the molecule doesn’t have time to translate, rotate, or vibrate to any significant extent during one cycle of tunneling.



What is the implication of such tunneling for the Hamiltonian matrix? The matrix we dealt with in equation (4.30) 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 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 $\{|1\rangle, |2\rangle\}$ basis has the form

$$\begin{pmatrix} H_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{pmatrix} = \begin{pmatrix} E & Ae^{i\phi} \\ Ae^{-i\phi} & E \end{pmatrix}. \quad (4.32)$$

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 $|1\rangle$ and $|2\rangle$ are mirror images and hence $\langle 1|\hat{H}|1\rangle = \langle 2|\hat{H}|2\rangle$.

For this Hamiltonian, the Schrödinger equation is

$$\frac{d}{dt} \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix} = -\frac{i}{\hbar} \begin{pmatrix} E & Ae^{i\phi} \\ Ae^{-i\phi} & E \end{pmatrix} \begin{pmatrix} \psi_1(t) \\ \psi_2(t) \end{pmatrix}. \quad (4.33)$$

It’s hard to see how to solve this pair of differential equations. The matrix is not diagonal, so the differential equation for $\psi_1(t)$ involves the unknown function $\psi_2(t)$, and the differential equation for $\psi_2(t)$ involves the unknown function $\psi_1(t)$. However, while it’s hard to solve in this initial basis, it would be easy to solve in a basis where the matrix is diagonal.

To diagonalize an $N \times N$ Hermitian matrix M :

1. In initial basis, the matrix representation of \hat{A} is 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

$$\det |M - \lambda I| = 0.$$

3. Find the representation \mathbf{e}_n of the eigenvector $|e_n\rangle$ by solving N simultaneous linear equations

$$\mathbf{M}\mathbf{e}_n = \lambda_n\mathbf{e}_n.$$

[In the above equation, \mathbf{M} is an $N \times N$ matrix, \mathbf{e}_n is an $N \times 1$ matrix (the N unknowns), and λ_n is a known number (determined in the previous step).]

4. In the basis $\{|e_1\rangle, |e_2\rangle, \dots, |e_N\rangle\}$, the matrix representation of \hat{A} is diagonal

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{pmatrix}.$$

Let's carry out these steps for the ammonia molecule problem.

1. The Hamiltonian is represented in the initial basis $\{|1\rangle, |2\rangle\}$ by

$$\mathbf{M} = \begin{pmatrix} E & Ae^{i\phi} \\ Ae^{-i\phi} & E \end{pmatrix}$$

2. Find the eigenvalues.

$$\begin{aligned} \det \begin{vmatrix} E - \lambda & Ae^{i\phi} \\ Ae^{-i\phi} & E - \lambda \end{vmatrix} &= 0 \\ (E - \lambda)^2 - A^2 &= 0 \\ (E - \lambda)^2 &= A^2 \\ E - \lambda &= \pm A \\ \lambda &= E \pm A \end{aligned} \tag{4.34}$$

$$\begin{aligned} \lambda_1 &= E + A \\ \lambda_2 &= E - A \end{aligned} \tag{4.35}$$

As required, the eigenvalues are real.

3. Find the eigenvectors.

We start with the eigenvector for $\lambda_1 = E + A$:

$$\begin{aligned} \mathbf{M}\mathbf{e}_1 &= \lambda_1\mathbf{e}_1 \\ (\mathbf{M} - \lambda_1\mathbf{I})\mathbf{e}_1 &= 0 \\ \begin{pmatrix} E - \lambda_1 & Ae^{i\phi} \\ Ae^{-i\phi} & E - \lambda_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} -A & Ae^{i\phi} \\ Ae^{-i\phi} & -A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \begin{pmatrix} -1 & e^{i\phi} \\ e^{-i\phi} & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ -x + e^{i\phi}y &= 0 \\ e^{-i\phi}x - y &= 0 \end{aligned}$$

These two are not independent equations! They cannot be. There are many eigenvectors because if, say

$$\begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

is an eigenvector, then so are

$$\begin{pmatrix} -1 \\ -5 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 10 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 3i \\ 15i \end{pmatrix},$$

and infinitely more eigenvectors.

The solution is $y = e^{-i\phi}x$, so

$$\mathbf{e}_1 = \begin{pmatrix} x \\ e^{-i\phi}x \end{pmatrix}.$$

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

$$\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{-i\phi} \end{pmatrix}.$$

This is the representation of $|e_1\rangle$ in the basis $\{|1\rangle, |2\rangle\}$.

Exercise. Show that an eigenvector associated with $\lambda_2 = E - A$ is

$$|e_2\rangle \doteq \mathbf{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ e^{-i\phi} \end{pmatrix}.$$

Exercise. Verify that $\langle e_1|e_2\rangle = 0$.

In summary,

$$\begin{aligned} |e_1\rangle &= \frac{1}{\sqrt{2}} [+|1\rangle + e^{-i\phi}|2\rangle] \\ |e_2\rangle &= \frac{1}{\sqrt{2}} [-|1\rangle + e^{-i\phi}|2\rangle]. \end{aligned} \quad (4.36)$$

Exercise. Show that $\{|e_1\rangle, |e_2\rangle\}$ constitute a spanning set by building $|1\rangle$ and $|2\rangle$ out of $|e_1\rangle$ and $|e_2\rangle$. (Answer: $|1\rangle = \frac{1}{\sqrt{2}}(|e_1\rangle - |e_2\rangle)$, $|2\rangle = \frac{1}{\sqrt{2}}e^{i\phi}(|e_1\rangle + |e_2\rangle)$.)

What are these states like?

- States $|1\rangle$ and $|2\rangle$ have definite positions for the nitrogen atom, namely “up” or “down”. But they don’t have definite energies. These states are sketched on page 53.
- 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 passing through “both branches” of an interferometer — the atom does not have a definite position.)

4. In the basis $\{|e_1\rangle, |e_2\rangle\}$, the matrix representation of the Hamiltonian is

$$\begin{pmatrix} E + A & 0 \\ 0 & E - A \end{pmatrix}.$$

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 development 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

$$|\psi(t)\rangle = e^{-(i/\hbar)Et} \left[e^{-(i/\hbar)At}\bar{\psi}_1(0)|e_1\rangle + e^{+(i/\hbar)At}\bar{\psi}_2(0)|e_2\rangle \right]. \quad (4.37)$$

(It is surprising 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 4.8.)

Let's try out this general solution for a particular initial condition. Suppose the nitrogen atom starts out “up” — that is,

$$|\psi(0)\rangle = |1\rangle, \quad (4.38)$$

and we ask for the probability of finding it “down” — that is, $|\langle 2|\psi(t)\rangle|^2$.

The initial expansion coefficients in the $\{|e_1\rangle, |e_2\rangle\}$ basis are (see equations (4.36))

$$\begin{aligned} \bar{\psi}_1(0) &= \langle e_1|\psi(0)\rangle = \langle e_1|1\rangle = \frac{1}{\sqrt{2}} \\ \bar{\psi}_2(0) &= \langle e_2|\psi(0)\rangle = \langle e_2|1\rangle = -\frac{1}{\sqrt{2}} \end{aligned}$$

so

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}e^{-(i/\hbar)Et} \left[e^{-(i/\hbar)At}|e_1\rangle - e^{+(i/\hbar)At}|e_2\rangle \right].$$

The amplitude to find the nitrogen atom “down” is

$$\begin{aligned} \langle 2|\psi(t)\rangle &= \frac{1}{\sqrt{2}}e^{-(i/\hbar)Et} \left[e^{-(i/\hbar)At}\langle 2|e_1\rangle - e^{+(i/\hbar)At}\langle 2|e_2\rangle \right] \\ &= \frac{1}{\sqrt{2}}e^{-(i/\hbar)Et} \left[e^{-(i/\hbar)At} \left(\frac{1}{\sqrt{2}}e^{-i\phi} \right) - e^{+(i/\hbar)At} \left(\frac{1}{\sqrt{2}}e^{-i\phi} \right) \right] \\ &= \frac{1}{2}e^{-i\phi}e^{-(i/\hbar)Et} \left[e^{-(i/\hbar)At} - e^{+(i/\hbar)At} \right] \\ &= \frac{1}{2}e^{-i\phi}e^{-(i/\hbar)Et} \left[-2i \sin \left((1/\hbar)At \right) \right] \\ &= -ie^{-i\phi}e^{-(i/\hbar)Et} \sin \left(\frac{A}{\hbar}t \right) \end{aligned}$$

and thus the probability of finding the nitrogen atom “down” is

$$|\langle 2|\psi(t)\rangle|^2 = \sin^2 \left(\frac{A}{\hbar}t \right). \quad (4.39)$$

GRAPH HERE

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 Stimulated Emission of Radiation).

Problems

4.1 Probability of no change

In equation (4.39) we found the probability that the nitrogen atom began in the “up” position (equation 4.38) 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.

4.2 Tunneling for small times

Equation (4.37) 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

$$|\psi(\Delta t)\rangle = \left[\hat{1} - \frac{i}{\hbar} \hat{H} \Delta t + \dots \right] |\psi(0)\rangle. \quad (4.40)$$

Show that this equation, represented in the $\{|1\rangle, |2\rangle\}$ basis, is

$$\begin{pmatrix} \psi_1(\Delta t) \\ \psi_2(\Delta t) \end{pmatrix} \approx \begin{pmatrix} 1 - (i/\hbar)E\Delta t & -(i/\hbar)Ae^{i\phi}\Delta t \\ -(i/\hbar)Ae^{-i\phi}\Delta t & 1 - (i/\hbar)E\Delta t \end{pmatrix} \begin{pmatrix} \psi_1(0) \\ \psi_2(0) \end{pmatrix}. \quad (4.41)$$

- b. Express the initial condition $|\psi(0)\rangle = |1\rangle$, used above at equation (4.38), in the $\{|1\rangle, |2\rangle\}$ basis, and show that, for small times,

$$\begin{pmatrix} \psi_1(\Delta t) \\ \psi_2(\Delta t) \end{pmatrix} \approx \begin{pmatrix} 1 - (i/\hbar)E\Delta t \\ -(i/\hbar)Ae^{i\phi}\Delta t \end{pmatrix}. \quad (4.42)$$

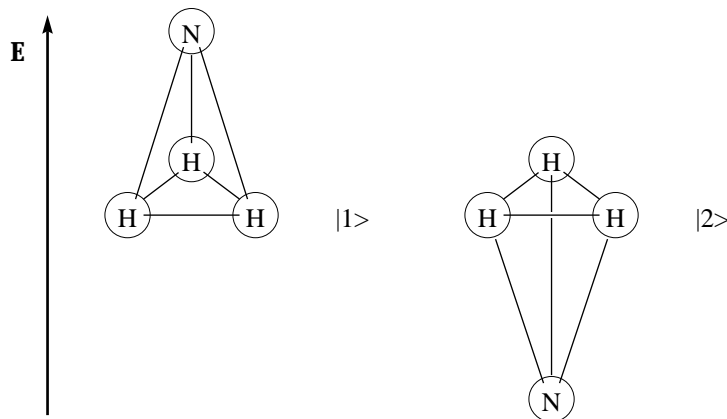
- c. This shows that the system starts with amplitude 1 for being in state $|1\rangle$, but that amplitude "seeps" (or "diffuses" or "hops") from $|1\rangle$ into $|2\rangle$. In fact, the amplitude to be found in $|2\rangle$ after a small time Δt has passed is $-(i/\hbar)Ae^{i\phi}\Delta t$. What is the probability of being found in $|2\rangle$? What is the condition for a "small" time?
- d. Show that the same probability results from approximating result (4.39) 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 (4.39).

- e. While this approach does indeed provide a lot of insight, it also raises a puzzle. What, according to equation (4.42), is the probability of being found in the initial state $|1\rangle$ after a short time has passed? Conclude that the total probability is greater than 1! We will resolve this paradox in problem 10.1

4.3 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 $|1\rangle$ and $|2\rangle$ are no longer symmetric, so we can no longer assume that $\langle 1|\hat{H}|1\rangle = \langle 2|\hat{H}|2\rangle$. Indeed, the proper matrix representation of \hat{H} in the $\{|1\rangle, |2\rangle\}$ 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. (Negative charge migrates toward the nitrogen atom.)

- Find the eigenvalues e_1 and e_2 of \hat{H} . (Check against the results (4.35) that apply when $\mathcal{E} = 0$.)
- Find the eigenvectors $|e_1\rangle$ and $|e_2\rangle$ in terms of $|1\rangle$ and $|2\rangle$. (Check against the results (4.36).)
- If a molecule is initially in state $|1\rangle$, find the probability that it will be found in state $|2\rangle$ as a function of time.

4.3 Formal properties of time evolution; Conservation laws

Quantal states evolve according to the Schrödinger time-development equation

$$\frac{d|\psi(t)\rangle}{dt} = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle. \quad (4.43)$$

The Hamiltonian operator \hat{H} is Hermitian, with eigenvectors $\{|e_n\rangle\}$ and eigenvalues e_n :

$$\hat{H}|e_n\rangle = e_n|e_n\rangle. \quad (4.44)$$

These are called the “energy eigenstates” or “states of definite energy”.

Theorem I: *Energy eigenstates are stationary states.*

If $|\psi(0)\rangle = (\text{number})|e_n\rangle$, then $|\psi(t)\rangle = (\text{number})'|e_n\rangle$, where both numbers have square modulus unity.

Because of this result, the energy eigenstates are also called “stationary states”: once you’re in one of them, you stay.

Proof: A formal proof will be given in the proof of theorem II. This informal proof provides less rigor and more insight.

Start at time $t = 0$ and step forward a small amount of time Δt :

$$\frac{\Delta|\psi\rangle}{\Delta t} \approx -\frac{i}{\hbar}\hat{H}|\psi(0)\rangle \quad (4.45)$$

$$= -\frac{i}{\hbar}\hat{H}(\text{number})|e_n\rangle \quad (4.46)$$

$$= (\text{stuff})|e_n\rangle. \quad (4.47)$$

$$\Delta|\psi\rangle = (\text{stuff})\Delta t|e_n\rangle. \quad (4.48)$$

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 desired. ■

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.

Theorem II: *Formal solution of the Schrödinger equation.*

$$\text{If } |\psi(0)\rangle = \sum_n \psi_n(0)|e_n\rangle, \text{ then } |\psi(t)\rangle = \sum_n \psi_n(0)e^{-(i/\hbar)e_n t}|e_n\rangle.$$

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,

$$H_{n,m} = \begin{cases} e_n & n = m \\ 0 & n \neq m \end{cases} = e_n \delta_{n,m}.$$

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

$$\psi_n(t) = \psi_n(0)e^{-(i/\hbar)e_n t}. \quad \blacksquare$$

So, this is how states change with time! What about measurements? We will first find how average values change with time, then look at “the whole shebang” – not just the average, but also the full distribution.

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

Theorem III: *Time evolution of averages.*

$$\frac{d\langle\hat{A}\rangle}{dt} = -\frac{i}{\hbar}\langle[\hat{A}, \hat{H}]\rangle.$$

Proof: (Using mathematical notation for inner products.)

$$\begin{aligned} \frac{d}{dt}\langle\hat{A}\rangle &= \frac{d}{dt}(\psi(t), \hat{A}\psi(t)) \\ &= \left(\frac{d\psi(t)}{dt}, \hat{A}\psi(t)\right) + \left(\psi(t), \hat{A}\frac{d\psi(t)}{dt}\right) \\ &= \left(-\frac{i}{\hbar}\hat{H}\psi(t), \hat{A}\psi(t)\right) + \left(\psi(t), \hat{A}\left[-\frac{i}{\hbar}\hat{H}\psi(t)\right]\right) \quad \text{[[use the fact that } \hat{H} \text{ is Hermitian]]} \\ &= \frac{i}{\hbar}(\psi(t), \hat{H}\hat{A}\psi(t)) - \frac{i}{\hbar}(\psi(t), \hat{A}\hat{H}\psi(t)) \\ &= -\frac{i}{\hbar}(\psi(t), [\hat{A}\hat{H} - \hat{H}\hat{A}]\psi(t)) \\ &= -\frac{i}{\hbar}\langle[\hat{A}, \hat{H}]\rangle \quad \blacksquare \end{aligned}$$

Corollary: If \hat{A} commutes with \hat{H} , then $\langle\hat{A}\rangle$ is constant.

However, just because the average 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 IV: *Time evolution of projection probabilities.*

If $|\phi\rangle$ is a time-independent state and $\hat{P}_\phi = |\phi\rangle\langle\phi|$ is its associated projection operator, then

$$\frac{d}{dt}|\langle\phi|\psi(t)\rangle|^2 = -\frac{i}{\hbar}\langle[\hat{P}_\phi, \hat{H}]\rangle. \quad (4.49)$$

Proof:

$$\begin{aligned} \frac{d}{dt}|\langle\phi|\psi(t)\rangle|^2 &= \frac{d}{dt}[\langle\phi|\psi(t)\rangle\langle\phi|\psi(t)\rangle^*] \\ &= \left[\langle\phi|\frac{d}{dt}|\psi(t)\rangle\right]\langle\phi|\psi(t)\rangle^* + \langle\phi|\psi(t)\rangle\left[\langle\phi|\frac{d}{dt}|\psi(t)\rangle\right]^* \end{aligned}$$

But $\langle\phi|\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}\langle\phi|\hat{H}|\psi(t)\rangle$, so

$$\frac{d}{dt}|\langle\phi|\psi(t)\rangle|^2 = -\frac{i}{\hbar}\left[\langle\phi|\hat{H}|\psi(t)\rangle\langle\phi|\psi(t)\rangle^* - \langle\phi|\psi(t)\rangle\langle\phi|\hat{H}|\psi(t)\rangle^*\right]$$

$$\begin{aligned}
&= -\frac{i}{\hbar} \left[\langle \psi(t) | \phi \rangle \langle \phi | \hat{H} | \psi(t) \rangle - \langle \psi(t) | \hat{H} | \phi \rangle \langle \phi | \psi(t) \rangle \right] \\
&= -\frac{i}{\hbar} \left[\langle \psi(t) | \left\{ | \phi \rangle \langle \phi | \hat{H} - \hat{H} | \phi \rangle \langle \phi | \right\} | \psi(t) \rangle \right] \\
&= -\frac{i}{\hbar} \langle \psi(t) | [\hat{P}_\phi, \hat{H}] | \psi(t) \rangle \quad \blacksquare
\end{aligned}$$

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: From the compatibility theorem, 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

$$\hat{B}|b_1\rangle\langle b_1| = \sum_n |b_n\rangle\langle b_n|b_1\rangle\langle b_1| = \sum_n |b_n\rangle\delta_{n,1}\langle b_1| = b_1|b_1\rangle\langle b_1|$$

while

$$|b_1\rangle\langle b_1|\hat{B} = \sum_n |b_1\rangle\langle b_1|b_n\rangle\langle b_n| = \sum_n |b_1\rangle\delta_{1,n}\langle b_n| = b_1|b_1\rangle\langle b_1|. \quad \blacksquare$$

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: All these results apply to time evolution uninterrupted by measurements.

4.4 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 $|\bar{K}^0\rangle$.

4.4 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 .

4.5 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.)

4.6 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.

4.7 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.]]

4.8 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.

Chapter 5

Continuum Systems

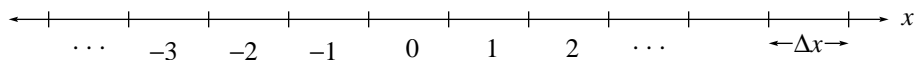
5.1 Describing states in continuum systems

At the start of this book 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. This led us to develop quantum mechanics for systems with two basis states. This was a very good thing, and we learned a lot about quantum mechanics, and also about practical applications like atomic clocks and MASERS.

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 continuum system. The system we'll pick is a particle moving 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 we'll treat both position and magnetic moment together.

Course-grained description

The situation is a point particle moving 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 definite position within bin 5 we say that it is in state $|5\rangle$.

I maintain that the set of states $\{|n\rangle\}$ with $n = 0, \pm 1, \pm 2, \pm 3, \dots$ 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

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

- 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

$$\sum_{n=-\infty}^{\infty} |n\rangle\langle n| = \hat{1}.$$

If the particle has no definite position, then it is in a state $|\psi\rangle$ that is a superposition of basis states (Have I used this word before?)

$$|\psi\rangle = \sum_{n=-\infty}^{\infty} \psi_n |n\rangle \quad (5.1)$$

where

$$\psi_n = \langle n|\psi\rangle \quad \text{so} \quad \sum_{n=-\infty}^{\infty} |\psi_n|^2 = 1. \quad (5.2)$$

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 I can come up with 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

$$\lim_{\Delta x \rightarrow 0} |\psi_k|^2 ?$$

In fact, this is not an interesting question. The answer to that question 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. (Put another way, the probability of finding a four-leaf clover at a point along the strip of lawn is zero, because that probability is

$$\frac{7}{\text{number of points}},$$

and the number of points along the strip is infinite.)

The interesting question concerns not the bin probability, which always goes to zero, but the *probability density*, that is, the probability of finding the particle per length.

Exercise. What is the probability density for finding a four-leaf clover in the strip of lawn described above? Be sure to include the dimensions in your answer.

The probability per length of finding the particle at x_0 , called the probability density at x_0 , is the finite quantity

$$\lim_{\Delta x \rightarrow 0} \frac{|\psi_k|^2}{\Delta x}. \quad (5.3)$$

(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 goes to a finite quantity. In other words, for small values of Δx , we have

$$|\psi_k|^2 \approx (\text{constant})\Delta x, \quad (5.4)$$

where that constant is the probability density for finding the particle at point x_0 .

We know that amplitudes are more general than probabilities, because probabilities give the results for measurement experiments, but amplitudes give the results for both interference and measurement experiments. What does equation (5.4) say about bin amplitudes? It says that for small values of Δx

$$\psi_k \approx (\text{constant})'\sqrt{\Delta x} \quad (5.5)$$

whence the limit

$$\lim_{\Delta x \rightarrow 0} \frac{\psi_k}{\sqrt{\Delta x}}$$

exists. This limit defines the quantity, a function of x_0 ,

$$\lim_{\Delta x \rightarrow 0} \frac{\psi_k}{\sqrt{\Delta x}} = \psi(x_0).$$

If I were naming this quantity, I would have named it “amplitude density”. But for historical reasons it has a different name, namely “the wavefunction”.

The wavefunction evaluated at x_0 is often 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”.

Exercise. Show that the wavefunction for a point particle in one dimension has the dimensions $1/\sqrt{\text{length}}$.

Working with wavefunctions

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 continuum 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 $\Delta x \rightarrow 0$ limit 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

$$\langle \phi | \psi \rangle = \lim_{\Delta x \rightarrow 0} \sum_{i=-\infty}^{\infty} \frac{\phi_i^*}{\sqrt{\Delta x}} \frac{\psi_i}{\sqrt{\Delta x}} \Delta x = \int_{-\infty}^{+\infty} \phi^*(x) \psi(x) dx.$$

Exercise. 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 $\Delta x \rightarrow 0$ 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 basis!¹ 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

$$\lim_{\Delta x \rightarrow 0} \psi_k = \lim_{\Delta x \rightarrow 0} \langle k | \psi \rangle = \left[\lim_{\Delta x \rightarrow 0} \langle k | \right] | \psi \rangle. \quad (5.6)$$

(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

$$\lim_{\Delta x \rightarrow 0} \frac{\psi_k}{\sqrt{\Delta x}} = \lim_{\Delta x \rightarrow 0} \frac{\langle k | \psi \rangle}{\sqrt{\Delta x}} = \left[\lim_{\Delta x \rightarrow 0} \frac{\langle k |}{\sqrt{\Delta x}} \right] | \psi \rangle = \psi(x_0). \quad (5.7)$$

This sequence motivates the definition of the “position basis state”

$$|x_0\rangle = \lim_{\Delta x \rightarrow 0} \frac{|k\rangle}{\sqrt{\Delta x}}. \quad (5.8)$$

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

¹You 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 that studies transfinite numbers.

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 again fascinating mathematical questions² but this is not a mathematics book, so we won’t make a big fuss over the distinction.

Completeness relation for continuum basis states:

$$\hat{1} = \sum_{i=-\infty}^{\infty} |i\rangle\langle i| = \lim_{\Delta x \rightarrow 0} \sum_{i=-\infty}^{\infty} \frac{|i\rangle}{\sqrt{\Delta x}} \frac{\langle i|}{\sqrt{\Delta x}} \Delta x = \int_{-\infty}^{+\infty} |x\rangle\langle x| dx. \quad (5.9)$$

Orthogonality relation for continuum basis states:

$$\begin{aligned} \langle i|j\rangle &= \delta_{i,j} \\ \langle x|y\rangle &= 0 \quad \text{when } x \neq y \\ \langle x|x\rangle &= \lim_{\Delta x \rightarrow 0} \frac{\langle i|i\rangle}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} = \infty \\ \langle x|y\rangle &= \delta(x-y). \end{aligned}$$

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 “generalized function” or a “Schwartz distribution”.

Comparison of discrete and continuous basis states

Discrete

basis states $|n\rangle$; dimensionless

$$\psi_n = \langle n|\psi\rangle$$

ψ_n is dimensionless

$$\sum_n |\psi_n|^2 = 1$$

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

$$\langle \phi|\psi\rangle = \sum_n \phi_n^* \psi_n$$

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

Continuous

basis states $|x\rangle$; dimensions $\frac{1}{\sqrt{\text{length}}}$

$$\psi(x) = \langle x|\psi\rangle$$

$\psi(x)$ has dimensions $\frac{1}{\sqrt{\text{length}}}$

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = 1$$

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

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

$$\int_{-\infty}^{+\infty} |x\rangle\langle x| dx = \hat{1}$$

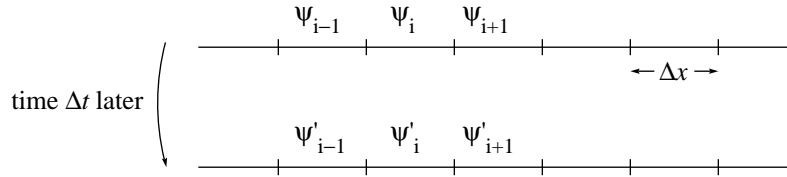
Exercise: Show that $\langle \phi|\psi\rangle = \int_{-\infty}^{+\infty} \phi^*(x)\psi(x) dx$. Hint: $\langle \phi|\psi\rangle = \langle \phi|\hat{1}|\psi\rangle$.

²If you find them interesting, take a course in rigged Hilbert spaces.

5.2 How does position amplitude change with time?

In classical mechanics, the equation telling us 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 motivate it.

The role of this section is to uncover the quantal equivalent of $\vec{F} = m\vec{a}$: namely the equation telling us 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 arguments in this section are suggestive, not definitive. Indeed, in some circumstances the arguments are false (e.g. for a single charged particle in a magnetic field, or for a pair of entangled particles).



Normalization requirement

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 $\Delta\psi$.) 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

$$1 = \sum_i \psi_i'^* \psi'_i = \sum_i (\psi_i^* + \Delta'\psi_i^*)(\psi_i + \Delta'\psi_i) = \sum_i (\psi_i^* \psi_i + \psi_i^* \Delta'\psi_i + \Delta'\psi_i^* \psi_i + \Delta'\psi_i^* \Delta'\psi_i).$$

The first term on the far right sums to exactly 1, due to initial normalization. The next two terms are of the form $z + z^* = 2\Re\{z\}$, so

$$0 = \sum_i 2\Re\{\psi_i^* \Delta'\psi_i\} + \Delta'\psi_i^* \Delta'\psi_i.$$

When we go to the limit of very small Δt , then $\Delta'\psi_i$ will be very small, and $\Delta'\psi_i^* \Delta'\psi_i$, as the product of two very small quantities, will be ultra small. Thus we neglect it and conclude that, due to normalization,

$$\Re\left\{\sum_i \psi_i^* \Delta'\psi_i\right\} = 0. \quad (5.10)$$

We can change this to a relation about wavefunction rather than bin amplitude by remembering that, if x_i is the point at the center of bin i , then

$$\psi(x_i) = \lim_{\Delta x \rightarrow 0} \frac{\psi_i}{\sqrt{\Delta x}} \quad (5.11)$$

For very small bins, equation (5.10) becomes

$$\Re e \left\{ \sum_i \psi^*(x_i) \sqrt{\Delta x} \Delta' \psi(x_i) \sqrt{\Delta x} \right\} = 0$$

or

$$\Re e \left\{ \int_{-\infty}^{+\infty} \psi^*(x) \Delta' \psi(x) dx \right\} = 0. \quad (5.12)$$

The flow of amplitude

Deductions from the preservation of normalization are important but purely formal...they don't tell us anything about the *physics* that's going on as time evolves. We begin with a very reasonable surmise:

$$\psi'_i = A_i \psi_{i-1} + B_i \psi_i + C_i \psi_{i+1}.$$

This says nothing more³ 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).

The only important 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".)

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 how much amplitude is originally in bin $i - 1$.

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 on a line containing a mountain and a valley, the flow will be more likely towards the valley than towards the mountain. However, this means 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 mountains and valleys — that "a particle in motion remains at motion in the absence of an external force" — and the flow amplitudes concern

³Compare the rules for combining amplitude on page 13.

this part of motion, the 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 mountain 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

$$\psi'_i = A\psi_{i-1} + B_i\psi_i + A\psi_{i+1}. \quad (5.13)$$

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

$$\Delta'\psi_i = \psi'_i - \psi_i = A(\psi_{i-1} - \psi_i) + D_i\psi_i + A(\psi_{i+1} - \psi_i),$$

which emphasizes amplitude differences rather than amplitude totals. In terms of the differences sketched below

this equation is

$$\Delta'\psi_i = -A\Delta\psi_L + D_i\psi_i + A\Delta\psi_R. \quad (5.14)$$

Writing this way, in terms of differences, sets us up 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

$$\Delta\psi_R - \Delta\psi_L = (\Delta x)^2 \left(\frac{\frac{\Delta\psi_R}{\Delta x} - \frac{\Delta\psi_L}{\Delta x}}{\Delta x} \right)$$

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 (5.14) by $\sqrt{\Delta x}$ and use equation (5.11) to write (in an approximation that grows increasingly accurate as $\Delta x \rightarrow 0$)

$$\Delta'\psi(x_i) \approx -A(\Delta x)^2 \left(\frac{\partial^2 \psi}{\partial x^2} \right)_{x=x_i} + D_i\psi(x_i)$$

While I have written this equation for the point at the center of bin i , of course it holds for any point. Defining $D(x_i) = D_i$ gives

$$\Delta'\psi(x) \approx -A(\Delta x)^2 \frac{\partial^2 \psi}{\partial x^2} + D(x)\psi(x). \quad (5.15)$$

Using the normalization equation

This is a good time to use result (5.12), the consequence of the normalization requirement. Applying (5.15) in (5.12) shows that

$$\int_{-\infty}^{+\infty} \psi^*(x) \Delta' \psi(x) dx = -A(\Delta x)^2 \int_{-\infty}^{+\infty} \psi^*(x) \frac{\partial^2 \psi}{\partial x^2} dx + \int_{-\infty}^{+\infty} \psi^*(x) D(x) \psi(x) dx \quad (5.16)$$

is 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. (That is, 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 (5.16) 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 (5.16) is

$$\int_{-\infty}^{+\infty} \psi^*(x) D(x) \psi(x) dx$$

which must be imaginary for *all wavefunctions* $\psi(x)$, even wavefunctions that are pure real. Thus $D(x)$ must be pure imaginary.

We have found that the amplitudes A and $D(x)$ must be pure imaginary, so we define the pure real quantities a and $d(x)$ through

$$A = ia \quad \text{and} \quad D(x) = id(x).$$

The discrete-time amplitude equation (5.15) becomes

$$\Delta' \psi(x) \approx i \left[-a(\Delta x)^2 \frac{\partial^2 \psi}{\partial x^2} + d(x) \psi(x) \right]. \quad (5.17)$$

Dimensional analysis

Let's find more about the quantity a , which is dimensionless. 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 . (It makes sense that a should depend on the inertia of the particle m , as we've already pointed out that this part of the Hamiltonian is involved with flow.)

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 $\Delta x \rightarrow 0$, so a must depend on Δx through the proportionality

$$a \propto \frac{1}{(\Delta x)^2} \quad \text{dimensions of right-hand side: } \frac{1}{[\ell]^2}.$$

To make a dimensionless we'll need to cancel the dimensions of length. The only way to do this is through \hbar :

$$a \propto \frac{\hbar}{(\Delta x)^2} \quad \text{dimensions of right-hand side: } \frac{[m]}{[t]}.$$

Now we need to cancel out the dimensions of mass and time. Again there is only one way to do this:

$$a \propto \frac{\hbar}{(\Delta x)^2} \frac{\Delta t}{m} \quad \text{dimensions of right-hand side: none.}$$

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 (5.13), the quantity a increases with Δt and decreases with Δx .

With our new understanding we write equation (5.17) 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 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 two limits $\Delta x \rightarrow 0$ and $\Delta t \rightarrow 0$, we find

$$\frac{\partial \psi(x, t)}{\partial t} = -\frac{i}{\hbar} \left[\frac{\hbar^2 n_d}{m} \frac{\partial^2 \psi(x, t)}{\partial x^2} - v(x) \psi(x, t) \right]. \quad (5.18)$$

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 5.4 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 see that the first part of the Hamiltonian corresponds to 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

$$\frac{\partial \psi(x, t)}{\partial t} = -\frac{i}{\hbar} \left[-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x) \psi(x, t) \right], \quad (5.19)$$

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. 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.

Bibliography:

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.

Gordon Baym, *Lectures on Quantum Mechanics* (Benjamin, Reading, Massachusetts, 1969) pages 46–53.

W. Moore, *Schrödinger: Life and Thought* (Cambridge University Press, 1989) page 194.

5.3 Operators and their representations; The momentum basis

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!

$$\hat{x}|x'\rangle = x'|x'\rangle.$$

Furthermore, we can find the action of \hat{x}^2 on every member of the $\{|x\rangle\}$ basis as follows:

$$\hat{x}^2|x'\rangle = \hat{x}(\hat{x}|x'\rangle) = \hat{x}(x'|x'\rangle) = x'(\hat{x}|x'\rangle) = x'(x'|x'\rangle) = (x')^2|x'\rangle.$$

Similarly, for any integer power n ,

$$\hat{x}^n|x'\rangle = (x')^n|x'\rangle.$$

Exercise: Prove this using mathematical induction.

If $f(x)$ is a scalar function with Taylor series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n, \quad (5.20)$$

then we define the *operator* $f(\hat{x})$ through

$$f(\hat{x}) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} \hat{x}^n. \quad (5.21)$$

(This enables us to find operators for quantities like e^x .) The upshot is that for these operators, the position basis states are eigenstates:

$$f(\hat{x})|x'\rangle = f(x')|x'\rangle.$$

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:

$$\begin{aligned} f(\hat{x})|\psi\rangle &= f(\hat{x})\hat{1}|\psi\rangle \\ &= f(\hat{x}) \int_{-\infty}^{+\infty} |x'\rangle \langle x'| dx' |\psi\rangle \\ &= \int_{-\infty}^{+\infty} f(\hat{x})|x'\rangle \langle x'|\psi\rangle dx' \\ &= \int_{-\infty}^{+\infty} |x'\rangle f(x') \langle x'|\psi\rangle dx'. \end{aligned}$$

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:

$$\begin{aligned}\langle x|f(\hat{x})|\psi\rangle &= \int_{-\infty}^{+\infty} \langle x|x'\rangle f(x')\langle x'|\psi\rangle dx' \\ &= \int_{-\infty}^{+\infty} \delta(x-x')f(x')\psi(x') dx' \\ &= f(x)\psi(x).\end{aligned}$$

IMPORTANT RESULT: The representation of an operator $f(\hat{x})$ in the position basis is

$$\langle x|f(\hat{x})|\psi\rangle = f(x)\langle x|\psi\rangle. \quad (5.22)$$

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.

The relation between a function-of-position operator and its position basis representation is simple: erase the hats!

$$|\phi\rangle = f(\hat{x})|\psi\rangle \iff \phi(x) = f(x)\psi(x). \quad (5.23)$$

Another application:

$$\langle \phi|f(\hat{x})|\psi\rangle = \int_{-\infty}^{+\infty} dx \langle \phi|x\rangle x|f(\hat{x})|\psi\rangle = \int_{-\infty}^{+\infty} \phi^*(x)f(x)\psi(x) dx. \quad (5.24)$$

So you might think we're home free. But no, because...

There are other operators

For example, the Hamiltonian operator, defined in terms of its components in the position basis, is

$$\langle x|\hat{H}|\psi\rangle = \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \langle x|\psi\rangle. \quad (5.25)$$

The logical definition of the momentum operator is through

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}), \quad (5.26)$$

so

$$\langle x|\hat{p}^2|\psi\rangle = -\hbar^2 \frac{\partial^2}{\partial x^2} \langle x|\psi\rangle. \quad (5.27)$$

IMPORTANT RESULT: Define the momentum operator \hat{p} in terms of its components / representation in the position basis as

$$\langle x|\hat{p}|\psi\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|\psi\rangle. \quad (5.28)$$

The operator with “+i” rather than “-i” out in front would have the same square, but would not have the correct classical limit. (See problem 5.7 and the second exercise below.)

Exercise: 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?

Exercise — 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 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.

Answer:

$$\begin{aligned} \langle x|\hat{p}_1|\psi_R\rangle &= -i\hbar \frac{\partial}{\partial x} Ae^{i(+kx-\omega t)} = -i\hbar(+ik)Ae^{i(+kx-\omega t)} = (+\hbar k)\psi_R(x;t) \\ \langle x|\hat{p}_1|\psi_L\rangle &= -i\hbar \frac{\partial}{\partial x} Ae^{i(-kx-\omega t)} = -i\hbar(-ik)Ae^{i(-kx-\omega t)} = (-\hbar k)\psi_L(x;t) \\ \langle x|\hat{p}_2|\psi_R\rangle &= +i\hbar \frac{\partial}{\partial x} Ae^{i(+kx-\omega t)} = +i\hbar(+ik)Ae^{i(+kx-\omega t)} = (-\hbar k)\psi_R(x;t) \\ \langle x|\hat{p}_2|\psi_L\rangle &= +i\hbar \frac{\partial}{\partial x} Ae^{i(-kx-\omega t)} = +i\hbar(-ik)Ae^{i(-kx-\omega t)} = (+\hbar k)\psi_L(x;t) \end{aligned}$$

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 intuitively associate rightward motion with positive momentum, candidate 1 is superior.

Check on \hat{p}^2 :

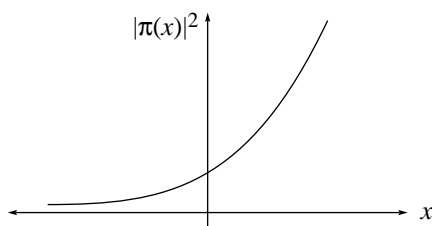
$$\begin{aligned} \langle x|\hat{p}^2|\psi\rangle &= \langle x|\hat{p}\hat{p}|\psi\rangle \quad \llbracket \text{define } |\phi\rangle = \hat{p}|\psi\rangle \rrbracket \\ &= \langle x|\hat{p}|\phi\rangle \\ &= -i\hbar \frac{\partial}{\partial x} \langle x|\phi\rangle \\ &= -i\hbar \frac{\partial}{\partial x} \left(\langle x|\hat{p}|\psi\rangle \right) \\ &= -i\hbar \frac{\partial}{\partial x} \left(-i\hbar \frac{\partial}{\partial x} \langle x|\psi\rangle \right) \\ &= -\hbar^2 \frac{\partial^2}{\partial x^2} \langle x|\psi\rangle \end{aligned}$$

Now that we have the momentum operator, we will of course want to find its eigenstates $|p\rangle$! (Purists will point out that these are not actually eigenstates, but rigged eigenstates.)

Find the position representation $\pi(x) = \langle x|p\rangle$ of the momentum eigenstates

$$\begin{aligned}
 \hat{p}|p\rangle &= \lambda|p\rangle \\
 \langle x|\hat{p}|p\rangle &= \lambda\langle x|p\rangle \\
 -i\hbar\frac{\partial}{\partial x}\langle x|p\rangle &= \lambda\langle x|p\rangle \\
 -i\hbar\frac{\partial\pi(x)}{\partial x} &= \lambda\pi(x) \\
 \frac{\partial\pi(x)}{\partial x} &= i\frac{\lambda}{\hbar}\pi(x) \\
 \pi(x) &= Ce^{i(\lambda/\hbar)x}
 \end{aligned} \tag{5.29}$$

That's funny. When we solve an eigenproblem, we expect that only a few eigenvalues λ 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 λ can be anything... positive, negative, even complex! A complex-valued λ will result in a probability density looking like this:



with an infinite pile of probability density off to the right or off to the left. This seems unphysical. Furthermore, complex values of λ would result in a non-Hermitian momentum operator, so we reject them. (Remember that in this section we are *not* making rigorous mathematical derivations, instead we are seeking sensible definitions.⁴ Complex-valued eigenvalues λ for the momentum operator are not sensible.⁵)

⁴“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.

⁵In exactly the same way, when you solve this classical trajectory problem —“Carol stands atop a 96 meter cliff and tosses

The constant C is just an overall normalization constant. The best convention is (see problem 5.1)

$$C = \frac{1}{\sqrt{2\pi\hbar}}. \quad (5.30)$$

In summary, the operator \hat{p} has eigenvectors $|p\rangle$ (technically, rigged vectors) satisfying

$$\hat{p}|p\rangle = p|p\rangle \quad (5.31)$$

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{i(p/\hbar)x}. \quad (5.32)$$

Exercise: Show that $|p\rangle$ has the dimensions of $1/\sqrt{\text{momentum}}$. What are the dimensions of $\langle x|p\rangle$?

Problem 5.1 will show that the momentum states are orthonormal

$$\langle p|p'\rangle = \delta(p - p') \quad (5.33)$$

and complete

$$\hat{1} = \int_{-\infty}^{+\infty} |p\rangle\langle p| dp, \quad (5.34)$$

and hence the set $\{|p\rangle\}$ constitutes a continuum (“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)

$$\psi(x) = \langle x|\psi\rangle. \quad (5.35)$$

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)

$$\tilde{\psi}(p) = \langle p|\psi\rangle. \quad (5.36)$$

Either representation carries complete information about the state $|\psi\rangle$, so you can obtain one from the other

$$\tilde{\psi}(p) = \langle p|\psi\rangle = \int_{-\infty}^{+\infty} \langle p|x\rangle\langle x|\psi\rangle dx = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-i(p/\hbar)x} \psi(x) dx \quad (5.37)$$

$$\psi(x) = \langle x|\psi\rangle = \int_{-\infty}^{+\infty} \langle x|p\rangle\langle p|\psi\rangle dp = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{+i(p/\hbar)x} \tilde{\psi}(p) dp. \quad (5.38)$$

In short, the position and momentum wavefunctions are related to each other through a Fourier transform!

a baseball at speed 45 m/s and angle 33° above the horizontal. When does the baseball hit the ground?” — you find two solutions: 7.6 s and -2.6 s. Even though the negative number is a mathematically correct solution, you reject it on physical grounds.

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,

$$\langle p|\hat{p}|\psi\rangle = [\langle\psi|\hat{p}|p\rangle]^* = [p\langle\psi|p\rangle]^* = p\langle p|\psi\rangle. \quad (5.39)$$

More generally, for any function of the momentum operator,

$$\langle p|f(\hat{p})|\psi\rangle = f(p)\langle p|\psi\rangle. \quad (5.40)$$

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!

$$\begin{aligned} \langle p|\hat{x}|\psi\rangle &= \langle p|\hat{x}\hat{1}|\psi\rangle \\ &= \int_{-\infty}^{+\infty} \langle p|\hat{x}|x\rangle\langle x|\psi\rangle dx \\ &= \int_{-\infty}^{+\infty} \langle p|x\rangle x\langle x|\psi\rangle dx \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-i(p/\hbar)x} x\langle x|\psi\rangle dx \\ &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \frac{\hbar}{-i} \frac{\partial}{\partial p} \left[e^{-i(p/\hbar)x} \right] \langle x|\psi\rangle dx \\ &= +i\hbar \frac{1}{\sqrt{2\pi\hbar}} \frac{\partial}{\partial p} \left[\int_{-\infty}^{+\infty} e^{-i(p/\hbar)x} x\langle x|\psi\rangle dx \right] \\ &= +i\hbar \frac{\partial}{\partial p} \left[\int_{-\infty}^{+\infty} \langle p|x\rangle\langle x|\psi\rangle dx \right] \\ &= +i\hbar \frac{\partial}{\partial p} \langle p|\psi\rangle \end{aligned} \quad (5.41)$$

There's a nice symmetry to this result, making it easy to remember: The momentum operator, represented in the position basis, is

$$\langle x|\hat{p}|\psi\rangle = -i\hbar \frac{\partial}{\partial x} \psi(x) \quad (5.42)$$

while the position operator, represented in the momentum basis, is

$$\langle p|\hat{x}|\psi\rangle = +i\hbar \frac{\partial}{\partial p} \tilde{\psi}(p) \quad (5.43)$$

Problems

5.1 The states $\{|p\rangle\}$ constitute a continuum basis

At equation (5.29) we showed that the inner product $\langle x|p\rangle$ must have the form

$$\langle x|p\rangle = C e^{i(p/\hbar)x} \quad (5.44)$$

where C may be chosen for convenience.

a. Show that the operator

$$\hat{A} = \int_{-\infty}^{\infty} |p\rangle\langle p| dp \quad (5.45)$$

is equal to

$$2\pi\hbar|C|^2\hat{1} \quad (5.46)$$

by evaluating

$$\langle\phi|\hat{A}|\psi\rangle = \langle\phi|\hat{1}\hat{A}\hat{1}|\psi\rangle \quad (5.47)$$

for arbitrary states $|\psi\rangle$ and $|\phi\rangle$. Hints: 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

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk \quad (5.48)$$

(see Griffiths equation [2.144] on page 77) 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

$$\langle p|p'\rangle = \delta(p - p'). \quad (5.49)$$

The expression (5.48) is again helpful.

5.2 Peculiarities of continuum basis states

Recall that the elements of a continuum 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

$$\int_{-\infty}^{\infty} \psi^*(x)\psi(x) dx = 1. \quad (5.50)$$

Show that the states $|x'\rangle$ and $|p\rangle$ cannot obey this normalization.

5.3 Hermiticity of the momentum operator

Show that the momentum operator is Hermitian over the space of states $|\psi\rangle$ that have wavefunction $\psi(x)$ which vanish at $x = \pm\infty$. Hint:

$$\langle\phi|\hat{p}|\psi\rangle = \int_{-\infty}^{\infty} \phi^*(x) \left(-i\hbar \frac{d\psi(x)}{dx} \right) dx. \quad (5.51)$$

Integrate by parts.

5.4 Commutator of \hat{x} and \hat{p}

Show that $[\hat{x}, \hat{p}] = i\hbar$ by showing that $\langle\phi|[\hat{x}, \hat{p}]|\psi\rangle = i\hbar\langle\phi|\psi\rangle$ for arbitrary $|\phi\rangle$ and $|\psi\rangle$. Hints: First evaluate $\langle x|\hat{p}\hat{x}|\psi\rangle$ and $\langle x|\hat{x}\hat{p}|\psi\rangle$. It helps to define $|\chi\rangle = \hat{x}|\psi\rangle$.

5.5 Momentum representation of the Schrödinger equation

You know that the Schrödinger equation

$$\frac{d|\psi(t)\rangle}{dt} = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle \quad (5.52)$$

has the position representation

$$\frac{\partial \langle x | \psi(t) \rangle}{\partial t} = -\frac{i}{\hbar} \langle x | \hat{H} | \psi(t) \rangle \quad (5.53)$$

or

$$i\hbar \frac{\partial \psi(x; t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x; t)}{\partial x^2} + V(x)\psi(x; t). \quad (5.54)$$

In this problem you will uncover the corresponding equation that governs the time development of

$$\tilde{\psi}(p; t) = \langle p | \psi(t) \rangle. \quad (5.55)$$

The left hand side of equation (5.52) is straightforward because

$$\langle p | \frac{d}{dt} | \psi(t) \rangle = \frac{\partial \tilde{\psi}(p; t)}{\partial t}. \quad (5.56)$$

To investigate the right hand side of equation (5.52) write

$$\hat{H} = \frac{1}{2m} \hat{p}^2 + \hat{V} \quad (5.57)$$

where \hat{p} is the momentum operator and \hat{V} the potential energy operator.

- a. Use the Hermiticity of \hat{p} to show that

$$\langle p | \hat{H} | \psi(t) \rangle = \frac{p^2}{2m} \tilde{\psi}(p; t) + \langle p | \hat{V} | \psi(t) \rangle. \quad (5.58)$$

Now we must investigate $\langle p | \hat{V} | \psi(t) \rangle$.

- b. Show that

$$\langle p | \hat{V} | \psi(t) \rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-i(p/\hbar)x} V(x) \psi(x; t) dx \quad (5.59)$$

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

$$\tilde{V}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-i(p/\hbar)x} V(x) dx \quad (5.60)$$

$$= \int_{-\infty}^{\infty} \langle p | x \rangle V(x) dx. \quad (5.61)$$

Note that $\tilde{V}(p)$ has funny dimensions. Show that

$$V(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{i(p/\hbar)x} \tilde{V}(p) dp \quad (5.62)$$

$$= \int_{-\infty}^{\infty} \langle x | p \rangle \tilde{V}(p) dp. \quad (5.63)$$

You may use either forms (5.60) and (5.62), in which case the proof employs equation (5.48), or forms (5.61) and (5.63), in which case the proof involves completeness and orthogonality of basis states.

d. Hence show that

$$\langle p | \hat{V} | \psi(t) \rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \tilde{V}(p-p') \tilde{\psi}(p'; t) dp'. \quad (5.64)$$

(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

$$\frac{\partial \tilde{\psi}(p; t)}{\partial t} = -\frac{i}{\hbar} \left[\frac{p^2}{2m} \tilde{\psi}(p; t) + \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp' \tilde{V}(p-p') \tilde{\psi}(p'; t) \right]. \quad (5.65)$$

This form of the Schrödinger equation is particularly useful in the study of superconductivity.

5.4 Time evolution of average quantities

In our general treatment of time evolution we found that for any measurable with operation \hat{A} , the average value $\langle \hat{A} \rangle_t$ changed with time according to

$$\frac{d\langle \hat{A} \rangle_t}{dt} = -\frac{i}{\hbar} \langle [\hat{A}, \hat{H}] \rangle_t. \quad (5.66)$$

For the systems of this chapter,

$$\hat{H} = \frac{1}{2m} \hat{p}^2 + V(\hat{x}), \quad (5.67)$$

where \hat{x} and \hat{p} satisfy the commutation relation

$$[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar. \quad (5.68)$$

Knowing this, let's see how the average 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

$$\frac{d\langle\hat{x}\rangle_t}{dt} = \frac{\langle\hat{p}\rangle_t}{m}, \quad (5.69)$$

a result that stirs our memories of classical mechanics!

Meanwhile, what happens for average 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!

$$F(x) = -\frac{\partial V(x)}{\partial x}. \quad (5.70)$$

Continuing the evaluation begun above,

$$\begin{aligned} \langle x|[\hat{p}, V(\hat{x})]|\psi\rangle &= i\hbar [F(x)\psi(x)] \\ &= i\hbar \langle x|F(\hat{x})|\psi\rangle. \end{aligned}$$

Because this relation holds for any $|x\rangle$ and for any $|\psi\rangle$, we know that the operators are related as

$$[\hat{p}, V(\hat{x})] = i\hbar F(\hat{x}). \quad (5.71)$$

Going back to the time evolution of average 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

$$\frac{d\langle\hat{p}\rangle_t}{dt} = \langle F(\hat{x})\rangle_t, \quad (5.72)$$

which is suspiciously close to Newton's second law!

These two results together,

$$\frac{d\langle\hat{x}\rangle_t}{dt} = \frac{\langle\hat{p}\rangle_t}{m} \quad (5.73)$$

$$\frac{d\langle\hat{p}\rangle_t}{dt} = \langle F(\hat{x})\rangle_t, \quad (5.74)$$

which tug so strongly on our classical heartstrings, are called the Ehrenfest⁶ equations. There are two things you should remember 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

5.6 Alternative derivation.

Derive result 5.71 by expanding $V(x)$ in a Taylor series.

5.7 Choice of sign for momentum operator

If we had taken the opposite sign choice for the momentum operator at equation (5.28) (call this choice \hat{p}_2), then what would have been the commutator $[\hat{x}\hat{p}_2]$? What would have been the result 5.69?

5.8 Quantities in the Hamiltonian

When we derived equation (5.18) 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.

⁶Paul Ehrenfest (1880–1933), Austrian and Dutch theoretical physicist. He was known particularly for asking probing questions that clarified the essence and delineated the unsolved problems of any matter under discussion. 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.

Chapter 6

The Free Particle

6.1 Problems

6.1 Energy eigenstates

In lecture we 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

$$|\rho(0)\rangle = A|+p_0\rangle + B|-p_0\rangle, \quad (6.1)$$

where $|A|^2 + |B|^2 = 1$, has definite energy $E(p_0) = p_0^2/2m$. (That is, $|\rho(0)\rangle$ is an energy eigenstate with eigenvalue $p_0^2/2m$).

- b. Show that the “wavefunction” corresponding to $|\rho(t)\rangle$ evolves in time as

$$\rho(x;t) = \frac{1}{\sqrt{2\pi\hbar}} \left[A e^{i(+p_0x - E(p_0)t)/\hbar} + B e^{i(-p_0x - E(p_0)t)/\hbar} \right]. \quad (6.2)$$

I use the term wavefunction in quotes because $\rho(x;t)$ is not $\langle x|\text{normal state}\rangle$ but rather a sum of two terms like $\langle x|\text{continuum basis state}\rangle$.

- c. Show that the “probability density” $|\rho(x;t)|^2$ is independent of time and given by

$$|\rho(x;t)|^2 = \frac{1}{2\pi\hbar} \left[1 + 2\Re\{A^*B\} \cos\left(\frac{2p_0x}{\hbar}\right) + 2\Im\{A^*B\} \sin\left(\frac{2p_0x}{\hbar}\right) \right]. \quad (6.3)$$

6.2 A useful integral

Using $\int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}$, show that

$$\int_{-\infty}^{\infty} e^{-u^2\alpha^2/2} e^{iuy} du = \frac{\sqrt{2\pi}}{\alpha} e^{-y^2/2\alpha^2} \quad (6.4)$$

where α may be complex, but $\Re\{\alpha^2\} > 0$. *Hint:* Complete the square by writing

$$-\frac{u^2\alpha^2}{2} + iuy = -\left(\frac{u\alpha}{\sqrt{2}} - i\frac{y}{\sqrt{2}\alpha}\right)^2 - \frac{y^2}{2\alpha^2}.$$

Note: If c is a real number independent of x , you know that

$$\lim_{x \rightarrow \infty} (x + c) = \infty.$$

You might think that a different limit would result if the additive constant c were complex, but in fact, that is not the case:

$$\lim_{x \rightarrow \infty} (x + ic) = \infty.$$

It is not unusual for the limit of a sequence of complex numbers to be real.

6.3 A somewhat less useful integral

Given $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$, show that

$$\int_{-\infty}^{\infty} x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{2}. \quad (6.5)$$

Hint: $\int_{-\infty}^{\infty} x^2 e^{-x^2} dx = 2 \int_0^{\infty} x^2 e^{-x^2} dx$, then integrate by parts.

6.4 Static properties of a Gaussian wavepacket

Consider the wavefunction

$$\psi(x; 0) = \frac{A}{\sqrt{\sigma}} e^{-x^2/2\sigma^2} e^{i(p_0/\hbar)x}. \quad (6.6)$$

- Show that the wavefunction is properly normalized when $A = 1/\sqrt[4]{\pi}$.
- Show that in this state $\langle \hat{x} \rangle = 0$ (trivial), and $\Delta x = \sqrt{\langle (\hat{x} - \langle \hat{x} \rangle)^2 \rangle} = \sigma/\sqrt{2}$ (easy).
- Use equation (6.4) to show that

$$\tilde{\psi}(p; 0) = A \sqrt{\frac{\sigma}{\hbar}} e^{-(p-p_0)^2 \sigma^2 / 2\hbar^2}. \quad (6.7)$$

- Hence show that $\langle \hat{p} \rangle = p_0$ and $\Delta p = \hbar/(\sqrt{2}\sigma)$.

6.5 Force-free motion of a Gaussian wavepacket

A particle with the initial wavefunction given in the previous problem evolves as

$$\tilde{\psi}(p; t) = e^{-(i/\hbar)E(p)t} \tilde{\psi}(p, 0) \quad (6.8)$$

so that

$$\psi(x; t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{i(px - E(p)t)/\hbar} \tilde{\psi}(p; 0) dp. \quad (6.9)$$

- a. Plug in $\tilde{\psi}(p; 0)$ and change the integration variable to k where $\hbar k = p - p_0$ in order to show that

$$\psi(x; t) = A \sqrt{\frac{\sigma}{2\pi}} e^{i(p_0 x - E(p_0)t)/\hbar} \int_{-\infty}^{\infty} e^{-k^2(\sigma^2 + i\frac{\hbar t}{m})/2} e^{ik(x - \frac{p_0}{m}t)} dk. \quad (6.10)$$

Hint: Change variable first to $p' = p - p_0$, then to $k = p'/\hbar$.

- b. Define the complex dimensionless quantity

$$\beta = 1 + i \frac{\hbar t}{m\sigma^2} \quad (6.11)$$

and evaluate the integral using equation (6.4), giving

$$\psi(x; t) = A \frac{1}{\sqrt{\sigma\beta}} e^{i(p_0 x - E(p_0)t)/\hbar} e^{-(x - \frac{p_0}{m}t)^2 / 2\sigma^2\beta}. \quad (6.12)$$

- c. Hence show that

$$|\psi(x; t)|^2 = \frac{1}{\sqrt{\pi}\sigma|\beta|} e^{-(x - \frac{p_0}{m}t)^2 / \sigma^2|\beta|^2}. \quad (6.13)$$

By comparing $|\psi(x; t)|^2$ with $|\psi(x; 0)|^2$, read off the results

$$\langle \hat{x} \rangle = \frac{p_0}{m}t, \quad \Delta x = \frac{\sigma|\beta|}{\sqrt{2}} = \frac{\sigma}{\sqrt{2}} \sqrt{1 + \left(\frac{\hbar}{m\sigma^2}t\right)^2}. \quad (6.14)$$

(No computation is required!)

Chapter 7

Square Wells

Will include both infinite and finite square wells. See the problem on “Scaling” below.

7.1 The infinite square well

You are no doubt familiar with the energy eigenproblem for the infinite square well, but a short review is never-the-less in order. Consider a well of width L , and place the origin at the left edge of the well. The mathematical problem is to find values E_n and functions $\eta_n(x)$ such that a solution to

$$-\frac{\hbar^2}{2m} \frac{d^2\eta_n}{dx^2} = E_n\eta_n(x) \quad (7.1)$$

has

$$\eta_n(0) = \eta_n(L) = 0. \quad (7.2)$$

Show that the eigenvalues are

$$E_n = n^2 \frac{\pi^2 \hbar^2}{2mL^2} \quad n = 1, 2, 3, 4, \dots \quad (7.3)$$

and the eigenfunctions are

$$\eta_n(x) = \sqrt{\frac{2}{L}} \sin\left(n\pi \frac{x}{L}\right). \quad (7.4)$$

Be sure to explain carefully why negative and zero values of n are not used. Note that $\eta_n(x)$ is even under reflection about the well center for n odd, odd for n even.

7.2 Ground state energy for the infinite square well

Make up a problem like problem 8.1.

7.3 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 $\hbar \rightarrow 0$? (*Clue:* Consider the classical limit.)
- b. ... this energy vanishes as $L \rightarrow \infty$? (*Clue:* Think about the Heisenberg indeterminacy relation. Compare problem 7.2.)
- c. ... this energy varies as $1/m$?

7.4 Scaling

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

Suppose $\eta(x)$ is a solution to

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

with a particular energy E . Now consider a different potential energy function $U(x) = s^2V(sx)$. For example, if $s = 2$ then $U(x)$ has the same shape as $V(x)$, but has half the width and four times the height. Call $\sigma(x) = \eta(sx)$. Show that $\sigma(x)$ solves the energy eigenproblem for potential energy $U(x)$, but with a different energy. Find that energy.

7.1 What does an electron look like?

One electron is in an infinite square well, in energy eigenstate $n = 2$, with a node right in the center of the well. The electron has probability $\frac{1}{2}$ of being found in the left half of the well, and probability $\frac{1}{2}$ of being found in the right half of the well, but no probability density at all of being found at the center. How can the electron move from the left half to the right half without passing through the center?

This question betrays a deepseated misconception about quantum mechanics. It arises from incorrectly thinking that the electron in energy eigenstate $n = 2$ has a definite position that we don't know (or that is changing erratically). That's a wrong picture: the electron 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 an electron. It is as silly and as unanswerable as the question "If love is blue and passion is red-hot, how can passionate love exist?"

A similar conundrum is this one: "Suppose I start with an electron in a state so that it has equal probability of being anywhere in a box. If I shine a strong light throughout the entire box I will find the electron at only one point. But what happens if I shine the light on only the left half of the box, and don't find the electron? I now know that the electron is somewhere in the right half. How could the light, shining where the electron isn't, affect the electron?" This conundrum (called the Renninger negative-result experiment) has the same resolution as the "passage through nodes" conundrum: namely, the conundrum arises from an incorrect visualization of the electron as a hard, tiny marble. Before the light shining, the electron didn't have a position. Instead the amplitude to be found in the left half had the same magnitude as the amplitude to be found in the right half.

Problems

7.5 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.

Chapter 8

The Simple Harmonic Oscillator

8.1 Resume of energy eigenproblem

The energy eigenproblem for the simple harmonic oscillator is

$$-\frac{\hbar^2}{2m} \frac{d^2 \eta_n(x)}{dx^2} + \frac{m\omega^2}{2} x^2 \eta_n(x) = E_n \eta_n(x). \quad (8.1)$$

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 (8.1) to find those few needles with finite norm.

8.2 Solution of the energy eigenproblem: Differential equation approach

Problem: Given m and ω , find values E_n such that the corresponding solutions $\eta_n(x)$ of

$$-\frac{\hbar^2}{2m} \frac{d^2 \eta_n(x)}{dx^2} + \frac{m\omega^2}{2} x^2 \eta_n(x) = E_n \eta_n(x) \quad (8.2)$$

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 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.

In this treatment, I'll play fast and loose with asymptotic analysis. But everything I'll do is both reasonable and rigorously justifiable. (C.M. Bender and S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* McGraw-Hill, New York, 1978.)

1. *Convert to dimensionless variable:* The only combination of m , ω , and \hbar with the dimensions of [length] is $\sqrt{\hbar/m\omega}$. Hence define the dimensionless variable proportional to length

$$q = \sqrt{\frac{m\omega}{\hbar}} x. \quad (8.3)$$

In terms of this variable, the ordinary differential equation (8.2) is

$$\frac{d^2 \eta_n(q)}{dq^2} + \left(\frac{2E_n}{\hbar\omega} - q^2 \right) \eta_n(q) = 0. \quad (8.4)$$

Exercise: 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?

2. *Remove asymptotic behavior of solutions:* Consider the limit as $q^2 \rightarrow \infty$. In this limit, the ODE (8.4) "becomes"

$$\frac{d^2 \eta_n(q)}{dq^2} - q^2 \eta_n(q) = 0, \quad (8.5)$$

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

$$f_n(q) = e^{-q^2/2}. \quad (8.6)$$

When we test to see whether this is a solution, we find

$$\begin{aligned} & \frac{d^2 f_n(q)}{dq^2} - q^2 f_n(q) \\ &= \left(q^2 e^{-q^2/2} - e^{-q^2/2} \right) - q^2 e^{-q^2/2} = -e^{-q^2/2} \end{aligned}$$

So the function (8.6) does *not* solve the ODE (8.5). On the other hand, the amount by which it "misses" solving (8.5) is small in the sense that

$$\lim_{q^2 \rightarrow \infty} \frac{d^2 f/dq^2 - q^2 f}{q^2 f} = \lim_{q^2 \rightarrow \infty} \frac{-e^{-q^2/2}}{q^2 e^{-q^2/2}} = \lim_{q^2 \rightarrow \infty} \frac{-1}{q^2} = 0.$$

A similar result holds for $g_n(x) = e^{+q^2/2}$.

Our conclusion is that, in the limit $q^2 \rightarrow \infty$, the solution $\eta_n(q)$ behaves like

$$\eta_n(q) \approx Ae^{-q^2/2} + Be^{+q^2/2}.$$

If $B \neq 0$, then $\eta_n(q)$ will not be normalizable because the probability density would become infinite as $q^2 \rightarrow \infty$. Thus the solutions we want — the normalizable solutions — behave like

$$\eta_n(q) \approx Ae^{-q^2/2}$$

in the limit that q^2 becomes very large.

The three paragraphs above motivate us to define a new function $v_n(q)$ through

$$\eta_n(q) = e^{-q^2/2}v_n(q). \quad (8.7)$$

(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 (8.4) becomes

$$\frac{d^2v_n(q)}{dq^2} - 2q\frac{dv_n(q)}{dq} + \left(\frac{2E_n}{\hbar\omega} - 1\right)v_n(q) = 0. \quad (8.8)$$

For brevity we introduce the shorthand notation

$$e_n = \frac{2E_n}{\hbar\omega} - 1. \quad (8.9)$$

3. *Find non-asymptotic behavior using the series method:* Okay, but how are we going to solve equation (8.8) for $v_n(q)$? Through the power series method!

Try a solution of the form

$$\begin{aligned} v(q) &= \sum_{k=0}^{\infty} a_k q^k \\ v'(q) &= \sum_{k=0}^{\infty} k a_k q^{k-1} & qv'(q) &= \sum_{k=0}^{\infty} k a_k q^k \\ v''(q) &= \sum_{k=0}^{\infty} k(k-1)a_k q^{k-2} & \llbracket \text{note that first two terms vanish ...} \rrbracket \\ &= \sum_{k=2}^{\infty} k(k-1)a_k q^{k-2} & \llbracket \text{change summation index to } k' = k - 2 \dots \rrbracket \\ &= \sum_{k'+2=2}^{\infty} (k'+2)(k'+1)a_{k'+2} q^{k'} & \llbracket \text{rename dummy index } k' \text{ to } k \dots \rrbracket \\ &= \sum_{k=0}^{\infty} (k+2)(k+1)a_{k+2} q^k \end{aligned}$$

Then equation (8.8) becomes

$$\sum_{k=0}^{\infty} [(k+2)(k+1)a_{k+2} - 2ka_k + e_n a_k] q^k = 0. \quad (8.10)$$

All of the terms in square brackets must vanish, whence the recursion relation

$$a_{k+2} = \frac{2k - e_n}{(k+2)(k+1)} a_k \quad k = 0, 1, 2, \dots \quad (8.11)$$

Like any second order ODE, equation (8.8) has two linearly independent solutions:

- An *even solution* of equation (8.8) comes by taking $a_0 = 1$, $a_1 = 0$. It is

$$v^{(e)}(q) = 1 - \frac{e_n}{2!} q^2 + \frac{(e_n - 4)e_n}{4!} q^4 - \frac{(e_n - 8)(e_n - 4)e_n}{6!} q^6 + \dots$$

- An *odd solution* of equation (8.8) comes by taking $a_0 = 0$, $a_1 = 1$. It is

$$v^{(o)}(q) = q - \frac{e_n - 2}{3!} q^3 + \frac{(e_n - 6)(e_n - 2)}{5!} q^5 - \frac{(e_n - 10)(e_n - 6)(e_n - 2)}{7!} q^7 + \dots$$

What is the asymptotic behavior of such solutions $v_n(q)$ as $q^2 \rightarrow \infty$? Well, the large q behavior will be dominated by the high-order terms of the series. Generally, as $k \rightarrow \infty$,

$$\frac{a_{k+2}}{a_k} = \frac{2k - e_n}{(k+2)(k+1)} \rightarrow \frac{2}{k}. \quad (8.12)$$

Compare this behavior to the expansion

$$e^{q^2} = b_0 + b_2 q^2 + b_4 q^4 + \dots \quad (8.13)$$

which has

$$\frac{b_{k+2}}{b_k} = \frac{1}{(k/2) + 1} \rightarrow \frac{2}{k}. \quad (8.14)$$

So whenever this happens,

$$v_n(q) \approx e^{q^2} \quad \text{and} \quad \eta_n(q) = e^{-q^2/2} v_n(q) \approx e^{q^2/2},$$

Thus giving us the very same unnormalizable behavior we've been trying so hard to avoid!

Is there no way to repair the situation?

4. *Invoke normalization to terminate the series as a polynomial:* The candidate wavefunction $\eta_n(q)$ is not normalizable when $a_{k+2}/a_k \rightarrow 2/k$ (see equation (8.12)). There is only one way to avoid this limit: when the series for $v_n(q)$ terminates as a polynomial. This termination occurs when, for some non-negative integer n , we have $2n = e_n$ whence (by recursion relation (8.11)), $a_k = 0$ for all $k > n$. Hence the only physical states correspond to energies with

$$2n = e_n = \frac{2E_n}{\hbar\omega} - 1.$$

Or, rephrasing,

Energy (eigen)states can exist only if they correspond to the energy (eigen)values

$$E_n = \hbar\omega(n + \frac{1}{2}) \quad n = 0, 1, 2, 3, \dots \quad (8.15)$$

What are the wavefunctions of the energy eigenstates?

For n even, $v_n^{(e)}(q)$ terminates and $v_n^{(o)}(q)$ doesn't.

For n odd, $v_n^{(o)}(q)$ terminates and $v_n^{(e)}(q)$ doesn't.

By tradition one defines the *Hermite*¹ polynomial of n th order $H_n(q)$:

$$n \text{ even: } H_n(q) = (-1)^{n/2} \frac{n!}{(n/2)!} v_n^{(e)}(q) \quad (8.16)$$

$$n \text{ odd: } H_n(q) = (-1)^{(n-1)/2} \frac{2n!}{((n-1)/2)!} v_n^{(o)}(q) \quad (8.17)$$

so that

$$\eta_n(x) = C_n e^{-q^2/2} H_n(q) \quad q = \sqrt{\frac{m\omega}{\hbar}} x \quad (8.18)$$

where C_n is a normalization factor.

8.3 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's dependent on a particular representation. Worst of all, it's hard to use. For example, suppose we wanted to find the expected value of the potential energy in the n -th energy eigenstate. We would find

$$\langle \hat{U} \rangle_n = \frac{m\omega^2}{2} \langle \eta_n | \hat{x}^2 | \eta_n \rangle = \frac{m\omega^2}{2} \int_{-\infty}^{+\infty} x^2 \eta_n^2(x) dx = \frac{m\omega^2}{2} \frac{\hbar}{m\omega} \frac{\int_{-\infty}^{+\infty} q^2 e^{-q^2} H_n^2(q) dq}{\int_{-\infty}^{+\infty} e^{-q^2} H_n^2(q) dq}.$$

Unless you happen to love integrating the Hermite polynomials, these last two integrals are intimidating.

I'll show you a method, invented by Dirac (or was it Schrödinger?), which avoids all these problems. On the other hand the method is hard to motivate: It clearly springs from the mind of genius.

Start with the Hamiltonian

$$\hat{H} = \frac{1}{2m} \hat{p}^2 + \frac{m\omega^2}{2} \hat{x}^2. \quad (8.19)$$

Since we're in a mathematical mode, it makes sense to define the dimensionless operators

$$\hat{X} = \sqrt{\frac{m\omega}{2\hbar}} \hat{x} \quad \text{and} \quad \hat{P} = \frac{1}{\sqrt{2m\hbar\omega}} \hat{p}, \quad (8.20)$$

¹Biographical information on Charles Hermite is given on page 38.

which satisfy

$$[\hat{X}, \hat{P}] = \sqrt{\frac{m\omega}{2\hbar}} \frac{1}{\sqrt{2m\hbar\omega}} [\hat{x}, \hat{p}] = \frac{i}{2} \hat{1}, \quad (8.21)$$

and write

$$\hat{H} = \hbar\omega(\hat{X}^2 + \hat{P}^2). \quad (8.22)$$

Now, one of the most fundamental tools of problem solving is to break something complex into its simpler pieces. (“All Gaul is divided into three parts.”) If we 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 concerning the expression

$$(\hat{X} - i\hat{P})(\hat{X} + i\hat{P}) = \hat{X}^2 + i\hat{X}\hat{P} - i\hat{P}\hat{X} + \hat{P}^2 = \hat{X}^2 + i[\hat{X}, \hat{P}] + \hat{P}^2 = \hat{X}^2 + \hat{P}^2 - \frac{1}{2}\hat{1}. \quad (8.23)$$

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

$$\hat{H} = \hbar\omega[(\hat{X} - i\hat{P})(\hat{X} + i\hat{P}) + \frac{1}{2}]. \quad (8.24)$$

From here, define

$$\hat{a} = \hat{X} + i\hat{P}. \quad (8.25)$$

The Hermitian adjoint of \hat{a} is

$$\hat{a}^\dagger = \hat{X} - i\hat{P}. \quad (8.26)$$

Note that the operators \hat{a} and \hat{a}^\dagger are not Hermitian. There is no observable corresponding to \hat{a} . The commutator is

$$[\hat{a}, \hat{a}^\dagger] = \hat{1}. \quad (8.27)$$

Exercise: Verify the above commutator.

And in terms of \hat{a} and \hat{a}^\dagger , the Hamiltonian is

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2}). \quad (8.28)$$

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 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

$$\hat{N}|n\rangle = n|n\rangle. \quad (8.29)$$

Because \hat{N} is Hermitian, its eigenvalues are real. Furthermore, they are positive because (where we define the vector $|\phi\rangle$ through $|\phi\rangle = \hat{a}|n\rangle$)

$$n = \langle n|\hat{N}|n\rangle = \langle n|\hat{a}^\dagger \hat{a}|n\rangle = \langle n|\hat{a}^\dagger|\phi\rangle = \langle \phi|\hat{a}|n\rangle^* = \langle \phi|\phi\rangle^* \geq 0. \quad (8.30)$$

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

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle \quad (8.31)$$

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 (8.30) that the eigenvalues were positive or zero. This seems present 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 any states of eigenvalue 2.3 to start off with.

However, if we start 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 (8.31), 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 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 $n+1$:

$$|\phi\rangle = C|n+1\rangle.$$

The operator \hat{a}^\dagger is a "raising operator"!

Exercise: Find the normalization constant C and conclude that

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{8.32}$$

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

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})$$

has eigenstates $|0\rangle, |1\rangle, |2\rangle, \dots$ with eigenvalues $\hbar\omega(\frac{1}{2}), \hbar\omega(1 + \frac{1}{2}), \hbar\omega(2 + \frac{1}{2}), \dots$. These eigenstates are related through

$$\begin{aligned}\hat{a}|n\rangle &= \sqrt{n}|n-1\rangle && \text{"raising operator"} \\ \hat{a}^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle && \text{"lowering operator"}\end{aligned}$$

The operators \hat{a} and \hat{a}^\dagger are collectively called "ladder operators" or "elevator operators".

8.4 Problems

8.1 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 exercise attempts to explain the non-zero ground state energy in seat-of-the-pants, semiclassical terms rather than in rigorous, formal, mathematical terms. It then goes on to use these ideas plus the uncertainty principle to guess at a value for the ground state energy. You may abhor such non-rigorous arguments, but you must be able to do them in order to make informed guesses about the behavior of systems that are too complicated to yield to rigorous mathematical methods.

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 kinetic energy (or, to be precise, a non-zero expectation value for kinetic energy). The 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 (expected value of the) 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 expected value of the momentum is zero.) Continuing in our non-rigorous 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 $E(d) = 0$!)

8.2 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 \hat{a}^\dagger .

8.3 Ladder operators for the simple harmonic oscillator

- Express \hat{x} and \hat{p} in terms of \hat{a} and \hat{a}^\dagger .
- Calculate the following simple harmonic oscillator matrix elements:

$$\begin{array}{ccc} \langle m|\hat{a}|n\rangle & \langle m|\hat{p}|n\rangle & \langle m|\hat{x}\hat{p}|n\rangle \\ \langle m|\hat{a}^\dagger|n\rangle & \langle m|\hat{x}^2|n\rangle & \langle m|\hat{p}\hat{x}|n\rangle \\ \langle m|\hat{x}|n\rangle & \langle m|\hat{p}^2|n\rangle & \langle m|\hat{H}|n\rangle \end{array}$$

- Show that the expectation value of the potential energy in a SHO energy eigenstate equals the expectation value of the kinetic energy in that state. (Recall that for a classical simple harmonic oscillator, the time averaged potential energy equals the time averaged kinetic energy.)
- Find Δx , Δp , and $\Delta x\Delta p$ for the energy eigenstate $|n\rangle$.

8.4 Simple harmonic oscillator states

Use scaled variables throughout this problem

- Concerning the ground energy state: What is $\eta_0(x)$ at $x = 0.5$? What is the probability density $\rho_0(x)$ there?
- Concerning the first excited energy state: What is $\eta_1(x)$ at $x = 0.5$? What is the probability density $\rho_1(x)$ there?
- Concerning the “50–50 combination” $\psi_A(x) = (\rho_0(x) + \rho_1(x))/\sqrt{2}$: What is $\psi_A(x)$ at $x = 0.5$? What is the probability density $\rho_A(x)$ there?
- Concerning another “50–50 combination” $\psi_B(x) = (\rho_0(x) - \rho_1(x))/\sqrt{2}$: What is $\psi_B(x)$ at $x = 0.5$? What is the probability density $\rho_B(x)$ there?
- 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?
- (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.)

Chapter 9

Qualitative Solution of Energy Eigenproblems

Chapter 10

Perturbation Theory

10.1 The \mathcal{O} notation

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

$$f(x) = \frac{e^{-x}}{1-x} \quad (10.1)$$

valid for small values of x , that is, for $x \ll 1$. You know that

$$e^{-x} = 1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3 + \dots \quad (10.2)$$

and that

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots, \quad (10.3)$$

so it seems that reasonable approximations are

$$e^{-x} \approx 1 - x \quad (10.4)$$

and

$$\frac{1}{1-x} \approx 1 + x, \quad (10.5)$$

whence

$$\frac{e^{-x}}{1-x} \approx (1-x)(1+x) = 1 - x^2. \quad (10.6)$$

Let’s try out this approximation at $x_0 = 0.01$. A calculator shows that

$$\frac{e^{-x_0}}{1-x_0} = 1.0000503\dots \quad (10.7)$$

while the value for the approximation is

$$1 - x_0^2 = 0.9999000. \quad (10.8)$$

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

$$e^{-x} = 1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3 + \dots \quad (10.9)$$

as

$$e^{-x} = 1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3), \quad (10.10)$$

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

$$\frac{1}{1-x} = 1 + x + x^2 + \mathcal{O}(x^3), \quad (10.11)$$

and find the product

$$f(x) = [1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)] \times [1 + x + x^2 + \mathcal{O}(x^3)] \quad (10.12)$$

$$= [1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)] \quad (10.13)$$

$$+ [1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)]x \quad (10.14)$$

$$+ [1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)]x^2 \quad (10.15)$$

$$+ [1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)]\mathcal{O}(x^3). \quad (10.16)$$

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

$$f(x) = [1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)] \quad (10.17)$$

$$+ [x - x^2 + \mathcal{O}(x^3)] \quad (10.18)$$

$$+ [x^2 + \mathcal{O}(x^3)] \quad (10.19)$$

$$+ \mathcal{O}(x^3) \quad (10.20)$$

$$= 1 + \frac{1}{2}x^2 + \mathcal{O}(x^3). \quad (10.21)$$

Thus we have the approximation

$$f(x) \approx 1 + \frac{1}{2}x^2. \quad (10.22)$$

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

$$1 + \frac{1}{2}x_0^2 = 1.0000500, \quad (10.23)$$

far superior to our old approximation.

What went wrong on our first try? The $-x^2$ in approximation (10.6) is the same as the $-x^2$ on line (10.18). However, lines (10.17) and (10.19) 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 \dots) 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

10.1 Tunneling for small times — \mathcal{O} notation version

Problem 4.2, 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

$$|\psi(\Delta t)\rangle = \left[\hat{1} - \frac{i}{\hbar} \hat{H} \Delta t - \frac{1}{2\hbar^2} \hat{H}^2 (\Delta t)^2 + \mathcal{O}(\Delta t^3) \right] |\psi(0)\rangle. \quad (10.24)$$

Find the representation of this equation in the $\{|1\rangle, |2\rangle\}$ basis.

- b. Conclude that for initial condition $|\psi(0)\rangle = |1\rangle$,

$$\begin{pmatrix} \psi_1(\Delta t) \\ \psi_2(\Delta t) \end{pmatrix} = \begin{pmatrix} 1 - (i/\hbar)E\Delta t - (1/2\hbar^2)(E^2 + A^2)(\Delta t)^2 + \mathcal{O}(\Delta t^3) \\ -(i/\hbar)Ae^{-i\phi}\Delta t - (1/\hbar^2)EAe^{-i\phi}(\Delta t)^2 + \mathcal{O}(\Delta t^3) \end{pmatrix}. \quad (10.25)$$

- 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 .

10.2 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

$$x^3 - 4.001x + 0.002 = 0. \quad (10.26)$$

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

$$x^3 - 4x = 0, \quad (10.27)$$

which *does* have straightforward exact solutions, namely

$$0, \pm 2. \quad (10.28)$$

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

$$x^3 - 4x + (-0.001x + 0.002) = 0. \quad (10.29)$$

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

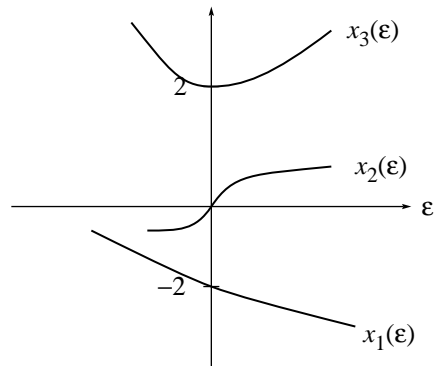
$$x^3 - 4x + T(-x + 2) = 0. \quad (10.30)$$

And now I'll generalize this problem by inserting a variable ϵ in front of the “small” part:

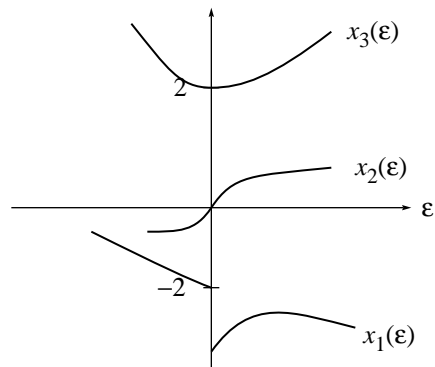
$$x^3 - 4x + \epsilon T(-x + 2) = 0. \quad (10.31)$$

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 $0 \leq \epsilon \leq 1$. 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:

$$x_1(\epsilon) = \sum_{i=1}^{\infty} a_i \epsilon^i \quad (10.32)$$

$$= -2 + a_1 \epsilon + a_2 \epsilon^2 + \mathcal{O}(\epsilon^3) \quad (10.33)$$

This function $x_1(\epsilon)$ has to satisfy

$$x_1^3(\epsilon) - (4 + \epsilon T)x_1(\epsilon) + 2\epsilon T = 0. \quad (10.34)$$

I can write the middle term above as an expansion in powers of ϵ using equation (10.33):

$$\begin{aligned} -4x_1(\epsilon) &= 8 - \epsilon(4a_1) - \epsilon^2(4a_2) + \mathcal{O}(\epsilon^3) \\ -\epsilon T x_1(\epsilon) &= + \epsilon(2T) - \epsilon^2(Ta_1) + \mathcal{O}(\epsilon^3) \\ -(4 + \epsilon T)x_1(\epsilon) &= 8 + \epsilon(-4a_1 + 2T) + \epsilon^2(-4a_2 - Ta_1) + \mathcal{O}(\epsilon^3) \end{aligned}$$

With just a bit more effort, I can work out the left-most term in equation (10.34) as an expansion:

$$\begin{aligned}x_1^2(\epsilon) &= 4 - \epsilon(4a_1) + \epsilon^2(-4a_2 + a_1^2) + \mathcal{O}(\epsilon^3) \\x_1^3(\epsilon) &= -8 - \epsilon(-12a_1) + \epsilon^2(12a_2 - 6a_1^2) + \mathcal{O}(\epsilon^3)\end{aligned}$$

So finally, I have worked out the expansion of every term in equation (10.34):

$$\begin{aligned}x_1^3(\epsilon) &= -8 - \epsilon(-12a_1) + \epsilon^2(12a_2 - 6a_1^2) + \mathcal{O}(\epsilon^3) \\-(4 + \epsilon T)x_1(\epsilon) &= 8 + \epsilon(-4a_1 + 2T) + \epsilon^2(-4a_2 - Ta_1) + \mathcal{O}(\epsilon^3) \\2\epsilon T &= + \epsilon(2T)\end{aligned}$$

Summing the three equations above must, according to equation (10.34), produce zero:

$$\begin{aligned}0 &= (-8 + 8) + \epsilon(12a_1 - 4a_1 + 4T) + \epsilon^2(12a_2 - 6a_1^2 - 4a_2 - Ta_1) + \mathcal{O}(\epsilon^3) \\0 &= (-8 + 8) + \epsilon(8a_1 + 4T) + \epsilon^2(8a_2 - 6a_1^2 - Ta_1) + \mathcal{O}(\epsilon^3)\end{aligned}$$

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

$$(8a_1 + 4T) = 0 \quad \text{whence} \quad a_1 = -\frac{1}{2}T.$$

And the term quadratic in ϵ must vanish, so

$$(8a_2 - 6a_1^2 - Ta_1) = 0 \quad \text{whence} \quad a_2 = \frac{3}{4}a_1^2 + \frac{1}{8}Ta_1 = \frac{1}{8}T^2.$$

The expansion for $x_1(\epsilon)$ is thus

$$x_1(\epsilon) = -2 - \frac{1}{2}T\epsilon + \frac{1}{8}T^2\epsilon^2 + \mathcal{O}(\epsilon^3)$$

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.

10.3 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

$$\hat{H}^{(0)}|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle \tag{10.35}$$

is known and where \hat{H}' is “small” compared with $\hat{H}^{(0)}$, we set

$$\hat{H}(\epsilon) = \hat{H}^{(0)} + \epsilon \hat{H}' \tag{10.36}$$

and then find $|n(\epsilon)\rangle$ and $E_n(\epsilon)$ such that

$$\hat{H}(\epsilon)|n(\epsilon)\rangle = E_n(\epsilon)|n(\epsilon)\rangle \tag{10.37}$$

and

$$\langle n(\epsilon)|n(\epsilon)\rangle = 1. \tag{10.38}$$

Intermediate goal

Find $|\bar{n}(\epsilon)\rangle$ and $E_n(\epsilon)$ such that

$$\hat{H}(\epsilon)|\bar{n}(\epsilon)\rangle = E_n(\epsilon)|\bar{n}(\epsilon)\rangle \tag{10.39}$$

and

$$\langle n^{(0)}|\bar{n}(\epsilon)\rangle = 1. \tag{10.40}$$

Then our final goal will be

$$|n(\epsilon)\rangle = \frac{|\bar{n}(\epsilon)\rangle}{\langle \bar{n}(\epsilon)|\bar{n}(\epsilon)\rangle^{1/2}}. \tag{10.41}$$

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:

$$|\bar{n}(\epsilon)\rangle = |n^{(0)}\rangle + \epsilon|\bar{n}^{(1)}\rangle + \epsilon^2|\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3) \tag{10.42}$$

$$E_n(\epsilon) = E_n^{(0)} + \epsilon E_n^{(1)} + \epsilon^2 E_n^{(2)} + \mathcal{O}(\epsilon^3) \tag{10.43}$$

[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 $\langle n^{(0)}|\bar{n}(\epsilon)\rangle = 1$ gives rise to interesting and useful consequences. First, take the inner product of $|n^{(0)}\rangle$ with equation (10.42)

$$\begin{aligned} \langle n^{(0)}|\bar{n}(\epsilon)\rangle &= \langle n^{(0)}|n^{(0)}\rangle + \epsilon\langle n^{(0)}|\bar{n}^{(1)}\rangle + \epsilon^2\langle n^{(0)}|\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3) \\ 1 &= 1 + \epsilon\langle n^{(0)}|\bar{n}^{(1)}\rangle + \epsilon^2\langle n^{(0)}|\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3) \end{aligned}$$

Because this relationship holds for *all* values of ϵ , the coefficient of each ϵ^m must vanish:

$$\langle n^{(0)}|\bar{n}^{(m)}\rangle = 0 \quad m = 1, 2, 3, \dots \tag{10.44}$$

Whence

$$\begin{aligned}
\langle \bar{n}(\epsilon) | \bar{n}(\epsilon) \rangle &= \left[\langle n^{(0)} | + \epsilon \langle \bar{n}^{(1)} | + \epsilon^2 \langle \bar{n}^{(2)} | + \mathcal{O}(\epsilon^3) \right] \left[|n^{(0)}\rangle + \epsilon |\bar{n}^{(1)}\rangle + \epsilon^2 |\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3) \right] \\
&= \langle n^{(0)} | n^{(0)} \rangle + \epsilon \left[\langle \bar{n}^{(1)} | n^{(0)} \rangle + \langle n^{(0)} | \bar{n}^{(1)} \rangle \right] + \epsilon^2 \left[\langle \bar{n}^{(2)} | n^{(0)} \rangle + \langle \bar{n}^{(1)} | \bar{n}^{(1)} \rangle + \langle n^{(0)} | \bar{n}^{(2)} \rangle \right] + \mathcal{O}(\epsilon^3) \\
&= 1 + \epsilon \left[0 + 0 \right] + \epsilon^2 \left[0 + \langle \bar{n}^{(1)} | \bar{n}^{(1)} \rangle + 0 \right] + \mathcal{O}(\epsilon^3) \\
&= 1 + \epsilon^2 \langle \bar{n}^{(1)} | \bar{n}^{(1)} \rangle + \mathcal{O}(\epsilon^3).
\end{aligned} \tag{10.45}$$

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 (10.42) and (10.43) into

$$\hat{H}(\epsilon) |n(\epsilon)\rangle = E_n(\epsilon) |n(\epsilon)\rangle \tag{10.46}$$

to find

$$\begin{aligned}
&\left[\hat{H}^{(0)} + \epsilon \hat{H}' \right] \left[|n^{(0)}\rangle + \epsilon |\bar{n}^{(1)}\rangle + \epsilon^2 |\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3) \right] \\
&= \left[E_n^{(0)} + \epsilon E_n^{(1)} + \epsilon^2 E_n^{(2)} + \mathcal{O}(\epsilon^3) \right] \left[|n^{(0)}\rangle + \epsilon |\bar{n}^{(1)}\rangle + \epsilon^2 |\bar{n}^{(2)}\rangle + \mathcal{O}(\epsilon^3) \right]
\end{aligned} \tag{10.47}$$

Separating out powers of ϵ gives

$$\hat{H}^{(0)} |n^{(0)}\rangle = E_n^{(0)} |n^{(0)}\rangle \tag{10.48}$$

$$\hat{H}^{(0)} |\bar{n}^{(1)}\rangle + \hat{H}' |n^{(0)}\rangle = E_n^{(1)} |n^{(0)}\rangle + E_n^{(0)} |\bar{n}^{(1)}\rangle \tag{10.49}$$

$$\hat{H}^{(0)} |\bar{n}^{(2)}\rangle + \hat{H}' |\bar{n}^{(1)}\rangle = E_n^{(2)} |n^{(0)}\rangle + E_n^{(1)} |\bar{n}^{(1)}\rangle + E_n^{(0)} |\bar{n}^{(2)}\rangle \tag{10.50}$$

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 (10.49) we will find $E_n^{(1)}$ and $|\bar{n}^{(1)}\rangle$. Knowing these, from equation (10.50) 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 (10.49) by $\langle n^{(0)} |$ to find

$$\begin{aligned}
\langle n^{(0)} | \hat{H}^{(0)} | \bar{n}^{(1)} \rangle + \langle n^{(0)} | \hat{H}' | n^{(0)} \rangle &= E_n^{(1)} \langle n^{(0)} | n^{(0)} \rangle + E_n^{(0)} \langle n^{(0)} | \bar{n}^{(1)} \rangle \\
E_n^{(0)} \langle n^{(0)} | \bar{n}^{(1)} \rangle + \langle n^{(0)} | \hat{H}' | n^{(0)} \rangle &= E_n^{(1)} + E_n^{(0)} \langle n^{(0)} | \bar{n}^{(1)} \rangle
\end{aligned} \tag{10.51}$$

Or,

$$E_n^{(1)} = \langle n^{(0)} | \hat{H}' | n^{(0)} \rangle. \tag{10.52}$$

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 (10.49) by $\langle m^{(0)}|$ ($m \neq n$) to find

$$\begin{aligned} \langle m^{(0)}|\hat{H}^{(0)}|\bar{n}^{(1)}\rangle + \langle m^{(0)}|\hat{H}'|n^{(0)}\rangle &= E_n^{(1)}\langle m^{(0)}|n^{(0)}\rangle + E_n^{(0)}\langle m^{(0)}|\bar{n}^{(1)}\rangle \\ E_m^{(0)}\langle m^{(0)}|\bar{n}^{(1)}\rangle + \langle m^{(0)}|\hat{H}'|n^{(0)}\rangle &= 0 + E_n^{(0)}\langle m^{(0)}|\bar{n}^{(1)}\rangle \\ \langle m^{(0)}|\hat{H}'|n^{(0)}\rangle &= (E_n^{(0)} - E_m^{(0)})\langle m^{(0)}|\bar{n}^{(1)}\rangle \end{aligned} \quad (10.53)$$

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

$$\langle m^{(0)}|\bar{n}^{(1)}\rangle = \frac{\langle m^{(0)}|\hat{H}'|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} \quad (m \neq n) \quad (10.54)$$

But we already know, from equation (10.44), that

$$\langle n^{(0)}|\bar{n}^{(1)}\rangle = 0. \quad (10.55)$$

So now all the projections $\langle m^{(0)}|\bar{n}^{(1)}\rangle$ are known, and therefore the vector is known:

$$|\bar{n}^{(1)}\rangle = \sum_m |m^{(0)}\rangle \langle m^{(0)}|\bar{n}^{(1)}\rangle \quad (10.56)$$

In conclusion — if $|n^{(0)}\rangle$ is non-degenerate

$$|\bar{n}^{(1)}\rangle = \sum_{m \neq n} |m^{(0)}\rangle \frac{\langle m^{(0)}|\hat{H}'|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}. \quad (10.57)$$

10.4 Perturbation theory for the energy eigenproblem: Summary of results

Given: Solution for the $\hat{H}^{(0)}$ eigenproblem:

$$\hat{H}^{(0)}|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle \quad \langle n^{(0)}|n^{(0)}\rangle = 1. \quad (10.58)$$

Find: Solution for the $\hat{H}^{(0)} + \epsilon\hat{H}'$ eigenproblem:

$$(\hat{H}^{(0)} + \epsilon\hat{H}')|n(\epsilon)\rangle = E_n(\epsilon)|n(\epsilon)\rangle \quad \langle n(\epsilon)|n(\epsilon)\rangle = 1. \quad (10.59)$$

Define the “matrix elements”

$$\langle n^{(0)}|\hat{H}'|m^{(0)}\rangle = H'_{nm}. \quad (10.60)$$

The solutions are (provided $|n^{(0)}\rangle$ is not degenerate):

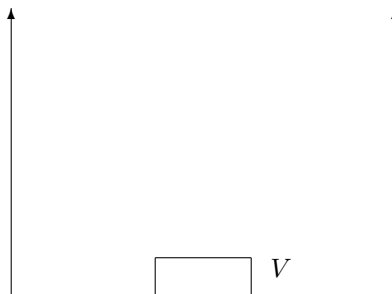
$$E_n(\epsilon) = E_n^{(0)} + \epsilon H'_{nn} + \epsilon^2 \sum_{m \neq n} \frac{H'_{nm} H'_{mn}}{E_n^{(0)} - E_m^{(0)}} + \mathcal{O}(\epsilon^3) \quad (10.61)$$

$$\begin{aligned} |n(\epsilon)\rangle &= |n^{(0)}\rangle \\ &+ \epsilon \sum_{m \neq n} |m^{(0)}\rangle \frac{H'_{mn}}{E_n^{(0)} - E_m^{(0)}} \\ &+ \epsilon^2 \left[\sum_{m \neq n} \sum_{\ell \neq n} |m^{(0)}\rangle \frac{H'_{m\ell} H'_{\ell n}}{(E_n^{(0)} - E_m^{(0)})(E_n^{(0)} - E_\ell^{(0)})} \right. \\ &\quad \left. - \sum_{m \neq n} |m^{(0)}\rangle \frac{H'_{nn} H'_{mn}}{(E_n^{(0)} - E_m^{(0)})^2} - |n^{(0)}\rangle \frac{1}{2} \sum_{m \neq n} \frac{H'_{nm} H'_{mn}}{(E_n^{(0)} - E_m^{(0)})^2} \right] \\ &+ \mathcal{O}(\epsilon^3) \end{aligned} \quad (10.62)$$

10.5 Problems

10.2 Square well with a bump

An infinite square well of width L (problem 7.1) 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.



10.3 Anharmonic oscillator

- a. Show that for the simple harmonic oscillator,

$$\langle m|\hat{x}^3|n\rangle = \sqrt{\left(\frac{\hbar}{2m\omega}\right)^3} \left[\sqrt{n(n-1)(n-2)} \delta_{m,n-3} + 3\sqrt{n^3} \delta_{m,n-1} \right. \\ \left. + 3\sqrt{(n+1)^3} \delta_{m,n+1} + \sqrt{(n+1)(n+2)(n+3)} \delta_{m,n+3} \right].$$

- b. Recall that the simple harmonic oscillator is always an approximation. The real problem always has a potential $V(x) = \frac{1}{2}kx^2 + bx^3 + cx^4 + \dots$. 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

$$-\frac{b^2}{\hbar\omega} \left(\frac{\hbar}{2m\omega}\right)^3 (30n^2 + 30n + 11). \quad (10.63)$$

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 .

10.4 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

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2} \hat{x}^2 - \frac{1}{8c^2 m^3} \hat{p}^4. \quad (10.64)$$

- b. Show that

$$\langle m|\hat{p}^4|0\rangle = \left(\frac{m\hbar\omega}{2}\right)^2 (3\delta_{m,0} - 6\sqrt{2}\delta_{m,2} + 2\sqrt{6}\delta_{m,4}). \quad (10.65)$$

- 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 $|0\rangle$.

10.5 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

$$a_0 I + a_1 \sigma_1 + a_3 \sigma_3 \quad (10.66)$$

where a_0 , a_1 , and a_3 are real numbers and the σ ’s are Pauli matrices. (See problem 48.)

- 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 $a_3 \ll a_0 \approx a_1$.

10.6 Degenerate perturbation theory in a two-state system

Consider a two state system with a Hamiltonian represented in some basis by

$$a_0 I + a_1 \sigma_1 + a_3 \sigma_3. \quad (10.67)$$

We shall call the basis for this representation the “initial basis”. This exercise shows how to use perturbation theory to solve (approximately) the energy eigenproblem in the case $a_0 \gg a_1 \approx a_3$.

$$\hat{H}^{(0)} = \begin{pmatrix} a_0 & 0 \\ 0 & a_0 \end{pmatrix} \quad \hat{H}' = \begin{pmatrix} a_3 & a_1 \\ a_1 & -a_3 \end{pmatrix} \quad (10.68)$$

In this case the unperturbed Hamiltonian is degenerate. The initial basis

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \quad (10.69)$$

is a perfectly acceptable energy eigenbasis (both states have energy a_0), but the basis

$$\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}, \quad (10.70)$$

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 (10.53), namely

$$(E_n^{(0)} - E_m^{(0)}) \langle m^{(0)} | \hat{n}^{(1)} \rangle = \langle m^{(0)} | \hat{H}' | n^{(0)} \rangle \quad \text{whenever } m \neq n. \quad (10.71)$$

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

$$\langle m^{(0)} | \hat{H}' | n^{(0)} \rangle = 0 \quad \text{whenever } m \neq n \text{ and } E_n^{(0)} = E_m^{(0)}. \quad (10.72)$$

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

$$\frac{1}{\sqrt{2}\sqrt{a_1^2 + a_3^2} - a_3\sqrt{a_1^2 + a_3^2}} \begin{pmatrix} +a_1 \\ -a_3 + \sqrt{a_1^2 + a_3^2} \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}, \quad (10.73)$$

$$\frac{1}{\sqrt{2}\sqrt{a_1^2 + a_3^2} + a_3\sqrt{a_1^2 + a_3^2}} \begin{pmatrix} -a_1 \\ +a_3 + \sqrt{a_1^2 + a_3^2} \end{pmatrix} = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}, \quad (10.74)$$

where

$$\tan \theta = \frac{a_1}{a_3 + \sqrt{a_1^2 + a_3^2}}. \quad (10.75)$$

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 11

Quantum Mechanics in Two and Three Dimensions

11.1 More degrees of freedom

Let's think of the process of adding degrees of freedom.

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 ∞ elements. (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 elements are orthonormal,

$$\langle x|x'\rangle = \delta(x - x'), \quad (11.1)$$

and complete

$$\hat{1} = \int_{-\infty}^{+\infty} dx |x\rangle\langle x|. \quad (11.2)$$

[[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

$$|\psi\rangle \doteq \langle x|\psi\rangle = \psi(x). \quad (11.3)$$

5. When the position is measured, the probability of finding the particle at a position within dx about x_0 is

$$|\psi(x_0)|^2 dx. \quad (11.4)$$

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 $\infty \times 2$ elements. (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 $|5, +\rangle$ represents a particle at position 5 with spin +. The state

$$\frac{1}{\sqrt{2}}[|5, +\rangle - |7, -\rangle]$$

represents a particle with amplitude $1/\sqrt{2}$ to be at position 5 with spin + and amplitude $-1/\sqrt{2}$ to be at position 7 with spin -, but with no amplitude to be at position 5 with spin -, and no amplitude to be at position 6 with any spin.]

3. These basis elements are orthonormal,

$$\begin{aligned} \langle x, +|x', +\rangle &= \delta(x - x') \\ \langle x, +|x', -\rangle &= 0 \\ \langle x, i|x', j\rangle &= \delta(x - x')\delta_{i,j} \end{aligned} \quad (11.5)$$

and complete

$$\begin{aligned} \hat{1} &= \int_{-\infty}^{+\infty} dx |x, +\rangle\langle x, +| + \int_{-\infty}^{+\infty} dx |x, -\rangle\langle x, -| \\ \hat{1} &= \sum_{i=+,-} \int_{-\infty}^{+\infty} dx |x, i\rangle\langle x, i| \end{aligned} \quad (11.6)$$

4. The state $|\psi\rangle$ is represented (in this basis) by the numbers

$$\begin{pmatrix} \langle x, +|\psi\rangle \\ \langle x, -|\psi\rangle \end{pmatrix} = \begin{pmatrix} \psi_+(x) \\ \psi_-(x) \end{pmatrix}. \quad (11.7)$$

5. When both the spin projection and the position are measured, the probability of finding the particle with spin up and at a position within dx about x_0 is

$$|\psi_+(x_0)|^2 dx. \quad (11.8)$$

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.

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 $\{|\mathbf{x}\rangle\}$ — has ∞^3 elements. (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 elements are orthonormal,

$$\langle x, y, z|x', y', z'\rangle = \delta(x - x')\delta(y - y')\delta(z - z'), \quad (11.9)$$

which is also written as

$$\langle \mathbf{x}|\mathbf{x}'\rangle = \delta(\mathbf{x} - \mathbf{x}'). \quad (11.10)$$

In addition, the basis elements are complete

$$\hat{1} = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz |x, y, z\rangle\langle x, y, z|, \quad (11.11)$$

which is also written as

$$\hat{1} = \int_{-\infty}^{+\infty} d^3x |\mathbf{x}\rangle\langle \mathbf{x}|. \quad (11.12)$$

4. The state $|\psi\rangle$ is represented (in the position basis) by the numbers $\langle \mathbf{x}|\psi\rangle = \psi(\mathbf{x})$ (a complex-valued function of a vector argument).
5. When the position is measured, the probability of finding the particle at a position within d^3x about \mathbf{x}_0 is

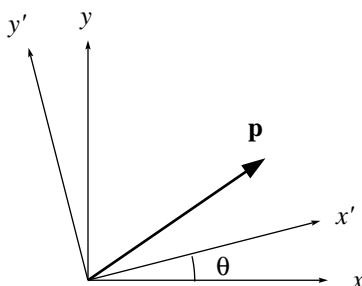
$$|\psi(\mathbf{x}_0)|^2 d^3x. \quad (11.13)$$

11.2 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 \mathbf{p} . If the coordinate axes are x and y , the components of the vector \mathbf{p} are p_x and p_y . But if the coordinate axes are x' and y' , then the components of the vector \mathbf{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

$$\begin{aligned} p_{x'} &= p_x \cos \theta + p_y \sin \theta \\ p_{y'} &= -p_x \sin \theta + p_y \cos \theta \end{aligned} \quad (11.14)$$

(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:

$$\begin{aligned} \hat{p}_{x'} &= \hat{p}_x \cos \theta + \hat{p}_y \sin \theta \\ \hat{p}_{y'} &= -\hat{p}_x \sin \theta + \hat{p}_y \cos \theta \end{aligned} \quad (11.15)$$

(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

$$\hat{p}^2 = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2. \quad (11.16)$$

11.3 Multiple particles

In section 11.1 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 14 — so in this section we'll consider non-identical particles.)

Consider a system of two spinless particles (call them red and green) moving 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 elements. (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 elements are orthonormal,

$$\langle x_R, x_G | x'_R, x'_G \rangle = \delta(x_R - x'_R) \delta(x_G - x'_G). \quad (11.17)$$

In addition, the basis elements are complete

$$\hat{1} = \int_{-\infty}^{+\infty} dx_R \int_{-\infty}^{+\infty} dx_G |x_R, x_G\rangle \langle x_R, x_G|. \quad (11.18)$$

4. The state $|\psi\rangle$ is represented (in the position basis) by the numbers $\langle x_R, x_G | \psi \rangle = \psi(x_R, x_G)$ (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

$$|\psi(x_A, x_B)|^2 dx_A dx_B. \quad (11.19)$$

11.4 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 continuum systems in three dimensions. It's time to return to the level of phenomena and ask how the phenomena of quantum mechanics generalize to continuum 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 12

Angular Momentum

12.1 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

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z, \quad \text{and cyclic permutations} \quad (12.1)$$

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° , will have exactly the same eigenvalues, and eigenvectors with the same structure.

Note that the 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, or 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

$$\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \quad (12.2)$$

and note that

$$[\hat{J}^2, \hat{J}_i] = 0 \quad \text{for } i = x, y, z. \quad (12.3)$$

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 $|\beta, \mu\rangle$, which satisfy

$$\hat{J}^2|\beta, \mu\rangle = \hbar^2\beta|\beta, \mu\rangle \quad (12.4)$$

$$\hat{J}_z|\beta, \mu\rangle = \hbar\mu|\beta, \mu\rangle \quad (12.5)$$

We define the values β and μ in this way so that β and μ will be dimensionless. (If this is not obvious to you, show it now. Also, if it's not obvious that the equations (12.3) follow from the equations (12.1), you should show that, too. What is the commutator $[\hat{J}^2, \hat{J}_{28^\circ}]$?)

Start off by noting that

$$(\hat{J}_x^2 + \hat{J}_y^2)|\beta, \mu\rangle = (\hat{J}^2 - \hat{J}_z^2)|\beta, \mu\rangle = \hbar^2(\beta - \mu^2)|\beta, \mu\rangle. \quad (12.6)$$

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 to inspire the definitions

$$\hat{J}_- = \hat{J}_x - i\hat{J}_y \quad \text{and} \quad \hat{J}_+ = \hat{J}_x + i\hat{J}_y \quad (12.7)$$

so that

$$\hat{J}_-\hat{J}_+ = \hat{J}_x^2 + \hat{J}_y^2 + i(\hat{J}_x\hat{J}_y - \hat{J}_y\hat{J}_x) = \hat{J}_x^2 + \hat{J}_y^2 + i[\hat{J}_x, \hat{J}_y] = \hat{J}_x^2 + \hat{J}_y^2 - \hbar\hat{J}_z. \quad (12.8)$$

This tells us that

$$\hat{J}_-\hat{J}_+|\beta, \mu\rangle = (\hbar^2\beta - \hbar^2\mu^2 - \hbar^2\mu)|\beta, \mu\rangle = \hbar^2(\beta - \mu(\mu + 1))|\beta, \mu\rangle. \quad (12.9)$$

We have immediately that

$$\langle\beta, \mu|\hat{J}_-\hat{J}_+|\beta, \mu\rangle = \hbar^2(\beta - \mu(\mu + 1)). \quad (12.10)$$

But if we define

$$|\phi\rangle = \hat{J}_+|\beta, \mu\rangle \quad \text{then} \quad \langle\phi| = \langle\beta, \mu|\hat{J}_-$$

then equation (12.10) is just the expression for $\langle\phi|\phi\rangle$, and we know that for any vector $\langle\phi|\phi\rangle \geq 0$. Thus

$$\beta \geq \mu(\mu + 1). \quad (12.11)$$

With these preliminaries out of the way, we investigate the operator \hat{J}_+ . First, its commutation relations:

$$[\hat{J}^2, \hat{J}_+] = 0, \quad (12.12)$$

$$[\hat{J}_z, \hat{J}_+] = [\hat{J}_z, \hat{J}_x] + i[\hat{J}_z, \hat{J}_y] = (i\hbar\hat{J}_y) + i(-i\hbar\hat{J}_x) = \hbar\hat{J}_+. \quad (12.13)$$

Then, use the commutation relations to find the effect of \hat{J}_+ on $|\beta, \mu\rangle$. If we again define $|\phi\rangle = \hat{J}_+|\beta, \mu\rangle$, then

$$\hat{J}^2|\phi\rangle = \hat{J}^2\hat{J}_+|\beta, \mu\rangle = \hat{J}_+\hat{J}^2|\beta, \mu\rangle = \hbar^2\beta\hat{J}_+|\beta, \mu\rangle = \hbar^2\beta|\phi\rangle, \quad (12.14)$$

$$\hat{J}_z|\phi\rangle = \hat{J}_z\hat{J}_+|\beta, \mu\rangle = (\hat{J}_+\hat{J}_z + \hbar\hat{J}_+)|\beta, \mu\rangle = \hbar\mu\hat{J}_+|\beta, \mu\rangle + \hbar\hat{J}_+|\beta, \mu\rangle = \hbar(\mu + 1)|\phi\rangle. \quad (12.15)$$

That is, the vector $|\phi\rangle$ is an eigenvector of \hat{J}^2 with eigenvalue β and an eigenvector of \hat{J}_z with eigenvalue $\mu + 1$. In other words,

$$\hat{J}_+|\beta, \mu\rangle = C|\beta, \mu + 1\rangle \quad (12.16)$$

where C is a normalization factor to be determined.

To find C , we contrast

$$\langle\phi|\phi\rangle = |C|^2\langle\beta, \mu|\beta, \mu\rangle = |C|^2 \quad (12.17)$$

with the result of equation (12.10), namely

$$\langle\phi|\phi\rangle = \langle\beta, \mu|\hat{J}_-\hat{J}_+|\beta, \mu\rangle = \hbar^2(\beta - \mu(\mu + 1)). \quad (12.18)$$

From this we may select $C = \hbar\sqrt{\beta - \mu(\mu + 1)}$ so that

$$\hat{J}_+|\beta, \mu\rangle = \hbar\sqrt{\beta - \mu(\mu + 1)}|\beta, \mu + 1\rangle. \quad (12.19)$$

In short, the operator \hat{J}_+ applied to $|\beta, \mu\rangle$ acts as a *raising operator*: it doesn't change the value of β , but it increases the value of μ by 1.

Parallel reasoning applied to \hat{J}_- shows that

$$\hat{J}_-|\beta, \mu\rangle = \hbar\sqrt{\beta - \mu(\mu - 1)}|\beta, \mu - 1\rangle. \quad (12.20)$$

In short, the operator \hat{J}_- applied to $|\beta, \mu\rangle$ acts as a *lowering operator*: it doesn't change the value of β , but it decreases the value of μ by 1.

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 (12.11),

$$\beta \geq \mu(\mu + 1), \quad (12.21)$$

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 $|\beta, \mu_{\max}\rangle$ results not in a vector proportional to $|\beta, \mu_{\max} + 1\rangle$, but results instead in 0. It is clear from equation (12.19) that this value of μ satisfies

$$\beta - \mu_{\max}(\mu_{\max} + 1) = 0. \quad (12.22)$$

And it's equally clear from equation (12.20) that there is a minimum value μ_{\min} satisfying

$$\beta - \mu_{\min}(\mu_{\min} - 1) = 0. \quad (12.23)$$

Solving these two equations simultaneously, we find that

$$\mu_{\max} = -\mu_{\min} \quad \text{with} \quad \mu_{\max} \geq 0 \quad (12.24)$$

and that

$$\beta = \mu_{\max}(\mu_{\max} + 1). \quad (12.25)$$

But there's more. Because we raise or lower μ by 1 with each application of \hat{J}_+ or \hat{J}_- , the value of μ_{\max} must be an integer above μ_{\min} :

$$\begin{aligned}\mu_{\max} &= \mu_{\min} + (\text{an integer}) \\ 2\mu_{\max} &= (\text{an integer}) \\ \mu_{\max} &= \frac{\text{an integer}}{2} \geq 0\end{aligned}\tag{12.26}$$

Common practice is to call the half-integer μ_{\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 $|\beta, \mu\rangle$ 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

$$\hbar^2 j(j+1) \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots\tag{12.27}$$

For a given j , the eigenvalues of \hat{J}_z are

$$\hbar m \quad m = -j, -j+1, \dots, j-1, j.\tag{12.28}$$

The eigenstates $|j, m\rangle$ are related through the operators

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \quad \hat{J}_- = \hat{J}_x - i\hat{J}_y\tag{12.29}$$

by

$$\hat{J}_+ |j, m\rangle = \hbar \sqrt{j(j+1) - m(m+1)} |j, m+1\rangle\tag{12.30}$$

$$\hat{J}_- |j, m\rangle = \hbar \sqrt{j(j+1) - m(m-1)} |j, m-1\rangle.\tag{12.31}$$

Problems:

1. Write out the “parallel reasoning” that results in equation (12.20).
2. The simultaneous solution of equations (12.22) and (12.23) results in two possible solutions, namely (12.24) and $\mu_{\min} = \mu_{\max} + 1$. Why do we reject this second solution? Why do we, in equation (12.24), insert the proviso $\mu_{\max} \geq 0$?

12.2 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

$$\hbar^2 j(j+1) \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$$

For a given j , the eigenvalues of \hat{J}_z are

$$\hbar m \quad m = -j, -j+1, \dots, j-1, j.$$

The eigenstates $|j, m\rangle$ are related through the operators

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \quad \hat{J}_- = \hat{J}_x - i\hat{J}_y$$

by

$$\begin{aligned} \hat{J}_+ |j, m\rangle &= \hbar \sqrt{j(j+1) - m(m+1)} |j, m+1\rangle \\ \hat{J}_- |j, m\rangle &= \hbar \sqrt{j(j+1) - m(m-1)} |j, m-1\rangle. \end{aligned}$$

12.3 Ordinary differential equations for the $d_{m,m'}^{(j)}(\theta)$

$$\begin{aligned} d_{m,m'}^{(j)}(\theta) &= \langle j, m | e^{-i(\hat{J}_y/\hbar)\theta} |j, m'\rangle \\ \frac{d}{d\theta} [d_{m,m'}^{(j)}(\theta)] &= \langle j, m | e^{-i(\hat{J}_y/\hbar)\theta} (-i\hat{J}_y/\hbar) |j, m'\rangle \\ &= -\frac{1}{2\hbar} \langle j, m | e^{-i(\hat{J}_y/\hbar)\theta} (\hat{J}_+ - \hat{J}_-) |j, m'\rangle \\ &= -\frac{1}{2} \sqrt{j(j+1) - m'(m'+1)} \langle j, m | e^{-i(\hat{J}_y/\hbar)\theta} |j, m'+1\rangle \\ &\quad + \frac{1}{2} \sqrt{j(j+1) - m'(m'-1)} \langle j, m | e^{-i(\hat{J}_y/\hbar)\theta} |j, m'-1\rangle \end{aligned}$$

$$\frac{d}{d\theta} [d_{m,m'}^{(j)}(\theta)] = -\frac{1}{2} \sqrt{j(j+1) - m'(m'+1)} d_{m,m'+1}^{(j)}(\theta) + \frac{1}{2} \sqrt{j(j+1) - m'(m'-1)} d_{m,m'-1}^{(j)}(\theta)$$

For a given j and m , these are $2j+1$ coupled first-order ODEs. In matrix form they are

$$\frac{d}{d\theta} \begin{pmatrix} d_{m,j}^{(j)}(\theta) \\ d_{m,j-1}^{(j)}(\theta) \\ d_{m,j-2}^{(j)}(\theta) \\ d_{m,j-3}^{(j)}(\theta) \\ \vdots \\ d_{m,-j+1}^{(j)}(\theta) \\ d_{m,-j}^{(j)}(\theta) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & \sqrt{j} & 0 & 0 & \dots & 0 & 0 \\ -\sqrt{j} & 0 & \sqrt{2j-1} & 0 & \dots & 0 & 0 \\ 0 & -\sqrt{2j-1} & 0 & \sqrt{3j-3} & \dots & 0 & 0 \\ 0 & 0 & -\sqrt{3j-3} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & \sqrt{j} \\ 0 & 0 & 0 & 0 & \dots & -\sqrt{j} & 0 \end{pmatrix} \begin{pmatrix} d_{m,j}^{(j)}(\theta) \\ d_{m,j-1}^{(j)}(\theta) \\ d_{m,j-2}^{(j)}(\theta) \\ d_{m,j-3}^{(j)}(\theta) \\ \vdots \\ d_{m,-j+1}^{(j)}(\theta) \\ d_{m,-j}^{(j)}(\theta) \end{pmatrix}.$$

12.4 Problems

12.1 Trivial pursuit

- Show that if an operator commutes with two components of an angular momentum vector, it commutes with the third as well.
- If \hat{J}_x and \hat{J}_z are represented by matrices with pure real entries (as is conventionally the case, see problem 12.2), show that \hat{J}_y is represented by a matrix with pure imaginary entries.

12.2 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

$$\begin{pmatrix} \langle\uparrow|\hat{A}|\uparrow\rangle & \langle\uparrow|\hat{A}|\downarrow\rangle \\ \langle\downarrow|\hat{A}|\uparrow\rangle & \langle\downarrow|\hat{A}|\downarrow\rangle \end{pmatrix}, \quad (12.32)$$

and recall also that this isn't always the easiest way to find a matrix representation.

- 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!
- Find normalized column matrix representations for the eigenstates of \hat{S}_x :

$$\hat{S}_x|\rightarrow\rangle = +\frac{\hbar}{2}|\rightarrow\rangle \quad (12.33)$$

$$\hat{S}_x|\leftarrow\rangle = -\frac{\hbar}{2}|\leftarrow\rangle. \quad (12.34)$$

12.3 Rotations and spin- $\frac{1}{2}$

Verify explicitly that

$$|\rightarrow\rangle = e^{-i(\hat{S}_y/\hbar)(+\pi/2)}|\uparrow\rangle, \quad (12.35)$$

$$|\leftarrow\rangle = e^{-i(\hat{S}_y/\hbar)(-\pi/2)}|\uparrow\rangle. \quad (12.36)$$

(Problems 2.16 through 2.18 are relevant here.)

12.4 Spin-1 projection amplitudes

- (Easy.) Prove that

$$d_{m,m'}^{(j)}(\theta) = [d_{m',m}^{(j)}(-\theta)]^*. \quad (12.37)$$

- Show that the $d_{m,m'}^{(j)}(\theta)$ with $j = 1$ are

$$\begin{array}{lll} d_{1,1}^{(1)}(\theta) = +\frac{1}{2}(\cos \theta + 1) & d_{1,0}^{(1)}(\theta) = -\frac{1}{\sqrt{2}}\sin \theta & d_{1,-1}^{(1)}(\theta) = -\frac{1}{2}(\cos \theta - 1) \\ d_{0,1}^{(1)}(\theta) = +\frac{1}{\sqrt{2}}\sin \theta & d_{0,0}^{(1)}(\theta) = \cos \theta & d_{0,-1}^{(1)}(\theta) = -\frac{1}{\sqrt{2}}\sin \theta \\ d_{-1,1}^{(1)}(\theta) = -\frac{1}{2}(\cos \theta - 1) & d_{-1,0}^{(1)}(\theta) = +\frac{1}{\sqrt{2}}\sin \theta & d_{-1,-1}^{(1)}(\theta) = +\frac{1}{2}(\cos \theta + 1) \end{array}$$

Chapter 13

Central Force Motion

13.1 Energy eigenproblem in two dimensions

In one dimension, the energy eigenproblem is

$$-\frac{\hbar^2}{2M} \frac{d^2 \eta_n(x)}{dx^2} + V(x) \eta_n(x) = E_n \eta_n(x). \quad (13.1)$$

The generalization to two dimensions is straightforward:

$$-\frac{\hbar^2}{2M} \left[\frac{\partial^2 \eta(x, y)}{\partial x^2} + \frac{\partial^2 \eta(x, y)}{\partial y^2} \right] + V(x, y) \eta(x, y) = E_n \eta(x, y). \quad (13.2)$$

The part in square brackets is called “the Laplacian of $\eta_n(x, y)$ ” and represented by the symbol “ ∇^2 ” as follows

$$\left[\frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} \right] \equiv \nabla^2 f(x, y). \quad (13.3)$$

Thus the “mathematical form” of the energy eigenproblem is

$$\nabla^2 \eta_n(\vec{r}) = -\frac{2M}{\hbar^2} [E_n - V(\vec{r})] \eta_n(\vec{r}). \quad (13.4)$$

Suppose $V(x, y)$ 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:

$$\nabla^2 f(\vec{r}) = \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f(r, \theta)}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f(r, \theta)}{\partial \theta^2} \right]. \quad (13.5)$$

Thus, the partial differential equation to be solved is

$$\left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \eta_n(r, \theta)}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \eta_n(r, \theta)}{\partial \theta^2} \right] = -\frac{2M}{\hbar^2} [E_n - V(r)] \eta_n(r, \theta) \quad (13.6)$$

or

$$\left\{ \frac{\partial^2}{\partial \theta^2} + r \frac{\partial}{\partial r} \left(r \frac{\partial}{\partial r} \right) + \frac{2M}{\hbar^2} r^2 [E_n - V(r)] \right\} \eta_n(r, \theta) = 0. \quad (13.7)$$

For convenience, we wrap up all the r dependence into one piece by defining the linear operator

$$\mathcal{L}_n(r) \equiv r \frac{\partial}{\partial r} \left(r \frac{\partial}{\partial r} \right) + \frac{2M}{\hbar^2} r^2 [E_n - V(r)] \quad (13.8)$$

and write the above as

$$\left\{ \frac{\partial^2}{\partial \theta^2} + \mathcal{L}_n(r) \right\} \eta_n(r, \theta) = 0. \quad (13.9)$$

There are at least two ways to approach the above equation: the Fourier series method and the separation of variables method. We'll try each one in turn.

Fourier series:

Because increasing the angle θ by 2π brings you to the same point where you started, the function $\eta_n(r, \theta)$ is periodic in θ with period 2π . And the theory of Fourier series teaches that any such function can be written in the form

$$a_0 + \sum_{n=1}^{\infty} (a_n \cos n\theta + b_n \sin n\theta). \quad (13.10)$$

But, because

$$\cos n\theta = \frac{e^{in\theta} + e^{-in\theta}}{2} \quad \text{and} \quad \sin n\theta = \frac{e^{in\theta} - e^{-in\theta}}{2i},$$

this series can also be written in the form

$$\sum_{\ell=-\infty}^{\infty} c_{\ell} e^{i\ell\theta}. \quad (13.11)$$

[When dealing with real functions, the form (13.10) has obvious attractions. But the form (13.11) is always more compact, and when dealing with complex-valued functions it's more natural as well.] The Fourier theorem (13.11) is just another way of saying that the set of functions

$$e^{i\ell\theta} \quad \text{with} \quad \ell = 0, \pm 1, \pm 2, \pm 3, \dots \quad (13.12)$$

constitute a basis for the set of continuous functions with periodicity 2π .

Applying this Fourier theorem to the problem at hand, we write the energy eigenfunction as

$$\eta_n(r, \theta) = \sum_{\ell=-\infty}^{\infty} R_{n,\ell}(r) e^{i\ell\theta} \quad (13.13)$$

and then take

$$\begin{aligned} 0 &= \left\{ \frac{\partial^2}{\partial \theta^2} + \mathcal{L}_n(r) \right\} \eta_n(r, \theta) \\ &= \left\{ \frac{\partial^2}{\partial \theta^2} + \mathcal{L}_n(r) \right\} \sum_{\ell=-\infty}^{\infty} R_{n,\ell}(r) e^{i\ell\theta} \end{aligned}$$

$$\begin{aligned}
&= \sum_{\ell=-\infty}^{\infty} \left\{ \frac{\partial^2}{\partial \theta^2} + \mathcal{L}_n(r) \right\} R_{n,\ell}(r) e^{i\ell\theta} \\
&= \sum_{\ell=-\infty}^{\infty} [-\ell^2 R_{n,\ell}(r) + \mathcal{L}_n(r) R_{n,\ell}(r)] e^{i\ell\theta}.
\end{aligned} \tag{13.14}$$

Because the basis set $\{e^{i\ell\theta}\}$ is complete (or, equivalently, because Fourier series are unique) the expression in square brackets above must vanish, for *every* value of ℓ , whence

$$\mathcal{L}_n(r) R_{n,\ell}(r) - \ell^2 R_{n,\ell}(r) = 0 \quad \text{for } \ell = 0, \pm 1, \pm 2, \pm 3, \dots \tag{13.15}$$

We no longer have a partial differential equation. Instead we have an infinite number of ordinary differential equations

$$\left[r \frac{d}{dr} \left(r \frac{d}{dr} \right) + \frac{2M}{\hbar^2} r^2 [E_n - V(r)] - \ell^2 \right] R_{n,\ell}(r) = 0. \tag{13.16}$$

Separation of variables:

Our equation

$$\left\{ \frac{\partial^2}{\partial \theta^2} + \mathcal{L}_n(r) \right\} \eta_n(r, \theta) = 0. \tag{13.17}$$

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 elements of a basis. We cast around using the technique of “separation of variables”, namely by looking for solutions of the form

$$\eta_n(r, \theta) = R(r)\Theta(\theta). \tag{13.18}$$

Plugging this form into the PDE gives

$$\begin{aligned}
R(r)\Theta''(\theta) + \Theta(\theta)\mathcal{L}_n(r)R(r) &= 0 \\
\frac{\Theta''(\theta)}{\Theta(\theta)} + \frac{\mathcal{L}_n(r)R(r)}{R(r)} &= 0
\end{aligned} \tag{13.19}$$

Through the usual separation-of-variables argument, we recognize that if a function of r alone plus a function of θ alone sum to zero, where r and θ are independent variables, then both functions must be equal to a constant:

$$\frac{\mathcal{L}_n(r)R(r)}{R(r)} = -\frac{\Theta''(\theta)}{\Theta(\theta)} = \text{const.} \tag{13.20}$$

First, look at the angular part:

$$\Theta''(\theta) = -\text{const}\Theta(\theta). \tag{13.21}$$

This is the differential equation for a mass on a spring! The two linearly independent solutions are

$$\Theta(\theta) = \sin(\sqrt{\text{const}}\theta) \quad \text{or} \quad \Theta(\theta) = \cos(\sqrt{\text{const}}\theta). \tag{13.22}$$

Now, the boundary condition for this ODE is just that the function must come back to itself if θ increases by 2π :

$$\Theta(\theta) = \Theta(2\pi + \theta). \tag{13.23}$$

If you think about this for a minute, you'll see that this means $\sqrt{\text{const}}$ must be an integer. The negative integers don't give us anything new, so we'll take

$$\sqrt{\text{const}} = \ell \quad \text{where} \quad \ell = 0, 1, 2, \dots \quad (13.24)$$

In summary, the solution to the angular problem is

$$\Theta(\theta) \quad \begin{array}{ccccccc} \ell = 0 & \ell = 1 & \ell = 2 & \ell = 3 & \dots \\ 1 & \sin \theta \text{ or } \cos \theta & \sin 2\theta \text{ or } \cos 2\theta & \sin 3\theta \text{ or } \cos 3\theta & \dots \end{array}$$

Alternatively, we can take linear combinations of the above to produce the set of solutions

$$\Theta(\theta) = e^{i\ell\theta} \quad \text{for} \quad \ell = 0, \pm 1, \pm 2, \pm 3, \dots \quad (13.25)$$

Now examine the radial part of the problem:

$$\frac{\mathcal{L}_n(r)R(r)}{R(r)} = \text{const} = \ell^2. \quad (13.26)$$

Because the radial operator depends on n , the solutions $R(r)$ will depend on both n and ℓ , so we denote them by $R_{n,\ell}(r)$. They solve the equation

$$\mathcal{L}_n(r)R_{n,\ell}(r) - \ell^2 R_{n,\ell}(r) = 0 \quad \text{for} \quad \ell = 0, \pm 1, \pm 2, \pm 3, \dots \quad (13.27)$$

We no longer have a partial differential equation. Instead we have an infinite number of ordinary differential equations

$$\left[r \frac{d}{dr} \left(r \frac{d}{dr} \right) + \frac{2M}{\hbar^2} r^2 [E_n - V(r)] - \ell^2 \right] R_{n,\ell}(r) = 0. \quad (13.28)$$

We are at the end of the bifurcation. Both the ‘‘Fourier series’’ route and the ‘‘Separation of variables’’ route have arrived at the same destination. This is not a coincidence. The mathematical theory of Sturm-Liouville problems assures us that for a wide class of partial differential equations, the result of separation of variables will give rise to a complete set of functions, like the trigonometric functions that form a basis for the continuous functions with period 2π .

Write the resulting one-variable ODE as

$$\begin{aligned} \left\{ \frac{1}{r} \frac{d}{dr} \left(r \frac{d}{dr} \right) + \frac{2M}{\hbar^2} [E_n - V(r)] - \frac{\ell^2}{r^2} \right\} R_{n,\ell}(r) &= 0 \\ \left\{ \frac{1}{r} \frac{d}{dr} \left(r \frac{d}{dr} \right) + \frac{2M}{\hbar^2} \left[E_n - V(r) - \frac{\hbar^2 \ell^2}{2M r^2} \right] \right\} R_{n,\ell}(r) &= 0 \end{aligned} \quad (13.29)$$

I want to compare this differential equation with another one-variable differential equation, namely the one for the energy eigenvalue problem in one dimension:

$$\left\{ \frac{d^2}{dx^2} + \frac{2M}{\hbar^2} [E - V(x)] \right\} \eta(x) = 0. \quad (13.30)$$

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

$$\int_{-\infty}^{\infty} |\eta(x)|^2 dx = 1, \quad (13.31)$$

whereas the two-dimensional energy eigenfunction satisfies the normalization

$$\begin{aligned} \int |\eta(x, y)|^2 dx dy &= 1 \\ \int_0^{\infty} dr \int_0^{2\pi} r d\theta |R_{n,\ell}(r) e^{i\ell\theta}|^2 &= 1 \\ 2\pi \int_0^{\infty} dr r |R_{n,\ell}(r)|^2 &= 1. \end{aligned} \quad (13.32)$$

This suggests that the true analog of the one-dimensional $\eta(x)$ is not $R_{n,\ell}(r)$, but rather

$$u_{n,\ell}(r) = \sqrt{r} R_{n,\ell}(r). \quad (13.33)$$

Furthermore,

$$\text{if } u(r) = \sqrt{r} R(r), \quad \text{then } \frac{1}{r} \frac{d}{dr} (r R'(r)) = \frac{1}{\sqrt{r}} \left(u''(r) + \frac{1}{4} \frac{u(r)}{r^2} \right). \quad (13.34)$$

Using this change of function, the radial equation (13.29) becomes

$$\begin{aligned} \left\{ \frac{d^2}{dr^2} + \frac{1}{4} \frac{1}{r^2} + \frac{2M}{\hbar^2} \left[E - V(r) - \frac{\hbar^2 \ell^2}{2M r^2} \right] \right\} u(r) &= 0 \\ \left\{ \frac{d^2}{dr^2} + \frac{2M}{\hbar^2} \left[E - V(r) - \frac{\hbar^2}{2M} \left(\ell^2 - \frac{1}{4} \right) \frac{1}{r^2} \right] \right\} u(r) &= 0 \\ \left\{ \frac{d^2}{dr^2} + \frac{2M}{\hbar^2} \left[E - V(r) - \frac{\hbar^2 (\ell - \frac{1}{2})(\ell + \frac{1}{2})}{2M} \frac{1}{r^2} \right] \right\} u(r) &= 0 \end{aligned} \quad (13.35)$$

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

$$V(r) + \hbar^2 (\ell - \frac{1}{2})(\ell + \frac{1}{2}) / 2Mr^2.$$

These two functions play parallel mathematical roles in the two problems. To emphasize these similar roles, we define an “effective potential energy” for the radial problem, namely

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2 (\ell - \frac{1}{2})(\ell + \frac{1}{2})}{2M} \frac{1}{r^2}. \quad (13.36)$$

Don't read too much into the term “effective potential energy.” No actual potential energy will depend upon \hbar , or 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 eigenproblem.

Now that the radial equation (13.35) is in exact correspondence with the one-dimensional equation (13.30), we can solve this eigenproblem using the same “curve toward or away from axis” techniques that we developed

for the one-dimensional problem in chapter 9. (Or any other technique which 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 $n = 1, 2, 3, \dots$ calling them $u_{n,\ell}(r)$ with eigenvalue $E_{n,\ell}$.

Summary:

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

$$-\frac{\hbar^2}{2M}\nabla^2\eta(\vec{r}) + V(r)\eta(\vec{r}) = E\eta(\vec{r}), \quad (13.37)$$

first solve the radial energy eigenproblem

$$-\frac{\hbar^2}{2M}\frac{d^2u(r)}{dr^2} + \left[V(r) + \frac{\hbar^2(\ell - \frac{1}{2})(\ell + \frac{1}{2})}{2M} \frac{1}{r^2} \right] u(r) = Eu(r) \quad (13.38)$$

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

$$\text{For } \ell = 0: \quad \eta(r, \theta) = \frac{u_{n,0}(r)}{\sqrt{r}} \quad \text{with energy } E_{n,0} \quad (13.39)$$

and

$$\text{For } \ell = 1, 2, 3, \dots: \quad \eta(r, \theta) = \frac{u_{n,\ell}(r)}{\sqrt{r}} e^{i\ell\theta} \quad \text{and} \quad \eta(r, \theta) = \frac{u_{n,\ell}(r)}{\sqrt{r}} e^{-i\ell\theta} \quad \text{with energy } E_{n,\ell}. \quad (13.40)$$

Alternatively, for the last equation we can use

$$\text{For } \ell = 1, 2, 3, \dots: \quad \eta(r, \theta) = \frac{u_{n,\ell}(r)}{\sqrt{r}} \sin(\ell\theta) \quad \text{and} \quad \eta(r, \theta) = \frac{u_{n,\ell}(r)}{\sqrt{r}} \cos(\ell\theta) \quad \text{with energy } E_{n,\ell}. \quad (13.41)$$

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

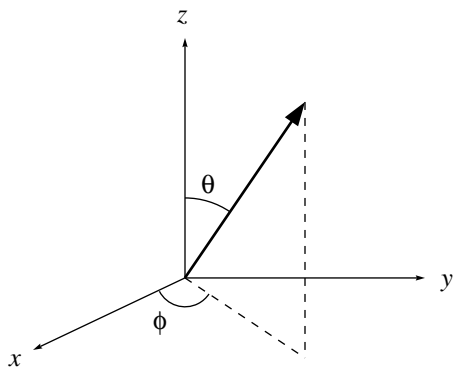
$$\{\text{angular operator} + \text{radial operator}\}\eta_n(r, \theta) = 0. \quad (13.42)$$

- There was an angular operator eigenbasis $\{\Phi_\ell(\theta)\}$ such that

$$\{\text{angular operator}\}\Phi_\ell(\theta) = \text{number } \Phi_\ell(\theta). \quad (13.43)$$

13.2 Energy eigenproblem in three dimensions

Can we get the same miracle to occur in three dimensions?



The energy eigenproblem is

$$-\frac{\hbar^2}{2M}\nabla^2\eta_n(\vec{x}) + V(r)\eta_n(\vec{x}) = E_n\eta_n(\vec{x}), \quad (13.44)$$

and the Laplacian in spherical coordinates is

$$\begin{aligned} \nabla^2 &= \frac{1}{r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right] \\ &= \frac{1}{r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{\partial}{\partial\mu} \left((1-\mu^2) \frac{\partial}{\partial\mu} \right) + \frac{1}{1-\mu^2} \frac{\partial^2}{\partial\phi^2} \right]. \end{aligned} \quad (13.45)$$

Here ϕ ranges from 0 to 2π , and θ ranges from 0 to π . It is often convenient to use, in place of θ , the variable

$$\mu = \cos\theta \quad \text{where } \mu \text{ ranges from } -1 \text{ to } +1. \quad (13.46)$$

(The situation $\mu = +1$ corresponds to the “north pole” of the spherical coordinate system – the positive z axis – while the situation $\mu = -1$ corresponds to the “south pole” – the negative z axis.) The energy eigenproblem is then

$$\left\{ \nabla^2 + \frac{2M}{\hbar^2} [E_n - V(r)] \right\} \eta_n(\vec{x}) = 0, \quad (13.47)$$

or

$$\left\{ \frac{\partial}{\partial\mu} \left((1-\mu^2) \frac{\partial}{\partial\mu} \right) + \frac{1}{1-\mu^2} \frac{\partial^2}{\partial\phi^2} + \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{2M}{\hbar^2} r^2 [E_n - V(r)] \right\} \eta_n(\vec{x}) = 0, \quad (13.48)$$

and this expression is indeed in the desired form (13.42).

Now, is there a basis of angular operator eigenfunctions as required by (13.43)?

We seek a complete set of functions on the unit sphere $\{y_\lambda(\mu, \phi)\}$ such that

$$\{\text{angular operator}\} y_\lambda(\mu, \phi) = \left\{ \frac{\partial}{\partial\mu} \left((1-\mu^2) \frac{\partial}{\partial\mu} \right) + \frac{1}{1-\mu^2} \frac{\partial^2}{\partial\phi^2} \right\} y_\lambda(\mu, \phi) = \lambda y_\lambda(\mu, \phi). \quad (13.49)$$

These functions, like any function of angle ϕ , must be periodic in ϕ with period 2π , so

$$y_\lambda(\mu, \phi) = \sum_{m=-\infty}^{\infty} p_{\lambda,m}(\mu) e^{im\phi}. \quad (13.50)$$

In these terms, the eigenproblem (13.49) becomes

$$\sum_{m=-\infty}^{\infty} \left\{ \left[\frac{\partial}{\partial \mu} \left((1-\mu^2) \frac{\partial}{\partial \mu} \right) - \frac{m^2}{1-\mu^2} - \lambda \right] p_{\lambda,m}(\mu) \right\} e^{im\phi} = 0. \quad (13.51)$$

Because of the uniqueness of Fourier series (compare the reasoning below equation (13.14)) each term in curly brackets must vanish individually. For each m

$$\left[\frac{d}{d\mu} \left((1-\mu^2) \frac{d}{d\mu} \right) - \frac{m^2}{1-\mu^2} - \lambda \right] p_{\lambda,m}(\mu) = 0 \quad m = 0, \pm 1, \pm 2, \dots \quad (13.52)$$

This differential equation is called the “generalized Legendre¹ equation.” It can be solved using the power-series method — details are left to problem X.Y. There are, of course, two linearly independent solutions for all values of m and λ . However, most of those solutions diverge at either $\mu = 1$, or $\mu = -1$, or both. (That is, either at the north or south pole, or both.) Solutions are finite for all μ from -1 to $+1$, inclusive, if and only if

$$\lambda = -\ell(\ell + 1) \quad \ell = 0, 1, 2, 3, \dots \quad (13.53)$$

and

$$m = -\ell, -\ell + 1, \dots, 0, \dots, \ell - 1, \ell. \quad (13.54)$$

In these cases, the finite solution is called the “associated Legendre function” $P_\ell^m(\mu)$.

These functions have a lot of interesting properties, but the main point is that we can use them to define a set of functions on the unit sphere called the *spherical harmonics*

$$Y_\ell^m(\theta, \phi) \quad \text{or} \quad Y_\ell^m(\mu, \phi) = \sqrt{\frac{2\ell + 1}{4\pi} \frac{(\ell - m)!}{(\ell + m)!}} P_\ell^m(\mu) e^{im\phi}. \quad (13.55)$$

(Some authors use different prefactors.)

The spherical harmonics satisfy

$$\{\text{angular operator}\} Y_\ell^m(\mu, \phi) = \left\{ \frac{\partial}{\partial \mu} \left((1-\mu^2) \frac{\partial}{\partial \mu} \right) + \frac{1}{1-\mu^2} \frac{\partial^2}{\partial \phi^2} \right\} Y_\ell^m(\mu, \phi) = -\ell(\ell + 1) Y_\ell^m(\mu, \phi) \quad (13.56)$$

and are complete in the sense that

Theorem: If $f(\mu, \phi)$ is a differentiable function on the unit sphere then

$$f(\mu, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell,m} Y_\ell^m(\mu, \phi) \quad \text{where} \quad f_{\ell,m} = \int_0^{2\pi} d\phi \int_{-1}^1 d\mu (Y_\ell^m(\mu, \phi))^* f(\mu, \phi). \quad (13.57)$$

¹Adrien-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.

The above paragraph is precisely analogous to the Fourier series result that the “trigonometric” functions $e^{i\ell\theta}$ satisfy

$$\left\{ \frac{\partial^2}{\partial\theta^2} \right\} e^{i\ell\theta} = -\ell^2 e^{i\ell\theta} \quad (13.58)$$

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

$$f(\theta) = \sum_{\ell=-\infty}^{\infty} f_{\ell} e^{i\ell\theta} \quad \text{where} \quad f_{\ell} = \frac{1}{2\pi} \int_0^{2\pi} d\theta (e^{i\ell\theta})^* f(\theta). \quad (13.59)$$

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.

Now that we have a complete set of eigenfunctions for the angular operator, we can carry out the “reduction to one dimension” process in three dimensions just as we did in two dimensions. The energy eigenvalue problem is

$$\left\{ -\frac{\hbar^2}{2M} \nabla^2 + V(r) - E_n \right\} \eta_n(\vec{r}) = 0. \quad (13.60)$$

Write

$$\eta_n(r, \mu, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} R_{n,\ell,m}(r) Y_{\ell}^m(\mu, \phi) \quad (13.61)$$

to prove that $R_{n,\ell,m}(r)$ satisfies

$$\left\{ \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) + \frac{2M}{\hbar^2} \left[E_n - V(r) - \frac{\hbar^2}{2M} \frac{\ell(\ell+1)}{r^2} \right] \right\} R_{n,\ell,m}(r) = 0. \quad (13.62)$$

Note that the differential equation is independent of m , so the solution must also be independent of m . Thus we drop the subscript m and write $R_{n,\ell}(r)$.

The energy eigenfunction satisfies the normalization

$$\begin{aligned} \int |\eta(x, y, z)|^2 dx dy dz &= 1 \\ \int_0^{\infty} dr \int_{-1}^1 d\mu \int_0^{2\pi} r^2 d\phi |R_{n,\ell}(r) Y_{\ell}^m(\mu, \phi)|^2 &= 1 \\ \int_0^{\infty} dr r^2 |R_{n,\ell}(r)|^2 &= 1. \end{aligned} \quad (13.63)$$

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

$$\left\{ -\frac{\hbar^2}{2M} \frac{d^2}{dr^2} + \left[V(r) + \frac{\hbar^2 \ell(\ell+1)}{2Mr^2} \right] \right\} u_{n,\ell}(r) = E_{n,\ell} u_{n,\ell}(r). \quad (13.64)$$

Summary:

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

$$-\frac{\hbar^2}{2M}\nabla^2\eta(\vec{r}) + V(r)\eta(\vec{r}) = E\eta(\vec{r}), \quad (13.65)$$

first solve the radial energy eigenproblem

$$-\frac{\hbar^2}{2M}\frac{d^2u(r)}{dr^2} + \left[V(r) + \frac{\hbar^2\ell(\ell+1)}{2M} \frac{1}{r^2} \right] u(r) = Eu(r) \quad (13.66)$$

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

$$\eta_{n,\ell,m}(r, \theta, \phi) = \frac{u_{n,\ell}(r)}{r} Y_\ell^m(\theta, \phi) \quad \text{with energy } E_{n,\ell}, \quad (13.67)$$

where m takes on the $2\ell + 1$ values $-\ell, -\ell + 1, \dots, 0, \dots, \ell - 1, \ell$. Notice that the $2\ell + 1$ different solutions for a given n and ℓ , but with different m , are degenerate.

Problem: 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 .

13.3 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

$$\left\{ -\frac{\hbar^2}{2M}\frac{d^2}{dr^2} + \left[-\frac{k}{r} + \frac{\hbar^2\ell(\ell+1)}{2Mr^2} \right] \right\} U_{n,\ell}(r) = E_{n,\ell}U_{n,\ell}(r) \quad (13.68)$$

are normalizable wavefunctions

$$\int_0^\infty |U_{n,\ell}(r)|^2 dr = 1. \quad (13.69)$$

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

$$a = \frac{\hbar^2}{kM}. \quad (13.70)$$

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

$$\frac{\hbar^2}{m_e} \frac{4\pi\epsilon_0}{e^2} \equiv a_0 \equiv \quad \text{“the Bohr radius”}. \quad (13.71)$$

Convert to the dimensionless variable

$$\rho = \frac{r}{a} \quad (13.72)$$

and the dimensionless wavefunction

$$u_{n,\ell}(\rho) = \sqrt{a} U_{n,\ell}(a\rho). \quad (13.73)$$

The resulting eigenproblem is

$$\left\{ -\frac{d^2}{d\rho^2} - \frac{2}{\rho} + \frac{\ell(\ell+1)}{\rho^2} \right\} u_{n,\ell}(\rho) = \frac{E_{n,\ell}}{k^2 M / 2\hbar^2} u_{n,\ell}(\rho) \quad (13.74)$$

with

$$\int_0^\infty |u_{n,\ell}(\rho)|^2 d\rho = 1. \quad (13.75)$$

It's clear that the energy

$$\frac{k^2 M}{2\hbar^2} \quad (13.76)$$

is the characteristic energy for this problem. For hydrogen, its value is approximately

$$\left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{m_e}{2\hbar^2} \equiv \text{Ry} \equiv \quad \text{“the Rydberg”} = 13.6 \text{ eV}.$$

Thus it is reasonable, for brevity, to define the dimensionless energy parameter

$$\mathcal{E}_{n,\ell} = \frac{E_{n,\ell}}{k^2 M / 2\hbar^2}. \quad (13.77)$$

Furthermore, for the bound state problem $\mathcal{E}_{n,\ell}$ is negative so we define

$$b_{n,\ell}^2 = -\mathcal{E}_{n,\ell} \quad (13.78)$$

and the eigenproblem becomes

$$\left\{ \frac{d^2}{d\rho^2} + \frac{2}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - b_{n,\ell}^2 \right\} u_{n,\ell}(\rho) = 0 \quad (13.79)$$

with

$$\int_0^\infty |u_{n,\ell}(\rho)|^2 d\rho = 1. \quad (13.80)$$

2. Remove asymptotic behavior of solutions:

Note: In this section we will show that as $\rho \rightarrow 0$,

$$u_{n,\ell}(\rho) \approx \rho^{\ell+1}, \quad (13.81)$$

and that as $\rho \rightarrow \infty$,

$$u_{n,\ell}(\rho) \approx e^{-b_{n,\ell}\rho}, \quad (13.82)$$

so we will set

$$u_{n,\ell}(\rho) = \rho^{\ell+1} e^{-b_{n,\ell}\rho} v_{n,\ell}(\rho) \quad (13.83)$$

and then solve an ODE for $v_{n,\ell}(\rho)$. As far as rigor is concerned we could have just pulled the change-of-function (13.83) out of a hat. Thus this section is motivational and doesn't need to be rigorous.

Because equation (13.79) has problems (or, formally, a “regular singular point”) at $\rho = 0$, it pays to find the asymptotic behavior when $\rho \rightarrow 0$ as well as when $\rho \rightarrow \infty$.

2A. Find asymptotic behavior as $\rho \rightarrow 0$: The ODE is

$$\left\{ \frac{d^2}{d\rho^2} + \left[\frac{2}{\rho} - \frac{\ell(\ell+1)}{\rho^2} - b_{n,\ell}^2 \right] \right\} u_{n,\ell}(\rho) = 0. \quad (13.84)$$

As $\rho \rightarrow 0$ the term in square brackets is dominated (unless $\ell = 0$) by $-\ell(\ell+1)/\rho^2$. The equation

$$\left\{ \frac{d^2}{d\rho^2} - \frac{\ell(\ell+1)}{\rho^2} \right\} u(\rho) = 0 \quad (13.85)$$

is solved by

$$u(\rho) = A\rho^{\ell+1} + B\rho^{-\ell}. \quad (13.86)$$

However, it's not healthy to keep factors like $\rho^{-\ell}$ around, because

$$\int_0^{\rho_0} \rho^{-2\ell} d\rho = \frac{1}{-2\ell+1} \left[\frac{1}{\rho^{2\ell-1}} \right]_0^{\rho_0} = \infty \quad [\text{for } \ell > \frac{1}{2}], \quad (13.87)$$

so wavefunctions with $\rho^{-\ell}$ prefactors tend to be unnormalizable. (Here ρ_0 is just any positive number.) Thus the wavefunction must behave as

$$u(\rho) \approx A\rho^{\ell+1} \quad (13.88)$$

as $\rho \rightarrow 0$.

Our arguments have relied upon $\ell \neq 0$, but it turns out that by stupid good luck the result (13.88) 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 $\rho \rightarrow \infty$: In this case, the square bracket term in equation (13.84) is dominated by $-b_{n,\ell}^2$, so the approximate ODE is

$$\left\{ \frac{d^2}{d\rho^2} - b_{n,\ell}^2 \right\} u_{n,\ell}(\rho) = 0 \quad (13.89)$$

with solutions

$$u_{n,\ell}(\rho) = Ae^{-b_{n,\ell}\rho} + be^{+b_{n,\ell}\rho}. \quad (13.90)$$

Clearly, normalization requires that $B = 0$, so the wavefunction has the expected exponential cutoff for large ρ .

In this way, we have justified the definition of $v_{n,\ell}(\rho)$ in equation (13.83). Plugging (13.83) into ODE (13.79), we find that $v_{n,\ell}(\rho)$ satisfies the ODE

$$\left\{ \rho \frac{d^2}{d\rho^2} + 2[\ell + 1 - b_{n,\ell}\rho] \frac{d}{d\rho} - 2[b_{n,\ell}\ell + b_{n,\ell} - 1] \right\} v_{n,\ell}(\rho) = 0 \quad (13.91)$$

3. *Find non-asymptotic behavior using the series method:* We try out the solution

$$v_{n,\ell}(\rho) = \sum_{k=0}^{\infty} a_k \rho^k \quad (13.92)$$

and readily find that

$$a_{k+1} = \frac{2b_{n,\ell}(k + \ell + 1) - 2}{(k + 1)(k + 2\ell + 2)} a_k \quad k = 0, 1, 2, \dots \quad (13.93)$$

(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

$$\frac{a_{k+1}}{a_k} \rightarrow \frac{2b_{n,\ell}}{k} \quad \text{as } k \rightarrow \infty. \quad (13.94)$$

As in the SHO, this leads to $v(\rho) \approx e^{2b_{n,\ell}\rho}$ as $\rho \rightarrow \infty$, which is pure disaster. To avoid catastrophe, we must truncate the series as a k th order polynomial by demanding

$$b_{n,\ell} = \frac{1}{k + \ell + 1} \quad k = 0, 1, 2, \dots \quad (13.95)$$

Thus $b_{n,\ell}$ is always the reciprocal of the integer

$$n = k + \ell + 1 \quad (13.96)$$

and

$$\mathcal{E}_{n,\ell} = -b_{n,\ell}^2 = -\frac{1}{n^2} \quad n = 1, 2, 3, \dots \quad (13.97)$$

We have found the permissible bound state energies!

What are the eigenfunctions? The solution $v_{n,\ell}(\rho)$ that is a polynomial of order $k = n - \ell - 1$ has a name: it is the Laguerre² polynomial

$$L_{n-\ell-1}^{2\ell+1}((2/n)\rho). \quad (13.98)$$

It would be nicer to have a more direct notation like our own $v_{n,\ell}(\rho)$, but Laguerre died before quantum mechanics was born, so he could not have known how to make his notation convenient for the quantum

²Edmond Laguerre (1834–1886), French artillery officer and mathematician, who made contributions to analysis and especially geometry.

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

$$\eta_{n,\ell,m}(\rho, \theta, \phi) = [\text{constant}] \rho^\ell e^{-\rho/n} L_{n-\ell-1}^{2\ell+1}((2/n)\rho) Y_\ell^m(\theta, \phi). \quad (13.99)$$

Degeneracy

Recall that each $v_{n,\ell}(\rho)$ already has an associated $2\ell + 1$ -fold degeneracy. In addition, each ℓ gives rise to an infinite number of eigenvalues:

$$\mathcal{E}_{n,\ell} = -\frac{1}{(k + \ell + 1)^2} \quad k = 0, 1, 2, \dots \quad (13.100)$$

In tabular form

$$\begin{array}{l} \ell = 0 \quad \text{gives} \quad n = \quad 1, \quad 2, \quad 3, \quad 4, \quad \dots \\ \ell = 1 \quad \text{gives} \quad n = \quad \quad 2, \quad 3, \quad 4, \quad \dots \\ \ell = 2 \quad \text{gives} \quad n = \quad \quad \quad 3, \quad 4, \quad \dots \\ \vdots \end{array}$$

So...

$$\begin{array}{l} \ell = 0 \quad (\text{degeneracy } 1) \quad \text{gives} \quad \mathcal{E}_{n,\ell} = \quad -1, \quad -\frac{1}{2^2}, \quad -\frac{1}{3^2}, \quad -\frac{1}{4^2}, \quad \dots \\ \ell = 1 \quad (\text{degeneracy } 3) \quad \text{gives} \quad \mathcal{E}_{n,\ell} = \quad \quad -\frac{1}{2^2}, \quad -\frac{1}{3^2}, \quad -\frac{1}{4^2}, \quad \dots \\ \ell = 2 \quad (\text{degeneracy } 5) \quad \text{gives} \quad \mathcal{E}_{n,\ell} = \quad \quad \quad -\frac{1}{3^2}, \quad -\frac{1}{4^2}, \quad \dots \\ \vdots \end{array}$$

Eigenenergies of $-1/n^2$ are associated with n different values of ℓ , namely $\ell = 0, 1, \dots, n - 1$. The total degeneracy is thus

$$\sum_{\ell=0}^{n-1} (2\ell + 1) = n^2. \quad (13.101)$$

13.4 Summary of the bound state energy eigenproblem for a Coulombic potential

A complete set of energy eigenfunctions is $\eta_{n,\ell,m}(r, \theta, \phi)$

$$\begin{array}{l} \text{where} \quad \quad \quad n = 1, 2, 3, \dots \\ \text{and for each } n \quad \quad \ell = 0, 1, 2, \dots, n - 1 \\ \text{and for each } n \text{ and } \ell \quad m = -\ell, -\ell + 1, \dots, \ell - 1, \ell. \end{array}$$

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

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

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.]

13.5 Problems

13.1 Positronium

The "atom" positronium is a bound state of an electron and a positron. Find the allowed energies for positronium.

13.2 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

$$\left[-\frac{d^2}{d\rho^2} + \frac{\ell(\ell+1)}{\rho^2} - \frac{2}{\rho} \right] u_{n,\ell}(\rho) \equiv h_\ell u_{n,\ell}(\rho) = \mathcal{E}_{n,\ell} u_{n,\ell}(\rho). \quad (13.102)$$

Introduce the operators

$$D_\pm^{(\ell)} \equiv \frac{d}{d\rho} \mp \frac{\ell}{\rho} \pm \frac{1}{\ell} \quad (13.103)$$

and show that

$$D_-^{(\ell+1)} D_+^{(\ell+1)} = -h_\ell - \frac{1}{(\ell+1)^2}, \quad D_+^{(\ell)} D_-^{(\ell)} = -h_\ell - \frac{1}{\ell^2}. \quad (13.104)$$

From this, conclude that

$$h_{\ell+1} D_+^{(\ell+1)} u_{n,\ell}(\rho) = \mathcal{E}_{n,\ell} D_+^{(\ell+1)} u_{n,\ell}(\rho) \quad (13.105)$$

whence

$$D_+^{(\ell+1)} u_{n,\ell}(\rho) \propto u_{n,\ell+1}(\rho) \quad (13.106)$$

and $\mathcal{E}_{n,\ell}$ is independent of ℓ .

Argue that for every $\mathcal{E}_{n,\ell} < 0$ there is a maximum ℓ . (Hint: examine the effective potential for radial motion.) Call this value ℓ_{\max} , and set $n = \ell_{\max} + 1$ to show that

$$\mathcal{E}_{n,\ell} = -\frac{1}{n^2}, \quad \ell = 0, \dots, n-1. \quad (13.107)$$

13.3 A non-Coulombic central force

The central potential

$$V(r) = -\frac{k}{r} + \frac{c}{r^2} \quad (13.108)$$

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.

- Sketch the potential, assuming that k and c are both positive.
- Following the method of section 13.3, convert the radial equation of the energy eigenproblem into

$$\left[-\frac{d^2}{d\rho^2} - \frac{2}{\rho} + \frac{\gamma + \ell(\ell+1)}{\rho^2} \right] u_{n,\ell}(\rho) = \mathcal{E}_{n,\ell} u_{n,\ell}(\rho). \quad (13.109)$$

where $\gamma = 2cM/\hbar^2$ and where ρ , $\mathcal{E}_{n,\ell}$, and $u_{n,\ell}(\rho)$ are to be identified.

- Find two values of x such that $x(x+1) = \gamma + \ell(\ell+1)$. Select whichever one will be most convenient for later use.
- Convince yourself that the solution described in section 13.3 does not depend upon ℓ being an integer, and conclude that the energy eigenvalues are

$$\mathcal{E}_{n,\ell} = \frac{-1}{[n - \ell + \frac{1}{2}(-1 + \sqrt{(2\ell+1)^2 + 4\gamma})]^2} \quad (13.110)$$

where $n = 1, 2, 3, \dots$ and where for each n , ℓ can take on values $\ell = 0, 1, 2, \dots, n-1$.

- Verify that this energy spectrum reduces to the Coulomb limit when $c = 0$.

13.4 The quantum mechanical virial theorem

- Argue that, in an energy eigenstate $|\eta(t)\rangle$, the expectation value $\langle \hat{\mathbf{r}} \cdot \hat{\mathbf{p}} \rangle$ does not change with time.
- Hence conclude that $\langle \eta(t) | [\hat{\mathbf{r}} \cdot \hat{\mathbf{p}}, \hat{H}] | \eta(t) \rangle = 0$.
- Show that $[\hat{\mathbf{r}} \cdot \hat{\mathbf{p}}, \hat{\mathbf{p}}^2] = 2i\hbar \hat{\mathbf{p}}^2$, while $[\hat{\mathbf{r}} \cdot \hat{\mathbf{p}}, V(\hat{\mathbf{r}})] = -i\hbar \hat{\mathbf{r}} \cdot \nabla V(\hat{\mathbf{r}})$, where $V(\mathbf{r})$ is any scalar function of the vector \mathbf{r} . (Hint: For the second commutator, use an explicit position basis representation.)
- Suppose the Hamiltonian is

$$\hat{H} = \frac{1}{2m} \hat{\mathbf{p}}^2 + V(\hat{\mathbf{r}}) = \hat{T} + \hat{V}. \quad (13.111)$$

Define the force function $\mathbf{F}(\mathbf{r}) = -\nabla V(\mathbf{r})$ and the force operator $\hat{\mathbf{F}} = \mathbf{F}(\hat{\mathbf{r}})$. Conclude that, for an energy eigenstate,

$$2\langle \hat{T} \rangle = -\langle \hat{\mathbf{r}} \cdot \hat{\mathbf{F}} \rangle. \quad (13.112)$$

This is the “virial theorem.”

- If $V(\mathbf{r}) = C/r^n$, show that $2\langle \hat{T} \rangle = -n\langle \hat{V} \rangle$ for any energy eigenstate, and that

$$\langle \hat{T} \rangle = \frac{n}{n-2} E, \quad \langle \hat{V} \rangle = \frac{-2}{n-2} E, \quad (13.113)$$

for the energy eigenstate with energy E .

13.5 Research project

Discuss the motion of wavepackets in a Coulombic potential. Does the expectation value of $\hat{\mathbf{r}}$ 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 14

Identical Particles

Note: Heap algorithm? Permutation groupie things in an appendix?

Identical particles *not* necessarily interacting, so two particles can be at same point.

14.1 Many-particle systems in quantum mechanics

One particle moves in one dimension. (Ignore spin.) How can we represent this system's state?

There are several ways: The ordinary wavefunction $\psi(x)$ represents the state in terms of the position basis. The momentum wavefunction

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dx e^{-i(p/\hbar)x} \psi(x)$$

represents the state in terms of the momentum basis. The energy expansion coefficients

$$c_n = \int_{-\infty}^{+\infty} \eta_n^*(x) \psi(x) dx$$

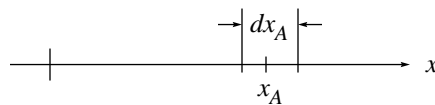
represent the state in terms of the energy basis. [Meaning that

$$\psi(x) = \sum_n c_n \eta_n(x). \quad]$$

Or we can represent the state in terms of the expectation of position $\langle x \rangle$, the expectation of momentum $\langle p \rangle$, the indeterminacy in position (which involves $\langle x^2 \rangle$), the indeterminacy in momentum (which involves $\langle p^2 \rangle$), the moments $\langle x^3 \rangle$ and $\langle p^3 \rangle$ and so forth, the correlation functions $\langle xp \rangle$ and $\langle xp^2 x^3 p \rangle$ and so forth. You can prove (it's not easy!) that if all these mean values are known then one can reconstruct the wavefunction.

Suppose we know the ordinary position wavefunction. Then if you measure the particle's location, the probability of finding it in a window of width dx_A about position x_A is $|\psi(x_A)|^2 dx_A$. This is *not* sufficient information to specify the particle's state: it tells you everything there is to know about position, but nothing about momentum or about position-momentum correlation functions.

One-particle window



Meanwhile the *amplitude* of finding the particle in this window is $\psi(x_A)\sqrt{dx_A}$. If you know the amplitude at every point x_A , then you *do* have full information about the state.

The normalization is of course

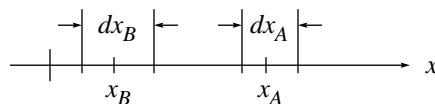
$$\int_{-\infty}^{+\infty} |\psi(x)|^2 dx = 1.$$

Variations: If a spin-zero particle moves in three dimensions, the wavefunction is $\psi(x, y, z)$. If a spin-half particle ($s_z = \pm\frac{1}{2}$) moves in three dimensions, the wavefunction is $\psi(x, y, z, s_z)$, or $\tilde{\psi}(p_x, p_y, p_z, s_x)$. In general, when I say things like “the variable x ”, you will have to generalize in different circumstances to, for example, “the variables p_x, p_y, p_z, s_x ”.

Two particles, say an electron and a neutron, move in one dimension. (Ignore spin.) How can we represent this system’s state?

There is now a wavefunction $\psi(x_A, x_B)$ with the interpretation that if you measure the location of both particles, then the probability of finding the electron in a window of width dx_A about position x_A , and the neutron in a window of width dx_B about position x_B , is $|\psi(x_A, x_B)|^2 dx_A dx_B$.

Two-particle windows



Note that the letters A and B refer to two different positions, not two different particles. The particles are represented by the sequence of arguments: the first argument pertains to the electron, the second argument pertains to the neutron.

I particularly emphasize that the wavefunction applying to the system “electron plus neutron” is *one* function of *two* variables, and is not *two* functions each of *one* variable:

$$\psi_{\text{electron}}(x_A) \quad \text{as well as} \quad \psi_{\text{neutron}}(x_B) \quad \mathbf{NO!}$$

The wavefunction of the system *might* happen to have the factorized form

$$\psi_{\text{electron}}(x_A)\psi_{\text{neutron}}(x_B) \quad \mathbf{PERHAPS}$$

but it does *not necessarily* have this form.

The difference feeds directly into this question: How many (real) numbers does it take to specify a state? In classical mechanics, the answer is straightforward. The state of any single particle is specified through two numbers: the position and momentum of that particle. The state of a collection of several particles is specified through the state of each particle. In summary

particles	real numbers needed to specify classical state
1	2
2	4
3	6
⋮	⋮
N	2N

In quantum mechanics, the answer is more subtle. To specify the state of even a single particle, one must give the wavefunction $\psi(x)$... an infinite number of complex numbers! For concreteness suppose we approximate this function on a computer, using a grid of 100 points. Then we need 100 complex numbers, that is $2(100)$ real numbers. But one of these numbers is fixed through the normalization condition, and one is an overall phase that can be set arbitrarily. The end result is that to specify a single-particle wavefunction to this degree of accuracy requires $2(100) - 2 = 198$ real numbers.

What about two particles? Now we have a wavefunction on a grid of 100×100 points, so specifying a two-particle wavefunction to this degree of accuracy requires $2(100)^2 - 2$ real numbers. This number (19998) is *much larger* than twice 198. 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 location, so in a two-particle system an individual particle might not have a state.

In summary

particles	real numbers needed to specify quantal state
1	$2(100) - 2 = 198$
2	$2(100)^2 - 2 = 19998$
3	$2(100)^3 - 2 = 1999998$
⋮	⋮
N	$2(100)^N - 2$

Much of the spectacular richness and complexity of the quantum world arises from this rapid increase of information with particle number. (Design of quantum computer.)

Two identical particles, say two neutrons, move in one dimension. (Ignore spin.) How can we represent this system's state?

Of course, there is a wavefunction $\psi(x_A, x_B)$, but the interpretation is somewhat different. The question is *not* "What is the probability of finding neutron α within window A and neutron β within window B ?"

These neutrons are identical, so there is no such thing as “neutron α ” or “neutron β .” The question instead is “What is the probability of finding a neutron within window A and a neutron within window B ?” The answer to this question is

$$\begin{cases} 2|\psi(x_A, x_B)|^2 dx_A dx_B & \text{if the windows don't overlap} \\ |\psi(x_A, x_B)|^2 dx_A dx_B & \text{if } x_B = x_A \end{cases}$$

The normalization condition is

$$2 \int_{-\infty}^{+\infty} dx_A \int_{-\infty}^{x_A} dx_B |\psi(x_A, x_B)|^2 = 1 \quad \text{or} \quad \int_{-\infty}^{+\infty} dx_A \int_{-\infty}^{\infty} dx_B |\psi(x_A, x_B)|^2 = 1.$$

If the two particles are identical, then it's certainly true that

$$|\psi(x_A, x_B)|^2 = |\psi(x_B, x_A)|^2.$$

But this condition insures only that the position probabilities are unaffected if you swap the windows. If the two particles are identical, then the same holds for momentum probabilities. In other words, the wavefunctions

$$\psi(x_A, x_B) \text{ and } \psi(x_B, x_A)$$

represent the same state, so

$$\psi(x_A, x_B) = e^{i\delta} \psi(x_B, x_A),$$

where δ is a number, not a function of x_A or x_B . Thus, for example,

$$\psi(5, 7) = e^{i\delta} \psi(7, 5).$$

But

$$\psi(7, 5) = e^{i\delta} \psi(5, 7),$$

so

$$\psi(5, 7) = (e^{i\delta})^2 \psi(5, 7).$$

We conclude that

$$e^{i\delta} = \pm 1.$$

In other words, when the wavefunction swaps arguments, it either remains the same or changes sign. In the first case, the wavefunction is called “symmetric under exchange,” (or “under swapping,” or “under interchange”)¹ in the second, “antisymmetric.”

What if there are **three identical particles**? The wavefunction is $\psi(x_A, x_B, x_C)$ and you can swap either the first and second arguments, or the second and third arguments, or the first and third arguments. The arguments of the next three paragraphs will show that the wavefunction must be either symmetric under each of these three interchanges or else antisymmetric under each of these three interchanges.

¹I prefer “swap” to emphasize that we're swapping mathematical windows, not exchanging physical particles, but the most commonly used term is “exchange.”

After any swapping, you must produce a wavefunction representing the same state, so any swapping can introduce at most a constant phase factor. Thus

$$\begin{aligned}\psi(x_A, x_B, x_C) &= e^{i\alpha}\psi(x_B, x_A, x_C) \\ &= e^{i\beta}\psi(x_A, x_C, x_B) \\ &= e^{i\gamma}\psi(x_C, x_B, x_A)\end{aligned}$$

The “double swap” argument above shows that $e^{i\alpha}$ is either $+1$ or -1 , that $e^{i\beta}$ is either $+1$ or -1 , and that $e^{i\gamma}$ is either $+1$ or -1 . We can gain more information through repeated swappings that return ultimately to the initial sequence. For example

$$\begin{aligned}\psi(x_A, x_B, x_C) &= e^{i\alpha}\psi(x_B, x_A, x_C) && \text{[[swapping first and second arguments]]} \\ &= e^{i\alpha}e^{i\beta}\psi(x_B, x_C, x_A) && \text{[[swapping second and third arguments]]} \\ &= e^{i\alpha}e^{i\beta}e^{i\gamma}\psi(x_A, x_C, x_B) && \text{[[swapping first and third arguments]]} \\ &= e^{i\alpha}e^{i\beta}e^{i\gamma}e^{i\beta}\psi(x_A, x_B, x_C) && \text{[[swapping second and third arguments]]}\end{aligned}$$

We already know that $(e^{i\beta})^2 = 1$, so this argument reveals that $e^{i\alpha}e^{i\gamma} = 1$, i.e., these two phase factors are either both $+1$ or both -1 .

Further arguments of this type will convince you that the three phase factors must either be all $+1$ or else all -1 . For suppose that

$$\begin{aligned}\psi(x_A, x_B, x_C) &= -\psi(x_B, x_A, x_C) \\ &= +\psi(x_A, x_C, x_B) \\ &= -\psi(x_C, x_B, x_A)\end{aligned}$$

(That is, antisymmetric under swaps of the first and second arguments or the first and third arguments, symmetric under swaps of the second and third arguments.) Then we can go from $\psi(x_A, x_B, x_C)$ to $\psi(x_B, x_C, x_A)$ via two different swapping routes:

$$\begin{aligned}\psi(x_A, x_B, x_C) &= (-1)\psi(x_B, x_A, x_C) && \text{[[swapping first and second arguments]]} \\ &= (-1)(+1)\psi(x_B, x_C, x_A) && \text{[[swapping second and third arguments]]}\end{aligned}$$

or

$$\begin{aligned}\psi(x_A, x_B, x_C) &= (-1)\psi(x_C, x_B, x_A) && \text{[[swapping first and third arguments]]} \\ &= (-1)(-1)\psi(x_B, x_C, x_A) && \text{[[swapping first and second arguments]]}\end{aligned}$$

The only function that satisfies both of conditions is $\psi(x_A, x_B, x_C) = 0$.

The other possible “mixed symmetric and antisymmetric” possibility is

$$\begin{aligned}\psi(x_A, x_B, x_C) &= +\psi(x_B, x_A, x_C) \\ &= -\psi(x_A, x_C, x_B) \\ &= +\psi(x_C, x_B, x_A)\end{aligned}$$

but this can be shown impossible by the “two route” argument of the previous paragraph.

14.1 Problem: Show that the same result applies for functions of four or more arguments by considering first swaps among the first, second, and third arguments; then swaps among the first, second, and fourth arguments; then swaps among the first, second, and fifth arguments; etc.

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.

14.2 Problem: 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.

14.3 Problem: Show that the momentum wavefunction has 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?)

14.4 Problem: 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.

Given what we’ve said 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. A collection of neutrons is *always* in an antisymmetric state. Furthermore, if two neutrons that have never encountered each other are brought close together, then the system goes into an antisymmetric state. The exchange symmetry turns out to have nothing to do with history, and depends only on the type of particle. Neutrons, protons, electrons, carbon-13 nuclei, and sigma baryons are always in antisymmetric states — they are called *fermions*. Photons, alpha particles, carbon-12 nuclei, and pi mesons are always in symmetric states — they are called *bosons*.

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. (See *Pauli and the Spin-Statistics Theorem* by Ian Duck and E.C.G. Sudarshan, and the review of this book by A.S. Wightman in *Am. J. Phys.* **67** (August 1999) 742–746.)

Given their obvious importance, it makes sense to spend some time on the **mathematics of completely symmetric and completely antisymmetric functions**. Given a garden-variety two-variable “seed” function $f(x_A, x_B)$, we can build a symmetric function

$$f_S(x_A, x_B) = f(x_A, x_B) + f(x_B, x_A),$$

and an antisymmetric function

$$f_A(x_A, x_B) = f(x_A, x_B) - f(x_B, x_A).$$

Note that these built functions are not necessarily normalized.

A generalized process works for three-variable functions: The built functions are sums over all $3!$ permutations of arguments. The function

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

is completely symmetric while the function

$$f_A(x_A, x_B, x_C) = f(x_A, x_B, x_C) - f(x_A, x_C, x_B) + f(x_C, x_A, x_B) - f(x_C, x_B, x_A) + f(x_B, x_C, x_A) - f(x_B, x_A, x_C)$$

is completely antisymmetric. This process of building an symmetric function f_S from arbitrary seed function f is called “symmetrization”. Similarly for “antisymmetrization”. If the function is a wavefunction, the (anti)symmetrization process is usually understood to include also normalizing the resulting wavefunction.

14.5 Problem: If the seed $f(x_A, x_B, x_C)$ happens to be symmetric to begin with, what are the symmetrized and antisymmetrized functions? What if the seed happens to be antisymmetric to begin with?

14.6 Problem: 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?

Exchange symmetry and position correlations

Symmetric implies “huddled together”, antisymmetric implies “spread apart”. (“cluster” / “avoid”)

This is not a result of repulsion. Two electrons, of course, repel each other electrically. This electrical repulsion is reflected through a term in the Hamiltonian of the pair. But the exchange symmetry effect holds even when there is no interaction term in the Hamiltonian. If the two particles are “independent” in that there is no interaction term in the Hamiltonian, in that they don’t interact through a repulsive or attractive force, they still have a tendency to “huddle together” or “spread apart” through the exchange symmetry requirement. (A pair of particles in this situation is said to have “no interaction” in the physics sense of the word “interaction,” even though they do affect each other in the everyday sense of the word “interaction”.)

Question: I can see how two electrons, repelling each other through an electric field, can affect each other. But you’ve just said that two identical particles which don’t exert a force on each other nevertheless

affect each other. The two particles are not in contact and don't exert a force. What is the mechanism through which one affects the other?

Answer: The two particles affect each other through “smelling out” the various positions available to each. Remember that these particles *don't have* positions.

Although this section is titled “Exchange symmetry and position correlations” remember that the symmetry requirement holds also for the momentum wavefunction and for the energy coefficients. Antisymmetric combinations are spread apart in momentum as well as position.

See problem 14.10.

Symmetric and antisymmetric bases

If we want to study identical particles, we'll need to build a basis of symmetric states and a basis of antisymmetric states. Here's how.

Start with a single particle subject to a potential, and solve the energy eigenproblem. Suppose the results are an energy basis of

$$\begin{array}{ll} \eta_1(x) & \epsilon_1 \\ \eta_2(x) & \epsilon_2 \\ \vdots & \vdots \\ \eta_M(x) & \epsilon_M \end{array}$$

In most cases the number of one-particle energy eigenstates M is infinite, but it's useful to keep that number as a variable anyway. There might or might not be some degeneracies in the system... it doesn't matter.

Three non-identical particles. Now suppose there is not one, but three particles moving in this potential, and that they're not identical. We can build a basis of product wavefunctions.

$$\begin{array}{lll} \eta_1(x_A)\eta_1(x_B)\eta_1(x_C) & |1, 1, 1\rangle & \epsilon_1 + \epsilon_1 + \epsilon_1 \\ \eta_1(x_A)\eta_1(x_B)\eta_2(x_C) & |1, 1, 2\rangle & \epsilon_1 + \epsilon_1 + \epsilon_2 \\ \eta_1(x_A)\eta_2(x_B)\eta_1(x_C) & |1, 2, 1\rangle & \epsilon_1 + \epsilon_2 + \epsilon_1 \\ \eta_2(x_A)\eta_1(x_B)\eta_1(x_C) & |2, 1, 1\rangle & \epsilon_2 + \epsilon_1 + \epsilon_1 \\ \vdots & \vdots & \vdots \\ \eta_1(x_A)\eta_7(x_B)\eta_3(x_C) & |1, 7, 3\rangle & \epsilon_1 + \epsilon_7 + \epsilon_3 \\ \eta_7(x_A)\eta_3(x_B)\eta_1(x_C) & |7, 3, 1\rangle & \epsilon_7 + \epsilon_3 + \epsilon_1 \\ \vdots & \vdots & \vdots \\ \eta_M(x_A)\eta_M(x_B)\eta_M(x_C) & |M, M, M\rangle & \epsilon_M + \epsilon_M + \epsilon_M \end{array}$$

A few remarks: (1) There are M^3 elements in the basis. (2) We have 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

elements. (3) It's tiring to write always the form in the left column so we abbreviate it through the form in the center column. (4) If the three particles don't interact, then this is an energy basis with the eigenvalues shown. But even if they do interact, it's a basis. (5) If the particles don't interact, then there is necessarily degeneracy in this basis. (6) To keep in mind the distinction between this basis for the three-particle system and the basis for the one-particle system from which it is built, we often call the three-particle basis elements "states" and the one-particle basis elements "levels". The levels are the building blocks out of which states are constructed.

Building an antisymmetric basis. We need a basis of states each of which is antisymmetric. Now any wavefunction can be expressed as a sum over the above basis,

$$\psi(x_A, x_B, x_C) = \sum_{r=1}^M \sum_{s=1}^M \sum_{t=1}^M c_{r,s,t} \eta_r(x_A) \eta_s(x_B) \eta_t(x_C) = \sum_{r,s,t} c_{r,s,t} |r, s, t\rangle,$$

but we're not interested in any wavefunction, we're interested in antisymmetric wavefunctions. To build an antisymmetric wavefunction, we execute the antisymmetrization process on $\psi(x_A, x_B, x_C)$. Doing so, we conclude that this antisymmetric wavefunction can be expressed as a sum over the antisymmetrization of each basis element.

Let's think a bit about the antisymmetrization of

$$\eta_r(x_A) \eta_s(x_B) \eta_t(x_C) \quad \text{also known as} \quad |r, s, t\rangle.$$

When we introduced the antisymmetrization process, we permuted the variables (representing position windows) x_A , x_B , and x_C . But if the seed function is a product like this, it's obviously the same thing to permute the level indices r , s , and t . We represent the antisymmetrization of $|r, s, t\rangle$ as

$$\hat{\mathcal{A}}|r, s, t\rangle = \text{const} (|r, s, t\rangle - |r, t, s\rangle + |t, r, s\rangle - |t, s, r\rangle + |s, t, r\rangle - |s, r, t\rangle)$$

where "const" is a normalization constant.

If we go through and antisymmetrize each element of the basis for three non-identical particles, we will find a basis for antisymmetric states. Let's start with $|1, 1, 1\rangle$. This antisymmetrizes to zero. Same with $|1, 1, 2\rangle$:

$$\hat{\mathcal{A}}|1, 1, 2\rangle = \text{const} (|1, 1, 2\rangle - |1, 2, 1\rangle + |2, 1, 1\rangle - |2, 1, 1\rangle + |1, 2, 1\rangle - |1, 1, 2\rangle) = 0.$$

It's clear, in fact, that *any* basis element with two indices the same will antisymmetrize to zero. The only way to avoid antisymmetrization to zero is for all the level indices to be zero. 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 elements $|1, 7, 3\rangle$, $|7, 3, 1\rangle$, $|3, 7, 1\rangle$, etc. all antisymmetrize to the same thing.

We are left with a basis of

$$\frac{M(M-1)(M-2)}{3!}$$

antisymmetric elements. One of the neat things about these elements is that they're long ... for example one of them is

$$\frac{1}{\sqrt{3!}} [\eta_1(x_A)\eta_7(x_B)\eta_3(x_C) - \eta_1(x_A)\eta_3(x_B)\eta_7(x_C) + \eta_3(x_A)\eta_1(x_B)\eta_7(x_C) \\ - \eta_3(x_A)\eta_7(x_B)\eta_1(x_C) + \eta_7(x_A)\eta_3(x_B)\eta_1(x_C) - \eta_7(x_A)\eta_1(x_B)\eta_3(x_C)]$$

but to specify them we need only state 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 1, a particle in level 7, and a particle in level 3”. This phrase is not absolutely 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 antisymmetric 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 believe it.

Building a symmetric basis. We can build a basis of states, each of which is symmetric, in a parallel manner by symmetrizing each element of the basis for non-identical particles and discarding duplicates.

The symmetrization of

$$\eta_r(x_A)\eta_s(x_B)\eta_t(x_C) \quad \text{also known as} \quad |r, s, t\rangle$$

results in

$$\hat{S}|r, s, t\rangle = \text{const} (|r, s, t\rangle + |r, t, s\rangle + |t, r, s\rangle + |t, s, r\rangle + |s, t, r\rangle + |s, r, t\rangle)$$

where “const” is again a normalization constant.

Let's start with $|1, 1, 1\rangle$. This symmetrizes to itself:

$$\hat{S}|1, 1, 1\rangle = |1, 1, 1\rangle.$$

Next comes $|1, 1, 2\rangle$:

$$\hat{S}|1, 1, 2\rangle = \text{const} (|1, 1, 2\rangle + |1, 2, 1\rangle + |2, 1, 1\rangle + |2, 1, 1\rangle + |1, 2, 1\rangle + |1, 1, 2\rangle) = 2 \text{const} (|1, 1, 2\rangle + |1, 2, 1\rangle + |2, 1, 1\rangle).$$

It's clear, then, that

$$\hat{S}|1, 1, 2\rangle = \hat{S}|1, 2, 1\rangle = \hat{S}|2, 1, 1\rangle.$$

and in fact that all states built out of any three given levels are the same state. The discarded duplicates result in a basis of

$$\frac{M(M+1)(M+2)}{3!}$$

symmetric elements.

Once again these states have long expressions like

$$\frac{1}{\sqrt{3!}} [\eta_1(x_A)\eta_7(x_B)\eta_3(x_C) + \eta_1(x_A)\eta_3(x_B)\eta_7(x_C) + \eta_3(x_A)\eta_1(x_B)\eta_7(x_C) \\ + \eta_3(x_A)\eta_7(x_B)\eta_1(x_C) + \eta_7(x_A)\eta_3(x_B)\eta_1(x_C) + \eta_7(x_A)\eta_1(x_B)\eta_3(x_C)]$$

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 antisymmetric wavefunctions, except that now multiple balls can go into the same bucket. 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 symmetrizing.

14.7 **Problem:** Count the elements in the antisymmetric and symmetric bases for N particles rather than three. (Continue to use M levels.) Does your expression have the proper limits when $N = 1$ and when $N = M$?

14.8 **Problem:** Find the normalization constant for $\hat{S}|7, 3, 7\rangle$.

14.9 **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.

14.10 Mean separation

(Be sure to read Griffiths section 5.1.2, “Exchange forces,” before attempting this problem.) Two noninteracting particles, each of mass m , are in an infinite square well of width L . The associated one-body energy eigenstates are $\eta_n(x)$ and $\eta_m(x)$, where

$$\eta_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi}{L}x\right).$$

Calculate the root-mean-square separation

$$\sqrt{\langle (x_A - x_B)^2 \rangle}$$

if these are

- two non-identical particles, one in state $\eta_n(x_A)$ and the other in state $\eta_m(x_B)$
- two identical bosons, in state

$$\frac{1}{\sqrt{2}} [\eta_n(x_A)\eta_m(x_B) + \eta_m(x_A)\eta_n(x_B)]$$

- two identical fermions, in state

$$\frac{1}{\sqrt{2}} [\eta_n(x_A)\eta_m(x_B) - \eta_m(x_A)\eta_n(x_B)]$$

14.11 Building basis states

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 and count all the possible three-particle states representing

- three non-identical particles

- b. three identical bosons
- c. three identical fermions

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.

14.2 An antisymmetric basis for the helium problem

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

$$\hat{H} = \underbrace{\widehat{\text{KE}}_A + \hat{U}_{nA}}_{\equiv \hat{H}_A} + \underbrace{\widehat{\text{KE}}_B + \hat{U}_{nB}}_{\equiv \hat{H}_B} + \underbrace{\hat{U}_{AB}}_{\equiv \hat{H}'}$$

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 appear as “point A” and “point B”.

We will 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 begin perturbation theory.

One-particle levels

We begin by finding the one-particle levels (or “orbitals”) for just the Hamiltonian \hat{H}_A . We combine these with levels for \hat{H}_B and antisymmetrize the result.

The problem \hat{H}_A is just the Hydrogen atom Coulomb problem with two changes:

$$\begin{aligned} \text{nuclear mass is } 4m_p &\implies \text{very small effect (“ignore nuclear motion”)} \\ \text{nuclear charge is } 2e &\implies \text{the Rydberg is } \text{Ry} = \frac{m_e}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2, \text{ so } E_{n_A}^{(A)} = -\frac{4\text{Ry}}{n_A^2} \end{aligned}$$

Similarly, the energy eigenstates for \hat{H}_A are represented by familiar functions like

$$\eta_{n\ell m}(\mathbf{r})|\uparrow\rangle \quad \text{or} \quad \eta_{n\ell m}(\mathbf{r})\chi_+$$

Soon we will need to keep track of \hat{H}_A versus \hat{H}_B . A notation like $\eta_{n\ell m}(\mathbf{r}_A)|\uparrow\rangle$ is fine for the space part of the eigenstate, but leaves the spin part ambiguous. We will hence 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

$$\frac{1}{\sqrt{2}}[\eta_{200}(\mathbf{r})\chi_+ + \eta_{210}(\mathbf{r})\chi_-]$$

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

Recall how we build an antisymmetrized wavefunction from a product of two one-particle levels, $\eta_n(A)$ and $\eta_m(B)$:

$$\hat{\mathcal{A}}\eta_n(A)\eta_m(B) = \frac{1}{\sqrt{2}}[\eta_n(A)\eta_m(B) - \eta_m(A)\eta_n(B)]$$

(The normalization factor $1/\sqrt{2}$ holds when $\eta_n(A)$ and $\eta_m(A)$ are orthogonal.)

Two theorems:

- If you antisymmetrize a product of the same two levels, you end up with zero:

$$\hat{\mathcal{A}}\eta_n(A)\eta_n(B) = 0.$$

- If you antisymmetrize two levels in the opposite sequence, you end up with the same state:

$$\hat{\mathcal{A}}\eta_n(A)\eta_m(B) = -\hat{\mathcal{A}}\eta_m(A)\eta_n(B).$$

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:

distinguishable

$$\begin{aligned} &\eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_+(B) \\ &\eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_-(B) \\ &\eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_+(B) \\ &\eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_-(B) \end{aligned}$$

But if you have identical fermions, two of these basis states antisymmetrize to zero, and the other two antisymmetrize to the same state:

distinguishable	antisymmetrized
$\eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_+(B)$	0
$\eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_-(B)$	$\frac{1}{\sqrt{2}}[\eta_{100}(A)\chi_+(A)\eta_{100}(B)\chi_-(B) - \eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_+(B)]$
$\eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_+(B)$	$-\frac{1}{\sqrt{2}}$ [above]
$\eta_{100}(A)\chi_-(A)\eta_{100}(B)\chi_-(B)$	0

Hence the Hamiltonian $\hat{H}_A + \hat{H}_B$ has a non-degenerate ground state, namely

$$\eta_{100}(A)\eta_{100}(B)\frac{1}{\sqrt{2}}[\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)].$$

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

$$\begin{aligned} &\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_+(B) \\ &\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_-(B) \\ &\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_+(B) \\ &\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_-(B) \\ &\eta_{n\ell m}(A)\chi_+(A)\eta_{100}(B)\chi_+(B) \\ &\eta_{n\ell m}(A)\chi_+(A)\eta_{100}(B)\chi_-(B) \\ &\eta_{n\ell m}(A)\chi_-(A)\eta_{100}(B)\chi_+(B) \\ &\eta_{n\ell m}(A)\chi_-(A)\eta_{100}(B)\chi_-(B) \end{aligned}$$

But if you have identical fermions, the “combination” means a multiplication followed by an antisymmetrization. Because of the second theorem concerning antisymmetrization, each of the last four products above antisymmetrize to the same state as one of the first four products. The first four products result in antisymmetrized states as follows:

distinguishable

antisymmetrized

$$\begin{aligned} \text{(a)} \quad &\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_+(B) && \frac{1}{\sqrt{2}}[\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_+(B) - \eta_{n\ell m}(A)\chi_+(A)\eta_{100}(B)\chi_+(B)] \\ \text{(b)} \quad &\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_-(B) && \frac{1}{\sqrt{2}}[\eta_{100}(A)\chi_+(A)\eta_{n\ell m}(B)\chi_-(B) - \eta_{n\ell m}(A)\chi_-(A)\eta_{100}(B)\chi_+(B)] \\ \text{(c)} \quad &\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_+(B) && \frac{1}{\sqrt{2}}[\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_+(B) - \eta_{n\ell m}(A)\chi_+(A)\eta_{100}(B)\chi_-(B)] \\ \text{(d)} \quad &\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_-(B) && \frac{1}{\sqrt{2}}[\eta_{100}(A)\chi_-(A)\eta_{n\ell m}(B)\chi_-(B) - \eta_{n\ell m}(A)\chi_-(A)\eta_{100}(B)\chi_-(B)] \end{aligned}$$

Antisymmetrized expressions (a) and (d) readily factor into a space part times a spin part:

$$\begin{aligned} \text{(a)} &\implies \frac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_+(A)\chi_+(B) \\ \text{(d)} &\implies \frac{1}{\sqrt{2}}[\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_-(A)\chi_-(B) \end{aligned}$$

But expressions (b) and (c) do *not* factor. One thing to do about this is nothing — after all, there’s no *requirement* that the wavefunctions factorize. But another approach is to look for a simple change of basis (remember, these four states all have the same energy of $\hat{H}_A + \hat{H}_B$). Someone (I don’t know who) thought about the favorite change of basis in planar geometry — a rotation of the axes by 45° :

Applying this transformation to the basis elements (“unit vectors”) given through (b) and (c) results in the new basis elements

$$\begin{aligned} \frac{1}{\sqrt{2}}[(b) + (c)] &= \frac{1}{2} \{ [\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_+(A)\chi_-(B) \\ &\quad + [\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_-(A)\chi_+(B) \} \\ &= \frac{1}{\sqrt{2}} [\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)] \frac{1}{\sqrt{2}} [\chi_+(A)\chi_-(B) + \chi_-(A)\chi_+(B)] \end{aligned}$$

and

$$\begin{aligned} \frac{1}{\sqrt{2}}[(b) - (c)] &= \frac{1}{2} \{ [\eta_{100}(A)\eta_{n\ell m}(B) + \eta_{n\ell m}(A)\eta_{100}(B)]\chi_+(A)\chi_-(B) \\ &\quad - [\eta_{100}(A)\eta_{n\ell m}(B) + \eta_{n\ell m}(A)\eta_{100}(B)]\chi_-(A)\chi_+(B) \} \\ &= \frac{1}{\sqrt{2}} [\eta_{100}(A)\eta_{n\ell m}(B) + \eta_{n\ell m}(A)\eta_{100}(B)] \frac{1}{\sqrt{2}} [\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)]. \end{aligned}$$

This process results in an antisymmetric basis of

$$\begin{aligned} &\frac{1}{\sqrt{2}} [\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_+(A)\chi_+(B) \\ &\frac{1}{\sqrt{2}} [\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)] \frac{1}{\sqrt{2}} [\chi_+(A)\chi_-(B) + \chi_-(A)\chi_+(B)] \\ &\frac{1}{\sqrt{2}} [\eta_{100}(A)\eta_{n\ell m}(B) - \eta_{n\ell m}(A)\eta_{100}(B)]\chi_-(A)\chi_-(B) \\ &\frac{1}{\sqrt{2}} [\eta_{100}(A)\eta_{n\ell m}(B) + \eta_{n\ell m}(A)\eta_{100}(B)] \frac{1}{\sqrt{2}} [\chi_+(A)\chi_-(B) - \chi_-(A)\chi_+(B)]. \end{aligned}$$

The first three elements are called a “triplet” (with “space antisymmetric, spin symmetric”). The last element is called a “singlet” (with “space symmetric, spin antisymmetric”). This particular basis has three nice properties: (1) Every basis element factorizes into a spatial part times a spin part. (2) Every basis element factorizes into a symmetric part times an antisymmetric part. (3) All three elements in 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 elements of any triplet remain degenerate even when the effects of the sub-Hamiltonian \hat{U}_{AB} are considered.

[[This process works for combining an arbitrary level $|n\ell m\rangle$ with a ground state level $|100\rangle$. Since this should work for any $|n\ell m\rangle$, what happens if we take $|n\ell m\rangle = |100\rangle$, the situation we first considered? In particular, what’s up with the normalization?]]

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 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.

14.12 Electron-electron repulsion

In class we wrote the (approximate) Hamiltonian for helium as $\hat{H}_A + \hat{H}_B + \hat{U}_{AB}$ and found antisymmetric energy eigenstates for $\hat{H}_A + \hat{H}_B$ that we called 1^1S , 2^3S , 2^1P , and so forth. Then we qualitatively discussed how the energy associated with such a state would change, under perturbation theory, through the electron-electron repulsion term \hat{U}_{AB} . Write down expressions for the first-order energy shifts due to \hat{U}_{AB} for 2^1S , 2^3S , and 2^1P . (That is, set up the integrals in terms of the one-particle eigenstates $\eta_{mlm}(\mathbf{r})$. Do *not* evaluate the integrals.) *Bonus*: Argue that the energy shift for 2^1P is greater than the shift for 2^1S .

14.13 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.

14.14 The meaning of two-particle wavefunctions (Old)

- a. The wavefunction $\psi(x_A, x_B)$ describes two non-identical particles in one dimension. Does

$$\int_{-\infty}^{\infty} dx_A \int_{-\infty}^{\infty} dx_B |\psi(x_A, x_B)|^2 \quad (14.1)$$

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

$$\int_{-\infty}^{\infty} dx_A \int_{-\infty}^{\infty} dx_B |\psi(x_A, x_B)|^2 \quad (14.2)$$

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”?

14.15 **Symmetric and close together, antisymmetric and far apart** (Old)

In lecture I argued that symmetric wavefunctions describe particles that huddle together while antisymmetric wavefunctions describe particles that avoid one another.

- a. Illustrate this principle as follows: Construct symmetric and antisymmetric two-particle wavefunctions out of the single-particle wavefunctions

$$\eta_1(x) = \sqrt{\frac{2}{L}} \sin\left(\pi \frac{x}{L}\right) \quad \text{and} \quad \eta_2(x) = \sqrt{\frac{2}{L}} \sin\left(2\pi \frac{x}{L}\right), \quad 0 \leq x \leq L, \quad (14.3)$$

which are the first and second energy eigenfunctions for the infinite square well of width L . For each (anti)symmetrized function make a plot of x_A and x_B and shade in regions of high probability density.

- b. Prove that if the two wavefunctions $\psi(\mathbf{x})$ and $\phi(\mathbf{x})$ are orthogonal, then the expectation value of $(\mathbf{x}_A - \mathbf{x}_B)^2$ for the antisymmetric combination of the two wavefunctions is greater than or equal to that for the symmetric combination.

14.16 **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. [Hint: 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.

Chapter 15

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.”

15.1 Problems

15.1 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.)

[*Note:* This problem raises deep and as yet unresolved questions about the classical limit of quantum mechanics. See D.F. Styer, “Quantum revivals versus classical periodicity in the infinite square well,” *American Journal of Physics* **69** (January 2001) 56–62.]

15.2 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, \dots, n_r evolves in time with a period of

$$\frac{h}{\text{Ry}} N^2,$$

where Ry is the Rydberg constant and the integer N is the least common multiple of n_1, n_2, \dots, n_r .

15.3 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 class, 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

$$a_0 = \frac{4\pi\epsilon_0 \hbar^2}{e^2 m}.$$

- 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. (In class, we found *one* way to perform this combination, but I want you to prove that this is the *only* way. *Hint:* Instead of the conventional basic dimensions of length, mass, and time, use the unconventional basic dimensions of length, mass, and energy.)
- Find the characteristic time τ_0 . What is its numerical value in terms of femtoseconds?
- Bonus:* Show that, in the Bohr model, the period of the innermost orbit is $2\pi\tau_0$. What is the period of the n th orbit?
- 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

$$\tilde{r} = \frac{r}{a_0} \quad \text{“lengths measured in atomic units”}$$

and

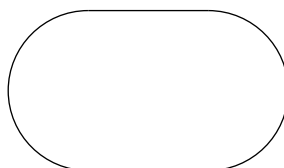
$$\tilde{t} = \frac{t}{\tau_0} \quad \text{“time measured in atomic units”}.$$

Be sure to use the dimensionless wavefunction

$$\tilde{\Psi}(\tilde{x}, \tilde{y}, \tilde{z}, \tilde{t}) = (a_0)^{3/2} \Psi(x, y, z, t).$$

15.4 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 (1984); S. Tomsovic and E.J. Heller, “Long-Time Semiclassical Dynamics of Chaos: The Stadium Billiard” *Phys. Rev. E*, **47**, 282 (1993); E.J. Heller and S. Tomsovic, “Postmodern Quantum Mechanics” *Physics Today*, **46**, 38 (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 scaled up three times.) And show that the energy eigenvalues of the larger stadium are one-ninth the energy eigenvalues of the smaller stadium.



15.5 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. *Hint:* 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.

15.6 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 . (Hint: Examine the effective potential for radial motion.) Call this ℓ value ℓ_n .

d. Define $n = \ell_n + 1$ and show that

$$\epsilon_{n,\ell} = -\frac{1}{2n^2} \quad \text{where} \quad \ell = 0, \dots, n-1.$$

(One can also continue this game to find the energy eigenfunctions.)

Chapter 16

Hydrogen

16.1 The Stark effect

This problem is based on Griffiths problems 6.36 and 6.37 and uses the notation and situation established there.

- In class we applied degenerate perturbation theory to the four $n = 2$ states, and found that all but two elements of the relevant 4×4 matrix were zero, and that those two remaining elements were equal. Find the value of these two elements.
- The “good” energy eigenstates for the $n = 2$ Stark effect are $|2, 1, +1\rangle$, $|2, 1, -1\rangle$, $(|2, 0, 0\rangle + |2, 1, 0\rangle)/\sqrt{2}$, and $(|2, 0, 0\rangle - |2, 1, 0\rangle)/\sqrt{2}$. Find the expectation value of the electric dipole moment ($\mathbf{p} = -e\mathbf{r}$) in each of these states.
- The expectation value of the dipole moment is zero in state $|2, 0, 0\rangle$ and zero in state $|2, 1, 0\rangle$, yet it is non-zero in state $(|2, 0, 0\rangle + |2, 1, 0\rangle)/\sqrt{2}$. At first glance, this might seem like a contradiction. What great principle of quantum mechanics allows this fact to escape the curse of contradiction? (Answer in one sentence.)
- (Bonus.)** Describe these four states qualitatively and explain why they are the “good” states for use in the Stark effect.
- 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. (*Hint:* Use a selection rule to determine the best order of basis elements before actually working any integrals.)
- Find the eigenvalues and degeneracies.

16.2 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 1.4) give us any assistance? Can you derive any inspiration from our proof of Kramers' relation (problem below)?

16.3 Kramers' relation

Kramers' relation states that for any energy eigenstate $\eta_{\ell m}(\vec{r})$ 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. *Hints:* 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 $\int(u(r)r^s u''(r)) dr$ in terms of $\langle r^s \rangle$, $\langle r^{s-1} \rangle$, and $\langle r^{s-2} \rangle$. Then integrate $\int(u(r)r^s u''(r)) dr$ by parts to find an integral involving $u'(r)$ as the highest derivative. Show that $\int(u(r)r^s u'(r)) dr = -[s/2]\langle r^{s-1} \rangle$ and that $\int(u'(r)r^s u'(r)) dr = -[2/(s+1)]\int(u''(r)r^{s+1}u'(r)) dr$.

- 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 17

Helium

17.1 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

$$E_g = -78.975 \text{ eV.}$$

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

$$\hat{H} = \hat{H}_A + \hat{H}_B + \hat{U}_{AB} \quad (17.1)$$

where

$$\hat{U}_{AB} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_A - \mathbf{r}_B|}. \quad (17.2)$$

The ground state wavefunction for H is

$$\eta_{100}(\mathbf{r}) = \frac{1}{\sqrt{\pi}a_0^{3/2}} e^{-r/a_0}. \quad (17.3)$$

But if the nucleus had charge $+Ze$, this would be

$$\eta_{100}(\mathbf{r}) = \frac{Z^{3/2}}{\sqrt{\pi}a_0^{3/2}} e^{-Zr/a_0}. \quad (17.4)$$

So the $\hat{U}_{AB} = 0$ ground state is

$$\eta_{100}(\mathbf{r}_A)\eta_{100}(\mathbf{r}_B) = \frac{Z^3}{\pi a_0^3} e^{-Z(r_A+r_B)/a_0} \quad \text{with} \quad Z = 2. \quad (17.5)$$

This state gives a ground state energy of $E_g = -8(\text{Ry}) = -109 \text{ eV}$.

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 $1 < Z_{\min} < 2$. And in fact minimizing $\langle H \rangle$ with over this class of trial wavefunctions gives $Z_{\min} = 1.69$ and $E_g = -77.5 \text{ eV}$. (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)

$$\psi(\mathbf{r}_A, \mathbf{r}_B) = e^{-Z(r_A+r_B)} \sum c_{nlm} (Z(r_A+r_B))^n (Z(r_A-r_B))^{2l} (Z|\mathbf{r}_A-\mathbf{r}_B|)^m.$$

[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)

$$\psi(\mathbf{r}_A, \mathbf{r}_B) = e^{-Z(r_A+r_B)} \sum c_{nlm} (Z(r_A+r_B))^n \left(Z \frac{r_A-r_B}{|\mathbf{r}_A-\mathbf{r}_B|} \right)^{2l} \left(\frac{|\mathbf{r}_A-\mathbf{r}_B|}{r_A+r_B} \right)^m.$$

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.

Chapter 18

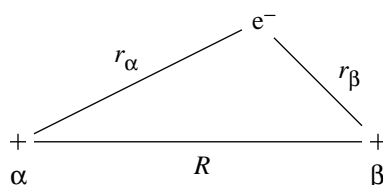
Atoms

Chapter 19

Molecules

19.1 The hydrogen molecule ion

The hydrogen molecule ion is two protons and 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

$$\hat{H} = \widehat{\text{KE}}_\alpha + \widehat{\text{KE}}_\beta + \widehat{\text{KE}}_e + \hat{U}_{\alpha\beta} + \hat{U}_{\alpha e} + \hat{U}_{\beta e} \quad (19.1)$$

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 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

$$\widehat{\text{KE}}_\alpha = 0; \quad \widehat{\text{KE}}_\beta = 0; \quad \hat{U}_{\alpha\beta} = U_{\alpha\beta} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{R}. \quad (19.2)$$

The remaining, quantum mechanical, piece of the full Hamiltonian is the electronic Hamiltonian

$$\hat{H}_e = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\epsilon_0}\left(\frac{1}{r_\alpha} + \frac{1}{r_\beta}\right). \quad (19.3)$$

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 :

$$\eta_\alpha(\vec{r}) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r_\alpha/a_0} \equiv |\alpha\rangle. \quad (19.4)$$

Similarly if nucleus α did not exist, the ground state wavefunction would be

$$\eta_\beta(\vec{r}) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r_\beta/a_0} \equiv |\beta\rangle. \quad (19.5)$$

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

$$\psi(\vec{r}) = A\eta_\alpha(\vec{r}) + B\eta_\beta(\vec{r}). \quad (19.6)$$

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

$$\psi(\vec{r}) = A_\pm[\eta_\alpha(\vec{r}) \pm \eta_\beta(\vec{r})], \quad (19.7)$$

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*

$$I(R) \equiv \langle \alpha | \beta \rangle = \int \eta_{\alpha}(\vec{r}) \eta_{\beta}(\vec{r}) d^3r. \quad (19.8)$$

In terms of this integral, we can select the normalization to be

$$A_{\pm} = \frac{1}{\sqrt{2(1 \pm I(R))}}. \quad (19.9)$$

Evaluating the electronic Hamiltonian in the trial wavefunction gives

$$\begin{aligned} \langle \hat{H}_e \rangle &= \frac{(\langle \alpha | \pm \langle \beta |) \hat{H}_e (| \alpha \rangle \pm | \beta \rangle)}{2(1 \pm I(R))} \\ &= \frac{\langle \alpha | \hat{H}_e | \alpha \rangle \pm \langle \alpha | \hat{H}_e | \beta \rangle \pm \langle \beta | \hat{H}_e | \alpha \rangle + \langle \beta | \hat{H}_e | \beta \rangle}{2(1 \pm I(R))} \\ &= \frac{\langle \alpha | \hat{H}_e | \alpha \rangle \pm \langle \beta | \hat{H}_e | \alpha \rangle}{1 \pm I(R)} \end{aligned} \quad (19.10)$$

But we have already done large parts of these two integrals:

$$\begin{aligned} \hat{H}_e | \alpha \rangle &= \left[\widehat{\text{KE}} - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{\alpha}} - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{\beta}} \right] | \alpha \rangle \\ &= \left[\widehat{\text{KE}} - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{\alpha}} \right] | \alpha \rangle - \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{\beta}} | \alpha \rangle \end{aligned}$$

$$\begin{aligned}
&= -\text{Ry} |\alpha\rangle - 2 \text{Ry} a_0 \frac{1}{r_\beta} |\alpha\rangle \\
&= -\text{Ry} \left[|\alpha\rangle + 2 \frac{a_0}{r_\beta} |\alpha\rangle \right]
\end{aligned} \tag{19.11}$$

whence

$$\langle \alpha | \hat{H}_e | \alpha \rangle = -\text{Ry} \left[1 + 2 \left\langle \alpha \left| \frac{a_0}{r_\beta} \right| \alpha \right\rangle \right] \tag{19.12}$$

$$\langle \beta | \hat{H}_e | \alpha \rangle = -\text{Ry} \left[\langle \beta | \alpha \rangle + 2 \left\langle \beta \left| \frac{a_0}{r_\beta} \right| \alpha \right\rangle \right]. \tag{19.13}$$

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*

$$D(R) \equiv \left\langle \alpha \left| \frac{a_0}{r_\beta} \right| \alpha \right\rangle \tag{19.14}$$

and the *exchange integral*

$$X(R) \equiv \left\langle \beta \left| \frac{a_0}{r_\beta} \right| \alpha \right\rangle. \tag{19.15}$$

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 expectation value of the electronic Hamiltonian in the trial wavefunction. It is

$$\begin{aligned}
\langle \hat{H}_e \rangle &= -\text{Ry} \frac{1 + 2D(R) \pm I(R) \pm 2X(R)}{1 \pm I(R)} \\
&= -\text{Ry} \left[1 + 2 \frac{D(R) \pm X(R)}{1 \pm I(R)} \right].
\end{aligned} \tag{19.16}$$

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

$$\frac{e^2}{4\pi\epsilon_0} \frac{1}{R} = 2 \text{Ry} \frac{a_0}{R}, \tag{19.17}$$

so the upper bound on the total ground state energy is

$$\text{Ry} \left[2 \frac{a_0}{R} - 1 - 2 \frac{D(R) \pm X(R)}{1 \pm I(R)} \right]. \tag{19.18}$$

19.2 Problems

19.1 The hydrogen molecule ion: Evaluation of integrals

Evaluate the direct and exchange integrals $D(R)$ and $X(R)$. (*Hint*: Remember that $\sqrt{x^2} = |x|$.) Plot as a function of R the overlap integral, $I(R)$, as well as $D(R)$ and $X(R)$.

19.2 The hydrogen molecule ion: Thinking about integrals

For the hydrogen molecule ion, find and plot the expectation 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?

19.3 Improved variational wavefunction

Everett Schlawin (‘09) suggested using “shielded” subwavefunctions like equation (17.4) in place of the subwavefunctions (19.4) and (19.5) that go into making trial wavefunction (19.7). 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 $R = 0$.

19.3 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)

$$(\text{two electrons in ground level}) \times (\text{spin singlet}) \quad (19.19)$$

while the helium atom first excited state had the form

$$(\text{one electron in ground level, one in first excited level}) \times (\text{spin triplet}). \quad (19.20)$$

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 tenuous candidate for a ground state, namely

$$\psi_+(\vec{r}) = A_+[\eta_\alpha(\vec{r}) + \eta_\beta(\vec{r})]. \quad (19.21)$$

The only other state we’ve considered is

$$\psi_-(\vec{r}) = A_+[\eta_\alpha(\vec{r}) - \eta_\beta(\vec{r})]. \quad (19.22)$$

With even less justification, we consider this to be a first excited state.

In the equations below, A and B denote the two electron windows, whereas α and β denote the two nuclei.

Chapter 20

WKB: The Quasiclassical Approximation

Chapter 21

Perturbation Theory for the Time Development Problem

21.1 Setup

Here's our problem:

Solve the initial value problem for the Hamiltonian

$$\hat{H}(t) = \hat{H}^{(0)} + \hat{H}'(t) \quad (21.1)$$

given the solution $\{|n\rangle\}$ of the unperturbed energy eigenproblem

$$\hat{H}^{(0)}|n\rangle = E_n|n\rangle. \quad (21.2)$$

Here we're thinking of $\hat{H}'(t)$ as being in some sense “small” compared to the unperturbed Hamiltonian $\hat{H}^{(0)}$. 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 $\{|n\rangle\}$:

$$|\psi(t)\rangle = \sum_n C_n(t)|n\rangle \quad \text{where} \quad C_n(t) = \langle n|\psi(t)\rangle. \quad (21.3)$$

Once we know the $C_n(t)$, we'll know the solution $|\psi(t)\rangle$. Now, the state vector evolves according to

$$\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle \quad (21.4)$$

so the expansion coefficients evolve according to

$$\begin{aligned}
\frac{dC_n(t)}{dt} &= -\frac{i}{\hbar} \langle n | \hat{H} | \psi(t) \rangle \\
&= -\frac{i}{\hbar} \sum_m \langle n | \hat{H} | m \rangle C_m(t) \\
&= -\frac{i}{\hbar} \sum_m \left[\langle n | \hat{H}^{(0)} | m \rangle + \langle n | \hat{H}' | m \rangle \right] C_m(t) \\
&= -\frac{i}{\hbar} \sum_m [E_m \delta_{m,n} + H'_{n,m}] C_m(t) \\
&= -\frac{i}{\hbar} \left[E_n C_n(t) + \sum_m H'_{n,m} C_m(t) \right]
\end{aligned} \tag{21.5}$$

This result is exact: we have yet to make any approximation.

Now, if $\hat{H}'(t)$ vanished, the solutions would be

$$C_n(t) = C_n(0) e^{-(i/\hbar)E_n t}, \tag{21.6}$$

which motivates us to define new variables $c_n(t)$ through

$$C_n(t) = c_n(t) e^{-(i/\hbar)E_n t}. \tag{21.7}$$

Because the “bulk of the time development” 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 development equation (21.5) gives

$$\frac{dc_n(t)}{dt} e^{-(i/\hbar)E_n t} + c_n(t) (-i/\hbar)E_n e^{-(i/\hbar)E_n t} = -\frac{i}{\hbar} \left[E_n c_n(t) e^{-(i/\hbar)E_n t} + \sum_m H'_{n,m} c_m(t) e^{-(i/\hbar)E_m t} \right] \tag{21.8}$$

or

$$\frac{dc_n(t)}{dt} = -\frac{i}{\hbar} \sum_m H'_{n,m} c_m(t) e^{+(i/\hbar)(E_n - E_m)t}. \tag{21.9}$$

Once again, this equation is exact. Its formal solution, given the initial values $c_n(0)$, is

$$c_n(t) = c_n(0) - \frac{i}{\hbar} \sum_m \int_0^t H'_{n,m}(t') c_m(t') e^{+(i/\hbar)(E_n - E_m)t'} dt'. \tag{21.10}$$

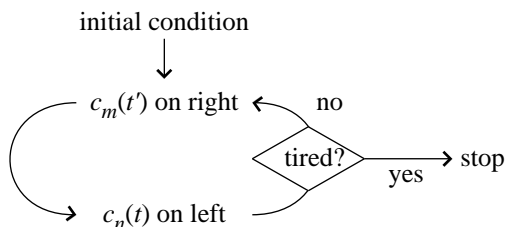
This set of equations (one for each basis element) is exact, but at first glance seems useless. The unknown quantities $c_n(t)$ are present on the left, but also the right-hand sides.

We can 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 (21.10) 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 (21.10), 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 (21.10), 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) 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.

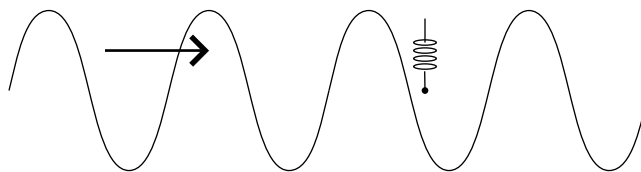
In particular, if the initial state is some eigenstate $|a\rangle$ of the unperturbed Hamiltonian $\hat{H}^{(0)}$, then to first order

$$\begin{aligned}
 c_n(t) &= -\frac{i}{\hbar} \int_0^t H'_{n,a}(t') e^{+(i/\hbar)(E_n - E_a)t'} dt' & \text{for } n \neq a \\
 c_a(t) &= 1 - \frac{i}{\hbar} \int_0^t H'_{a,a}(t') dt'
 \end{aligned}
 \tag{21.11}$$

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*

$$P_{a \rightarrow b}(t) = |C_b(t)|^2 = |c_b(t)|^2.
 \tag{21.12}$$

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 $|b\rangle$?



Solution part A — *What is the Hamiltonian?* If it were a classical particle of charge $-e$ exposed to electric field $E_0 \cos \omega t$, it would experience a force $-eE_0 \cos \omega t$ and hence have a potential energy of $eE_0 x \cos \omega t$. (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

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega_0^2}{2}\hat{x}^2 + eE_0\hat{x} \cos \omega t. \quad (21.13)$$

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

$$H_{n,0}(t) = \langle n | \hat{H}'(t) | 0 \rangle = eE_0 \cos \omega t \langle n | \hat{x} | 0 \rangle = eE_0 \cos \omega t \sqrt{\frac{\hbar}{2m\omega_0}} \delta_{n,1}. \quad (21.14)$$

(Remember your raising and lowering operators! See equation (D.31).) Invoking equations (21.11), we obtain

$$c_n(t) = 0 \quad \text{for } n \neq 0, 1 \quad (21.15)$$

$$c_1(t) = -\frac{i}{\hbar} eE_0 \sqrt{\frac{\hbar}{2m\omega_0}} \int_0^t \cos \omega t' e^{i\omega_0 t'} dt' \quad (21.16)$$

$$c_0(t) = 1 \quad (21.17)$$

We will eventually need to perform the time integral in equation (21.16), 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 (21.16) will be evaluated at equation (21.29). For now, let’s just call it $I(t)$. In terms of this integral, the transition probabilities are

$$P_{0 \rightarrow b}(t) = 0 \quad \text{for } b \neq 0, 1 \quad (21.18)$$

$$P_{0 \rightarrow 1}(t) = \frac{e^2 E_0^2}{2m\hbar\omega_0} I(t) I^*(t) \quad (21.19)$$

$$P_{0 \rightarrow 0}(t) = 1 \quad (21.20)$$

21.2 Fermi's golden rule

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

$$\vec{E}(\vec{r}, t) = E_0 \hat{k} \cos(\vec{k} \cdot \vec{r} - \omega t). \quad (21.21)$$

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

$$\vec{F}(t) = -eE_0 \hat{k} \cos(\omega t), \quad (21.22)$$

so the associated potential energy is

$$U(t) = eE_0 z \cos(\omega t). \quad (21.23)$$

Turning this classical potential energy into a quantal operator gives

$$\hat{H}'(t) = eE_0 \hat{z} \cos(\omega t). \quad (21.24)$$

(Note that the hat \hat{k} in equation (21.22) signifies unit vector, whereas the hat \hat{z} in equation (21.24) 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 (21.11). (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,

$$H'_{a,a}(t) = \langle a | H'(t) | a \rangle = eE_0 \langle a | \hat{z} | a \rangle \cos(\omega t) = 0, \quad (21.25)$$

whence $c_a(t) = 1$ and $P_{a \rightarrow a} = 1$. Most of the atoms *don't* make transitions.

But what about those that do? For these we need to find the matrix elements

$$H'_{b,a}(t) = \langle b|H'(t)|a\rangle = eE_0\langle b|\hat{z}|a\rangle \cos(\omega t). \quad (21.26)$$

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

$$c_b(t) = -\frac{i}{\hbar}eE_0\langle b|\hat{z}|a\rangle \int_0^t \cos(\omega t')e^{+(i/\hbar)(E_b-E_a)t'} dt'. \quad (21.27)$$

It is convenient (and conventional!) to follow the lead of Einstein's $\Delta E = \hbar\omega$ and define

$$E_b - E_a = \hbar\omega_0. \quad (21.28)$$

The time integral is then

$$\begin{aligned} \int_0^t \cos(\omega t')e^{i\omega_0 t'} dt' &= \int_0^t \frac{e^{+i\omega t'} + e^{-i\omega t'}}{2} e^{i\omega_0 t'} dt' \\ &= \frac{1}{2} \left[\int_0^t e^{i(\omega_0+\omega)t'} dt' + \int_0^t e^{i(\omega_0-\omega)t'} dt' \right] \\ &= \frac{1}{2} \left[\frac{e^{i(\omega_0+\omega)t'} - 1}{i(\omega_0+\omega)} + \frac{e^{i(\omega_0-\omega)t'} - 1}{i(\omega_0-\omega)} \right] \\ &= -\frac{i}{2} \left[\frac{e^{i(\omega_0+\omega)t} - 1}{\omega_0+\omega} + \frac{e^{i(\omega_0-\omega)t} - 1}{\omega_0-\omega} \right] \\ &= -\frac{i}{2} \left[e^{i(\omega_0+\omega)t/2} \frac{e^{i(\omega_0+\omega)t/2} - e^{-i(\omega_0+\omega)t/2}}{\omega_0+\omega} + e^{i(\omega_0-\omega)t/2} \frac{e^{i(\omega_0-\omega)t/2} - e^{-i(\omega_0-\omega)t/2}}{\omega_0-\omega} \right] \\ &= -\frac{i}{2} \left[e^{i(\omega_0+\omega)t/2} \frac{2i \sin((\omega_0+\omega)t/2)}{\omega_0+\omega} + e^{i(\omega_0-\omega)t/2} \frac{2i \sin((\omega_0-\omega)t/2)}{\omega_0-\omega} \right] \\ &= e^{i(\omega_0+\omega)t/2} \frac{\sin((\omega_0+\omega)t/2)}{\omega_0+\omega} + e^{i(\omega_0-\omega)t/2} \frac{\sin((\omega_0-\omega)t/2)}{\omega_0-\omega}. \end{aligned} \quad (21.29)$$

Enrico Fermi thought about this integral and realized that most of the time it would not be substantial. The only time it's 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

$$c_b(t) = -\frac{i}{\hbar}eE_0\langle b|\hat{z}|a\rangle e^{i(\omega_0-\omega)t/2} \frac{\sin((\omega_0-\omega)t/2)}{\omega_0-\omega}. \quad (21.30)$$

The transition probability is then

$$P_{a \rightarrow b} = \frac{e^2 E_0^2 |\langle b|\hat{z}|a\rangle|^2 \sin^2((\omega_0-\omega)t/2)}{\hbar^2 (\omega_0-\omega)^2}. \quad (21.31)$$

This rule, like all rules, 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."

21.3 Problems

21.1 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.

21.2 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)

$$H'(x; t) = A_0 \sin\left(\frac{2\pi x}{L}\right) \delta(t), \quad (21.32)$$

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?

21.3 Second-order time-dependent perturbation theory

In lecture 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 F_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:

$$c_0^{(2)}(t) = 1 - \frac{i}{\hbar} \int_0^t H'_{01}(t') e^{-i\omega t'} c_1^{(1)}(t') dt', \quad (21.33)$$

$$c_1^{(2)}(t) = -\frac{i}{\hbar} \int_0^t H'_{10}(t') e^{+i\omega t'} c_0^{(1)}(t') dt', \quad (21.34)$$

$$c_2^{(2)}(t) = -\frac{i}{\hbar} \int_0^t H'_{21}(t') e^{+i\omega t'} c_1^{(1)}(t') dt', \quad (21.35)$$

where

$$H'_{01}(t) = H'_{10}(t) = F_0 \sqrt{\frac{\hbar}{2m\omega}} \cos(\omega' t), \quad (21.36)$$

and

$$H'_{21}(t) = F_0 \sqrt{\frac{2\hbar}{2m\omega}} \cos(\omega' t). \quad (21.37)$$

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 $(F_0)^2$ (where F_0 is in some sense “small”), and that $c_1^{(2)}(t) = c_1^{(1)}(t)$.

Chapter 22

The Interaction of Matter and Radiation

Chapter 23

Quantization of the Electromagnetic Field

23.1 Overview

The classical electromagnetic field is a pair of vector functions, $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$, that evolve in time according to the Maxwell equations

$$\vec{\nabla} \cdot \vec{E}(\vec{r}, t) = \rho(\vec{r}, t)/\epsilon_0 \quad (23.1)$$

$$\vec{\nabla} \cdot \vec{B}(\vec{r}, t) = 0 \quad (23.2)$$

$$\vec{\nabla} \times \vec{E}(\vec{r}, t) = -\frac{\partial \vec{B}(\vec{r}, t)}{\partial t} \quad (23.3)$$

$$\vec{\nabla} \times \vec{B}(\vec{r}, t) = \mu_0 \vec{J}(\vec{r}, t) + \mu_0 \epsilon_0 \frac{\partial \vec{E}(\vec{r}, t)}{\partial t} \quad (23.4)$$

The energy of the classical EM field interacting with charged particles is

$$H(t) = \frac{1}{2} \int \left(\epsilon_0 \vec{E}^2(\vec{r}, t) + \frac{1}{\mu_0} \vec{B}^2(\vec{r}, t) \right) d^3r + \text{Hamiltonian of particles.} \quad (23.5)$$

This chapter considers the “free field” case, in which there are no particles, no charge, and no current, so the only source of $\vec{E}(\vec{r}, t)$ is the change in $\vec{B}(\vec{r}, t)$, and the only source of $\vec{B}(\vec{r}, t)$ is the change in $\vec{E}(\vec{r}, t)$.

In this case the (classical) EM field Hamiltonian already looks much like the (classical) simple harmonic oscillator Hamiltonian

$$H = \frac{1}{2m} p^2(t) + \frac{m\omega^2}{2} x^2(t)$$

in that it's the sum of two quadratic functions. The EM field is more complicated in that there's “an infinite number of oscillators”, one at each point in space, and you have to sum (integrate) over all of them to

find the total energy. This analogy is correct but not immediately useful, because the analogous oscillators adjacent in space are coupled: if there is EM field at one point only, it will spread out into adjacent points.

In this chapter we first look at the classical EM field to represent it in terms of oscillators that *aren't* coupled: The answer turns out to involve oscillators in \vec{k} -space rather than in \vec{r} -space. Once we've represented the EM field as a sum of oscillators it's easy to write a quantum field theory of the EM field, because we already know so much about the quantum properties of oscillators.

23.2 Classical electromagnetism

Free field solutions

One solution of the free field Maxwell equations is the linearly polarized plane wave

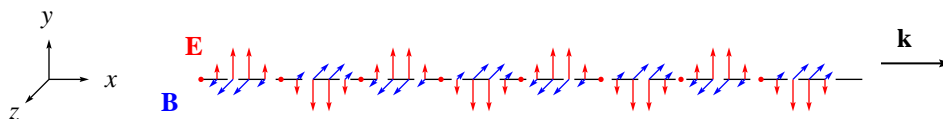
$$\vec{E}(\vec{r}, t) = E_0 \hat{y} \sin(\vec{k} \cdot \vec{r} - \omega t) \quad (23.6)$$

$$\vec{B}(\vec{r}, t) = B_0 \hat{z} \sin(\vec{k} \cdot \vec{r} - \omega t) \quad (23.7)$$

where

$$E_0 \text{ and } \vec{k} \text{ are arbitrary, but } \omega = kc \text{ and } B_0 = E_0/c.$$

This plane wave has the familiar snapshot visualization below. (You should be cautious of this visualization — and of all others as well! For example, the visualization suggests strongly that the wave is restricted to the x -axis, whereas in fact it spreads out over all space ... but if I represented that, the picture would be completely covered with ink! You're supposed to multiply the image in your mind and realize that this is going on at all lines pointing to the right, not just the x -axis. You're also supposed to reject that strong suggestion that the light travels in wavy lines.)



Of course, any superposition of plane waves also solves Maxwell's equations, and these superpositions give rise to the extraordinary variety of free-field electromagnetic phenomena: circular and elliptical polarization; standing waves (optical cavities); maroon, white, and all other colors not part of rainbow; beams of finite width (such as the beams from a flashlight or a laser); pulses; circular waves.

It is not obvious, but nevertheless true, that in fact *every* free-field solution is a superposition of plane waves. This is a consequence of Fourier's theorem.

Fourier analysis

We'll use the conventions

$$F(\vec{r}) = \int f(\vec{k}) e^{+i\vec{k}\cdot\vec{r}} d^3k \quad (23.8)$$

$$f(\vec{k}) = \int F(\vec{r}) e^{-i\vec{k}\cdot\vec{r}} \frac{d^3r}{(2\pi)^3} \quad (23.9)$$

where both \vec{r} and \vec{k} integrals range over all space. One consequence is that if $F(\vec{r})$ is real, then $f^*(\vec{k}) = f(-\vec{k})$. Another (Parseval's theorem) is that

$$\int F^2(\vec{r}) d^3r = \int f(\vec{k}) f(-\vec{k}) \frac{d^3k}{(2\pi)^3}. \quad (23.10)$$

Potentials

The fields $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$ encode the electromagnetic information through six functions of space and time (three components of electric field, three components of magnetic field). The same information is encoded more compactly, through four functions of space and time, in the scalar potential $\phi(\vec{r}, t)$ and the vector potential $\vec{A}(\vec{r}, t)$. These functions satisfy

$$\vec{E}(\vec{r}, t) = -\vec{\nabla}\phi(\vec{r}, t) - \frac{\partial\vec{A}(\vec{r}, t)}{\partial t} \quad (23.11)$$

$$\vec{B}(\vec{r}, t) = \vec{\nabla} \times \vec{A}(\vec{r}, t). \quad (23.12)$$

For given fields $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$ there are many possible potentials $\phi(\vec{r}, t)$ and $\vec{A}(\vec{r}, t)$ that satisfy these equations. For historical reasons, each choice is called a "gauge". In the free-field situation, the easiest choice is the "Coulomb gauge" in which

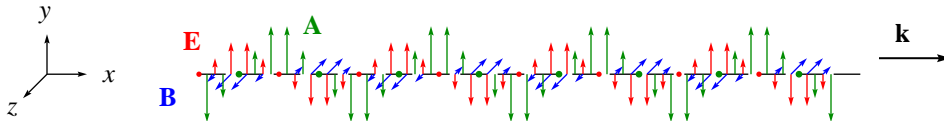
$$\vec{\nabla} \cdot \vec{A}(\vec{r}, t) = 0 \text{ and } \phi(\vec{r}, t) = 0. \quad (23.13)$$

The Maxwell equations for a free field in the Coulomb gauge are equivalent to

$$\vec{\nabla}^2 \vec{A}(\vec{r}, t) - \mu_0 \epsilon_0 \frac{\partial^2 \vec{A}(\vec{r}, t)}{\partial t^2} = 0. \quad (23.14)$$

For the linearly polarized plane wave shown above, the vector potential in the Coulomb gauge is

$$\vec{A}(\vec{r}, t) = -\frac{E_0}{\omega} \hat{y} \cos(\vec{k} \cdot \vec{r} - \omega t). \quad (23.15)$$



Fourier transform of the vector potential

Let

$$\vec{A}(\vec{r}, t) = \int \vec{a}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \frac{d^3k}{(2\pi)^3}. \quad (23.16)$$

What can we say about $\vec{a}(\vec{k}, t)$ in the Coulomb gauge?

1. We know

$$\begin{aligned} 0 &= \vec{\nabla} \cdot \vec{A}(\vec{r}, t) \\ &= \int \vec{\nabla} \cdot \left(\vec{a}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \right) \frac{d^3k}{(2\pi)^3} \\ &= \int i\vec{k} \cdot \vec{a}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \frac{d^3k}{(2\pi)^3} \end{aligned}$$

and so, for all \vec{k} ,

$$0 = \vec{k} \cdot \vec{a}(\vec{k}, t) \quad (23.17)$$

(because Fourier series are unique). That is, $\vec{a}(\vec{k}, t)$ is perpendicular to \vec{k} .

2. **Electric field.** The electric field is

$$\begin{aligned} \vec{E}(\vec{r}, t) &= -\frac{\partial \vec{A}(\vec{r}, t)}{\partial t} \\ &= -\int \dot{\vec{a}}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \frac{d^3k}{(2\pi)^3}. \end{aligned}$$

Thus the Fourier transform of the electric field is

$$\vec{e}(\vec{k}, t) = -\dot{\vec{a}}(\vec{k}, t). \quad (23.18)$$

3. **Magnetic field.** The magnetic field is

$$\begin{aligned} \vec{B}(\vec{r}, t) &= \vec{\nabla} \times \vec{A}(\vec{r}, t) \\ &= \int \vec{\nabla} \times \left(\vec{a}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \right) \frac{d^3k}{(2\pi)^3} \\ &= \int i\vec{k} \times \vec{a}(\vec{k}, t) e^{+i\vec{k}\cdot\vec{r}} \frac{d^3k}{(2\pi)^3}. \end{aligned}$$

Thus the Fourier transform of the magnetic field is

$$\vec{b}(\vec{k}, t) = i\vec{k} \times \vec{a}(\vec{k}, t). \quad (23.19)$$

4. **Maxwell equations.** The Maxwell equations, applied to $\vec{A}(\vec{r}, t)$ in the Coulomb gauge, result in

$$0 = \vec{\nabla}^2 \vec{A}(\vec{r}, t) - \mu_0 \epsilon_0 \frac{\partial^2 \vec{A}(\vec{r}, t)}{\partial t^2}$$

$$\begin{aligned}
&= \nabla^2 \vec{A}(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 \vec{A}(\vec{r}, t)}{\partial t^2} \\
&= \int \left[\vec{a}(\vec{k}, t) \nabla^2 e^{+i\vec{k}\cdot\vec{r}} - \frac{1}{c^2} \frac{d^2 \vec{a}(\vec{k}, t)}{dt^2} e^{+i\vec{k}\cdot\vec{r}} \right] \frac{d^3 k}{(2\pi)^3} \\
&= \int \left[\vec{a}(\vec{k}, t) (-k^2) - \frac{1}{c^2} \ddot{\vec{a}}(\vec{k}, t) \right] e^{+i\vec{k}\cdot\vec{r}} \frac{d^3 k}{(2\pi)^3}.
\end{aligned}$$

Thus, for all \vec{k} ,

$$\ddot{\vec{a}}(\vec{k}, t) = -(kc)^2 \vec{a}(\vec{k}, t). \quad (23.20)$$

The general solution is

$$\vec{a}(\vec{k}, t) = \vec{a}(\vec{k}, 0) e^{-i\omega t}. \quad (23.21)$$

Combining this with equation (23.18) results in

$$\vec{e}(\vec{k}, t) = i\omega \vec{a}(\vec{k}, t). \quad (23.22)$$

Hamiltonian in terms of the Fourier transform of vector potential

Using the property (23.10), the Hamiltonian becomes

$$\begin{aligned}
H(t) &= \frac{1}{2} \int \left(\epsilon_0 \vec{E}^2(\vec{r}, t) + \frac{1}{\mu_0} \vec{B}^2(\vec{r}, t) \right) d^3 r \\
&= \frac{1}{2} \int \left(\epsilon_0 \vec{e}(\vec{k}, t) \cdot \vec{e}(-\vec{k}, t) + \frac{1}{\mu_0} \vec{b}(\vec{k}, t) \cdot \vec{b}(-\vec{k}, t) \right) \frac{d^3 k}{(2\pi)^3} \\
&= \frac{1}{2} \int \left(\epsilon_0 \vec{e}(\vec{k}, t) \cdot \vec{e}^*(\vec{k}, t) + \frac{1}{\mu_0} \vec{b}(\vec{k}, t) \cdot \vec{b}^*(\vec{k}, t) \right) \frac{d^3 k}{(2\pi)^3}
\end{aligned}$$

where in the last step we have used the fact that $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$ are real.

Now, using first equation (23.22) and then equation (23.21),

$$\vec{e}(\vec{k}, t) \cdot \vec{e}^*(\vec{k}, t) = \omega^2 \vec{a}(\vec{k}, t) \cdot \vec{a}^*(\vec{k}, t) = \omega^2 \vec{a}(\vec{k}, 0) \cdot \vec{a}^*(\vec{k}, 0).$$

Similarly, using both equation (23.19) and the transverse property (23.17),

$$\vec{b}(\vec{k}, t) \cdot \vec{b}^*(\vec{k}, t) = k^2 \vec{a}(\vec{k}, t) \cdot \vec{a}^*(\vec{k}, t) = k^2 \vec{a}(\vec{k}, 0) \cdot \vec{a}^*(\vec{k}, 0).$$

Thus

$$H(t) = \frac{1}{2} \int \left(\epsilon_0 \omega^2 + \frac{k^2}{\mu_0} \right) \left(\vec{a}(\vec{k}, 0) \cdot \vec{a}^*(\vec{k}, 0) \right) \frac{d^3 k}{(2\pi)^3} \quad (23.23)$$

and we find what we've suspected all along, namely that $H(t)$ is independent of time. We write $\vec{a}(\vec{k}, t) = \vec{a}_{\mathbf{k}}$ and allow the symbol $\vec{a}_{\mathbf{k}}$ to either be a function of time or a constant, depending on context. A little manipulation shows that $\epsilon_0 \omega^2 + k^2/\mu_0 = 2k^2/\mu_0$, whence

$$H = \frac{1}{\mu_0} \int k^2 \vec{a}_{\mathbf{k}} \cdot \vec{a}_{\mathbf{k}}^* \frac{d^3 k}{(2\pi)^3} = \frac{1}{\mu_0} \sum_{\alpha} \int k^2 a_{\mathbf{k},\alpha} a_{\mathbf{k},\alpha}^* \frac{d^3 k}{(2\pi)^3} \quad (23.24)$$

where the index α stands for one of the two polarization directions. This is our formula for the field Hamiltonian as an integral over all \vec{k} and sum over each of the two polarization directions at each \vec{k} . We call this a “sum over modes”.

Each mode as a simple harmonic oscillator

The Hamiltonian is the sum over all modes of the energy contribution from each mode, namely

$$\frac{1}{\mu_0} k^2 a_{\mathbf{k},\alpha} a_{\mathbf{k},\alpha}^* \quad (23.25)$$

Our goal is to write this energy in the form of a simple harmonic oscillator. Since we’re dealing with only one particular mode, I’ll drop the elaborate subscript and write our desired form as

$$\frac{1}{\mu_0} k^2 a(t) a^*(t) = \frac{1}{2} [p^2(t) + \omega^2 x^2(t)] \quad (23.26)$$

where $x(t)$ and $p(t)$ are real and $\dot{x}(t) = p(t)$. This is the Hamiltonian for a simple harmonic oscillator of mass 1.

There is more than one way to define $x(t)$ and $p(t)$ that satisfy these conditions. My favorite choice starts with

$$x(t) = D (a(t) + a^*(t)),$$

where D is to be determined. Then (remembering that $\dot{a}(t) = -i\omega a(t)$)

$$p(t) = \dot{x}(t) = -i\omega D (a(t) - a^*(t))$$

so that

$$\begin{aligned} \frac{1}{\mu_0} k^2 a(t) a^*(t) &= \frac{1}{2} [-\omega^2 D^2 (a^2 - 2aa^* + a^{*2}) + \omega^2 D^2 (a^2 + 2aa^* + a^{*2})] \\ &= \frac{1}{2} [4\omega^2 D^2 aa^*] \\ &= 2k^2 c^2 D^2 aa^*. \end{aligned}$$

This choice works when $D = 1/\sqrt{2\mu_0 c^2}$.

We’ve done it. We’ve written the Maxwell equations for the free electromagnetic field in Hamiltonian form, and the Hamiltonian turns out to be nothing but a sum of simple harmonic oscillators, one for each mode (“mode” meaning a given one of the two polarizations at a given wavevector \vec{k}). The correspondence between the vector potential strength $a(t)$ of a given mode and the position and momentum of this analogous simple harmonic oscillator is given through

$$x(t) = \frac{1}{\sqrt{2\mu_0 c^2}} (a(t) + a^*(t)) \quad (23.27)$$

$$p(t) = -i \frac{\omega}{\sqrt{2\mu_0 c^2}} (a(t) - a^*(t)) \quad (23.28)$$

$$a(t) = \frac{1}{2} \sqrt{2\mu_0 c^2} \left(x(t) + \frac{i}{\omega} p(t) \right). \quad (23.29)$$

For completeness, we add that the field amplitudes in this mode are

$$e(t) = i\omega a(t) \quad (23.30)$$

$$b(t) = ika(t) = e(t)/c. \quad (23.31)$$

Everything we've done in this section has been classical. In fact, what we've done is just like solving a coupled oscillator problem in classical mechanics: We've found the independent normal modes of the coupled system, each one of which behaves like a single harmonic oscillator. *Any* time development of the whole system can be expressed as a sum over (superposition of) the normal modes. You will recall from your classical mechanics course that sometimes it's not obvious, looking at the motion of the whole system, that this motion can be expressed as nothing more than the sum of simple harmonic oscillators, but it's true whether obvious or not.

23.3 Quantal electromagnetism — What is a photon?

Once classical electrodynamics has been cast into this “sum over oscillators” form, it's easy to see how to invent quantum electrodynamics: Change the functions $x(t)$ and $p(t)$ into operators \hat{x} and \hat{p} , subject to the commutation rule $[\hat{x}, \hat{p}] = i\hbar$. Here are a few notable features of this quantization:

The traditional SHO lowering operator is

$$\bar{a} = \sqrt{\frac{\omega}{2\hbar}} \left(\hat{x} + \frac{i}{\omega} \hat{p} \right),$$

so this operator corresponds to the vector potential amplitude function $a(t)$ through

$$\bar{a} \iff \sqrt{\frac{\omega}{\hbar\mu_0 c^2}} a(t).$$

The classical mode Hamiltonian

$$\frac{1}{\mu_0} k^2 a(t) a^*(t)$$

thus corresponds to

$$\hbar\omega \bar{a} \bar{a}^\dagger.$$

(I'm not sure how the traditional $\frac{1}{2}\hbar\omega$ ground state energy fits into this prescription.) And the field amplitude operators are

$$\begin{aligned} \hat{e} &= i\sqrt{\hbar\omega} \sqrt{\mu_0 c^2} \bar{a} \\ \hat{b} &= \hat{e}/c \end{aligned}$$

The outline is clear: States with definite energy are eigenstates of $\bar{a} \bar{a}^\dagger$. States with definite field amplitude are eigenstates of \bar{a} . These operators don't commute:

$$[\bar{a}, \bar{a} \bar{a}^\dagger] = \bar{a}[\bar{a}, \bar{a}^\dagger] + [\bar{a}, \bar{a}] \bar{a}^\dagger = \bar{a},$$

so you can have an energy, or you can have a field strength, but you can't have both.

To say “the \mathbf{k}, α mode is in energy eigenstate n ” is pronounced “there are n photons of mode \mathbf{k}, α ”. Don't think that photons are hard round marbles. Don't think that photons are wavepackets of EM field.

Remember from classical mechanics that it can be hard to visualize the motion of a set of coupled oscillators as a superposition of normal modes. And remember from earlier in quantum mechanics that it can be hard to visualize the time development of a quantal state in a simple harmonic oscillator as a superposition of energy states. For the quantized EM field, we must do both at once, so it's no surprise that photons are hard to visualize.

There exist states of the EM field that have definite position (the “hard round marble” idea) — but these states do not have energy at all, much less energy $E = \hbar\omega$. There exist states of the EM field that have definite electric and magnetic field — but these states do not have either energy or position.

When he was my teacher, I heard Kurt Gottfried call this principle of “energy or field strength, but not both” the most astounding, “in your face”, instance of complementarity that he knew of. His written comment is a bit more reserved (pages 14–15): “It is impossible to determine the field strengths and the number of photons simultaneously. This shows the mutual exclusiveness of the photon and field-strength descriptions of electromagnetic phenomena and is a very illuminating example of Bohr's principle of complementarity.”

Chapter 24

The Vistas Open to Us

How to extend what's in this book:

- Relativistic quantum mechanics. (Don't make t an operator, instead turn x back to a variable and introduce creation and annihilation operators.)
- Quantum field theory.
- Quantal chaos and the classical limit of quantum mechanics.
- Friction and decay to ground state.
- Atomic, molecular, and solid state physics.

All of these fall solidly within the amplitude framework!

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” is 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.

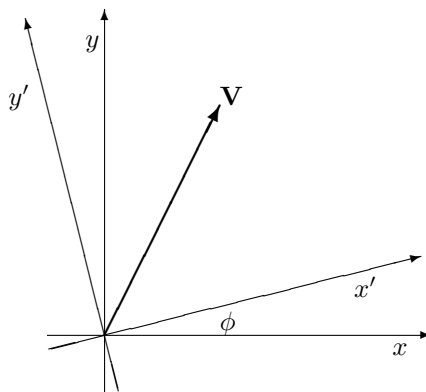
A.2 Vectors in two dimensions

One way of describing a two-dimensional vector \mathbf{V} is by giving its x and y components in the form of a 2×1 column matrix

$$\begin{pmatrix} V_x \\ V_y \end{pmatrix}. \quad (\text{A.1})$$

Indeed it is sometimes said that the vector \mathbf{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

$$\begin{pmatrix} V_{x'} \\ V_{y'} \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} V_x \\ V_y \end{pmatrix}. \quad (\text{A.2})$$



The 2×2 matrix above is called the “rotation matrix” and is usually denoted by $R(\phi)$:

$$R(\phi) \equiv \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix}. \quad (\text{A.3})$$

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*.¹ You could prove this by working a matrix multiplication,

¹Although all rotation matrices are orthogonal, there are orthogonal matrices that are not rotation matrices: see problem A.4.

but it is easier to simply realize that the inverse of a rotation by ϕ is simply a rotation by $-\phi$, and noting that

$$\mathbf{R}^{-1}(\phi) = \mathbf{R}(-\phi) = \mathbf{R}^\dagger(\phi). \quad (\text{A.4})$$

(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

$$\begin{pmatrix} V \\ 0 \end{pmatrix}, \quad (\text{A.5})$$

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 $\phi = 90^\circ$, $\phi = 180^\circ$, $\phi = 270^\circ$, and $\phi = 360^\circ$.

A.2 Problem: The rotation matrix

- Derive equation (A.2) through purely geometrical arguments.
- 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

$$V_{x'} = \mathbf{V} \cdot \hat{\mathbf{i}}' \quad \text{and} \quad V_{y'} = \mathbf{V} \cdot \hat{\mathbf{j}}' \quad (\text{A.6})$$

to derive equation (A.2).

- 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

$$\tan \phi = \frac{V_y}{V_x}. \quad (\text{A.7})$$

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.3 Tensors in two dimensions

A tensor, like a vector, is a geometrical entity that may be described 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 located² at (x, y) has components

$$\mathbf{T} = \begin{pmatrix} my^2 & -mxy \\ -mxy & mx^2 \end{pmatrix}. \quad (\text{A.8})$$

(Note the distinction between the tensor \mathbf{T} and its matrix of components \mathbf{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 200, the tensor components are of course

$$\mathbf{T}' = \begin{pmatrix} my'^2 & -mx'y' \\ -mx'y' & mx'^2 \end{pmatrix}. \quad (\text{A.9})$$

A little calculation shows that the components of the inertia tensor in two different coordinate systems are related through

$$\mathbf{T}' = \mathbf{R}(\phi)\mathbf{T}\mathbf{R}^{-1}(\phi). \quad (\text{A.10})$$

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.6). Therefore the symmetry is a property of the tensor, not of its

²Or, to be absolutely precise, the particle located at the point represented by the vector with components (x, y) .

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”.³ 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.6 Problem: Representations of symmetric tensors*

Show that if the matrix \mathbf{S} representing a tensor is symmetric, and if \mathbf{B} is any orthogonal matrix, then all of the representations

$$\mathbf{B}\mathbf{S}\mathbf{B}^\dagger \tag{A.11}$$

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 to prove the result in any number of dimensions, and for any orthogonal matrix \mathbf{B} , not just rotation matrices. This more general problem is considerably easier to solve.)

A.7 Problem: Diagonal inertia tensor

The matrix (A.8) 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 at the particle, one in which the y' -axis points directly at the particle, one in which the x' -axis points directly away from the particle, and one in which the y' -axis points directly away from the particle. Find the matrix representation in each of these four coordinate systems.

A.8 Problem: Representations of a certain tensor

Show that a tensor represented in one coordinate system by a diagonal matrix with equal elements, namely

$$\begin{pmatrix} d_0 & 0 \\ 0 & d_0 \end{pmatrix}, \tag{A.12}$$

has the same representation in *all* orthogonal coordinate systems.

³An 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.

A.9 Problem: Rotation to the preferred coordinate system*

A tensor is represented in the initial coordinate system by

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}. \quad (\text{A.13})$$

Show that the tensor is diagonal in a preferred coordinate system which is rotated from the initial system by an angle ϕ with

$$\tan(2\phi) = \frac{2b}{a-c}. \quad (\text{A.14})$$

This equation has four solutions. Find the rotation matrix for $\phi = 90^\circ$, 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

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}. \quad (\text{A.15})$$

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

$$\mathbf{T}' = \mathbf{R}\mathbf{T}\mathbf{R}^\dagger, \quad (\text{A.16})$$

where \mathbf{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

$$\mathbf{T}' = \mathbf{R}\mathbf{T}\mathbf{R}^\dagger, \quad (\text{A.17})$$

where \mathbf{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

$$A_d = d - 1 + A_{d-1} = d(d-1)/2. \quad (\text{A.18})$$

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.10 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

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (\text{A.19})$$

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 \mathbf{B} such that

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \mathbf{B} \begin{pmatrix} u \\ v \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{B} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (\text{A.20})$$

(For example, if the new coordinate system is the primed coordinate system of the figure on page 200, then the matrix \mathbf{B} that translates from the original to the new coordinates is the rotation matrix $\mathbf{R}(\phi)$.) Given this “translation dictionary”, we have

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \mathbf{B} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (\text{A.21})$$

But \mathbf{B} is invertible, so

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{B}^{-1} \begin{pmatrix} x' \\ y' \end{pmatrix} \quad (\text{A.22})$$

whence

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \mathbf{B} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mathbf{B}^{-1} \begin{pmatrix} x' \\ y' \end{pmatrix}. \quad (\text{A.23})$$

Thus the representation of the transformation in the primed coordinate system is

$$\mathbf{B} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mathbf{B}^{-1} \quad (\text{A.24})$$

(compare equation A.10). This equation has a very direct physical meaning. Remember that the matrix \mathbf{B} translates from the old (x, y) coordinates to the new (x', y') coordinates, while the matrix \mathbf{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

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (\text{A.25})$$

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:

$$x^2 + y^2 = 1. \quad (\text{A.26})$$

A brief calculation shows that the length of the output vector is then

$$\sqrt{u^2 + v^2} = \sqrt{a^2x^2 + b^2 + c^2y^2 + 2b(a+c)xy}, \quad (\text{A.27})$$

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

$$\tan \varphi = \frac{b + c \tan \theta}{a + b \tan \theta}, \quad (\text{A.28})$$

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,

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} d_1x' \\ d_2y' \end{pmatrix}. \quad (\text{A.29})$$

The unit circle is still

$$x'^2 + y'^2 = 1, \quad (\text{A.30})$$

so the image of the unit circle is

$$\left(\frac{u'}{d_1}\right)^2 + \left(\frac{v'}{d_2}\right)^2 = 1, \quad (\text{A.31})$$

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

$$\begin{pmatrix} d_1 \\ 0 \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (\text{A.32})$$

In other words, 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 \mathbf{B} , then the output vector

$$\mathbf{x}' = \mathbf{B}\mathbf{x} \tag{A.33}$$

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

$$\mathbf{B}\mathbf{x} = \lambda\mathbf{x}, \tag{A.34}$$

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

$$\mathbf{B}\mathbf{0} = \lambda\mathbf{0}, \tag{A.35}$$

for any value of λ for any linear transformation. On the other hand, if

$$\mathbf{B}\mathbf{x} = 0\mathbf{x} = \mathbf{0} \tag{A.36}$$

for some non-zero vector \mathbf{x} , then \mathbf{x} is an eigenvector with eigenvalue $\lambda = 0$.

A.11 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,

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}, \tag{A.37}$$

a simple rotation of axes would produce a new coordinate system in which the matrix representation is diagonal:

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}. \quad (\text{A.38})$$

These two matrices are related through

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} = \mathbf{R}(\phi) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \mathbf{R}^{-1}(\phi), \quad (\text{A.39})$$

where $\mathbf{R}(\phi)$ is the rotation matrix (A.3). Problem A.9 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

$$\begin{pmatrix} 5 \\ 0 \end{pmatrix}. \quad (\text{A.40})$$

Multiplying this vector by the matrix in question gives

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} 5 \\ 0 \end{pmatrix} = d_1 \begin{pmatrix} 5 \\ 0 \end{pmatrix}, \quad (\text{A.41})$$

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

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \quad (\text{A.42})$$

so that the inverse (transpose) matrix is

$$\mathbf{B}^{-1} = \mathbf{B}^\dagger = \begin{pmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{pmatrix}. \quad (\text{A.43})$$

The representation of $\hat{\mathbf{i}}'$ in the preferred basis is

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad (\text{A.44})$$

so its representation in the initial basis is (see equation A.2)

$$\mathbf{B}^\dagger \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} b_{11} \\ b_{12} \end{pmatrix}. \quad (\text{A.45})$$

Similarly, the representation of $\hat{\mathbf{j}}'$ in the initial basis is

$$\mathbf{B}^\dagger \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b_{21} \\ b_{22} \end{pmatrix}. \quad (\text{A.46})$$

Thus the rotation matrix is

$$\mathbf{B} = \begin{pmatrix} \text{initial rep. of } \hat{\mathbf{i}}', \text{ on its side} \\ \text{initial rep. of } \hat{\mathbf{j}}', \text{ on its side} \end{pmatrix}. \quad (\text{A.47})$$

Example

Suppose we need to find a diagonal representation for the matrix

$$\mathbf{T} = \begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix}. \quad (\text{A.48})$$

First we search for the special vectors—the eigenvectors—such that

$$\begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}. \quad (\text{A.49})$$

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:

$$\begin{aligned} 7x + 3y &= \lambda x \\ 3x + 7y &= \lambda y \end{aligned} \quad (\text{A.50})$$

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.49) and write it as

$$\begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \lambda \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (\text{A.51})$$

Then

$$\begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (\text{A.52})$$

or

$$\begin{pmatrix} 7 - \lambda & 3 \\ 3 & 7 - \lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (\text{A.53})$$

Let's think about this. It says that for some matrix $\mathbf{M} = \mathbf{T} - \lambda\mathbf{1}$, we have

$$\mathbf{M} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (\text{A.54})$$

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

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{M}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (\text{A.55})$$

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 \mathbf{M} is *not invertible*, then there will be other solutions to

$$\mathbf{M} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (\text{A.56})$$

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* \mathbf{M} is not invertible. These values come if and only if the determinant of \mathbf{M} vanishes. For this example, we have to find values of λ such that

$$\det \begin{pmatrix} 7 - \lambda & 3 \\ 3 & 7 - \lambda \end{pmatrix} = 0. \quad (\text{A.57})$$

This is a quadratic equation in λ

$$(7 - \lambda)^2 - 3^2 = 0 \quad (\text{A.58})$$

called the *characteristic equation*. Its two solutions are

$$7 - \lambda = \pm 3 \quad (\text{A.59})$$

or

$$\lambda = 7 \pm 3 = 10 \text{ or } 4. \quad (\text{A.60})$$

We have found the two eigenvalues of our matrix!

Finding eigenvectors

Let's look now for the eigenvector associated with $\lambda = 4$. Equation (A.50)

$$\begin{aligned} 7x + 3y &= \lambda x \\ 3x + 7y &= \lambda y \end{aligned}$$

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

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (\text{A.61})$$

An entirely analogous search for the eigenvectors associated with $\lambda = 10$ finds any multiple of

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (\text{A.62})$$

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

$$\hat{\mathbf{i}}' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \hat{\mathbf{j}}' = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \quad (\text{A.63})$$

(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.47)

$$\mathbf{B} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad (\text{A.64})$$

and the diagonalized matrix (or, more properly, the representation of the matrix in the preferred coordinate system) is

$$\begin{pmatrix} 10 & 0 \\ 0 & 4 \end{pmatrix}. \quad (\text{A.65})$$

You don't believe me? Then multiply out

$$\mathbf{B} \begin{pmatrix} 7 & 3 \\ 3 & 7 \end{pmatrix} \mathbf{B}^\dagger \quad (\text{A.66})$$

and see for yourself.

Problems

A.12 Problem: Diagonalize a 2×2 matrix*

Diagonalize the matrix

$$\begin{pmatrix} 26 & 12 \\ 12 & 19 \end{pmatrix}. \quad (\text{A.67})$$

- Find its eigenvalues.
- Find its eigenvectors, and verify that they are orthogonal.
- 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 $\hat{\mathbf{j}}'$?)
- Find the matrix that translates from the initial basis to the basis of eigenvectors produced in part (c.).
- Verify that the matrix produced in part (d.) is orthogonal.
- Verify that the representation of the matrix above in the basis of eigenvectors is diagonal.
- (Optional.) What is the rotation angle?

A.13 Problem: Eigenvalues of a 2×2 matrix

Show that the eigenvalues of

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad (\text{A.68})$$

are

$$\lambda = \frac{1}{2} \left[(a + c) \pm \sqrt{(a - c)^2 + 4b^2} \right]. \quad (\text{A.69})$$

Under what circumstances is an eigenvalue complex valued? Under what circumstances are the two eigenvalues the same?

A.14 Problem: Diagonalize a 3×3 matrix

Diagonalize the matrix

$$\frac{1}{625} \begin{pmatrix} 1182 & -924 & 540 \\ -924 & 643 & 720 \\ 540 & 720 & -575 \end{pmatrix}. \quad (\text{A.70})$$

- a. Find its eigenvalues by showing that the characteristic equation is

$$\lambda^3 - 2\lambda^2 - 5\lambda + 6 = (\lambda - 3)(\lambda + 2)(\lambda - 1) = 0. \quad (\text{A.71})$$

- b. Find its eigenvectors, and verify that they are orthogonal.
 c. Show that the translation matrix can be chosen to be

$$\mathbf{B} = \frac{1}{25} \begin{pmatrix} 20 & -15 & 0 \\ 9 & 12 & -20 \\ 12 & 16 & 15 \end{pmatrix}. \quad (\text{A.72})$$

Why did I use the phrase “the translation matrix can be chosen to be” rather than “the translation matrix is”?

A.15 Problem: A 3×3 matrix eigenproblem

Find the eigenvalues and associated eigenvectors for the matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}. \quad (\text{A.73})$$

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 \mathbf{B} such that the off-diagonal elements of

$$\mathbf{B}\mathbf{A}\mathbf{B}^{-1} \quad (\text{A.74})$$

are smaller than the off-diagonal elements of \mathbf{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 \mathbf{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 LAPACK⁵ (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.10:

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}. \quad (\text{A.75})$$

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

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (\text{A.76})$$

Are there any other eigenvectors? The equation

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \quad (\text{A.77})$$

tells us that

$$\begin{aligned} ax + by &= \lambda x \\ cy &= \lambda y \end{aligned}$$

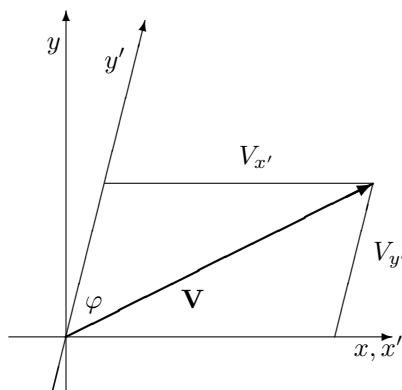
whence $\lambda = c$ and the eigenvector has polar angle θ where

$$\tan \theta = \frac{c - a}{b}. \quad (\text{A.78})$$

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

$$\mathbf{V} = V_{x'} \hat{\mathbf{i}}' + V_{y'} \hat{\mathbf{j}}' \quad (\text{A.79})$$

but, because $\hat{\mathbf{i}}'$ and $\hat{\mathbf{j}}'$ are not perpendicular, it is *not true* that

$$V_{x'} = \mathbf{V} \cdot \hat{\mathbf{i}}'. \quad \text{NO!} \quad (\text{A.80})$$

A little bit of geometry will convince you that the name of the vector \mathbf{V} changes according to

$$\begin{pmatrix} V_{x'} \\ V_{y'} \end{pmatrix} = \mathbf{B} \begin{pmatrix} V_x \\ V_y \end{pmatrix}, \quad (\text{A.81})$$

where

$$\mathbf{B} = \frac{1}{\sin \varphi} \begin{pmatrix} \sin \varphi & -\cos \varphi \\ 0 & 1 \end{pmatrix}. \quad (\text{A.82})$$

This matrix is *not* orthogonal. In fact its inverse is

$$\mathbf{B}^{-1} = \begin{pmatrix} 1 & \cos \varphi \\ 0 & \sin \varphi \end{pmatrix}. \quad (\text{A.83})$$

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.24) still applies, so

$$\mathbf{T}' = \mathbf{B} \mathbf{T} \mathbf{B}^{-1}. \quad (\text{A.84})$$

Using the explicit matrices already given, this says

$$\mathbf{T}' = \frac{1}{\sin \varphi} \begin{pmatrix} \sin \varphi & -\cos \varphi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} 1 & \cos \varphi \\ 0 & \sin \varphi \end{pmatrix} = \begin{pmatrix} a & (a-c)\cos \varphi + b\sin \varphi \\ 0 & c \end{pmatrix}. \quad (\text{A.85})$$

To make this diagonal, we need only choose a skew coordinate system where the angle φ gives

$$(a-c)\cos \varphi + b\sin \varphi = 0, \quad (\text{A.86})$$

that is, one with

$$\tan \varphi = \frac{c-a}{b}. \quad (\text{A.87})$$

Comparison with equation (A.78) 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

$$\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}. \quad (\text{A.88})$$

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.88), 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

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1/12 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 12 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \quad (\text{A.89})$$

This process is sometimes called a “stretching” or a “scaling” of the y -axis.

The transformation represented by matrix (A.88) in the initial coordinate system is represented in the new coordinate system by

$$\begin{pmatrix} 1 & 0 \\ 0 & 1/12 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 12 \end{pmatrix} = \begin{pmatrix} a & 12b \\ 0 & a \end{pmatrix}. \quad (\text{A.90})$$

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

$$\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}. \quad (\text{A.91})$$

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.10 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

¹For example, Kenneth Hoffman and Ray Kunze, *Linear Algebra*, second edition (Prentice-Hall, Englewood Cliffs, New Jersey, 1971).

²For example, Jerry Marion and Stephen Thornton, *Classical Dynamics of Particles and Systems*, fourth edition (Saunders College Publishing, Fort Worth, Texas, 1995) section 11.2.

³For example, Jerry Marion and Stephen Thornton, *Classical Dynamics of Particles and Systems*, fourth edition (Saunders College Publishing, Fort Worth, Texas, 1995) section 11.7.

⁴W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes* (Cambridge University Press, Cambridge, U.K., 1992).

⁵E. Anderson, *et al.*, *LAPACK Users' Guide* (SIAM, Philadelphia, 1992).

⁶B.T. Smith, *et al.*, *Matrix Eigensystem Routines—EISPACK Guide* (Springer-Verlag, Berlin, 1976).

Appendix B

The Spherical Harmonics

A “function on the unit sphere” is a function $f(\theta, \phi)$. Another convenient variable is $\mu = \cos \theta$. “Integration over the unit sphere” means

$$\int d\Omega f(\theta, \phi) = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi f(\theta, \phi) = \int_{-1}^{+1} d\mu \int_0^{2\pi} d\phi f(\theta, \phi).$$

$$\nabla^2 Y_\ell^m(\theta, \phi) = -\frac{1}{r^2} \ell(\ell+1) Y_\ell^m(\theta, \phi) \quad (\text{B.1})$$

$$\int Y_{\ell'}^{m'*}(\theta, \phi) Y_\ell^m(\theta, \phi) d\Omega = \delta_{\ell', \ell} \delta_{m', m} \quad (\text{B.2})$$

$$f(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell, m} Y_\ell^m(\theta, \phi) \quad \text{where} \quad (\text{B.3})$$

$$f_{\ell, m} = \int Y_\ell^{m*}(\theta, \phi) f(\theta, \phi) d\Omega \quad (\text{B.4})$$

In the table, square roots are always taken to be positive.

$$\begin{aligned}
Y_0^0(\mu, \phi) &= \left(\frac{1}{2^2\pi}\right)^{1/2} \\
Y_1^0(\mu, \phi) &= \left(\frac{3}{2^2\pi}\right)^{1/2} \mu &= \left(\frac{3}{2^2\pi}\right)^{1/2} \frac{z}{r} \\
Y_1^{\pm 1}(\mu, \phi) &= \mp \left(\frac{3}{2^3\pi}\right)^{1/2} \sqrt{1-\mu^2} e^{\pm i\phi} &= \mp \left(\frac{3}{2^3\pi}\right)^{1/2} \frac{1}{r} (x \pm iy) \\
Y_2^0(\mu, \phi) &= \left(\frac{5}{2^4\pi}\right)^{1/2} (3\mu^2 - 1) &= \left(\frac{5}{2^4\pi}\right)^{1/2} \left(3\frac{z^2}{r^2} - 1\right) \\
Y_2^{\pm 1}(\mu, \phi) &= \mp \left(\frac{3 \cdot 5}{2^3\pi}\right)^{1/2} \mu \sqrt{1-\mu^2} e^{\pm i\phi} &= \mp \left(\frac{3 \cdot 5}{2^3\pi}\right)^{1/2} \frac{z}{r^2} (x \pm iy) \\
Y_2^{\pm 2}(\mu, \phi) &= \left(\frac{3 \cdot 5}{2^5\pi}\right)^{1/2} (1-\mu^2) e^{\pm 2i\phi} &= \left(\frac{3 \cdot 5}{2^5\pi}\right)^{1/2} \frac{1}{r^2} (x \pm iy)^2 \\
Y_3^0(\mu, \phi) &= \left(\frac{7}{2^4\pi}\right)^{1/2} (5\mu^3 - 3\mu) &= \left(\frac{7}{2^4\pi}\right)^{1/2} \left(5\frac{z^3}{r^3} - 3\frac{z}{r}\right) \\
Y_3^{\pm 1}(\mu, \phi) &= \mp \left(\frac{3 \cdot 7}{2^6\pi}\right)^{1/2} (5\mu^2 - 1) \sqrt{1-\mu^2} e^{\pm i\phi} &= \mp \left(\frac{3 \cdot 7}{2^6\pi}\right)^{1/2} \left(5\frac{z^2}{r^2} - 1\right) \frac{1}{r} (x \pm iy) \\
Y_3^{\pm 2}(\mu, \phi) &= \left(\frac{3 \cdot 5 \cdot 7}{2^5\pi}\right)^{1/2} \mu (1-\mu^2) e^{\pm 2i\phi} &= \left(\frac{3 \cdot 5 \cdot 7}{2^5\pi}\right)^{1/2} \frac{z}{r^3} (x \pm iy)^2 \\
Y_3^{\pm 3}(\mu, \phi) &= \mp \left(\frac{5 \cdot 7}{2^6\pi}\right)^{1/2} (1-\mu^2) \sqrt{1-\mu^2} e^{\pm 3i\phi} &= \mp \left(\frac{5 \cdot 7}{2^6\pi}\right)^{1/2} \frac{1}{r^3} (x \pm iy)^3
\end{aligned}$$

Appendix C

Radial Wavefunctions for the Coulomb Problem

Based on Griffiths, page 154, but with scaled variables and with integers factorized.

$$\begin{aligned}R_{10}(r) &= 2e^{-r} \\R_{20}(r) &= \frac{1}{\sqrt{2}} \left(1 - \frac{1}{2}r\right) e^{-r/2} \\R_{21}(r) &= \frac{1}{\sqrt{2^3 \cdot 3}} r e^{-r/2} \\R_{30}(r) &= \frac{2}{\sqrt{3^3}} \left(1 - \frac{2}{3}r + \frac{2}{3^3}r^2\right) e^{-r/3} \\R_{31}(r) &= \frac{2^3}{3^3 \sqrt{2 \cdot 3}} \left(1 - \frac{1}{2 \cdot 3}r\right) r e^{-r/3} \\R_{32}(r) &= \frac{2^2}{3^4 \sqrt{2 \cdot 3 \cdot 5}} r^2 e^{-r/3} \\R_{40}(r) &= \frac{1}{2^2} \left(1 - \frac{3}{2^2}r + \frac{1}{2^3}r^2 - \frac{1}{2^6 \cdot 3}r^3\right) e^{-r/4} \\R_{41}(r) &= \frac{\sqrt{5}}{2^4 \sqrt{3}} \left(1 - \frac{1}{2^2}r + \frac{1}{2^4 \cdot 5}r^2\right) r e^{-r/4} \\R_{42}(r) &= \frac{1}{2^6 \sqrt{5}} \left(1 - \frac{1}{2^2 \cdot 3}r\right) r^2 e^{-r/4} \\R_{43}(r) &= \frac{1}{2^8 \cdot 3 \sqrt{5 \cdot 7}} r^3 e^{-r/4}\end{aligned}$$

Appendix D

Quantum Mechanics Cheat Sheet

Delta functions:

$$\int_{-\infty}^{+\infty} e^{ikx} dk = 2\pi\delta(x) \quad (\text{D.1})$$

$$\int_{-\infty}^{+\infty} e^{i(p/\hbar)x} dp = 2\pi\hbar\delta(x) \quad (\text{D.2})$$

$$\int_{-\infty}^{+\infty} e^{i\omega t} d\omega = 2\pi\delta(t) \quad (\text{D.3})$$

Fourier transforms:

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi(x) e^{-i(p/\hbar)x} dx \quad (\text{D.4})$$

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \tilde{\psi}(p) e^{+i(p/\hbar)x} dp \quad (\text{D.5})$$

$$\tilde{f}(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt \quad (\text{D.6})$$

$$f(t) = \int_{-\infty}^{+\infty} \tilde{f}(\omega) e^{+i\omega t} \frac{d\omega}{2\pi} \quad (\text{D.7})$$

Gaussian integrals:

$$\int_{-\infty}^{+\infty} e^{ax^2+bx} dx = \sqrt{\frac{\pi}{-a}} e^{-b^2/4a} \quad \Re\{a\} \leq 0 \quad (\text{D.8})$$

$$\frac{\int_{-\infty}^{+\infty} x^2 e^{-x^2/2\sigma^2} dx}{\int_{-\infty}^{+\infty} e^{-x^2/2\sigma^2} dx} = \sigma^2 \quad (\text{D.9})$$

Time development:

$$\frac{d|\psi(t)\rangle}{dt} = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle \quad (\text{D.10})$$

$$\frac{\partial\psi(\mathbf{x}, t)}{\partial t} = -\frac{i}{\hbar}\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x})\right]\psi(\mathbf{x}, t) \quad (\text{D.11})$$

$$|\psi(t)\rangle = \sum_n e^{-(i/\hbar)E_n t} c_n |\eta_n\rangle \quad (\text{D.12})$$

$$\frac{d\langle\hat{A}\rangle}{dt} = -\frac{i}{\hbar}\langle[\hat{A}, \hat{H}]\rangle \quad (\text{D.13})$$

Momentum:

$$\hat{p} \iff -i\hbar\frac{\partial}{\partial x} \quad (\text{D.14})$$

$$[\hat{x}, \hat{p}] = i\hbar \quad (\text{D.15})$$

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{i(p/\hbar)x} \quad (\text{D.16})$$

Dimensions:

$$\psi(x) \text{ has dimensions } [\text{length}]^{-1/2} \quad (\text{D.17})$$

$$\psi(\mathbf{x}_1, \mathbf{x}_2) \text{ has dimensions } [\text{length}]^{-6/2} \quad (\text{D.18})$$

$$\tilde{\psi}(p) \text{ has dimensions } [\text{momentum}]^{-1/2} \quad (\text{D.19})$$

$$\hbar \text{ has dimensions } [\text{length} \times \text{momentum}] \text{ or } [\text{energy} \times \text{time}] \quad (\text{D.20})$$

Energy eigenfunction sketching: (one dimension)

$$n\text{th excited state has } n \text{ nodes} \quad (\text{D.21})$$

$$\text{if classically allowed: regions of high } V(x) \text{ have large amplitude and long wavelength} \quad (\text{D.22})$$

$$\text{if classically forbidden: regions of high } V(x) \text{ have faster cutoff} \quad (\text{D.23})$$

Infinite square well: (width L)

$$\eta_n(x) = \sqrt{2/L} \sin kx \quad k = n\pi/L \quad n = 1, 2, 3, \dots \quad (\text{D.24})$$

$$E_n = \frac{\hbar^2 k^2}{2m} = n^2 \frac{\pi^2 \hbar^2}{2mL^2} \quad (\text{D.25})$$

Simple harmonic oscillator: ($V(x) = \frac{1}{2}Kx^2$, $\omega = \sqrt{K/m}$)

$$E_n = (n + \frac{1}{2})\hbar\omega \quad n = 0, 1, 2, \dots \quad (\text{D.26})$$

$$[\hat{a}, \hat{a}^\dagger] = \hat{1} \quad (\text{D.27})$$

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2}) \quad (\text{D.28})$$

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle \quad (\text{D.29})$$

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad (\text{D.30})$$

$$\hat{x} = \sqrt{\hbar/2m\omega}(\hat{a} + \hat{a}^\dagger) \quad (\text{D.31})$$

$$\hat{p} = -i\sqrt{m\hbar\omega/2}(\hat{a} - \hat{a}^\dagger) \quad (\text{D.32})$$

Coulomb problem:

$$E_n = -\frac{\text{Ry}}{n^2} \quad \text{Ry} = \frac{m}{2\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 = 13.6 \text{ eV} \quad (\text{D.33})$$

$$a_0 = \frac{\hbar^2}{m} \left(\frac{4\pi\epsilon_0}{e^2} \right) = 0.0529 \text{ nm} \quad (\text{Bohr radius}) \quad (\text{D.34})$$

$$\tau_0 = \frac{\hbar}{2\text{Ry}} = 0.0242 \text{ fsec} \quad (\text{characteristic time}) \quad (\text{D.35})$$

Angular momentum:

$$[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z, \text{ and cyclic permutations} \quad (\text{D.36})$$

The eigenvalues of \hat{J}^2 are

$$\hbar^2 j(j+1) \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots \quad (\text{D.37})$$

For a given j , the eigenvalues of \hat{J}_z are

$$\hbar m \quad m = -j, -j+1, \dots, j-1, j. \quad (\text{D.38})$$

The eigenstates $|j, m\rangle$ are related through the operators

$$\hat{J}_+ = \hat{J}_x + i\hat{J}_y \quad \hat{J}_- = \hat{J}_x - i\hat{J}_y \quad (\text{D.39})$$

by

$$\hat{J}_+ |j, m\rangle = \hbar\sqrt{j(j+1) - m(m+1)} |j, m+1\rangle \quad (\text{D.40})$$

$$\hat{J}_- |j, m\rangle = \hbar\sqrt{j(j+1) - m(m-1)} |j, m-1\rangle. \quad (\text{D.41})$$

Spherical harmonics:

A “function on the unit sphere” is a function $f(\theta, \phi)$. Another convenient variable is $\mu = \cos\theta$. “Integration over the unit sphere” means

$$\int d\Omega f(\theta, \phi) = \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi f(\theta, \phi) = \int_{-1}^{+1} d\mu \int_0^{2\pi} d\phi f(\theta, \phi).$$

$$\nabla^2 Y_\ell^m(\theta, \phi) = -\frac{1}{r^2} \ell(\ell+1) Y_\ell^m(\theta, \phi) \quad (\text{D.42})$$

$$\int Y_\ell^{m'*}(\theta, \phi) Y_\ell^m(\theta, \phi) d\Omega = \delta_{\ell', \ell} \delta_{m', m} \quad (\text{D.43})$$

$$f(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell, m} Y_\ell^m(\theta, \phi) \quad \text{where} \quad (\text{D.44})$$

$$f_{\ell, m} = \int Y_\ell^{m*}(\theta, \phi) f(\theta, \phi) d\Omega \quad (\text{D.45})$$

Index

approximation
 controlled, 107
 uncontrolled, 107

baum, 199

characteristic equation, 211
characteristic matrix, 211
controlled approximation, 107

definite value
 use of, 93

eigenvalues, 208, 209
eigenvectors, 207–209

love, 92

marble, 92, 93
matrix mathematics, 199–218

orthogonal matrix, 200

Renninger negative-result experiment, 92
rotation matrix, 200

Sturm-Liouville problems, 134

tree, 199

uncontrolled approximation, 107

visualization, 93