Second Chance: Unorthodox but Personally Effective Explanations in Probability and Physics

Blake C. Stacey

July 5, 2026

Contents

1 Overview

This is a book about heat and unpredictability. I am writing it during a pandemic, while humanity goes about destroying the Earth's climate and our democratic institutions, flawed to begin with, splinter and fall. I began this project during the first COVID spring, because I was teaching statistical physics to graduate students and wanted to provide as much for them as I could. I continue it because teaching science feels at least a little important and, unlike so many important things, may not be entirely beyond my capabilities.

As I told the students in my graduate quantum mechanics course this past semester, "I don't know whether I can make this class a reason to get up in the morning, but I will try. That requires, I think, a reason to study quantum physics which goes beyond the fact that the department requires it, maybe even a better reason than its being a useful theory in various practical applications. Perhaps, in a time of confusion, clarity can be a source of renewal; perhaps, when lies are commonplace and legitimized, feeling an argument unfold in logical sequence is a moment of respite. I don't know."

1.1 Opening Remarks

The beginning of a textbook can often be harder to write than the middle. In the middle, the author can presume that the reader has a sense of what the subject is about, and moreover the author can focus upon explaining how to solve specific problems. But the beginning of a book is supposed to introduce the basic ideas, and this forces the author to devise clear statements about big concepts. Nobody knows how to write exam problems about big concepts.

Mathematicians have a tradition going back at least to Euclid: Start with a set of postulates that we can all agree upon as established, and deduce from them. This can make for a difficult arrangement, when what is logically primitive does not coincide with what is familiar. An example that has slipped into legend is Russell and Whitehead taking 379 pages to prove that $1 + 1 = 2$. In physics, we face the additional challenge that we are not merely manipulating concepts for the pleasure of their abstract interplay, but also subjecting them to the Darwinian pressure of explaining the natural world. We deal in approximations of provisional validity, and that which is intuitive may by the same token be unrepresentative.

One of my earliest memories is about thermodynamics. This memory is from the same stratum as the shock of being separated from my plush Snoopy when we moved houses (and how big that moving truck loomed from my vantage point!). At the time, I had an assortment of toys, and I noticed that when I first took them from my toybox, they were cool to the touch, but after I played with them for a bit, they were warm.

My father explained that it was *me* who warmed them up.

The laws of thermodynamics distill centuries of experience, from the unsophisticated like mine to the quantitative in laboratories. We burned things and boiled other things and froze still other things; we dozed in the summer sun and we split solar radiation into spectra. For quite a while, the idea of heat as a fluid enjoyed considerable appeal — and it does make sense to speak of heat *flowing* from coffee into ice, or from an eccentric child into a toy saw. But the heat-as-fluid notion ran into trouble, most crucially because in the right circumstances, that fluid could be made in arbitrarily large amounts, from nothing but motion. After much tribulation, we learned that hotness *is* motion at the microscopic level. Warmth is the tangible manifestation of invisible movement.

In order to understand heat, then, we have to find a way to reason about aggregates of matter that we can only characterize imprecisely. We know that an object has internal parts, we know that the behavior of those parts is important, but what can we say about that behavior when we only have information about the bulk characteristics? This is the domain of statistical physics, which relies upon the mathematics of probability theory. To understand heat, we must not only understand motion, but also *information.* James Clerk Maxwell provided a presentiment of this discovery, as long ago as 1878 [1]:

> The idea of dissipation of energy depends on the extent of our knowledge. Available energy is energy which we can direct into any desired channel. Dissipated energy is energy which we cannot lay hold of and direct at pleasure, such as the energy of the confused agitation of molecules which we call heat. Now confusion, like the correlative term order, is not a property of material things in themselves, but only in relation to the mind which perceives them. A memorandum-book does not, provided it is neatly written, appear confused to an illiterate person, or the owner who understands it thoroughly, but to any other person able to read it appears to be inextricably confused. Similarly the notion of dissipated energy could not occur to a being who could not turn any of the energies of nature to his account, or to one who could trace the motion of every molecule and seize it at the right moment. It is only to a being in the intermediate stage, who can lay hold of some forms of energy while others elude his grasp, that energy appears to be passing from the available to the dissipated state.

Yet trying to apply probability theory to the physics of Newton and Maxwell didn't work: The calculations kept giving predictions that did not match up with what nature actually does. Thus, to do things right, we have to incorporate quantum mechanics. The big picture of that subject is something like the following:

- Quantum physics is a theory about calculating *probabilities.*
- The *rules* by which we calculate probabilities express *algebraic relationships* between one experiment and another.
- Closely studying those rules reveals that quantum probabilities cannot be "about" hidden properties within physical systems.

That last point is conceptually rather demanding, so before we do the math, it might help to provide an illustration, or at least a thing to contrast against. If I roll a pair of dice in a backgammon cup, and then I snap the cup down onto the game board, you don't know which way the dice came up, do you? You could guess, and even make some numerical statements about the situation using probability theory. But your basic assumption is that there is *some fact of the matter* about that dice roll: You don't know which faces are up, but *some* faces *are* up. I could resolve your uncertainty by lifting the cup and uncovering the fact that was already there, waiting for you.

Quantum physics is stranger than that. Our predictions about what an electron might do, for example, are not probabilistic because we are ignorant of a fact already dwelling inside the electron. Something more exotic is happening: We resort to probabilities not because we are ignorant of results waiting to be uncovered, but because nature itself has yet to make up its mind. We can even be quantitative about this. The *hypothesis* that our uncertainty is due to ignorance of hidden facts implies a relation between the predictions for one experiment and the predictions for another. We can then do the calculation in quantum theory and show that this relation is *violated.* Then we go and do the experiments in real life, and we find that nature agrees with the quantum prediction, not the hypothesis of hidden facts!

Doing that calculation in quantum theory requires mastering at least a portion of the *mathematical formalism* of quantum physics, the general abstract rules by which the quantum game is played. These rules are "algebraic" not just in the grade-school sense of involving math with x's and y's, but in the more professional sense of the word. In particular, we will be making heavy use of *linear algebra,* and we will even get into *group theory.*

I'll be honest with you: I spent a lot of time working to learn quantum and statistical physics, and not everything clicked on the first attempt. If anything, that understates the situation. Even if we start the clock when I got to college, thereby neglecting all the years of informally "studying" texts of ... highly variable quality, many topics took two or three grapples before I could claim any competence. My goal in this book is to present the lessons that finally *worked* for me. I make no promise that they will be the lessons that bring about the "click" for anyone else, but in my own experience teaching, sometimes they've done the trick. Satisfaction is possible, not guaranteed.

1.2 Notational Conventions

Most of the vector and matrix indices will start at zero, which (I am no stranger to controversy) will be regarded as a natural number. In many places, it will be convenient to designate an element of a matrix by writing the symbol for that matrix in brackets and supplying the indices in a subscript; thus, $[M]_{34}$ is the element in row 3 and column 4 of the matrix M.

The generic physicist who operates engines, writes down quantum states, etc., will be named Alice. The second physicist in a scenario or a protocol will be named Bob. The use of she/her pronouns for Alice and he/him pronouns for Bob is for ease of reference and flexibility of phrasing. Like in the rest of life, gender is merely a simplification to

be trusted only as far as it is convenient.

Numerical values of some handy constants are provided here for quick reference:

$$c = 2.99792458 \times 10^{8}\ \mathrm{m\,s^{-1}}\,. \tag{1.1}$$
$$\hbar = 1.054571817 \times 10^{-34}\ \mathrm{J\,s}\,. \tag{1.2}$$
$$q_e = 1.602176634 \times 10^{-19}\ \mathrm{C}\,. \tag{1.3}$$
$$k_B = 1.380649 \times 10^{-23}\ \mathrm{J\,K^{-1}}\,. \tag{1.4}$$
$$m_e = 9.1093837139(28) \times 10^{-31}\ \mathrm{kg}\,. \tag{1.5}$$
$$m_p = 1.67262192595(52) \times 10^{-27}\ \mathrm{kg}\,. \tag{1.6}$$
$$\epsilon_0 = 8.8541878188 \times 10^{-12}\ \mathrm{C^2\,kg^{-1}\,m^{-3}\,s^2}\,. \tag{1.7}$$
$$G = 6.67430(15) \times 10^{-11}\ \mathrm{m^3\,kg^{-1}\,s^{-2}}\,. \tag{1.8}$$

1.3 Readings and Recommendations

In addition to the specific citations inserted into the text, I have been guided by some general references. These include Mehran Kardar's *Statistical Physics of Particles* (Cambridge University Press, 2007); Enrico Fermi's *Thermodynamics* (Dover, 1956); David Tong's lecture notes on statistical physics (2012); and Asher Peres' *Quantum Theory: Concepts and Methods* (1993). Section 19.5 is based on notes I took during Kardar's lectures for a "statistical physics in biology" course. (My typescript was the original source for several of the documents posted on MIT's OpenCourseWare page for 8.592, though both my text and theirs evolved independently after diverging from that common ancestor.)

Exercises will be distributed under separate cover. And on the subject of getting one's practice in, I find that it helps to read physics with pen and paper in hand so that I can work through the equations actively as I go. Without this, it is too easy to gloss over any unclear points and fool myself into thinking I understand more than I do. This also helps me fix the steps of a calculation into my memory and build confidence before going off the trail and trying to calculate new things. I wish someone had suggested that to me a long time ago.

My younger self could have benefited from many morsels of advice. Among others: If possible, learn a "job skill" of day-to-day physics (writing a technical report, fitting a curve to data, etc.) *before* the curriculum requires it, so that you don't have to acquire the skill at the same time that you're struggling with new physics. Know that many homework problems are to physics what practicing chords and scales are to music, i.e., not valued for their own sake, but instead to make your fingers ready when they are needed. Be wary of autodidacts who pontificate on everything and calculate nothing. Expect the history of any subject to be more complicated than first presented. Don't let the boundaries of a class define the boundaries of a subject, but also, appreciate when the structure of a class drives you to learn something you wouldn't have sought out on your own. Listening to lectures as you fall asleep is probably not a useful way to study, but it can lead to strange dreams.

2 Tools for Probability

Probability theory is a way of managing uncertainty and making choices when equipped with incomplete information. In a sense, it is a theory of knowledge and learning, one which is applicable when our attitude about each proposition we consider, or each event we might encounter, can be encapsulated in a real number.

A physicist's intuition about probability will habitually go something like this: If an experiment has a probability p of producing some result, and I repeat that experiment N times, the number of times I get that result will likely be around Np. This is a useful intuition. It may even be all you need to get through the day. But as a *definition* of probability, it falls rather short, in ways that you may already have noticed.

For one thing, no experiment is *exactly* repeatable. If your lab equipment is sensitive to noise and vibration, then you will get different results at night after the trains have stopped running. The positions of the planets in the sky are slightly different from one night to the next, and you yourself will be older. How close does one experiment have to be to another for it to count as a "repetition"? Worse yet, our slogan for our intuition about "probability" says that we will "likely" get "around" Np results. How near is "around", and what is "likely"? We can't use the concept of probability as part of the definition of probability!

So, the people who by nature get a bug in their brain about fundamentals have tried all sorts of ways to ground the notion of *probability* more solidly. One crowd have tried to make the concept of a "collective" or "ensemble" of experiments logically sound. This has proven exceptionally difficult, perhaps impossible, and it starts to feel like a diversion: Why should we have to consider an arbitrarily large, perhaps infinite collection of experiments in order to understand the one single thing sitting on our lab bench? Others have tried to understand probability as generalized logic, and say that the probability of a proposition quantifies the extent to which it follows from evidence. Still others have postulated that within each situation, there dwells some physical property, something like a ratio of "probabilitons" coming in "up" and "down" flavors. But how does one measure that property? Try to work with that concept for a while, and it starts to feel like a crutch — something that we posit in order to have an "objective" fact, or rather the suggestion of such a fact, hanging about to soothe our qualms even though our *calculations* do not require it. Yet another approach has the flavor of game theory: It says that a probability is something that a specific agent asserts in order to manage their own expectations in a quantitative way. This is perhaps the most shocking approach of them all. We take it up in the next section.

2.1 Basic Properties

One way (but not the only way) to motivate the basic postulates on which this theory relies is a parable: the story of *the android and the Ferengi bartender*. The android walks into a bar, and the bartender, after a quick assessment of the situation, offers a bit of friendly sport. The bartender proposes to *buy* or *sell* a lottery ticket for any event the android chooses, at any price the android desires. For any event E, the android assigns a numerical weight $p(E)$, which the android deems to be the fair price of the lottery ticket

$$\boxed{\text{Pays \$1 if the event } E \text{ occurs.}} \tag{2.1}$$

The android will *buy* or *sell* this ticket for $\$p(E)$.

The bartender offers the android the opportunity to strike a deal on any combination of events, hoping to catch the android in an inconsistency, forcing a *sure loss*. The goal of the android is to avoid this eventuality.

For example, if the android assigns a price $p(E) < 0$ for some event E, then the android will sell a ticket for a negative amount of money, and the bartender wins. Likewise if the android assigns $p(E) > 1$: this indicates a willingness to sell a ticket for more than it could ever be worth.

Consider two events, E and F, which the android believes are mutually exclusive. The bartender proposes the following three lottery tickets:

$$\boxed{\text{Worth \$1 if } (E \text{ or } F)}, \tag{2.2}$$

along with

$$\boxed{\text{Worth \$1 if } E} \tag{2.3}$$

and finally

$$\boxed{\text{Worth \$1 if } F}. \tag{2.4}$$

The value of ticket (2.2) should be the same as the total value of tickets (2.3) and (2.4). If the android professes $p(E \text{ or } F) > p(E) + p(F)$, then the bartender wins. The bartender sells ticket (2.2) and buys tickets (2.3) and (2.4) leaving the android in the red, and whatever happens, the android can never recoup the loss. For example, if event E takes place, the android gains \$1 for the first ticket, but has to pay it back out again by the second.

So, for mutually exclusive events,

$$p(E \text{ or } F) = p(E) + p(F). \tag{2.5}$$

By a similar argument, the android must price the ticket

$$\boxed{\text{Worth } \$\tfrac{m}{n} \text{ if } E} \tag{2.6}$$

at $\$\frac{m}{n}p(E)$, if m and n are positive integers. A continuity argument shows that a similar statement holds for a positive real number x in place of the rational number m/n.

Furthermore, if "not E" is the event that E does not happen, then we have two mutually exclusive events, and the sum of the fair prices for their lottery tickets must be the fair price for a ticket that is worth \$1 no matter what happens. Therefore,

$$p(E) + p(\text{not } E) = 1. \tag{2.7}$$

The weights of an event and its complementary event must, to be consistent, sum up to unity.

Now, the bartender proposes a new game: *conditional* lottery tickets. For any two events E and F, a conditional ticket will be worth \$1 if both E and F occur, but the cost of the ticket will be returned if the event E does *not* occur.

$$\boxed{\text{Worth \$1 if } (E \text{ and } F). \text{ But money back if (not } E).} \tag{2.8}$$

This is a gamble on the event "F given E," which we can denote $F|E$. The price of this ticket is, by definition, $\$p(F|E)$. So, this ticket is the same as one written

$$\boxed{\text{Worth \$1 if } (E \text{ and } F). \text{ But refund } \$p(F|E) \text{ if (not } E).} \tag{2.9}$$

To avoid a sure loss at the bartender's hands, the android must price *this* ticket as the total price of these:

$$\boxed{\text{Worth \$1 if } (E \text{ and } F)}, \text{ and } \boxed{\text{Worth } \$p(F|E) \text{ if (not } E).} \tag{2.10}$$

The price of the final ticket must be $\$p(F|E)p(\text{not } E)$. So, consistency requires that

$$p(F|E) = p(E \text{ and } F) + p(F|E)p(\text{not } E). \tag{2.11}$$

The weights of E and its complement add up to 1, which implies

$$p(F|E) = p(E \text{ and } F) + p(F|E)(1 - p(E)). \tag{2.12}$$

We cancel the $p(F|E)$ from both sides and rearrange to find that

$$p(E \text{ and } F) = p(F|E)p(E). \tag{2.13}$$

We saw earlier that the ticket prices for events deemed mutually exclusive must add. What if the android contemplates two events, E and F, and decides that they are *not* mutually exclusive? In this case, the events "E and (not F)" and "F" are still automatically mutually exclusive in the android's judgment. Therefore,

$$p\left([E \text{ and } (\text{not } F)] \text{ or } F\right) = p\left(E \text{ and } (\text{not } F)\right) + p\left(F\right). \tag{2.14}$$

The "distributive rule" of Boolean logic, which is the subject of manipulating propositions hooked together with the logical connectives "and", "or" and "not", tells us that

$$[E \text{ and } (\text{not } F)] \text{ or } F = [E \text{ or } F] \text{ and } [(\text{not } F) \text{ or } F]. \tag{2.15}$$

This simplifies to

$$[E \text{ and } (\text{not } F)] \text{ or } F = [E \text{ or } F]. \tag{2.16}$$

Substituting this into the left-hand side of Eq. (2.14) yields

$$p\left(E \text{ or } F\right) = p\left(E \text{ and } (\text{not } F)\right) + p\left(F\right). \tag{2.17}$$

Using the definition of conditional lotteries, we have

$$p(E \text{ and } (\text{not } F)) = p(\text{not } F|E)p(E). \tag{2.18}$$

By normalization, this is

$$p(E \text{ and } (\text{not } F)) = \left(1 - p(F|E)\right) p(E). \tag{2.19}$$

Distributing $p(E)$ over the sum and identifying $p(F|E)p(E) = p(F \text{ and } E)$, we find that

$$p(E \text{ and } (\text{not } F)) = p(E) - p(F \text{ and } E) = p(E) - p(E \text{ and } F). \tag{2.20}$$

We can substitute this back into Eq. (2.17), yielding

$$p\left(E \text{ or } F\right) = p(E) + p(F) - p(E \text{ and } F). \tag{2.21}$$

One way to remember this relationship is in terms of areas: if each event is represented by a geometrical shape, then the total area of the shape representing the event "E or F" is the area of the E-shape, plus the area of the F-shape, minus the area of the region where they overlap, which would otherwise be double-counted.

In summary, the requirement that the android avoid a *sure loss* in *any single deal* imposes a rather intricate set of constraints on the fair-price function! Specifically, we have shown that prices must be bounded:

$$0 \leq p(E) \leq 1. \tag{2.22}$$

Also, for events believed to be mutually exclusive, the prices add:

$$p(E \text{ or } F) = p(E) + p(F). \tag{2.23}$$

For an event C which the android believes is *certain* to occur,

$$p(C) = 1. \tag{2.24}$$

If the set $\{E_i\}$ is an *exhaustive* set of mutually exclusive events, then

$$\sum_i p(E_i) = 1. \tag{2.25}$$

And *conditional* prices are related to *joint* prices, in accordance with

$$p(F|E) = \frac{p(E \text{ and } F)}{p(E)}. \tag{2.26}$$

It is known [2, 3, 4] that the android can avoid defeat at the hands of the bartender if and only if the price assignment function satisfies Eqs. (2.22), (2.23) and (2.24). An assignment which meets these criteria is called *coherent.*

Note that the bartender does not have to pick any numbers or make any price assignments. All the choices are made by the android; the bartender merely exposes inconsistencies in the numerical weightings which the android attaches to events. Indeed, we could restate the whole scenario as an inner challenge the android poses in the pursuit of self-consistency.[1] Furthermore, note that we have not mentioned *repeated experiments* or *sequences of trials* at all. Avoiding loss at the bartender's hands imposes consistency conditions on numerical weightings, even for events which can only happen once.

The coherence conditions can detect inconsistencies, but they themselves do not tell the gambler how to *resolve* them. If the gambler arrives at a number $p(E \text{ or } F)$ from one source, a number $p(E)$ from other considerations, and a number $p(F)$ from still a third wellspring in their soul, it may be that $p(E \text{ or } F) \neq p(E) + p(F)$. It is their job to realign themselves and achieve coherence; the equations cannot do it for them.

The chain of inferences leading to Eqs. (2.22)–(2.24) is typically known as a *Dutch book argument.* The players whom we have designated as the android and the bartender are, in the standard parlance, the bettor and the bookie, and the bettor strives to achieve Dutch-book coherence. The choice of terminology in this section emphasizes the importance of these ideas for machine learning [5, 6, 7], while in addition underscoring the distinction between a mathematicized standard of behavior and the way human beings actually conduct themselves in gambling establishments. Furthermore, the historical origin of the "Dutch book" term is rather obscure, anyway [8, 9].

The coherence conditions (2.22), (2.23) and (2.24) are familiar: together, they say that p is a *probability distribution.* Dutch- or Ferengi-book coherence provides an operational meaning to probability, which as we saw earlier is relevant even for experiments which can be performed only once. We could, on a purely mathematical level, have defined probability axiomatically, declaring that a probability distribution is a set with a particular kind of additional structure. Such a definition [10] might read like the following:

> A probability space $(\Omega, \mathcal{B}, \mathbf{P})$ is a tuple comprising a set Ω, a nonempty collection $\mathcal{B}$ of subsets of Ω such that $\mathcal{B}$ satisfies the properties of a σ-algebra, and a function $\mathbf{P} : \mathcal{B} \to [0, 1]$ which sends each element of $\mathcal{B}$ to a real number in the unit interval. The function $\mathbf{P}$ is countably additive, and $\mathbf{P}(\Omega) = 1$.

In fact, a mathematician might elect to abstract further from this definition and work with an algebra of random variables and expectation values, eliding the basic set Ω, the event set $\mathcal{B}$ and the mapping $\mathbf{P}$. This turns out to be useful for some subjects, such as random matrix theory [10], but it is not a perspective we will need to pursue for this chapter.

[1] Just as topology is the study of properties that remain invariant under continuous transformations, probability can be seen as the study of quantities that remain invariant across alternative specifications of events, like rewriting "E or F" as "(E and not F) or F".

Moreover, we will not have to stress greatly over questions of *continuous* versus *discrete* sets of propositions for events. We will casually switch between discrete probability distributions, normalized as

$$\sum_i p(E_i) = 1, \tag{2.27}$$

and continuous probability densities, which are normalized as

$$\int_{-\infty}^{\infty} dx\, p(x) = 1. \tag{2.28}$$

The interpretation of the function $p(x)$ is as follows: for any real number x_0, the quantity $p(x = x_0)dx$ is the probability of the event that x lies between x_0 and $x_0 + dx$. We will not have to be more exacting in our definitions than this. In other words, we will be living by the physicists' standard, which is indifferent to mathematical subtleties until they become unavoidable, treating them as niceties which are no more relevant than the question of how best to construct the real numbers from the integers in the first place [11].

Discussions of this topic are typically reliant upon examples like throwing a dart at a board. One habitually says that there is probability zero of the dart position being exactly x for any value of x, but a small yet nonzero probability that it will lie in a tiny neighborhood near x. Now, it is both impossible and irrelevant in practice to define an infinitely sharp tip of a dart, or to read out its position with infinite precision. And if we think about how we might actually locate where a dart has pierced a dartboard, we find there are multiple types of measurement at different scales that employ different methods. If the goal is to assign a score to the throw, we might only need the color of the region where the dart has interrupted the surface. Alternatively, to get information about a smaller distance scale, we might use a wooden ruler. Going still further, we could break out a steel micrometer. Eventually, we could find ourselves employing a magnifying glass, and so on.

In order to get on with life, we make an assumption of nonperversity and say that all of these techniques are in accord. Where their regimes of validity overlap, they give equivalent answers. (This seems simple for a dart and a dartboard, but it bears its teeth in astronomy with the "distance ladder" and how it must be calibrated. Relating parallax measurements to the oscillations of Cepheid variables and the observed brightnesses of type IA supernovae is a bit complicated.) We use a probability density function $p(x)$ to bundle together expectations for all these different measurements — eyeballing the color from across the room, laying down a meter-stick, et cetera — in a concise way. The "continuum counterpart" is a conceptual *addition* to our reckoning about measurements that each have *finitely* many distinguishable outcomes.

Let us briefly examine this idea of making a conceptual addition in more detail. What should it mean that we can consistently "bundle together" our predictions for a multiplicity of disparate experiments and techniques? We must somehow be fitting those predictions together into the same mathematical object. We do not yet know how to associate something meaningfully probability-like to a single point, but what

we *can* do is ask, "Will the measured value of the position be less than x?" Seen this way, we are talking about a binary result. Let's write $F(x)$ for the probability that the result of the position measurement is a value less than or equal to x. Depending on which value of x we consider, we may need to employ different techniques to make that measurement, but the spirit of consistency suggests that we can talk about the whole problem using the one overall function F.

Pick two real numbers a and b, with $a < b$. Consider two events: "the measured position is less than or equal to a" and "the measured position is less than or equal to b". The respective probabilities of these events are $F(a)$ and $F(b)$. We know that if $x < a$, then $x < b$ as well; there is no way to have the former without the latter, and so $F(b) \geq F(a)$. In other words, F is monotonically nondecreasing. It follows from the ordinary rules of coherence that the probability of the measured position falling in between a and b is $F(b) - F(a)$.

If the possible values of the position measurement are confined to a finite interval, then this is already enough to say that F is differentiable "almost everywhere", that is, everywhere except a set of measure zero. This follows from the monotonicity of F.

It is reasonable to ask a little more of F, even if we do not restrict our attention to a finite interval of $\mathbb{R}$. Remember, we are talking about comparing and combining measurement techniques that offer different levels of precision. We may posit that, for these techniques to fit together properly, it must be that providing an "approximately correct" value of the input x gives an "approximately correct" value of the output $F(x)$. Limiting the precision of the input limits the precision of the output but does not spoil it. This amounts to saying that F should be *continuous.* And we can go further than that. Let us consider experiment outcomes that correspond to very tiny intervals of the real line, like an event specified by the measured position falling within one of the disjoint intervals $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)$. If the total length of all these intervals is small, then the probability of this event should be small, no matter where the intervals are chosen. Translating this condition into the language of real analysis and imposing it on our function F means that F is *absolutely continuous.* This means that the derivative $p = F'$ can be defined almost everywhere, and the change in F is the Lebesgue integral of p:

$$F(b) - F(a) = \int_a^b dx\, p(x)\,. \tag{2.29}$$

Our capital F is the *cumulative distribution function,* and its derivative is the probability density.

2.2 Updating and the Reflection Principle

All of our considerations so far in this section have concerned the android's probability ascriptions at *a single time.* Even the conditional probabilities, as in Eq. (2.26), are statements of how the android is willing, at a particular moment, to gamble on events, even if one of those events may chronologically precede the other. We have yet to

address how the android might make changes to probability assignments in a self-consistent way.

When we consider probabilities changing with time, our android's probability ascriptions gain a time index. Let p_0 be a function from events to the interval $[0, 1]$ which expresses the android's gambling commitments at the time $t = 0$. Similarly, p_τ denotes the android's gambling commitments at a later time, $t = \tau$. Our problem is to relate p_0 and p_τ according to some standard of consistency.

Some probability assignments may carry two time indices. For example, the bartender can suggest a wager on how many drinks the bar patrons will order in an hour. We can define an event $E(N, T)$ as the event that N drinks will be bought during the interval from time T until one hour later. A probability ascription for this event has one time label in the event itself, and another index which denotes the time at which the ascription is made: $p_t(E(N, T))$. Our concern here is to relate probabilities with different subscripts, which is a problem distinct from the question of how $p_t(E(N, T))$ relates to $p_t(E(N, T'))$. Starting in section §2.8, we will address the latter question to a larger extent.

Suppose that at $t = 0$, the android is willing to gamble on two events, E and D, in such a way that

$$p_0(E|D) = \frac{p_0(E \text{ and } D)}{p_0(D)} = q. \tag{2.30}$$

Now, if the event D occurs between $t = 0$ and $t = \tau$, what should $p_\tau(E)$ be? The simplest choice is to use the number we already have on hand, and force it equal to q:

$$p_\tau(E) = p_0(E|D) = q, \text{ if } D \text{ occurs.} \tag{2.31}$$

However, we *cannot deduce this* from a coherence argument like we have used so far, because all those coherence arguments concern probability assignments *at a single time.*

In general, figuring out consistency conditions for updating probabilities is harder than doing so for probabilities that are ascribed simultaneously. How could it be otherwise? Life can intrude upon the agent with such force that even the sample space is not maintained. It doesn't make sense to write an equation for how to change my probability for an event E if the event E stops being well-defined! What if something happens that makes all bets about E impossible to settle?

There is a hint of this problem even in the early literature on Dutch-book arguments. In his "Truth and Probability" (1926), F. P. Ramsey says that the probability $P(p|q)$ is what an agent would bet on p in a conditional bet only valid if q is true. And he adds, "This is not the same as the degree to which he would believe p, if he believed q for certain; for knowledge of q might for psychological reasons profoundly alter his whole system of beliefs." I've taken to calling events that force a wholesale reevaluation of beliefs *Ramsey shocks* to emphasize that these departures from naive Bayes-rule clockwork are not a new concept. They were hinted at nearly a century ago. (To be fair, Ramsey's discussion of updating is mostly focused on the Bayes rule. But then again, his tragic death at a young age means that we don't have a finished book of his to go on.)

What we *can* discuss is how an agent should *expect* in the present to potentially change her beliefs in the future. We will now investigate the conditions under which Eq. (2.31) is a reasonable updating scheme. Suppose that at time $t = 0$, the android regards $\$p_0(E)$ as the fair price for a lottery ticket worth \$1 if the event E occurs. Then, at some later time $t = \tau$, the android's evaluation of the world has changed, perhaps in response to new information, and the new fair price is $\$p_\tau(E)$, where $p_\tau(E) < p_0(E)$. There is nothing *irrational* or inconsistent about this: Being willing to sell a ticket at a lower price is just a matter of cutting one's losses.

However, after the android explains this, the bartender proposes a new gamble, this time wagering on the android's own future beliefs. The bartender offers to buy or sell a ticket of the form

$$\boxed{\text{Worth \$1 if } p_\tau(E) = q}. \tag{2.32}$$

Already at $t = 0$, the android can assign a fair price for this ticket, which would be $\$p_0(p_\tau(E) = q)$. If the android is supremely confident that $p_\tau(E)$ will be q, then the fair price of this ticket is \$1.

Now, if $q < p_0(E)$, then the android will be willing to buy a ticket for $\$p_0(E)$ and then sell that ticket for a lower price later. The android faces a sure loss, *one that is already apparent at* $t = 0$, and the bartender has won. Likewise, if $q > p_0(E)$, the android will sell a ticket for $\$p_0(E)$ and buy it back later at a higher price, ensuring a sure loss again.

What can the android do to avoid this eventuality? Defeating the bartender in this scenario requires adhering to the condition

$$p_0(E|p_\tau(E) = q) = q. \tag{2.33}$$

Like all the consistency conditions derived so far, this is a requirement placed upon the current gambling commitments, though in this case, the space of events include the android's future declarations of belief. This condition is an example of the *reflection principle,* an idea due to van Fraassen [12, 4]. It can be derived from a more relaxed assumption: instead of $p_0(p_\tau(E) = q) = 1$, we can deduce Eq. (2.33) if

$$p_0(p_\tau(E) = q) > 0. \tag{2.34}$$

We are now in a position to apply the reflection principle to the task of choosing a probability-updating scheme. Let E be an event, and let Q_q be the event that at time $t = \tau$, we will have $p_\tau(E) = q$. Next, suppose that there exists a set $\{D_q\}$ of possible *data-acquisition events,* such that there is a bijection between $\{D_q\}$ and the possible values for $p_\tau(E)$. Then

$$p_0(E|D_q) = p_0(E|D_q, Q_q) = p_0(E|Q_q). \tag{2.35}$$

The reflection principle states that

$$p_0(E|Q_q) = p_0(E|p_\tau(E) = q) = q. \tag{2.36}$$

Therefore,

$$p_0(E|D_q) = q. \tag{2.37}$$

This is a statement of a gambling commitment made at time $t = 0$: the android is, at $t = 0$, willing to pay $\$q$ for the lottery ticket

$$\boxed{\text{Worth \$1 if } (E \text{ and } D_q). \text{ But money back if (not } D_q).} \tag{2.38}$$

By definition, q is $p_\tau(E)$ if the event Q_q occurs. Consequently, if the android goes ahead and does what the android had been confident about doing, then

$$p_\tau(E) = p_0(E|D_q). \tag{2.39}$$

This is just the rule we guessed in the first place, Eq. (2.31). It is known as the *Bayes rule.*

We can, again, arrive at this point from a more relaxed assumption. The essential requirement is that the android be able to identify at $t = 0$ an event which, the android expects, can determine the future gambling commitments. That is, there must be some event D such that

$$p_0(p_\tau(E) = q|D) = 1. \tag{2.40}$$

What if no such event D exists, and Eq. (2.40) does not hold? For example, the android may expect that whatever happens between $t = 0$ and $t = \tau$, the data will be too ambiguous to warrant upgrading the confidence in any proposition to 100%. Instead, the android expects that some uncertainty will necessarily remain, which can be represented by a probability distribution over data-acquisition events:

$$\sum_D p_\tau(D) = 1. \tag{2.41}$$

In this case, we can use a more general probability-updating scheme, the *Jeffrey rule*:

$$p_\tau(E) = \sum_{D'} p_0(E|D')p_\tau(D'). \tag{2.42}$$

This reduces to the Bayes rule, Eq. (2.31), if

$$p_\tau(D') = \delta_{D,D'}, \tag{2.43}$$

for some event D. And, like Eq. (2.31), the Jeffrey rule (2.42) can be justified using the reflection principle [13].

Diaconis and Zabell [14] give an example where this more general scheme of updating would be applicable:

> suppose we are about to hear one of two recordings of Shakespeare on the radio, to be read by either Olivier or Gielgud, but are uncertain as to which, and have a prior with mass $\frac{1}{2}$ on Olivier and $\frac{1}{2}$ on Gielgud. After hearing the recording, one might judge it fairly likely, but by no means certain, to be by Olivier. The change in belief takes place by direct recognition of the voice; all the integration of sensory stimuli has already taken place at a subconscious level. To demand a list of objective vocal features that we condition on in order to affect the change would be a logician's parody of a complex psychological process.

Furthermore [15],

> If the *only* impact of hearing the recording is to change the odds on Olivier and Gielgud, in the sense that for any A, $P_\tau(A|O) = P_0(A|O)$ and $P_\tau(A|G) = P_0(A|G)$, then after assessing $P_\tau(O)$ we may proceed to apply Jeffrey's rule. (Of course, the former might well *not* be the case; for example the quality of the recording might convey additional information as to its date or manufacture.)[2]

The Jeffrey rule, Eq. (2.42), has the form of a *convex combination*: We are taking a weighted sum of the $p_0(E|D')$, where the weights add up to 1. Convex combinations start showing up all over the place when one considers the reflection principle. Suppose our gambler contemplates a set of events $\{D_i\}$ that are exhaustive and mutually exclusive. For any event E that lies further in the future, plain old coherence requires

$$p_0(E) = \sum_i p_0(E|D_i)p_0(D_i)\,. \tag{2.44}$$

Now, we posit that the gambler expects that the event D_i comes to pass, they will assign the probability q_i to the event E. "D_i" is just another way of saying "$p_\tau(E) = q_i$". The sum becomes

$$p_0(E) = \sum_i p_0(E|p_\tau(E) = q_i)p_0(p_\tau(E) = q_i)\,. \tag{2.45}$$

The reflection principle now kicks in to simplify the last factor:

$$p_0(E) = \sum_i p_0(E|p_\tau(E) = q_i)q_i\,. \tag{2.46}$$

We haven't fixed a definite updating rule here. What we have concluded is that the gambler's probability for E now must be, on pain of incoherence, a convex combination of the probabilities for E that they can foresee updating to.

Example: My local ice cream shop has a freezer cabinet from which they sell individual pints of ice cream. Let's say that the have a shelf of the good ol' standby choices, vanilla and chocolate. On each day, the shelf is stocked with 10 pints. If there are 5 vanilla and 5 chocolate, and I pick a pint at random, then my probability of picking chocolate is 1/2. The shelf can be stocked differently from day to day. I might expect today that tomorrow, it will be restocked *either* with 4 chocolate pints *or* with 6 chocolate pints. Consequently, I will commit today to predicting that my probability of grabbing a chocolate pint tomorrow will be either 0.4 or 0.6. My probability now of grabbing a chocolate pint tomorrow will be a point somewhere in this interval:

$$p_0(E) = 0.4q + 0.6(1 - q)\,, \tag{2.47}$$

where q is my gambling commitment for how I believe the freezer will be stocked. Note that $p_0(E)$ is *not* necessarily what I *will* come to believe after I see the freezer

[2] I have adjusted the notation slightly in this passage to be consistent with the rest of this section.

restocked: That will be either 0.4 or 0.6. Nor is it necessarily how I will gamble on grabbing a pint of ice cream *today*. The event E is off in the future. Pulling a pint from the freezer today is a different experiment, for which my probability of chocolate could be 0.4, 0.6, 0, 1, whatever.

In general, we may say that the updating of probabilities is a subtle subject. The next section will relate this question, which seems to live in the realm of machine learning, to mathematical biology. When we establish this connection, the idea of mesoscale environmental structure will make an appearance.

2.3 An Analogy with Evolutionary Dynamics

Let $\{H_i\}$ be a set of n mutually exclusive and exhaustive hypotheses, so that at any time t,

$$\sum_{i=1}^{n} p_t(H_i) = 1. \tag{2.48}$$

In the previous section, we showed that at least under some circumstances, if an event E happens in between times $t = 0$ and $t = \tau$, we are justified in updating the probabilities according to the following rule:

$$p_\tau(H_i) = \frac{p_0(E|H_i)p_0(H_i)}{p_0(E)}. \tag{2.49}$$

This follows from the simple method for updating probabilities, the Bayes rule that we wrote down in Eq. (2.31).

Now, we jump sideways and consider a simple model of evolutionary dynamics in a panmictic population [16]. We suppose there are n types of organism. These could be different species, different genotypes in the same species, or in principle, genetically identical individuals who adhere to different social behaviors. We represent the configuration of the population by an n-tuple of nonnegative real numbers:

$$x = (x_1, x_2, \ldots, x_n). \tag{2.50}$$

By assuming panmixia, we deliberately blur over all spatial organization or other kinds of population structure. We neglect stochasticity, and we assume that at each instant, the population configuration is definitely specified. In other words, for this problem we consider evolution as a *deterministic* dynamical system, one which we will take to operate in discrete time. We establish the dynamics by writing an update rule, which yields a new tuple x' when given x.

That a model which neglects all stochasticity should relate to probability theory may be surprising, but we shall soon see that it is the case, and the relationship is quite direct.

To implement the idea of natural selection in this context, we introduce a *fitness function* which maps population tuples to real numbers. Each of the n types has its own fitness function:

$$f_i = f_i(x_1, x_2, \ldots, x_n). \tag{2.51}$$

Types which are more fit should be represented more strongly in the next generation. Our update rule for x_i should, consequently, have the form

$$x_i' \propto x_i f_i(x). \tag{2.52}$$

It is convenient to keep the overall population size constant, and in that case, we might as well use the x_i to represent proportions:

$$\sum_{i=1}^{n} x_i = 1. \tag{2.53}$$

To ensure that this normalization is maintained, we introduce the average fitness

$$\bar{f}(x) = \sum_{i=1}^{n} x_i f_i(x). \tag{2.54}$$

Our dynamical update law, the *replicator equation,* is

$$x_i' = \frac{x_i f_i(x)}{\bar{f}(x)}. \tag{2.55}$$

This is formally analogous [17] to the rule for updating probabilities by conditionalization, Eq. (2.49). To see the relationship, we make the following substitutions:

$$\{p_0(H_i)\} \to x = (x_1, \ldots, x_n), \tag{2.56}$$
$$p_0(E|H_i) \to f_i(x), \tag{2.57}$$
$$p_0(E) \to \bar{f}(x), \tag{2.58}$$
$$p_\tau(H_i) \to x' = (x_1', \ldots, x_n'). \tag{2.59}$$

Natural selection, in the situation modeled by the replicator equation, can be thought of as a learning process, in which the population gains information about the fitness landscape.

What about the more flexible Jeffrey rule? Adapting Eq. (2.42) to this context, we find that

$$p_\tau(H_i) = \sum_{E'} p_0(H_i|E')p_\tau(E') = \sum_{E'} \frac{p_0(H_i)p_0(E'|H_i)}{p_0(E')} p_\tau(E'). \tag{2.60}$$

This reduces to the simpler updating scheme, Eq. (2.49), if $p_\tau(E') = \delta_{E,E'}$. If $p_\tau(E')$ is not a delta function, then multiple potential events E' are relevant. The evolutionary analogue of this is the possibility of *multiple fitness landscapes.*[3]

Imagine that we have a test tube filled with various types of bacteria. We pipette a fraction w_j of its contents into each of m new test tubes. The conditions in these tubes can differ from one another: perhaps they are illuminated under different frequencies

[3] To my knowledge, the question of an evolutionary analogue of the Jeffrey rule hasn't been raised before. Harper [17] and Baez [18], for example, stop with the Bayes rule, Eq. (2.49).

of light, or they contain varying nutrient mixtures. The bacteria interact, and their population proportions change due to natural selection. We then pour the contents of the m tubes all back together again. The total population size remains constant throughout the whole process.

Let $f_i^{(j)}$ be the fitness function for type i in tube j. As before, the fitness depends on the proportions of the different types present in the environment. If all of the populations are scaled down by the same factor w_j, the relative proportions remain the same. Therefore, the initial proportions in each of the new tubes are the same as those in the original source, and we can write the fitnesses as $f_i^{(j)}(x)$. The mean fitness in tube number j is

$$\bar{f}^{(j)}(x) = \sum_i x_i f_i^{(j)}(x), \tag{2.61}$$

and this is true no matter what the value of w_j.

Because a fraction w_j of each x_i experiences the environment in the j^{th} tube, the new value x_i' will be

$$x_i' = x_i \sum_{j=1}^{m} \frac{f_i^{(j)}(x)}{\bar{f}^{(j)}(x)} w_j, \text{ with } \sum_{j=1}^{m} w_j = 1. \tag{2.62}$$

This is the analogue of the Jeffrey rule, Eq. (2.60).

2.4 Biased Coin-Flips and Urn Models

Consider the prototypical probability problem of repeatedly flipping a coin. Let p be the probability ascribed to the outcome that the coin comes up heads. What probability should we ascribe to the event E_m of seeing the coin land on heads exactly m times out of N? For this event to transpire, we must see heads m times and tails $N-m$ times. The probability of first seeing m instances of heads, followed by $N-m$ instances of tails, is $p^m(1-p)^{N-m}$. However, this is not the only way the event E_m can happen: Our specification of the event E_m coarse-grains over the details of the order in which the desired outcomes arrive. So, the probability $p(E_m)$ is the value we computed before, multiplied by the number of ways we can distribute the m heads throughout the total of N flips:

$$p(E_m) = \binom{N}{m} p^m (1-p)^{N-m}, \tag{2.63}$$

This defines the *binomial distribution,* in which we have used the binomial coefficients [19] more explicitly given as

$$\binom{N}{m} = \frac{N!}{m!(N-m)!}. \tag{2.64}$$

We can think of this scenario in a slightly different way, which leads to some useful modifications. Consider an urn filled with a total of N_B balls, and let $R = N_B p$, where

p is the parameter we used before. Exactly R of the balls in the urn are red, the other $G = N_B - R$ being green. We draw a ball from the urn at random, check its color and drop it back into the urn. The probability that the ball we pick will be red is just p. If we repeat the draw-and-replace operation N times, the probability that we see red in exactly m trials is given by Eq. (2.63).

What happens if we do *not* replace a ball after we withdraw it? Then, the population of the urn changes from one trial to the next. Call E_m the event of seeing red in exactly m trials out of N. The probability of this event is

$$p(E_m) = \frac{(\#\text{ of ways to get } m \text{ red balls}) \times (\#\text{ of ways to get } N-m \text{ green balls})}{(\text{total } \#\text{ of ways to make a selection})}. \tag{2.65}$$

Each factor in this expression can be found using combinatorics. In fact, each quantity which enters this formula is a binomial coefficient:

$$\begin{aligned}(\#\text{ of ways to get } m \text{ red balls}) &= \binom{R}{m}, \\ (\#\text{ of ways to get } N-m \text{ green balls}) &= \binom{G}{N-m}, \\ (\text{total } \#\text{ of ways to make a selection}) &= \binom{R+G}{N}.\end{aligned} \tag{2.66}$$

Putting these together, we have

$$p(E_m) = \frac{\binom{R}{m}\binom{G}{N-m}}{\binom{R+G}{N}}. \tag{2.67}$$

This defines the *hypergeometric distribution.* Expanding out the binomial coefficients per Eq. (2.64),

$$p(E_m) = \frac{R!G!}{m!(R-m)!(N-m)!(G-N+m)!}\frac{N!(R+G-N)!}{(R+G)!}. \tag{2.68}$$

The third variation of the urn problem is the following: Every time we draw out a ball, we check its color, and we return *two* balls of that color to the urn. This is the *Pólya urn model* [20, 21], and in this case, if we begin with R red balls and G green balls,

$$p(E_m) = \frac{N!}{m!(N-m)!}\frac{(m+R-1)!(N-m+G-1)!}{(N+R+G-1)!}\frac{(R+G-1)!}{(R-1)!(G-1)!}. \tag{2.69}$$

In terms of the gamma function,

$$p(E_m) = \frac{\Gamma(N+1)}{\Gamma(m+1)\Gamma(N-m+1)}\frac{\Gamma(m+R)\Gamma(N-m+G)}{\Gamma(N+R+G)}\frac{\Gamma(R+G)}{\Gamma(R)\Gamma(G)}. \tag{2.70}$$

This is known as the beta-binomial distribution.

These probability distributions are related in another way, other than their common appearance in variations of the urn scenario [22]. We can find this relationship by revisiting the conditionalization rules that we studied in the previous sections. Take the Bayes rule (2.49), which effects a map from one probability distribution to another:

$$p(\theta) \to \frac{p(y|\theta)p(\theta)}{p(y)}. \tag{2.71}$$

When confronted with a transformation, we typically like to know what it leaves unchanged. For example, the eigenvectors of a matrix are the vectors which the transformation represented by that matrix changes only by a scaling factor. Likewise, it is useful to identify probability distributions which, under the conditionalization map, keep the same form. The input to the map is known as the *prior,* and its output is the *posterior.* If the prior and posterior have the same functional form, then they are *conjugate.* This relation is dependent on the likelihood function $p(y|\theta)$ used in the mapping, but that can often be taken as fixed on other grounds.

Suppose that

$$p(y|\theta) = \binom{N}{y}\theta^y(1-\theta)^{N-y}, \tag{2.72}$$

for a positive integer N and all integers y satisfying $0 \leq y \leq N$. Then, the conjugate prior of the binomial distribution (2.63) is the Beta distribution,

$$p(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}, \tag{2.73}$$

and the conjugate prior of the hypergeometric distribution (2.67) is the beta-binomial, Eq. (2.70).

We now have the tools to dig a little more deeply. Consider a scenario in which we wish to *repeat* an experiment. That is, we have the budget to carry out a long experiment, made up of many successive trials [23]. We can represent the outcome of each trial by a random variable x_j, and we can assign a joint probability distribution $p(x_1, x_2, \ldots, x_N)$ over the possible outcomes of an N-trial experiment. Motivated by common experiences in the workaday life of a scientist, we now impose two conditions on this joint probability distribution. These conditions will help dramatically in narrowing down the possible forms of the joint distribution $p(x_1, x_2, \ldots, x_N)$. First, we require that it be *finitely exchangeable*: its value is invariant under permutations of its arguments. If π is any permutation of the numbers $\{1, \ldots, N\}$, then

$$p(x_1, \ldots, x_N) = p(x_{\pi(1)}, \ldots, x_{\pi(N)}). \tag{2.74}$$

Second, we require that $p(x_1, \ldots, x_N)$ be extensible, in the following manner. For any integer $M > 0$, there is a finitely exchangeable distribution with more arguments, p_{N+M}, such that

$$p(x_1, \ldots, x_N) = \sum_{x_{N+1}, \ldots, x_{N+M}} p_{N+M}(x_1, \ldots, x_N, x_{N+1}, \ldots, x_{N+M}). \tag{2.75}$$

These two requirements make precise the idea that our probability assignment p derives from an arbitrarily long sequence of random variables, the order of which is inconsequential. We say that a p which satisfies both conditions, finite exchangeability and extensibility, is *exchangeable.*

Let us say we have an exchangeable probability assignment for M binary random variables $x_1, \ldots, x_M$. Define $p(n, N)$ to be the probability of obtaining n 1's in N trials. Because the n appearances of the outcome 1 can arrive in any order, this is

$$p(n, N) = \binom{N}{n} p(x_1 = 1, \ldots, x_n = 1, x_{n+1} = 0, \ldots, x_N = 0). \tag{2.76}$$

This is equal to

$$p(n, N) = \binom{N}{n} \sum_{m=0}^{M} p(x_1 = 1, \ldots, x_n = 1, x_{n+1} = 0, \ldots, x_M = 0 | m, M) p(m, M). \tag{2.77}$$

Exchangeability means that all sequences with m occurrences of "1" in M trials are equiprobable. So, we have again an urn problem, specifically, the one that led us to the hypergeometric distribution. If we stock an urn with M balls, m of which are red, then the hypergeometric distribution tells us how likely we are to find n red balls in N draws made without replacement.

$$p(x_1 = 1, \ldots, x_n = 1, x_{n+1} = 0, \ldots, x_M = 0 | m, M) = \frac{\binom{m}{n}\binom{M-m}{N-n}}{\binom{M}{N}}. \tag{2.78}$$

Defining

$$(r)_q = \prod_{j=0}^{q-1} (r - j) = \frac{r!}{(r-q)!}, \tag{2.79}$$

we have after a spot of algebra that

$$p(n, N) = \binom{N}{n} \sum_{m=0}^{M} \frac{(m)_n (M-m)_{N-n}}{(M)_N} p(m, M). \tag{2.80}$$

We can rewrite this sum as an integral:

$$p(n, N) = \binom{N}{n} \int_0^1 dz \, \frac{(zM)_n [(1-z)M]_{N-n}}{(M)_N} P_M(z), \tag{2.81}$$

where we have defined

$$P_M(z) = \sum_{m=0}^{M} p(zM, M) \delta(z - m/M). \tag{2.82}$$

Thanks to the extendibility property, we can take the limit $M \to \infty$. In this limit, we find that $P_M(z)$ approaches a continuous curve, $P_\infty(z)$, and the rest of the integrand

goes to $z^n(1-z)^{N-n}$.

$$p(n,N) = \binom{N}{n} \int_0^1 dz\, z^n (1-z)^{N-n} P_\infty(z). \tag{2.83}$$

Interestingly, the result has the form of an integral over what we might call a "meta-probability." Note that $z^n(1-z)^{N-n}$, times the binomial coefficient out front, has the form of a binomial distribution with probability equal to z. The curve $P_\infty(z)$ has, by construction, the normalization property of a probability density:

$$\int_0^1 dz\, P_\infty(z) = 1. \tag{2.84}$$

It is as if we can write the function $p(n,N)$ in terms of "the meta-probability $P_\infty(z)$ that the probability is z."

Eq. (2.83) can be generalized readily enough beyond the case of binary random variables. The result is *de Finetti's theorem.* Let Δ_k denote the space of valid probability assignments over k outcomes:

$$\Delta_k = \left\{ \vec{p} : p_j \geq 0 \text{ for all } j \text{ and } \sum_j p_j = 1 \right\}. \tag{2.85}$$

Then, exchangeability implies that

$$p(x_1, \ldots, x_N) = \int_{\Delta_k} d\vec{p}\, P(\vec{p})\, p_{x_1} \cdots p_{x_N} = \int_{\Delta_k} d\vec{p}\, P(\vec{p})\, p_1^{n_1} \cdots p_k^{n_k}, \tag{2.86}$$

where $P(\vec{p})$ is properly normalized over Δ_k:

$$\int_{\Delta_k} d\vec{p}\, P(\vec{p}) = 1. \tag{2.87}$$

When we used the parable of the Ferengi bartender to motivate the basic rules of probability, the story left no room for an "unknown probability," per se. The android's probability assignments at any time are known to the android, and the idea of an "unknown known" sounds like a contradiction in terms. However, de Finetti's theorem gives meaning to it. The locution "unknown probability" is a *shorthand for a scenario in which the android is gambling on a sequence of trials that the android judges to be exchangeable.*

2.5 Shannon Information

Intuitively speaking, *information* is that which removes our uncertainty. We can quantify an amount of information by specifying how many questions are necessary to overcome the uncertainty we have about something. For example, suppose we have

an experiment which yields one of M different outcomes, stochastically. If we repeat this experiment N times, the number of different possible sequences of outcomes is M to the N^{th} power. Each of these sequences can be represented as a string of symbols, drawn from an alphabet of size M. The number of yes-or-no questions we would need to narrow down this set of possibilities to a single result is $\log_2 M^N = N \log_2 M$. However, we may have *expectations* about the experiment, meaning that we will find some strings of outcomes less surprising than others. Suppose that we ascribe a probability p_i to each of the M possible outcomes of an individual trial in the sequence. We have a set of nonnegative numbers which together satisfy

$$\sum_{i=1}^{M} p_i = 1. \tag{2.88}$$

This set defines a *random variable.* We do not know in advance exactly how many times the i^{th} outcome will occur, but the *number of such occurrences we will find least surprising* is easily computed:

$$N_i = N p_i. \tag{2.89}$$

How many of the M^N possible strings are, in terms of our probability assignment, unsurprising? A minimally surprising string is one where each of the M symbols in our alphabet occurs N_i times. The number of such strings having total length N is, by elementary combinatorics,

$$K = \frac{N!}{\prod_{i=1}^{M} N_i!} = \frac{N!}{\prod_{i=1}^{M} (N p_i)!}. \tag{2.90}$$

How many bits do we need to specify one message out of a set of *this* size? Applying the algebraic properties of logarithms, this is

$$\log_2 K = \log_2 N! - \sum_{i=1}^{M} \log_2 (N p_i)!. \tag{2.91}$$

Here, it is useful to invoke Stirling's approximation,

$$\log N! \approx N \log N - N. \tag{2.92}$$

With this, we can evaluate the natural log of K as

$$\log K \approx N \log N - N - \sum_i \left[N p_i \log(N p_i) - N p_i \right]. \tag{2.93}$$

We have assumed here that each of the N_i is large enough for Stirling's approximation to be viable. Now, we simplify. By normalization, the p_i must sum up to one, so we can cancel the second and final terms, leaving

$$\log K \approx N \log N - N \sum_i p_i \left[\log N + \log p_i \right]. \tag{2.94}$$

Invoking normalization again, we see that we are left with

$$\log K \approx -N \sum_i p_i \log p_i, \tag{2.95}$$

which we can easily convert to a base-2 logarithm by division. Then, if we divide by N, we can obtain a result in terms of bits per symbol.

This result provides us with a *measure of information* associated with the probability distribution $\{p_i\}$. If the probabilities p_i express our expectations of the stochastic experiment, then the number of yes-or-no questions which we should expect to require in order to remove our uncertainty about each iteration of that experiment is

$$H[\{p_i\}] = -\sum_i p_i \log_2 p_i. \tag{2.96}$$

This is the *Shannon information* of the probability distribution, also known variously as the Shannon entropy and the Shannon index. The base of the logarithm is a convention which depends on the subfield of science that one is currently studying. The quantity $H[\{p_i\}]$ vanishes if $p_i = \delta_{ij}$ for some j, and it is maximized by the uniform probability distribution over all i. This expresses the fact that if our expectations favor all possibilities equally, we cannot expect to gain any advantage by asking questions cleverly (as we could, for example, with English text: "Is the next letter an E?"). Contrariwise, if we are supremely confident that a particular outcome will obtain, we expect that we will require zero questions to ascertain the result.

One way to think of the Shannon information is as an "expected surprise". If Alice puts a low probability $p(E)$ upon an event E, she is saying that the event E happening would be a surprise to her. Accordingly, the *surprisal* of an event is defined to be $-\log p(E)$. This map turns the unit interval $[0, 1]$ into the much wider interval $[0, \infty)$, emphasizing the drama inherent in ascribing probabilities of 0 or 1. Given a probability distribution p_i, the expectation value of the surprisal is

$$\langle -\log p_i \rangle = \sum_i p_i (-\log p_i), \tag{2.97}$$

which is just the Shannon index again.

This perspective suggests a generalization. Suppose that Alice has decided she is willing to gamble according to the probability distribution p_i, while Bob has declared his gambling commitments are a distribution q_i, which may be different from Alice's. What is *Alice's expectation for how suprised Bob will be?* She can calculate this by taking a weighted average of Bob's surprisals, using her probabilities as the weights:

$$\langle -\log q_i \rangle = \sum_i p_i (-\log q_i). \tag{2.98}$$

Alice then compares this to how she expects her own level of surprise to go, taking the difference

$$\langle -\log q_i \rangle - \langle -\log p_i \rangle = \sum_i p_i \log \frac{p_i}{q_i}. \tag{2.99}$$

This defines the *relative entropy* between the probability distributions p_i and q_i. It vanishes when they are identical.

We can further the concept of Shannon information if we consider not just one random variable, but combinations of them. If we have two random variables X and Y and a joint probability distribution $p(x, y)$, the total joint information of X and Y considered together is, by the Shannon formula,

$$H(X, Y) = -\sum_{x,y} p(x, y) \log p(x, y). \tag{2.100}$$

If we only care about one of the two random variables, we can sum over the other:

$$p_X(x) = \sum_y p(x, y), \; p_Y(y) = \sum_x p(x, y). \tag{2.101}$$

In turn, we can use these probability distributions in the Shannon formula just as well.

It is helpful to define a measure of information shared between two random variables, following the relation

$$I(X; Y) = H(X) + H(X) - H(X, Y). \tag{2.102}$$

To find a formula for this *mutual information* in terms of Shannon indices, we start by re-expressing the Shannon indices of the individual variables in terms of their probability distributions:

$$I(X; Y) = -H(X, Y) - \sum_x p_X(x) \log p_X(x) - \sum_y p_Y(y) \log p_Y(y). \tag{2.103}$$

This, in turn, is the same as

$$\begin{aligned} I(X; Y) &= -H(X, Y) - \sum_{x,y} p(x, y) \log p_X(x) - \sum_{x,y} p(x, y) \log p_Y(y) \\ &= -H(X, Y) + \sum_{x,y} p(x, y)[-\log p_X(x) - \log p_Y(y)]. \end{aligned} \tag{2.104}$$

Using the definition of $H(X, Y)$ and the familiar properties of logarithms, we arrive at the mutual information between X and Y:

$$I(X; Y) = \sum_{x,y} p(x, y) \log \left(\frac{p(x, y)}{p_X(x) p_Y(y)} \right). \tag{2.105}$$

Note that $I(X; Y)$ vanishes if $p(x, y)$ factors neatly into $p_X(x)$ and $p_Y(y)$. In this case, knowing the value of x does not reduce the number of questions we require to ascertain the value of y, and vice versa. We can also see that the mutual information is the relative entropy of $p(x, y)$ and its "decorrelated" version $p_X(x)p_Y(y)$.

The Shannon index is a *natural* information function in the context of probabilities and random variables, because it is the unique functional which satisfies certain basic

and fairly intuitive desiderata [24]. First, the value of information function applied to a probability distribution should not change if we expand the set of outcomes with new elements which are assigned probability zero. Second, an information function should be invariant under permutations of the probabilities: we are just as uncertain if our expectations are

$$p_0 = x,\ p_1 = 1 - x, \tag{2.106}$$

as we are if

$$p_0 = 1 - x,\ p_1 = x. \tag{2.107}$$

Third, for two experiments represented by random variables X and Y, an information function should always satisfy

$$H(X \cup Y) \leq H(X) + H(Y). \tag{2.108}$$

Fourth, if the random variables X and Y are *independent,* we should have equality in the previous relation.

If we impose these four desiderata, then the only possible information functions are linear combinations of the Shannon index and the quantity

$$K[\{p_i\}] = \log_2 |\{p_i | p_i \neq 0\}|, \tag{2.109}$$

which counts the number of outcomes assigned greater than zero probability. This is known as the *Hartley entropy.* We can rule out a contribution of this form if we require in addition that for a random variable which has two possible outcomes, the information is small if the probability of one outcome is near zero.

Many other derivations of the Shannon index from sets of desiderata have been made. For a recent example, see Baez, Fritz and Leinster [25].

2.6 Moments and Cumulants

The mean of a probability distribution $p(x)$ is

$$\langle x \rangle \equiv \sum_{x=\infty}^{\infty} p(x) x, \tag{2.110}$$

or for a continuous distribution,

$$\langle x \rangle \equiv \int dx\, p(x) x. \tag{2.111}$$

This expression readily generalizes to that for the *expectation value* of a function f,

$$\langle f(x) \rangle \equiv \int dx\, f(x) p(x). \tag{2.112}$$

Expectation values of powers of x are called *moments* of x:

$$\langle x^n \rangle = \int dx\, p(x) x^n. \tag{2.113}$$

Moments are interesting things to know about a probability distribution, but they are not always the most convenient values for characterizing how such distributions work. For example, suppose we have two random variables, X and Y, which independently of one another take values according to probability distributions $p_X(x)$ and $p_Y(y)$. What can we say about a third random variable $Z \equiv X + Y$? And, is there some characterization of p_X and p_Y such that we can just add properties of p_X and p_Y to obtain the corresponding property of p_Z?

If Z takes the value z and X takes the value x, then Y must have taken the value $z - x$. Following the basic rules of probability, we deduce that because X and Y are independent, the probability of X taking the value x and Y the complementary value $z - x$ is $p_X(x)p_Y(z - x)$. But this is only one way for a measurement of Z to give the value z; the sum $x + y$ can work out to z for any value of x. So, adding up the probabilities, we find

$$p_Z(z) = \int dx\, p_X(x) p_Y(z - x). \tag{2.114}$$

The probability distribution for the sum of two random variables is the *convolution* of the distributions for the random variables being added. This indicates that the higher moments of p_X and p_Y will not combine neatly when we construct p_Z.

A useful result, the *convolution theorem*, says that the Fourier transform of a convolution is the product of the Fourier transforms of the functions convolved, up to a constant depending on how the transformation was normalized. (The proof is both standard and fairly direct, requiring only the ability to recognize a delta function; see a good text on mathematical methods [26].) This suggests a way to proceed.

The *characteristic function* of a probability distribution is just its Fourier transform, which we can write as an expectation value:

$$\tilde{p}(k) \equiv \left\langle e^{-ikx} \right\rangle = \int dx\, p(x) e^{-ikx}. \tag{2.115}$$

Thus,

$$p(x) = \frac{1}{2\pi} \int dk\, \tilde{p}(k) e^{ikx}. \tag{2.116}$$

When we add random variables, their characteristic functions multiply, which means that the *logarithms* of their characteristic functions *add.* So, if we want to define quantities which are like moments but which add conveniently when we combine random variables, we should look at logarithms of characteristic functions.

Recall that

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots, \tag{2.117}$$

so we can expand the characteristic function as

$$\tilde{p}(k) = \left\langle \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} x^n \right\rangle = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle. \tag{2.118}$$

This relates the characteristic function to the moments.

As indicated, we're interested in logarithms of characteristic functions, because characteristic functions multiply under convolution, and logarithms turn products into sums. Expanding the logarithm of $\tilde{p}(k)$ as we expanded $\tilde{p}(k)$ itself, we define the *cumulants* of x, written with the notation $\langle x^n \rangle_c$.

$$\log \tilde{p}(k) \equiv \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle_c \,. \tag{2.119}$$

When manipulating logarithms, it is sometimes useful to recall the Taylor expansion of the logarithm function for arguments near 1:

$$\log(1+\epsilon) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\epsilon^n}{n}. \tag{2.120}$$

How do the cumulants relate to the moments?[4] Can we find a simple formula for one set of numbers in terms of the other? We have two formulas involving the characteristic function $\tilde{p}(k)$, one using moments and the other using cumulants. If we equate the expressions for $\tilde{p}(k)$ in Eqs. (2.118) and (2.119), we can derive the relationship between the moments $\{\langle x^n \rangle\}$ and the cumulants $\{\langle x^n \rangle_c\}$. From Eqs. (2.118) and (2.119), we see that

$$\begin{aligned} \sum_{m=0}^{\infty} \frac{(-ik)^m}{m!} \langle x^m \rangle &= \exp\left[\sum_{l=1}^{\infty} \frac{(-ik)^l}{l!} \langle x^l \rangle_c \right] \\ &= \prod_{l=1}^{\infty} \exp\left[\frac{(-ik)^l}{l!} \langle x^l \rangle_c \right] \\ &= \prod_{l=1}^{\infty} \sum_{n_l=0}^{\infty} \left[\frac{(-ik)^{l n_l}}{n_l!} \left(\frac{\langle x^l \rangle_c}{l!} \right)^{n_l} \right]. \end{aligned} \tag{2.121}$$

The powers of $(-ik)^m$ on both sides have to be equal, so

$$\frac{\langle x^m \rangle}{m!} = {\sum_{\{n_l\}}}' \prod_l \frac{\langle x^l \rangle_c^{n_l}}{n_l! (l!)^{n_l}}, \tag{2.122}$$

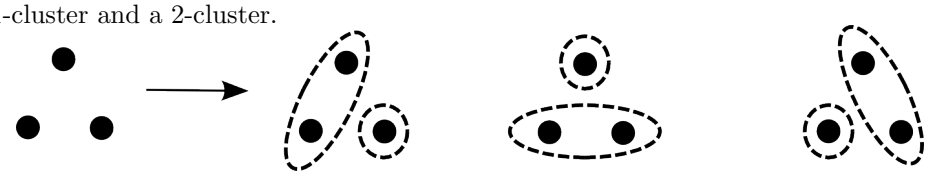
in which the prime on the Σ means we are restricting the sum such that $\sum l n_l = m$. Moving the $m!$ to the right-hand side, we find

$$\langle x^m \rangle = {\sum_{\{n_l\}}}' m! \prod_l \frac{1}{n_l! (l!)^{n_l}} \langle x^l \rangle_c^{n_l} \,. \tag{2.123}$$

The numerical factor multiplying $\langle x^l \rangle_c^{n_l}$ turns out to have a fairly simple combinatorial interpretation: it is the number of ways of breaking m points into clusters, satisfying the condition that for each value of l, the assignment has n_l clusters of l points.

[4] This discussion is based on Kardar [27, 28], but with more intermediate steps worked out.

This is easiest to see if we work out a simple example first. Suppose we have three points, which we wish to regroup into a set of two points and a third point left to itself. We can do this in three distinct ways, grouping together 1 and 2, 1 and 3 or 2 and 3. In other words, there exist three ways to assign a total of three particles to a 1-cluster and a 2-cluster.



If we write n_1 for the number of 1-clusters and n_2 for the number of 2-clusters in our decomposition, we can say that the number of ways W to organize our 3-point set is

$$W(n_1 = 1, n_2 = 1) = 3. \tag{2.124}$$

We now explore the behavior of W for a general set of m points. First, because every point has to be placed within one cluster or another, the cluster sizes times the number of clusters used have to add up to the total size of the set:

$$\sum_l l n_l = m. \tag{2.125}$$

We know that we can order an m-element set in $m!$ ways. We can find a general formula for W if we divide these $m!$ total permutations by the number of equivalent ones, that is, by the number of permutations which give indistinguishable results. Inside each subset of l elements, we can permute the labels in $l!$ different ways and still get an equivalent grouping. Furthermore, we can rearrange n_l subsets in $n_l!$ ways. The former gives us a factor of $(l!)^{n_l}$, and the latter a factor of $n_l!$. Thus,

$$W(\{n_l\}) = \frac{m!}{\prod_l n_l!(l!)^{n_l}}. \tag{2.126}$$

To sanity-check this formula, note that for $n_1 = n_2 = 1$,

$$W(1,1) = \frac{3!}{(1!1!^1)(1!2!^1)} = \frac{3 \cdot 2}{2} = 3. \tag{2.127}$$

Putting everything together, we use this combinatorial notation to rewrite Eq. (2.123):

$$\langle x^m \rangle = \sum_{\{n_l\}}{}' W(\{n_l\}) \prod_l \left\langle x^l \right\rangle_c^{n_l}, \text{ where } \sum_l l n_l = m. \tag{2.128}$$

Given any sequence $\{\langle x^l \rangle_c\}$ of cumulants, we can compute the moments. This means that if we have two independent random variables X and Y, described by the probability assignments p_X and p_Y, we can compute the moments of the new random variable $X + Y$. We will apply this to very good effect in section §2.8.

It is a straightforward matter to work out the first few moments in terms of the first few cumulants. (And these relations are convenient to have on hand for reference.)

First, because there's only one way to make a linked cluster out of one point, the first moment is the same as the first cumulant:

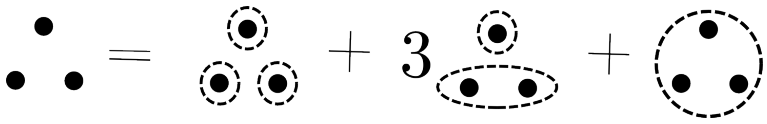
$$\langle x \rangle = \langle x \rangle_c \,. \tag{2.129}$$

The second moment, $\langle x^2 \rangle$, will have two terms, because we can make a diagram with two disconnected points (that is, two one-point clusters) and a diagram with two connected points (a 2-cluster).

$$\langle x^2 \rangle = \langle x^2 \rangle_c + \langle x \rangle_c^2 \tag{2.130}$$

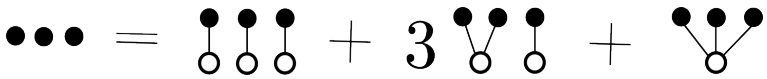
The third moment $\langle x^3 \rangle$ is slightly more complicated, since we can chop up the three-vertex diagram into a 2-cluster and a 1-cluster in three ways. (We worked this out explicitly earlier.) This means that our formula will have a symmetry factor.

In diagram form,



(2.131)

An alternate way to draw this picture is as a forest of trees:



(2.132)

Here, the filled circles on the upper row indicate factors of x, and they belong to the same group if they are linked to the same circle on the bottom row. Each group, which stands for a cumulant, is a rooted tree, and each little forest is a term in the expansion of $\langle x^3 \rangle$. Either way, the equivalent in algebraic symbols is

$$\langle x^3 \rangle = \langle x^3 \rangle_c + 3 \langle x^2 \rangle_c \langle x \rangle_c + \langle x \rangle_c^3 \,. \tag{2.133}$$

The expectation value of x^4 works analogously. Expanded out as products of cumulants, $\langle x^4 \rangle$ is a sum of five terms:

$$\langle x^4 \rangle = \langle x^4 \rangle_c + 4 \langle x^3 \rangle_c \langle x \rangle_c + 3 \langle x^2 \rangle_c^2 + 6 \langle x^2 \rangle_c \langle x \rangle_c^2 + \langle x \rangle_c^4 \,. \tag{2.134}$$

Note that the total degree of each term on the right-hand side is equal to 4. If x is a quantity which carries units, this is necessary for dimensional consistency. The coefficients are combinatorial symmetry factors that come from the number of ways a set of four points can be partitioned. There are four ways to pull out a single point, three ways to divide the set into two pairs, and six ways to group together a pair of points while leaving the other two points out. As before, we have multiple equivalent ways to derive this: by counting ways to circle points, by counting our options to make forests out of trees, or by plugging into our algebraic formula for $W(\{n_l\})$. In fact, we

can also build up these formulas recursively, by considering the operation of merging a new tree into a forest in all possible ways.

We have derived the first four moments in terms of the first four cumulants, using Eq. (2.128). Inverting these relationships yields the first four cumulants in terms of the first four moments:

$$\begin{aligned}
\langle x \rangle_c &= \langle x \rangle \\
\langle x^2 \rangle_c &= \langle x^2 \rangle - \langle x \rangle^2 \\
\langle x^3 \rangle_c &= \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 2\langle x \rangle^3 \\
\langle x^4 \rangle_c &= \langle x^4 \rangle - 4\langle x^3 \rangle \langle x \rangle - 3\langle x^2 \rangle^2 + 12\langle x^2 \rangle \langle x \rangle^2 - 6\langle x \rangle^4 .
\end{aligned} \tag{2.135}$$

The first cumulant is just the mean. The second cumulant, the *variance,* is a measure of spread about the mean. Together, they are the cumulants referred to and made use of most regularly. Soon, in §2.8, we will see why this is the case. The relations in Eq. (2.135) are, in practice, about as high-order as one needs to calculate. Cumulants $\langle x^n \rangle_c$ with $n > 4$ aren't invoked as often.

The coefficients in Eq. (2.134) show up somewhere else, too. Consider two functions, $f(x)$ and $g(x)$. Composing them yields a new function, $f(g(x))$, and the chain rule tells us its derivative:

$$\frac{d}{dx}\left[f(g(x))\right] = \frac{df}{dg}\frac{dg}{dx}. \tag{2.136}$$

Now, let's differentiate again. Using the product rule,

$$\frac{d^2}{dx^2}\left[f(g(x))\right] = \frac{d^2 f}{dg^2}\left(\frac{dg}{dx}\right)^2 + \frac{df}{dg}\frac{d^2 g}{dx^2}. \tag{2.137}$$

We can repeat this process, cranking through the algebra:

$$\frac{d^3}{dx^3}\left[f(g(x))\right] = 3\left(\frac{dg}{dx}\right)\frac{d^2 f}{dg^2}\frac{d^2 g}{dx^2} + \left(\frac{dg}{dx}\right)^3 \frac{d^3 f}{dg^3} + \frac{df}{dg}\frac{d^3 g}{dx^3}. \tag{2.138}$$

And, taking it one more step,

$$\begin{aligned}
\frac{d^3}{dx^3}\left[f(g(x))\right] = &\frac{d^4 f}{dg^4}\left(\frac{dg}{dx}\right)^4 \\
&+ 6\frac{d^3 f}{dg^3}\frac{d^2 g}{dx^2}\left(\frac{dg}{dx}\right)^2 \\
&+ 3\frac{d^2 f}{dg^2}\left(\frac{d^2 g}{dx^2}\right)^2 \\
&+ 4\frac{d^2 f}{dg^2}\frac{d^3 g}{dx^3}\frac{dg}{dx} \\
&+ \frac{df}{dg}\frac{d^4 g}{dx^4}.
\end{aligned} \tag{2.139}$$

This is sufficient to see a pattern take shape.

The coefficients in each sum are the same numbers, $W(\{n_l\})$, which we encountered when converting between moments and cumulants. In the ordinary chain rule, we have one term, and its coefficient is 1. For the second derivative, we get two terms, each of which has a coefficient of 1. In the third derivative, one term has a coefficient of 3, and when we go to the fourth derivative, the coefficients are the same as those in the expression for the moment $\langle x^4 \rangle$ in terms of the first four cumulants.

We can make the constructions exactly analogous if we identify a *connected cluster of k points* with the k^{th} derivative of $g(x)$. Each term is a derivative of f, times one or more factors which are derivatives of g. The diagram we used to demonstrate Eq. (2.124) also tells us the coefficients in the third derivative of $f(g(x))$.

Why should this be the case? We shall see in the next section.

2.7 Generating Functions

Our diagrammatic methods assign a *weight* to a graph, by interpreting that graph as an integral over a probability distribution. By construction, these graph weightings satisfy two properties. First, the weight of any graph is the product of the weights of its pieces, and second, the total weight of any collection of graphs is the sum of the weights of the individual graphs. These properties combine to yield an interesting and quite general result.

The moment $\langle x^n \rangle$ is the summed weight of all graphs having n vertices. To understand this series better, we "hang it on a clothesline" [29] by writing its *generating function*:

$$\mathcal{Q}(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \langle x^n \rangle. \tag{2.140}$$

Using Eq. (2.128), we can rewrite this generating function in terms of the cumulants $\langle x^n \rangle_c$:

$$\mathcal{Q}(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} {\sum_{\{n_l\}}}' W(\{n_l\}) \prod_l \langle x^l \rangle_c^{n_l}, \text{ where } \sum_l l n_l = n. \tag{2.141}$$

The l-th cumulant $\langle x^l \rangle_c$ is the weight of an l-vertex connected graph.

Because we are summing over all n, we can remove the restriction on the summation over $\{n_l\}$. First, we recall the definition of $W(\{n_l\})$.

$$\mathcal{Q}(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} {\sum_{\{n_l\}}}' \frac{n!}{\prod_l n_l! (l!)^{n_l}} \prod_l \langle x^l \rangle_c^{n_l}. \tag{2.142}$$

One way to think of this is as follows: having "primed" the summation symbol means that we've inserted an implicit delta function. We can really sum over all sets $\{n_l\}$, as long as it's understood that each term is multiplied by a Kronecker delta which is zero whenever $\sum_l l n_l$ is not equal to n. But the outer summation symbol means that

we have terms for all values of n, meaning that we can really take $\{n_l\}$ to be anything we like, as long as in that term, n is set equal to $\sum_l l n_l$.

In frighteningly abstracted notation, we are saying that

$$\sum_{n=0}^{\infty} \sum_{\{n_l\}} \delta_{\sum_l l n_l, n} = \sum_{\{n_l\}}, \tag{2.143}$$

so we can rewrite the generating function as

$$\begin{aligned} \mathcal{Q}(z) &= \sum_{\{n_l\}} z^{\sum_l l n_l} \prod_l \frac{\langle x^l \rangle_c^{n_l}}{n_l! (l!)^{n_l}} \\ &= \sum_{\{n_l\}} \prod_l \frac{1}{n_l!} \left(\frac{z^l \langle x^l \rangle_c}{l!} \right)^{n_l}. \end{aligned} \tag{2.144}$$

Interchanging the product and summation symbols and employing the definition of the exponential,

$$\mathcal{Q}(z) = \prod_l \exp\left(z^l \frac{\langle x^l \rangle_c}{l!} \right), \tag{2.145}$$

or,

$$\mathcal{Q}(z) = \exp\left(\sum_{l=1}^{\infty} \frac{\langle x^l \rangle_c}{l!} z^l \right). \tag{2.146}$$

The generating function over all graphs turns out to be the exponential of the generating function for *connected* graphs. This is known as the *linked-cluster theorem.*

We used only very general properties of graph weights in order to derive Eq. (2.146). As long as those basic conditions are met, the same relationship will hold, no matter what interpretation we give to the vertices and edges of our graphs.

Another application occurs in statistical physics. If the vertices of our graphs represent particles and the graph weighting is determined by pairwise interactions between particles, then the generating function $\mathcal{Q}(z)$ has the form of the grand partition function, with the formal variable z playing the role of the fugacity $e^{\beta\mu}/\lambda^3$, which represents the difficulty of changing particle number by thermal fluctuations [27]. The grand partition function includes all "Mayer graphs", connected or not; Eq. (2.146) states that the sum of all graphs is given by the exponential of the sum over connected graphs.

Now, we can see why the pattern of coefficients we saw in the iterated chain rule relates to moments and cumulants. If $f(z)$ is a generating function over some formal variable z,

$$f(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} f_n, \tag{2.147}$$

then we can extract the coefficient f_n by differentiating n times and evaluating the result at zero:

$$f_n = \left. \frac{d^n}{dz^n} f(z) \right|_{z=0}. \tag{2.148}$$

The differentiations remove all terms of lower order than n, and setting $z = 0$ removes all terms of higher order, leaving only f_n.

Consider the generating function defined by

$$f(z) = \exp\left[g(z)\right], \tag{2.149}$$

where $g(z)$ is itself a generating function,

$$g(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} g_n. \tag{2.150}$$

If we associate the expansion coefficients g_n with connected diagrams, then we are back to the linked-cluster theorem again. The values of f_n depend on the g_n. But we know that $f(z)$ is the *composition* of $g(z)$ with the exponential function, and we know that the f_n relate to the derivatives of $f(z)$, so we can compute the f_n by iterating the chain rule. The pictorial interpretation is that the operation of merging a new tree into a forest in all possible ways represents the product rule, which says that a derivative acts on a product to create a sum of new products, each of which has one factor modified. In probability theory, we also merge a new tree into a forest, when we go from the cumulant expansion of $\langle x^n \rangle$ to that of $\langle x^{n+1} \rangle$.

2.8 Central Limit Theorem

What happens when we add together many independent random variables? The same logic we used in §2.6 still applies: cumulants of their distributions add. If the random variable Y is given by

$$Y = \sum_{i=1}^{N} X_i, \tag{2.151}$$

where X_i are independent random variables described by distributions $p(x_i)$, then the nth cumulant of Y, $\langle y^n \rangle_c$, is simply the sum

$$\langle y^n \rangle_c = \sum_{i=1}^{N} \langle x_i^n \rangle_c . \tag{2.152}$$

To take the simplest case, suppose that the distributions of all the X_i are identical. Then $\langle y^n \rangle_c = N \langle x^n \rangle_c$, meaning that the new random variable

$$Z \equiv \frac{Y - N \langle x \rangle_c}{\sqrt{N}} \tag{2.153}$$

has a mean of zero, and higher cumulants which scale as $\langle z^n \rangle_c \propto N^{1-n/2}$. For N sufficiently big, all cumulants higher than the second die off, and Z becomes *Gaussian.*

$$\lim_{N\to\infty} p\left(z = \frac{\sum_i x_i - N \langle x \rangle_c}{\sqrt{N}}\right) = \frac{1}{2\pi \langle x^2 \rangle_c} \exp\left(-\frac{z^2}{2 \langle x^2 \rangle_c}\right). \tag{2.154}$$

We have arrived at the *Central Limit Theorem*: The sum of many uncorrelated random variables is itself a random variable, whose probability distribution is approximately Gaussian. A Gaussian distribution is specified by two parameters, which we denote here as μ and σ:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \tag{2.155}$$

The Fourier transform of a Gaussian curve is itself a Gaussian:

$$\tilde{p}(k) = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi\sigma^2}} e^{-ikx} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) = \exp\left(-ik\mu - \frac{k^2\sigma^2}{2}\right). \tag{2.156}$$

We can read off the cumulants:

$$\langle x \rangle_c = \mu, \ \langle x^2 \rangle_c = \sigma^2, \tag{2.157}$$

and $\langle x^n \rangle_c = 0$ for all $n > 2$. The Gaussian distribution is characterized entirely by its first two cumulants, namely the mean μ and the variance σ^2.

A neat thing happens when we take Eqs. (2.129) through (2.134) and set all cumulants of higher order than the variance to zero. Carrying this out, we get,

$$\begin{aligned} \langle x \rangle &= \mu \\ \langle x^2 \rangle &= \sigma^2 + \mu^2 \\ \langle x^3 \rangle &= 3\sigma^2\mu + \mu^3 \\ \langle x^4 \rangle &= 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4. \end{aligned} \tag{2.158}$$

In diagram language, our graphs are allowed to have clusters of one point and clusters of two points, but clusters of three or more points cause the graph value to vanish. Alternatively, in the forest picture, trees with more than two leaf nodes evaluate to zero. The numerical value of a forest is the product of the values of its constitutent trees, so any forest containing a tree with too many leaf nodes will, likewise, contribute nothing to the sum total.

One of the many applications of the Central Limit Theorem is to *random walks*. Consider a walker that starts at $x = 0$ takes steps whose direction and length are chosen at random, according to a probability distribution $p_1(l)$. We can choose, conventionally, that negative increments are motion toward the left, and positive increments are steps toward the right. What is the probability to be at position x after N steps? How do we characterize this probability distribution, and how does it change over time, as the walker takes more steps?

The position after N steps is

$$x = l_1 + l_2 + \cdots + l_N. \tag{2.159}$$

Because the l_i come from identical independent distributions, the average displacement is

$$\langle x \rangle = \langle l_1 \rangle + \cdots + \langle l_N \rangle = N\langle l \rangle. \tag{2.160}$$

However, the final displacement of the walker can be greater or smaller than this value. The probability distribution for the net displacement will have some variance,

$$\left\langle x^2 \right\rangle_c = \sum_{i,j=1}^{N} \left(\langle l_i l_j \rangle - \langle l_i \rangle \langle l_j \rangle \right) \tag{2.161}$$

$$= \sum_{i,j=1}^{n} \left\langle l^2 \right\rangle_c \delta_{ij} \tag{2.162}$$

$$= N \left\langle l^2 \right\rangle_c . \tag{2.163}$$

All the cumulants follow this general pattern: After N steps, their value is multiplied by N. This is a specific example of the general property which we established for cumulants of independent random variables added together. By the Central Limit Theorem, we can neglect the higher-order cumulants for $N \gg 1$. Thus, the final probability distribution for large N will be a Gaussian, which we can write

$$p(x, N) = \exp\left[-\frac{(x - N\langle l \rangle)^2}{2N \left\langle l^2 \right\rangle_c} \right] \frac{1}{\sqrt{2\pi N \left\langle l^2 \right\rangle_c}}. \tag{2.164}$$

Because $p(x, N)$ is an exponential function, taking its derivative yields back $p(x, N)$ again. We work out the first two derivatives with respect to position, and the first derivative with respect to N:

$$\frac{\partial p}{\partial x} = -\frac{x - N\langle l \rangle}{N \left\langle l^2 \right\rangle_c} p, \tag{2.165}$$

$$\frac{\partial^2 p}{\partial x^2} = \frac{(x - N\langle l \rangle)^2}{N^2 \left\langle l^2 \right\rangle_c^2} p - \frac{1}{N \left\langle l^2 \right\rangle_c} p, \tag{2.166}$$

and, formally,

$$\frac{\partial p}{\partial N} = \left[\frac{\langle l \rangle (x - N\langle l \rangle)}{N \left\langle l^2 \right\rangle_c} + \frac{(x - N\langle l \rangle)^2}{2N^2 \left\langle l^2 \right\rangle_c} - \frac{1}{2N} \right] p. \tag{2.167}$$

We can use the first two formulae to simplify the third:

$$\frac{\partial p}{\partial N} = -\langle l \rangle \frac{\partial p}{\partial x} + \frac{\left\langle l^2 \right\rangle_c}{2} \frac{\partial^2 p}{\partial x^2}. \tag{2.168}$$

Instead of expressing how p changes with the number of steps N, we can say how p changes over time, if we introduce a timestep τ. Define

$$v = \frac{\langle l \rangle}{\tau}, \; D = \frac{\left\langle l^2 \right\rangle_c}{\tau}. \tag{2.169}$$

Then, we obtain a *diffusion equation,* which relates the spatial and temporal partial derivatives of the probability density:

$$\frac{\partial p(x, t)}{\partial t} = -v \frac{\partial p}{\partial x} + \frac{D}{2} \frac{\partial^2 p}{\partial x^2}, \tag{2.170}$$

If the random walk is *unbiased,* then $v = 0$, and the diffusion equation simplifies to

$$\frac{\partial p(x,t)}{\partial t} = \frac{D}{2}\frac{\partial^2 p}{\partial x^2}. \tag{2.171}$$

We derived this diffusion equation in terms of the probability that a single random walker will arrive at position x. The equation codifies how our expectations for the walker at different times relate to each other. If we have many walkers moving along the same axis but not interacting with each other, then the fraction of those walkers which we expect to fall within a given interval scales with the integral of the probability density $p(x,t)$ over that interval. That is, the probability density governs our expectation for the number density.

This picture would become significantly more involved if the different random walkers in the population interacted with one another. In that case, we would have no reason to assume that the joint probability distribution over all walker positions should factor nicely into a product of single-walker probability densities.

The left-hand side of the diffusion equation (2.171) is a time derivative, so its units are those of p divided by time. On the right-hand side, we have a second derivative with respect to position, which introduces units of length^{-2}. In order for the dimensions on both sides to match, the constant D must have units of length^2 per time. This is one way to remember the intuition that for a diffusive process, variance grows linearly with time, or to say it another way, a characteristic length scale increases with the square root of the time elapsed.

The central limit theorem is perhaps the most elementary result about what those in the biz call *universality.* It takes a few broad strokes about an assemblage of random variables and deduces a quantitative result, and so it applies to any system that meets those general criteria. A *universality class* is a set of models that, despite differing perhaps widely in their details, exhibit quantitatively identical behavior in key aspects because they all meet the same broad-stroke conditions. We will meet other universality classes as we go along.

2.9 Large Deviations

Let's consider again the binomial distribution, Eq. (2.63):

$$P(m) = \binom{N}{m} p^m (1-p)^{N-m}\,. \tag{2.172}$$

If N is big, we can invoke Stirling's approximation,

$$\log N! \approx N\log N - N\,, \tag{2.173}$$

and if both m and $N-m$ are also big, we can approximate the binomial coefficient by

$$\binom{N}{m} \approx \frac{N^N e^{-N}}{m^m e^{-m}(N-m)^{(N-m)}e^{-(N-m)}}\,. \tag{2.174}$$

Canceling common factors to simplify,

$$\binom{N}{m} \approx \frac{N^N}{m^m (N-m)^{(N-m)}}\,, \tag{2.175}$$

and inserting this into the binomial distribution,

$$P(m) \approx N^N \frac{p^m}{m^m} \frac{(1-p)^{N-m}}{(N-m)^{(N-m)}}\,. \tag{2.176}$$

Anything we have neglected by taking Stirling's approximation will appear as a multiplicative factor, a possible slow variation with N hiding in that $\approx$ symbol. Define

$$q = \frac{m}{N}\,, \tag{2.177}$$

and so

$$P(q) \approx N^N \left(\frac{p}{Nq}\right)^{Nq} \left(\frac{1-p}{N-Nq}\right)^{N-Nq}\,. \tag{2.178}$$

The powers of N all cancel, because we have

$$N^N \left(\frac{1}{N}\right)^{Nq} \left(\frac{1}{N}\right)^{N-Nq} = 1\,. \tag{2.179}$$

Accordingly,

$$P(q) \approx \left(\frac{p}{q}\right)^{Nq} \left(\frac{1-p}{1-q}\right)^{N(1-q)}\,. \tag{2.180}$$

The right-hand side is the exponential of the quantity

$$N\left[q \log\left(\frac{p}{q}\right) + (1-q)\log\left(\frac{1-p}{1-q}\right)\right] = -N\left[q\log\left(\frac{q}{p}\right) + (1-q)\log\left(\frac{1-q}{1-p}\right)\right]. \tag{2.181}$$

This has the form of the relative entropy that we introduced in Eq. (2.99). Call the expression in brackets $I(q,p)$. Then

$$P(q) \approx e^{-NI(q,p)}\,. \tag{2.182}$$

The function $I(q,p)$ has a minimum at $q=p$, and going to second order in its Taylor series gives

$$I(q,p) \approx \frac{1}{2}\frac{1}{p(1-p)}(q-p)^2\,. \tag{2.183}$$

Inserting this into our expression for the probability gives a Gaussian curve of mean p and variance $\sigma^2 = 1/N$.

The function $I(q,p)$ is an example of a *rate function*, a central concept in *large deviation theory*.

In Eq. (2.86), we stated de Finetti's theorem, which tells us how to treat sequences of random events as mixtures of binomial distributions. This theorem can also be applied to calculate rate functions [30].

2.10 Conceptual Afterwords

We have come quite a way from the rather concept-oriented, foundationally-flavored beginnings of this chapter. For the most part, in this book we will only care about the philosophy of probability so far as it opens up freedom to maneuver in ways that clarify matters of physics.

One school of thought departs from all the reflection-based considerations we have developed here and says, "When you learn a new piece of evidence, you should change your probability distribution to be the one that is *least informative* given the constraints of the evidence you have gathered." Making the notion of "least informative" mathematically precise requires the tool of Shannon information. We will respect this school of thought but not stress it very much. The prototypical example of "evidence" that one learns in the scenarios they consider is something like the average value of a random variable. The idea would then be to write the most featureless function $p_X(x)$ that is consistent with normalization and with the condition that the average $\langle x \rangle$ is fixed to the desired value. This is a dangerously alluring oversimplification: In what experiment does one learn a true average? We have a system that we model probabilistically, and a laboratory apparatus that we also model probabilistically, and we couple them through some stochastic process. If a dial on the apparatus then points at a number, can we really equate that number with $\langle x \rangle$ exactly? But suppose that we do make the commitment: We go ahead and declare that $\langle x \rangle = c$, where c is what we read off the dial. Locking in $\langle x \rangle = c$ might leave us in a state of incoherence... but as we noted, the axioms can only tell us when we are incoherent, not how to regain coherence. How we escape the Dutch bookie is up to us.

When reading and working through the Dutch-book arguments earlier in this chapter, you may have felt that various limiting assumptions were being made. Perhaps it seemed that we were presuming that all the outcomes of an experiment could be neatly tabulated in advance. Or, when we spoke of combining propositions by Boolean connectives, it might have seemed that just because E and F are both meaningful statements, "E or F" doesn't have to be. Going in a slightly different direction, maybe propositions (or events) are meaningful when considered together, but one's states of belief regarding the two are simply not comparable. Concerns like these are *entirely fair.* The same process of reading, thinking and conversation that gradually convinced me that Ramseyish probability is valuable for physics also convinced me that it is not going to provide a complete theory of learning.

Occasionally, one meets people who talk as though the Bayes update rule is the cure for all that ails you. Upon closer investigation, these people are not trying to make justifiable decisions in the face of incomplete information, or detect inconsistencies among their own beliefs, or anything of the sort. It's not even a case of "GIGO" — "Garbage In, Garbage Out" implies that the process between input and output is legitimate at least. Instead, they're just emitting math-shaped noises to justify a lack of introspection, or even of conscience. They use language nominally about changing one's beliefs in order to keep their own beliefs locked in place.

Setting aside the cases where conversation is pointless, can we do interesting mathematics of physical relevance by relaxing assumptions like those we mentioned a moment

ago? History suggests that this is difficult. Even in quantum mechanics, we still compute probabilities that satisfy the rules we deduced from the bartender parable: Once we fix a given experimental context, the probabilities for the various potential outcomes of that experiment are real numbers in the good old unit interval. People have tried making sense of quantum physics by modifying the rules by which propositions are combined, relaxing the requirements of Boolean logic, but this has proved of limited usefulness; indeed, it seems faintly old-fashioned now. We will consider in a later chapter what quantum theory says about weaving disparate experimental contexts together. Plenty of puzzles remain, but modifying the rules of Boolean logic has not (so far) demonstrated itself to be the tool that can break them open.

What about representing an agent's mesh of beliefs — their *doxastic* state, to be fancy about it — by mathematical entities other than real numbers? Again, as we go along, we will see that even quantum theory presents less of a departure from this than one might have anticipated from the hype and the folklore. But perhaps the past is not prologue here! Some of the arguments that this avenue is a blind alley are unconvincing. For example, E. T. Jaynes wrote, "it appears to us that any computer designed to carry out the operations of a comparative probability theory must at some stage represent the ordering relations as inequalities of real numbers" [31]. No, they don't! A mathematical representation that only indicates which propositions are more plausible than which others can be implemented, for example, by links or pointers, each proposition being stored in memory along with the addresses of the next-least-plausible propositions. One way to think of a probability assignment is as a structure-preserving map between a Boolean lattice of statements and a totally ordered set, i.e., the unit interval. Perhaps some clever category theorist can devise a method that treats an agent's doxastic state as a functor between enriched categories. It seems to me that a major part of the challenge in any such project will be identifying what scientific question the generalization is meant to resolve.

3 Diffusive Processes

3.1 Variations on Diffusion

Diffusion becomes more complicated when multiple types of particle are moving together through the same space, and the overall amount of space available to move through is limited. By considering the effect of volume limitations at the rate-equation level, we can derive in a simpler way some interesting modifications to the ordinary diffusion equation, alterations which have been seen to emerge from a highly involved analysis [32, 33].

We consider three sites, with occupation numbers a_1, a_2 and a_3. Particles jump away from a site at an overall rate k, and jumps are equally likely in either direction. We make the approximation that the *actual* flow is the same as the *most probable* flow, so we can quantify everything in terms of average rates. The change of the population at site 2 is a result of flows into site 2 from sites 1 and 3, and the flows out of site 2 into its neighbors:

$$\frac{da_2}{dt} = \frac{k}{2}a_1 - ka_2 + \frac{k}{2}a_3\,. \tag{3.1}$$

We rearrange this as follows:

$$\frac{da_2}{dt} = \frac{k}{2}(a_1 - 2a_2 + a_3) \tag{3.2}$$

$$= \frac{k}{2}(a_1 - a_2 + a_3 - a_2) \tag{3.3}$$

$$= \frac{k}{2}[(a_3 - a_2) - (a_2 - a_1)]\,. \tag{3.4}$$

The time derivative of a_2 is the difference of two differences, which we can express as

$$\frac{da_2}{dt} = \frac{k}{2}[(\Delta a)_2 - (\Delta a)_1]\,, \tag{3.5}$$

or in terms of the *discrete Laplacian operator* as

$$\frac{da_2}{dt} = \frac{k}{2}(\Delta^2 a)_2\,. \tag{3.6}$$

This is exactly the discrete analogue of the diffusion equation for a continuous line, Eq. (2.171).

A more intricate case is the diffusion of two substances through a common space, with a limit imposed on how much total substance can be at any point:

$$a_i + b_i \leq N,\ \forall\ i\,. \tag{3.7}$$

Writing e_i for the amount of empty space at location i, we can turn this inequality into an equality:

$$a_i + b_i + e_i = N\,. \tag{3.8}$$

Because flow can only happen into empty space, the population of type a at site 2 changes as

$$\frac{da_2}{dt} = -k(e_1 + e_3)a_2 + \frac{k}{2}a_1 e_2 + \frac{k}{2}a_3 e_2\,. \tag{3.9}$$

Solving for e_i in terms of a_i and b_i, this becomes

$$\frac{da_2}{dt} = \frac{kN}{2}(a_1 - 2a_2 + a_3) + \frac{k}{2}\left[(b_1 + b_3)a_2 - b_2(a_1 + a_3)\right]\,. \tag{3.10}$$

The first term, which carries a factor of N, has the same functional dependence on the $\{a_i\}$ that we saw before, with the ordinary diffusion process. The second term is more complicated, having an interdependence between the $\{a_i\}$ and the $\{b_i\}$. A bit of algebra reveals a convenient way to state that interdependence:

$$\begin{aligned}(b_1 + b_3)a_2 - b_2(a_1 + a_3) &= a_2(b_1 + b_3) - 2a_2 b_2 - b_2(a_1 + a_3) + 2b_2 a_2 &(3.11)\\ &= a_2(b_1 - 2b_2 + b_3) - b_2(a_1 - 2a_2 + a_3)\,. &(3.12)\end{aligned}$$

This has the structure of "self times the Laplacian of the other, minus the other times the Laplacian of the self." The same holds true for the time evolution of b_i, with the roles of the a- and b-variables interchanged.

Now, we express the $\{a_i\}$ and the $\{b_i\}$ as *fractions* of the volume at each site, N:

$$a_i = Nf_i,\ b_i = Ng_i\,. \tag{3.13}$$

The time derivative of a_2 tells us the time derivative of f_2, which in terms of the f- and g-variables works out to be

$$\frac{df_2}{dt} = \frac{kN}{2}(f_1 - 2f_2 + f_3) + \frac{kN}{2}\left[(g_1 + g_3)f_2 - g_2(f_1 + f_3)\right]\,. \tag{3.14}$$

Using the discrete Laplacian operator,

$$\frac{df_i}{dt} = \frac{kN}{2}(\Delta^2 f)_i + \frac{kN}{2}\left[f_i(\Delta^2 g)_i - g_i(\Delta^2 f)_i\right]\,. \tag{3.15}$$

The differential equation we had before is modified by *cross-diffusion* terms, which couple the time evolutions of the two fields. These new terms arise from the fact that limiting the total available volume at each site curtails the maximum possible flow rate.

What if spreading happens by *reproduction* instead of by hopping? (This is, for example, characteristic of many models in mathematical biology.) As before, we require empty space to move into, so the growth rate from site i into site j should go like

$$ka_i e_j = ka_i(N - a_j - b_j)\,. \tag{3.16}$$

There is no hopping out of a site, only the possibility of budding into it. So, the rate of change at site a_2 must be given by the rate at which budding events can happen, which depends upon the number of particles present at the neighboring sites, modulated by the amount of empty space availabe at a_2. Implementing this idea in equations,

$$\frac{da_2}{dt} = ka_1(N - a_2 - b_2) + ka_3(N - a_2 - b_2) \tag{3.17}$$

$$= k(N - a_2 - b_2)(a_1 + a_3) \tag{3.18}$$

$$= k(N - a_2 - b_2)\left[(\Delta^2 a)_2 + 2a_2\right] . \tag{3.19}$$

Examining this result, we see an *effective diffusion* term appear: As before, the time derivative of a_2 can be written using $(\Delta^2 a)_2$.

Cross-diffusion terms like those in Eq. (3.15) become important in the study of *reaction-diffusion* systems. Suppose we have two substances, whose densities are given by the functions $a(x,t)$ and $b(x,t)$. The particles of both substances can spread diffusively, but when they bump into each other, reactions can take place, which introduce the possibility of creating or destroying particles. How many such reactions happen depends on the concentration of particles present. So, we write a system of equations to express how the densities should change:

$$\frac{\partial a(x,t)}{\partial t} = D_1 \frac{\partial^2 a}{\partial x^2} + f_1(a,b), \tag{3.20}$$

$$\frac{\partial b(x,t)}{\partial t} = D_2 \frac{\partial^2 b}{\partial x^2} + f_2(a,b) . \tag{3.21}$$

It is a common practice to start with a model defined with a single location, writing functions f_1 and f_2 to represent the local dynamics, and then promote this construction to a spatial model by making a and b position-dependent and adding the diffusion terms. However, if the effective available volume at each position is limited, then this is not correct [32, 33]. We have in that case to include the cross-diffusion terms:

$$\frac{\partial a(x,t)}{\partial t} = D_1 \left[\frac{\partial^2 a}{\partial x^2} + a\frac{\partial^2 b}{\partial x^2} - b\frac{\partial^2 a}{\partial x^2}\right] + f_1(a,b) , \tag{3.22}$$

$$\frac{\partial b(x,t)}{\partial t} = D_2 \left[\frac{\partial^2 b}{\partial x^2} + b\frac{\partial^2 a}{\partial x^2} - a\frac{\partial^2 b}{\partial x^2}\right] + f_2(a,b) . \tag{3.23}$$

The cross-diffusion terms introduce an extra interdependence between the $a(x,t)$ and $b(x,t)$, above and beyond that which might be defined in the reaction terms $f_1(a,b)$ and $f_2(a,b)$.

One very common way to study diffusion-like processes on a discrete set of locations is to organize those locations into a graph. The *adjacency matrix* A of a graph is defined by setting $[A]_{ij}$ to 1 if nodes i and j are linked, and leaving it 0 otherwise. (Weights and directionality on the edges can be incorporated by allowing the elements of A_{ij} to take other values in $\mathbb{R}$.) The *Laplacian matrix* L is defined by subtracting A

from the matrix which has the degrees k_i of the nodes for its diagonal entries and is zero otherwise:

$$L = K - A\,. \tag{3.24}$$

Eigenvalues and eigenvectors of these matrices have graph-theoretical significance. Consider, first, the adjacency matrix A. Its elements enumerate the number of length-1 walks between vertices i and j. If vertex i is connected to j which is in turn linked with k, then the product of A_{ij} with A_{jk} is 1; the total number of two-step paths between i and k is found by summing over the intermediate vertex j, which we can write with the Einstein summation convention as $A_{ij}A_{jk}$. But this is just the definition of matrix multiplication: raising the adjacency matrix to the second power gives the number of length-2 walks. One can show by induction that this relationship holds to any power, i.e., that the ij-th entry of the matrix A^l is the number of length-l walks between vertices i and j. Because traces of matrices can be rewritten as sums over eigenvalues, moments of the eigenvalue spectrum of A are related to the walks possible across the graph which A encodes.

Two graphs with the same degree distribution $\{k_i\}$ can have different topological structure, meaning that the walks connecting two vertices i and j can differ in length and number. This possibility is reflected in the sensitivity of the eigenvalue spectrum to rewiring: By swapping edges in the network, we preserve its degree distribution but change its topology. Figure 3.1 shows this effect for a random spatial graph, that is, a graph built by throwing vertices into a box and drawing an edge between each two vertices which fall within a specified threshold distance of each other.

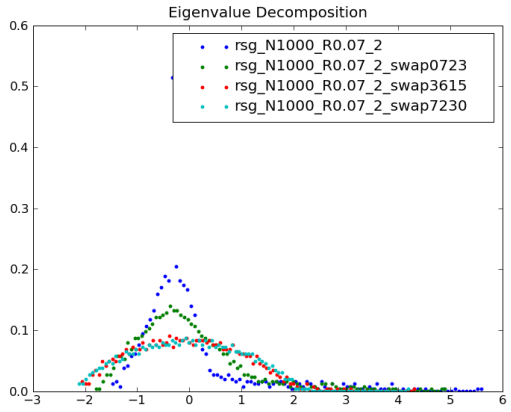


Figure 3.1: Eigenvalue spectra showing the effect of rewiring to destroy spatial structure. The histograms change from the triangular peak characteristic of random spatial graphs to the Wigner semicircle familiar from random matrix theory.

Eigenvalues of the Laplacian matrix reflect the behavior of stochastic processes tak-

ing place on the graph. To see why this is true, consider a random walk on a graph: A walker moves from vertex to vertex, in such a fashion that each possible step at a given location is equally likely. The number of possible steps the walker can take from vertex i is just the degree k_i, while the ability to move from i to j is given by the adjacency matrix element A_{ij}. Thus, the probability of moving to j from i is

$$\pi(i \to j) = \frac{A_{ij}}{k_i} \,. \tag{3.25}$$

Because the random walker has no memory of previous steps taken, its motion is a *Markov process* — a stochastic process in which what might happen in the next instant depends only on the current position, not the past history. The transition matrix for this Markov process can be deduced from the transition probability rule:

$$\Pi = K^{-1} A \,. \tag{3.26}$$

Multiplying the Laplacian by the inverse of the degree matrix K,

$$K^{-1} L = K^{-1} K - K^{-1} A \,, \tag{3.27}$$

shows that the transition matrix Π is closely related to the Laplacian:

$$\Pi = I - K^{-1} L \,. \tag{3.28}$$

Given some initial probability distribution $p_i(0)$ for random-walker positions, then this distribution evolves over time as

$$p_i(t) = (\Pi^T)^t p_i(0) \,. \tag{3.29}$$

If the graph is connected, and if it is not bipartite, then a stationary distribution p_i^* exists. The inverse of the smallest nonzero eigenvalue of Π gives the characteristic timescale for approaching this steady state.

3.2 The Fokker–Planck Equation

Consider a random walk over a set of sites, in which the walker moves from site i to site j at a rate Π_{ji}. That is, the probability to jump in a short time dt is $\Pi_{ji} dt$. How do the occupation probabilites $\{P_i(t)\}$ change over time? A given occupation probability can change because the walker is likely to jump *to* that site or because a walker at that site is likely to jump *away*. In mathematics,

$$\frac{dP_i}{dt} = -\sum_j \Pi_{ji} P_i + \sum_j \Pi_{ij} P_j \,. \tag{3.30}$$

This relation is known as the *master equation*. By taking the continuüm limit, in which the sites are very closely squished, we can arrive at another useful relation. We essentially make the transformations

$$i \to x, \;\; P_i \to P(x,t), \;\; \Pi_{ji} \to \Pi(x' - x, x) \,. \tag{3.31}$$

Note that we have re-parametrized the jump probability to use the difference of positions, rather than the positions themselves. (The two parametrizations are, of course, entirely equivalent.) The result of transforming Eq. (3.30) in this fashion is

$$\frac{\partial}{\partial t}P(x,t) = -\int dx'\,\Pi(x'-x,x)P(x,t) + \int dx'\,\Pi(x-x',x')P(x',t)\,. \tag{3.32}$$

The transition rates depend on the separation between the old position and the new position, $y = x' - x$. A small change in the new position x' implies a small change in the separation: $dx' = dy$. Using the new variable y, Eq. (3.32) becomes

$$\frac{\partial}{\partial t}P(x,t) = -\int dy\,\Pi(y,x)P(x,t) + \int dy\,\Pi(y,x-y)P(x-y,t)\,. \tag{3.33}$$

For most physical applications, we expect typical changes to be local, *i.e.,* that Π is dominated by small y. We can then expand the second integral in Eq. (3.33) as a Taylor series. The first term in this expansion cancels with the first integral in Eq. (3.33). The next two terms yield

$$\frac{\partial}{\partial t}P(x,t) = -\int dy\,y\frac{\partial}{\partial x}(\Pi(y,x)P(x,t)) + \frac{1}{2}\int dy\,y^2\frac{\partial^2}{\partial x^2}(\Pi(y,x)P(x,t))\,. \tag{3.34}$$

We can take the derivatives outside of the integrals, like so:

$$\frac{\partial}{\partial t}P(x,t) = -\frac{\partial}{\partial x}\left[P(x)\int dy\,y\Pi(y,x)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\int dy\,y^2(\Pi(y,x)P(x,t))\right]\,. \tag{3.35}$$

This becomes a close analogue of the diffusion equation, which we derived in Eq. (2.171). Each term involves a factor which is essentially a moment of $\Pi(y,x)$ with respect to y. Define

$$F(x) = \int dy\,y\Pi(y,x)\,, \tag{3.36}$$

and

$$D(x) = \int dy\,y^2\Pi(y,x)\,. \tag{3.37}$$

Then

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[F(x)P(x,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[D(x)P(x,t)\right]\,. \tag{3.38}$$

Eq. (3.38) is known as the *Fokker–Planck equation.* For many applications, $D(x)$ can be taken as a constant, and so,

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[F(x)P(x,t)\right] + \frac{D}{2}\frac{\partial^2 P(x,t)}{\partial x^2}\,. \tag{3.39}$$

An appropriate change of variables can in fact transform away the position dependence of $D(x)$, so this special case can be used even when the ease of diffusion is position-dependent [34].

If we apply the product rule, we can expand the derivatives in Eq. (3.38), obtaining

$$\frac{\partial P(x,t)}{\partial t} = -\frac{D}{2}P'' + (D' - F)P' + \left(\frac{D''}{2} - F'\right)P\,. \tag{3.40}$$

Note that this expression includes both the first derivative of $F(x)$ and the second derivative of $D(x)$. This will be important soon.

To find the steady-state solution, $P^*(x)$, for the Fokker–Planck equation with constant diffusion, set the derivative with respect to time equal to zero. This implies

$$\frac{\partial}{\partial x}\left[F(x)P^*(x)\right] = \frac{D}{2}\frac{\partial^2 P^*(x)}{\partial x^2}\,, \tag{3.41}$$

or

$$\frac{\partial^2 P^*(x)}{\partial x^2} = \frac{2}{D}\frac{\partial}{\partial x}\left[F(x)P^*(x)\right]\,. \tag{3.42}$$

Integrating once over x,

$$\frac{\partial P^*(x)}{\partial x} = \frac{2F(x)}{D}P^*(x)\,. \tag{3.43}$$

When the derivative of a function gives back the function itself, the function is an exponential. The question is, an exponential of what? The derivative of its argument with respect to x must be $2F(x)/D$. Recall that the derivative of an integral with respect to its upper limit is the integrand evaluated at that limit. So,

$$P^*(x) \propto \exp\left(\frac{2}{D}\int^x dx'\,F(x')\right)\,. \tag{3.44}$$

The choice of the lower limit and the constant of integration glossed over earlier can be absorbed into the prefactor which establishes the proper normalization,

$$\int dx\,P^*(x) = 1\,. \tag{3.45}$$

If the diffusion function $D(x)$ is not constant, then we can find the steady-state distribution in much the same way. The result is the slightly modified formula

$$P^*(x) \propto \frac{1}{D(x)}\exp\left(2\int^x dx'\frac{F(x')}{D(x')}\right)\,. \tag{3.46}$$

One can prove by going through some calculus that the relative entropy is a *Lyapunov function* for the Fokker–Planck equation. This means that if we start with two different probability densities and time-evolve both of them according to the Fokker–Planck equation, their relative entropy will never increase. Particularly of note is the relative entropy of an arbitrary probability density and the steady-state solution. The relative entropy diminishes as the probability density approaches the steady-state curve.

3.3 Application: An Integrate-and-Fire Neuron

We can apply the Fokker–Planck equation not just to diffusion in position, but to stochastic motion along any axis. In this section, we will consider diffusion in *voltage,* as seen in a *leaky integrate-and-fire neuron.* This is a model of a living cell in which the key dynamical quantity is the membrane potential, $V(t)$. The membrane potential changes over time as ions flow in and out of the cell. We can write a differential equation for it:

$$\tau_m \frac{dV}{dt} = V_L - V + IR_m + \sigma w(t)\,, \tag{3.47}$$

where the $w(t)$ term is a white noise contribution that satisfies $\langle w(t)\rangle = 0$ and $\langle w(t)w(t')\rangle = \delta(t-t')$. Reasonable numerical values for the constant parameters are $\tau_m = 5\,\text{ms}$, $V_L = -70\,\text{mV}$, $R_m = 10\,\text{M}\Omega$ and $\sigma = 10\,(\text{mV})(\text{ms})^{0.5}$.

First, we consider how the membrane potential $V(t)$ varies in the absence of noise and if the neuron has no threshold to fire. Working in the absence of noise implies setting $\sigma = 0$. Then Eq. (3.47) becomes

$$\tau_m \frac{dV}{dt} = V_L - V + IR_m, \tag{3.48}$$

or

$$\frac{dV}{dt} = -\frac{1}{\tau_m}\left(V - (V_L + IR_m)\right). \tag{3.49}$$

This differential equation has a steady-state solution if the voltage V is equal to

$$V^* = V_L + IR_m. \tag{3.50}$$

Define a new variable U which is the membrane potential measured relative to this steady-state value:

$$U(t) = V(t) - (V_L + IR_m). \tag{3.51}$$

Because the shift is independent of time,

$$\frac{dU}{dt} = \frac{dV}{dt} = -\frac{1}{\tau_m}U(t). \tag{3.52}$$

Clearly, $U(t)$ is an exponential decay with characteristic timescale τ_m. Translating this back in terms of V,

$$V(t) = (V(0) - V^*)\,e^{-t/\tau_m} + V^*. \tag{3.53}$$

For the given value of V_L and a reasonable membrane current of $I = 1.2\,\text{nA}$, the steady-state membrane potential is

$$V^* = -70\,\text{mV} + (1.2\,\text{nA})(1\times 10^7\,\Omega) = -58\,\text{mV}. \tag{3.54}$$

If the initial membrane potential is $-70\,\text{mV}$, then

$$V(t) = (-12\,\text{mV})\exp\left(-\frac{t}{5\,\text{ms}}\right) - 58\,\text{mV}. \tag{3.55}$$

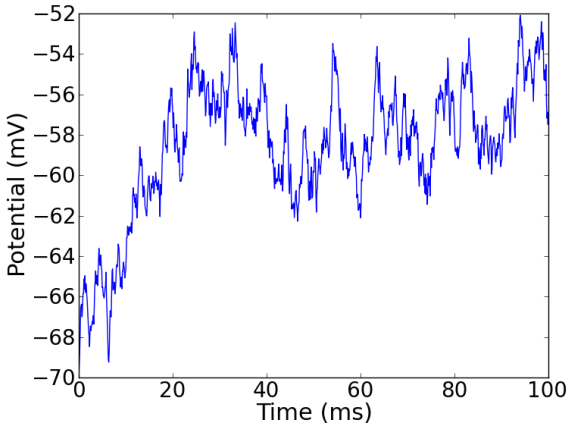


Figure 3.2: The membrane potential $V(t)$ as a function of time.

To see what happens when we incorporate noise, we first try a numerical simulation. A typical result is shown in Figure 3.2.

Next, we make an empirical stab at finding the probability distribution of V by histogramming our simulation results (running for 100 seconds with a time step of 0.1 ms, in this case). A typical result is shown in Figure 3.3. The solid line in Figure 3.3 is a Gaussian with mean $V_L + IR_m$ and variance $\frac{\sigma^2}{2\tau_m}$.

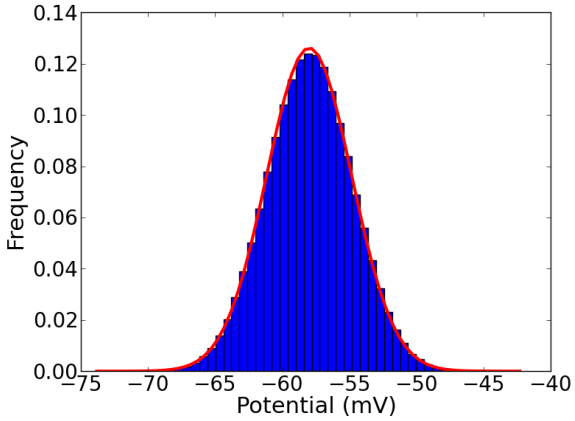


Figure 3.3: Probability distribution for the membrane potential $V(t)$. Bars: numerical results found by Monte Carlo simulation. Line: analytical solution.

The solid line, the analytic answer for the probability density curve, is the steady-state solution for the Fokker–Planck equation

$$P(V) \propto \exp\left(\frac{2}{D}\int_0^V dV'\, F(V')\right), \tag{3.56}$$

where we set

$$F(V) = -\frac{V - (V_L + IR_m)}{\tau_m}, \tag{3.57}$$

and

$$D = \frac{\sigma^2}{\tau_m^2}, \tag{3.58}$$

resulting in

$$P(V) \propto \exp\left[\frac{2\tau_m^2}{\sigma^2}\left(-\frac{1}{2\tau_m}V^2 + \frac{V_L + IR_m}{\tau_m}V\right)\right]. \tag{3.59}$$

Factoring out a $-1/(2\tau_m)$ gives

$$P(V) \propto \exp\left[-\frac{\tau_m}{\sigma^2}\left(V^2 - 2(V_L + IR_m)V\right)\right]. \tag{3.60}$$

Completing the square shows that this is a Gaussian with mean $V_L + IR_m$ and variance $\frac{\sigma^2}{2\tau_m}$. Including the proper normalization prefactor,

$$P(V) = \sqrt{\frac{\tau_m}{\pi\sigma^2}}\exp\left[-\frac{(V - (V_L + IR_m))^2}{\sigma^2/\tau_m}\right]. \tag{3.61}$$

This solution for $P(V)$ is the solid line in Figure 3.3.

If the noise vanishes, that is, if $\sigma = 0$, then the diffusion term in the Fokker–Planck equation disappears, and we have only

$$\frac{\partial P(V,t)}{\partial t} = -\frac{\partial}{\partial V}\left[\left(\frac{V^* - V}{\tau_m}\right)P(V,t)\right]. \tag{3.62}$$

By the product rule, this is

$$\frac{\partial P(V,t)}{\partial t} = \frac{1}{\tau_m}\left[P(V,t) - (V^* - V)\frac{\partial P(V,t)}{\partial V}\right]. \tag{3.63}$$

The first term in the brackets is always positive. The sign of the second term depends on whether the voltage V is above or below the steady-state value V^* and on whether the probability density $P(V,t)$ is increasing or decreasing along the V axis. For a given value of V, the probability density $P(V,t)$ will be increasing with time if

$$\frac{P(V,t)}{V^* - V} > \frac{\partial P(V,t)}{\partial V}. \tag{3.64}$$

Suppose that the probability distribution $P(V,t)$ is a lump localized below the steady-state voltage V^*. Then, for any V where the probability density is nonzero, $V^* - V > 0$. So, the left-hand side of the inequality is positive. On the downward slope of the lump, the right-hand side is negative, so $P(V,t)$ will increase with time. On the upward slope of the lump, the right-hand side is positive, and so the probability density can decrease with time. The lump will therefore move in the direction of increasing V. If the lump is instead localized above V^*, then the situation is reversed, and the lump will move to lower V. This shows that the stochastic dynamics of the Fokker–Planck equation are roughly consistent with the deterministic dynamics of Eq. (3.47).

3.4 Application: Evolutionary Dynamics

In what follows, we will use the Fokker–Planck equation in the following way. The variable x will denote the *value of a genetic trait in a population.* We will assume that the population is uniform: each time we survey it, all the organisms have the same genotype. However, in between observations, mutations can occur, creating individuals with slightly different genotypes. If the offspring of a mutant take over the population, then the value of x will change by a small amount. If the mutants fail to supplant the native population, then x stays the same. The timescale of this ecological competition will, in our analysis, be much shorter than the timescale over which we follow the variable x.

Even though the population is genetically uniform at each observation, we might not be *certain* about its genetic composition. For example, in a practical context we might only be able to carry out a few observations, and the measurements we take might be confounded by environmental factors, meaning that multiple values of x could be consistent with the data we gather. We therefore summarize our knowledge of the population by a probability distribution over x. If we know how mutations can happen, then we can say something about what x might be at a later time. However, because the outcomes of mutation events and competition are not certain, our statements about what x might be in the future must also be probabilistic. The function $P(x,t)$ expresses what we can deduce about what x might be at time t, given the information currently at hand. When the time t actually rolls around, we might go out and gather new data, allowing us to refine our information about what genotypes could be present. (This is the kind of problem for which we derived probability-update rules in §2.2.) For the most part, we will not be concerned here with that side of the problem; we will in this chapter focus on the computation of expectations in the absence of new information.

We define an evolutionary process in the following way [35, 36, 37]: At any time, a mutant can arise in an otherwise uniform population. The mutation rate can in general depend upon the current trait value, and the mutated trait value is near that of the resident population, but displaced by a random small amount. Whether the mutant variety takes over the population or not depends upon how the two varieties interact, which we codify as a game. The payoff for a mutant individual with trait value x_M playing against another individual drawn from a population having trait value x_R is

$A(x_M; x_R)$.

As time progresses, the population's trait value can change. Let

$$a_{ij} = A(x_i; x_j),\ i, j \in \{M, R\}. \tag{3.65}$$

We write the payoff matrix for mutant- and resident-type individuals interacting via two-player games as

$$G = \begin{pmatrix} a_{MM} & a_{MR} \\ a_{RM} & a_{RR} \end{pmatrix}. \tag{3.66}$$

The probability that a variety with trait value x' takes over a population with trait value x is $\rho(x'; x)$, which is some function of the matrix G.

Because we are considering two-player games with a set of two strategies, we can make use of a considerable amount of theory that has been developed for that case. It can be proved that, in a two-player game with two strategies, if the effect of selection is not too strong, then a strategy R is favored over another strategy M if the following inequality is satisfied:

$$\sigma(a_{RR} - a_{MM}) + a_{RM} - a_{MR} > 0. \tag{3.67}$$

Here, σ denotes the Tarnita *structure coefficient,* a number that depends on the population structure and the update rule, but is independent of the payoff matrix [38]. If this inequality is satisfied, then the fixation probability of R is greater than that of M: a single R-type individual in a population of type M is more likely to take over the ecosystem than a single M-type individual in the reverse scenario. Structure coefficients have been calculated for several different cases of interest [38, 36, 39].

Denote the mutation rate per individual by $u(x)$, and let the mutation step size be described by a probability density $\mu(z)$, which we take to be centered at zero. We assume that $\mu(z)$ is a narrow distribution, such that the expectation value of $|z|^2$ is much greater than that of $|z|^3$. Because many of our expressions will involve the variance of $\mu(z)$, we'll abbreviate it as ν:

$$\nu = \int dz\, z^2 \mu(z). \tag{3.68}$$

Suppose the current trait value is established to be x. By how much can we expect the trait value to change? We find this by integrating over the step size z:

$$\langle \Delta x \rangle = \int dz\, zNu(x)\mu(z)\rho(x+z; x) = Nu(x) \int dz\, z\mu(z)\rho(x+z; x). \tag{3.69}$$

If we expand the probability $\rho(x+z; x)$ to first order in the jump distance z,

$$\rho(x+z; x) = \rho(x; x) + z \left.\frac{\partial \rho(x'; x)}{\partial x'}\right|_{x'=x} + \mathcal{O}(|z|^2), \tag{3.70}$$

then our expectation value $\langle \Delta x \rangle$ becomes

$$\begin{aligned}\langle \Delta x \rangle = {} & N u(x) \rho(x;x) \int dz\, z \mu(z) \\ & + N u(x) \int dz\, z^2 \mu(z)\, \frac{\partial \rho(x';x)}{\partial x'}\bigg|_{x'=x} \\ & + \mathcal{O}\left[\int dz\, \mu(z) |z|^3 \right].\end{aligned} \tag{3.71}$$

The first term vanishes by symmetry, and the third is negligible by assumption. Therefore,

$$\langle \Delta x \rangle = N u(x) \int dz\, z^2 \mu(z)\, \frac{\partial \rho(x';x)}{\partial x'}\bigg|_{x'=x} = N u(x) \nu\, \frac{\partial \rho(x';x)}{\partial x'}\bigg|_{x'=x}. \tag{3.72}$$

This relates the expected change in x to the derivative of the fixation probability $\rho(x';x)$.

The simplest way to turn this into a dynamical system is to impose the condition that the *actual* change in x is given by this *expected* change in x. Doing so, we arrive at the *canonical equation* for the time evolution of the trait value:

$$\frac{dx}{dt} = N u(x) \nu\, \frac{\partial \rho(x';x)}{\partial x'}\bigg|_{x'=x}. \tag{3.73}$$

The derivative on the right-hand side can be written using the chain rule as

$$\frac{\partial \rho(x';x)}{\partial x'}\bigg|_{x'=x} = \frac{1}{A(x;x)} \sum_{j,k} \frac{\partial \rho}{\partial a_{jk}}\bigg|_{G=J} \frac{\partial a_{jk}}{\partial x'}\bigg|_{x'=x}. \tag{3.74}$$

Having derived Eq. (3.73) for the deterministic limit, we now push our understanding into new territory by including the effects of stochastic fluctuations. Instead of a single coordinate value which depends on time, $x(t)$, we will consider a time-varying probability density over the possible population states, $P(x,t)$.

Previously, we equated expected change in x with actual change, thereby establishing a deterministic dynamic. Now, we relax that requirement. Our calculation of the expected change, Eq. (3.72), is still valid, but it no longer tells the full story. Looking back over our derivation of the Fokker–Planck equation, we see that need a quantity which represents how much the probability density $P(x,t)$ spreads out over time. If we define

$$\left\langle (\Delta x)^2 \right\rangle = \int dz\, z^2 u(x) \mu(z) \rho(x+z;x), \tag{3.75}$$

then we find that

$$\left\langle (\Delta x)^2 \right\rangle = u(x) \nu \rho(x;x). \tag{3.76}$$

We can use Eqs. (3.72) and (3.76) to construct a Fokker–Planck equation for the time evolution of $P(x,t)$. The result is

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[u(x) \nu\, \frac{\partial \rho(x';x)}{\partial x'}\bigg|_{x'=x} P(x,t) \right] + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left[u(x) \nu \rho(x;x) P(x,t) \right]. \tag{3.77}$$

Note that the time evolution of the probability density depends on *second* derivatives of ρ. In contrast, the deterministic system defined by the canonical equation, Eq. (3.73), depends only on *first* derivatives of ρ.

The Fokker–Planck equation (3.77), which we can think of as a "canonical diffusion equation for adaptive dynamics," is equivalent to the stochastic differential equation derived by Champagnat and Lambert [35]. The route we have taken is more in line with a physicist's approach to stochastic processes. A similar derivation was recently given by Van Cleve [37].

From the steady-state solution to the Fokker–Planck equation, (3.46), we can derive the mutation-selection equilibrium for stochastic adaptive dynamics:

$$P^*(x) \propto \frac{1}{u(x)\nu\rho(x;x)} \exp\left(2\int^x dx' \frac{1}{\rho(x';x')} \left.\frac{\partial\rho(y;x')}{\partial y}\right|_{y=x'}\right). \tag{3.78}$$

The integrand in the exponential includes a first derivative of ρ, but not any higher derivatives.

We can write the payoff function for a continuous Prisoner's Dilemma as

$$A(x;y) = -C(x) + B(y). \tag{3.79}$$

The value $C(x)$ is the cost which an individual having genotype x pays in order to benefit another, and the amount of benefit which the other obtains is given by the function B. We set

$$C(0) = B(0) = 0, \tag{3.80}$$

and we require that both $C(x)$ and $B(x)$ are strictly increasing, with the inequality $C(x) < B(x)$ satisfied for $x > 0$. For the mathematics to work smoothly, we also posit that $C(x)$ and $B(x)$ are twice differentiable.[1]

Assuming that the neutral-drift takeover probability can be given in terms of an effective population size N_e, we find that

$$P^*(x) \propto \frac{N_e}{u(x)\nu} \exp\left(2N_e \int_0^x dx' \left.\frac{\partial\rho(y;x')}{\partial y}\right|_{y=x'}\right). \tag{3.81}$$

Plugging in the Prisoner's Dilemma, this becomes

$$P^*(x) \propto \frac{N_e}{u(x)\nu} \exp\left(2N_e \int_0^x \frac{dx'}{A(x';x')}\left[-C'(x') + \frac{\sigma-1}{\sigma+1}B'(x')\right]\right), \tag{3.82}$$

where σ is the Tarnita structure coefficient, which as we said earlier depends on the population structure and on the update rule [38]. To simplify matters, take the mutation rate $u(x)$ to be constant. Carrying out the integral in the exponential is difficult; however, we can understand a great deal about the solution by using the fact that the

[1] We note here a difference between our approach and that of Van Cleve [37]. In this chapter, the continuously variable trait x is a *strength of cooperation,* while in Van Cleve's study, the evolvable trait is the *fraction of time* in which an agent cooperates in a *discrete* game.

exponential will be peaked where its argument is maximized. To find this extremum, we can differentiate the argument with respect to x, and by the Fundamental Theorem of Calculus, this just extracts the integrand, evaluated at $x' = x$.

If $\sigma \leq 1$, then the argument of the exponential is always nonpositive, so the probability density $P^*(x)$ piles up at $x = 0$. On the other hand, if $\sigma > 1$, then $P^*(x)$ is peaked at the value of x which satisfies

$$\frac{B'(x)}{C'(x)} = \frac{\sigma + 1}{\sigma - 1}. \tag{3.83}$$

Ratios of benefits to costs are commonplace in the study of how social behaviors evolve [40]. Eq. (3.83) is an interesting variant, in that it is a comparison of *marginal* quantities, *i.e.,* of the derivatives of the benefit and cost functions. It tells us that the peak of $P^*(x)$ depends on the population structure and the details of the organisms' life cycles, through the structure coefficient σ.

Note that if the mutation rate u is not constant but instead varies with x, then the shape of the mutation-selection equilibrium curve $P^*(x)$ will change. This is another example of a theme noted by Allen and Tarnita: Standards of evolutionary success depend upon mutation rates [41].

Another scenario of evolutionary interest is the *Snowdrift game.* Doebeli et al. [42] provide a good description:

> [T]wo drivers are trapped on either side of a snowdrift and have the option of staying in their cars or removing the snowdrift. Letting the opponent do all the work is the best option, but if both players refuse to shovel they can't get home. The essential feature of the snowdrift game is that defection is better than cooperation if the opponent cooperates but worse if the opponent defects.

If the amount of cooperation is a continuously variable quantity, then this fits into the framework we have developed in this chapter. The payoff to an agent who cooperates an amount x when met with another agent who cooperates to the extent y is

$$A(x; y) = -C(x) + B(x + y). \tag{3.84}$$

Note that if we take a partial derivative with respect to x or with respect to y, we pick up a B' either way.

The mutation-selection equilibrium distribution for the Snowdrift game is

$$P^*(x) \propto \frac{N_e}{u(x)\nu} \exp\left(2N_e \int_0^x \frac{dx'}{A(x'; x')} \left[-C'(x') + \left(1 + \frac{\sigma - 1}{\sigma + 1}\right) B'(2x')\right]\right). \tag{3.85}$$

Again, simplifying the integral is not straightforward, but we can find the position where $P^*(x)$ is peaked.

Defining the convenient abbreviation

$$r = \frac{\sigma - 1}{\sigma + 1}, \tag{3.86}$$

the argument of the exponential is extremized where

$$\frac{B'(2x)}{C'(x)} = \frac{1}{1+r}. \tag{3.87}$$

In the deterministic limit, the Snowdrift game is known to exhibit interesting behavior with the quadratic cost and benefit functions

$$C(x) = c_2 x^2 + c_1 x, \; B(x) = b_2 x^2 + b_1 x. \tag{3.88}$$

Given these forms for the cost and benefit curves, then Eq. (3.87) is satisfied when

$$x = \frac{c_1 - (1+r)b_1}{4(1+r)b_2 - 2c_2}. \tag{3.89}$$

At $\sigma = 1$, we have $r = 0$, and our expression for x reduces to

$$x = \frac{c_1 - b_1}{4b_2 - 2c_2}, \tag{3.90}$$

which is the equilibrium value that Doebeli *et al.* derive for deterministic dynamics in a panmictic population [42].

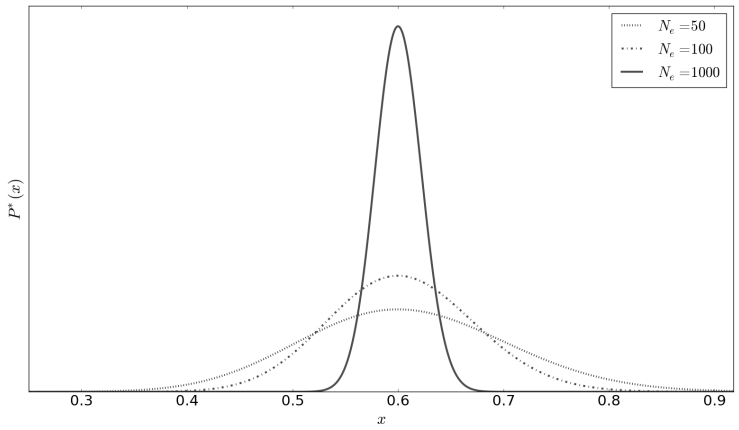


Figure 3.4: Steady-state probability distributions for the Snowdrift game, computed with three different choices of the effective population size N_e. As we increase N_e from 1 to 10 and then 100, the distribution becomes narrower. The peak of the distribution is, in each case, located at the "Evolutionary Stable Strategy" value of the deterministic dynamics, $x = 0.6$ ($b_1 = 7$, $b_2 = -1.5$, $c_1 = 4.6$, $c_2 = -1.0$, $\sigma = 1.0$).

The values of b_1, b_2, c_1 and c_2 govern whether or not the fixed point of the deterministic dynamics is stable. By adjusting these parameters appropriately, we can have

fixed points at the same value of x but opposite stabilities. For example,

$$b_1 = 7,\ b_2 = -1.5,\ c_1 = 4.6,\ c_2 = -1.0 \tag{3.91}$$

yields a fixed point at $x = 0.6$, as does

$$b_1 = 3.4,\ b_2 = -0.5,\ c_1 = 4.0,\ c_2 = -1.5. \tag{3.92}$$

However, in the former case, $x = 0.6$ is a stable strategy, while in the latter, it is a repellor point [42]. The analogue of this in the stochastic case is the fact that two probability distributions can have extrema in the same positions, one having a minimum and the other a maximum in that location. Now, we investigate this issue in more detail. In particular, we'd like to know how varying the structure coefficient σ affects stability.

Whether an extremum of $P^*(x)$ is a minimum or a maximum depends on the second derivative of the integral inside the exponential, which is the first derivative of the integrand. Applying the quotient rule to the integrand yields a ratio whose denominator is $A(x;x)^2$. This is always positive, so the sign can only depend on the sign of the numerator. Furthermore, the numerator itself simplifies at an extremum, leaving us with the condition

$$2(1+r)B''(2x) - C''(x) < 0. \tag{3.93}$$

For quadratic cost and benefit functions, this becomes

$$2(1+r)b_2 - c_2 < 0. \tag{3.94}$$

If this inequality is satisfied, then the extremum x is a maximum, and thus is evolutionarily stable.

Incorporating stochastic effects by way of our Fokker–Planck equation goes beyond prior work on the continuous Snowdrift game. We can go further by varying the structure coefficient σ, which corresponds to implementing the Snowdrift game with different short-term ecological dynamics. We plot typical results in Figure 3.5. Generally, as we increase σ, the peak of the steady-state distribution moves to higher x, indicating an increased disposition to cooperation.

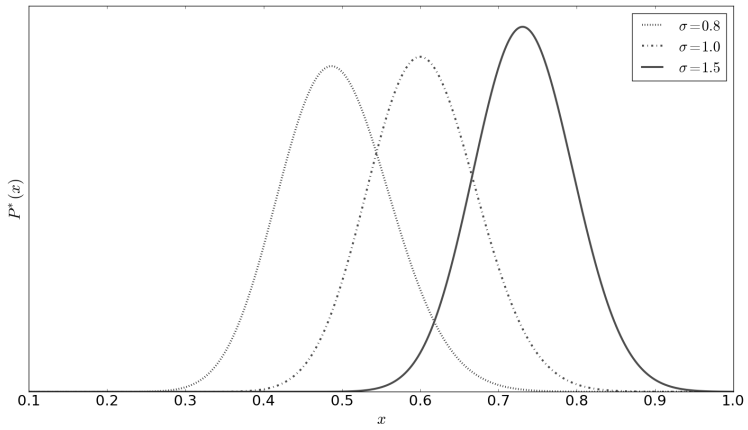


Figure 3.5: Steady-state probability distributions for the Snowdrift game, computed with three different choices of the Tarnita structure coefficient σ. As we increase σ, the peak of the steady-state distribution moves to higher x, indicating an increased disposition to cooperation ($b_1 = 7$, $b_2 = -1.5$, $c_1 = 4.6$, $c_2 = -1.0$, $N_e = 100$).

4 Renormalization

Our subject for the next few sections will be "the renormalization group." This term is a little like the Holy Roman Empire, in that the latter, as Voltaire observed, is neither holy, Roman nor an empire. The word *group* in this context is a technical term, but mathematically speaking, it is not really the *right* technical term. *Renormalization* is a bit of jargon that hails from quantum electrodynamics, and it dates to a time many years before people figured out what they were doing and why their techniques actually worked. Consequently, it is not very indicative. Even *the* isn't quite right: it carries the implication that "*the* renormalization group" is a single, specific construction, like "the quadratic formula."

Since the full name of the subject is a misnomer, it is convenient to elide that name as much as possible, and so we will use the common abbreviation *RG*. We begin, therefore, our discussion of *RG theory* and *RG transformations*.

4.1 The Central Limit Theorem by RG Transformations

To ease ourselves into RG theory, we will revisit a topic we addressed before, this time approaching it from a different, albeit related perspective. Our first example of RG theory is the central limit theorem, which we first proved in §2.8. A loose but useful statement of this result is that the sum total of many uncorrelated small influences is statistically distributed following a Gaussian curve.

Consider a system with many components, each of which is a random variable. We describe the system probabilistically by ascribing a joint distribution to the set of random variables. In this case, we postulate that all higher-order complexities vanish, meaning that the probability distribution factors neatly into single-variable distributions:

$$p(X_1 = x_1, X_2 = x_2, X_3 = x_3, \ldots) = p(X_1 = x_1)p(X_2 = x_2)p(X_3 = x_3)\cdots \tag{4.1}$$

To start with, let the components of our system be coins, which we say are independent of each other and equally likely to come up either heads ($+1$) or tails (-1). A snapshot of this system will be, we expect, a speckling of pluses and minuses in roughly equal proportions (we might get a large fluctuation one way or the other, but we're not gambling on that). What happens if we blur our view, grouping together pairs of components?

The total value of a pair of coins can be -1, 0 or $+1$. For a single coin, we have

$$p_1(-1) = \frac{1}{2},\ p_1(+1) = \frac{1}{2}. \tag{4.2}$$

There are two ways the total could sum to 0, and so

$$p_2(-2) = \frac{1}{4},\ p_2(0) = \frac{1}{2},\ p_2(+2) = \frac{1}{4}. \tag{4.3}$$

We see that the distribution stays symmetrical and becomes peaked in the middle.

Following the same logic, we can repeat the blurring procedure. The arithmetic becomes more involved, but the concept is straightforward. After just a few repetitions, the symmetrical, peaked distribution comes to resemble a Gaussian curve, as shown in Figure 4.1.

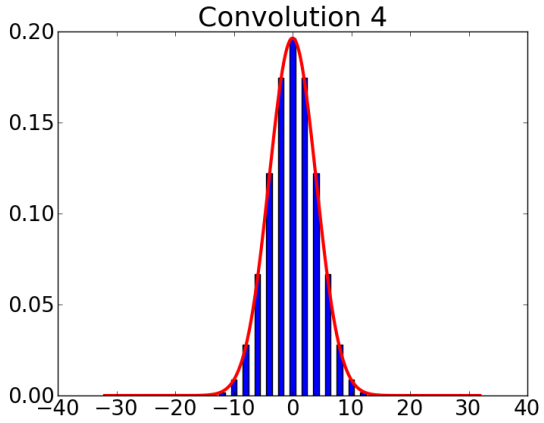


Figure 4.1: Discrete probability distribution obtained by coarse-graining uncorrelated fair coins four successive times. The solid curve is a Gaussian with mean zero and variance $\sigma^2 = 16$.

What happens if we start our gambling on the assumption that the coins are biased? Say, for example, that we had chosen

$$p_1(-1) = 0.1,\ p_1(+1) = 0.9. \tag{4.4}$$

Then we would have

$$p_2(-2) = 0.01,\ p_2(0) = 0.18,\ p_2(+2) = 0.81. \tag{4.5}$$

This p_2 is not symmetrical at all! However, as we repeat the blurring procedure, the asymmetry partially washes out. Examinining Figure 4.2, we see that the statistics for the coarse-grained view are still approximately Gaussian, but the mean of the Gaussian curve has shifted in the direction of the original tilt.

Let's explore the effects of successive coarse-graining in more generality. Take $p(x)$ to be the probability distribution for each of our original variables, and suppose for convenience that its mean is zero. When we convolve $p(x)$ with itself, we double all its

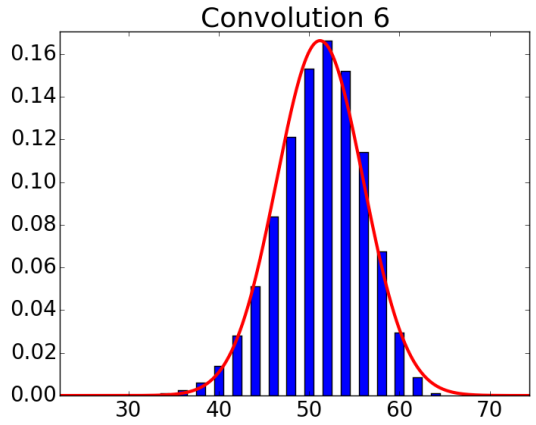


Figure 4.2: Discrete probability distribution obtained by coarse-graining uncorrelated loaded coins six successive times. The solid curve is a Gaussian with mean 51.2 and variance $\sigma^2 = 23.0$.

cumulants. Now, after we coarse-grain, let us *rescale the axis* by a factor $1/\sqrt{2}$. This restores the variance to its original value, because the variance involves two powers of x. We repeat the operations of coarse-graining and rescaling K times in succession. The mean and the variance remain unchanged, but what about the higher cumulants?

The cumulants of the original distribution are $\langle x^n \rangle_c$. Each coarse-graining doubles $\langle x^n \rangle_c$, and each rescaling multiplies $\langle x^n \rangle_c$ by a factor $(1/\sqrt{2})^n$. So, if $\langle z_K^n \rangle_c$ denotes the result of K iterations, then

$$\langle z_K^n \rangle_c = \langle x^n \rangle_c \left(\frac{2}{2^{n/2}} \right)^K = \langle x^n \rangle_c \left(2^{1-n/2} \right)^K . \tag{4.6}$$

For $n > 2$, repeating the operations of coarse-graining and rescaling will send $\langle z_K^n \rangle_c$ to zero. This is just the conclusion we drew in §2.8, phrased in terms of an iterative procedure. We say that the higher cumulants are *irrelevant*, because they shrink under the iteration.

If we hadn't invested time in developing our understanding of cumulants, we could have arrived at the same conclusion by going back to the basics of Fourier transforms. (This is the approach taken, for example, in Sethna's textbook [43].) We define the Fourier transform of a probability density function $p(x)$ as

$$\mathcal{F}[p](k) = \int_{-\infty}^{\infty} dx\, e^{-ikx} p(x) = \tilde{p}(k). \tag{4.7}$$

The inverse transformation is

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, e^{ikx} \tilde{p}(k). \tag{4.8}$$

We coarse-grain by convolution:

$$(p \star p)(x) = \int_{-\infty}^{\infty} dy\, p(x-y)p(y). \tag{4.9}$$

By the convolution theorem, the Fourier transform of a convolution is the product of the Fourier transforms of the functions convolved. In this case, convolving a curve with itself,

$$\mathcal{F}[p \star p](k) = \left(\tilde{p}(k)\right)^2. \tag{4.10}$$

The next step is to re-scale the curve, under the assumption that the mean is zero. This has the effect of creating a new function,

$$\mathcal{S}_{\sqrt{2}}[p](x) = \sqrt{2}p(\sqrt{2}x). \tag{4.11}$$

The prefactor of $\sqrt{2}$ is necessary to preserve normalization, as we can see by integrating:

$$\begin{aligned}\int_{-\infty}^{\infty} dx\, \mathcal{S}_{\sqrt{2}}[p](x) &= \sqrt{2}\int_{-\infty}^{\infty} dx\, p(\sqrt{2}x) \\ &= \sqrt{2}\int_{-\infty}^{\infty} \frac{dy}{\sqrt{2}} p(y) \\ &= 1.\end{aligned} \tag{4.12}$$

The change of length scale is equivalent to an inverse change of frequency:

$$\begin{aligned}\mathcal{F}\left[\mathcal{S}_{\sqrt{2}}[p]\right](k) &= \int_{-\infty}^{\infty} dx\, e^{-ikx}\sqrt{2}p(\sqrt{2}x) \\ &= \sqrt{2}\int_{-\infty}^{\infty} \frac{dy}{\sqrt{2}} e^{-iky/\sqrt{2}} p(y) \\ &= \tilde{p}(k/\sqrt{2}).\end{aligned} \tag{4.13}$$

It follows that the Gaussian curve

$$p^*(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{4.14}$$

is a fixed point of the coarse-grain-and-rescale transformation. We shall now prove this explicitly.

For brevity, we denote this composite transformation by $\mathcal{T}$. The $\mathcal{T}$ operator effects a convolution and a rescaling:

$$\mathcal{T}[p](x) = \mathcal{S}_{\sqrt{2}}[p \star p](x) = \sqrt{2}\int_{-\infty}^{\infty} dy\, p(\sqrt{2}x - y)p(y). \tag{4.15}$$

Explicitly,

$$\mathcal{T}[p^*](x) = \sqrt{2}\int_{-\infty}^{\infty} dy\, \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(\sqrt{2}x-y)^2}{2\sigma^2}\right) \exp\left(-\frac{y^2}{2\sigma^2}\right). \tag{4.16}$$

Expanding out and combining the arguments of the exponentials,

$$\mathcal{T}[p^*](x) = \frac{\sqrt{2}}{2\pi\sigma^2}\int_{-\infty}^{\infty} dy\, \exp\left(-\frac{2x^2 - 2\sqrt{2}xy + 2y^2}{2\sigma^2}\right). \tag{4.17}$$

With a bit of algebra,

$$\begin{aligned} 2x^2 - 2\sqrt{2}xy + 2y^2 &= 2y^2 - 2(\sqrt{2}y)x + 2x^2 \\ &= 2y^2 - 2(\sqrt{2}y)x + x^2 + x^2 \\ &= (\sqrt{2}y - x)^2 + x^2, \end{aligned} \tag{4.18}$$

we can make the integral over y a Gaussian integral, which we know how to evaluate. This eliminates the y dependence, leaving us with

$$\mathcal{T}[p^*](x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right) = p^*(x), \tag{4.19}$$

as desired.

The Fourier transform of p^* is

$$\tilde{p}^*(k) = \mathcal{F}[p^*] = \int_{-\infty}^{\infty}\frac{dx}{\sqrt{2\pi\sigma^2}}e^{-ikx - x^2/(2\sigma^2)} = \exp\left(-\frac{k^2\sigma^2}{2}\right). \tag{4.20}$$

Generally, the Fourier transform of a $\mathcal{T}$-transformed function is

$$\tilde{\mathcal{T}}[\tilde{p}](k) = \mathcal{F}\left[\mathcal{T}[p]\right](k) = \left(\tilde{p}(k/\sqrt{2})\right)^2, \tag{4.21}$$

with which it is easy to verify that

$$\tilde{\mathcal{T}}[\tilde{p^*}](k) = \mathcal{F}\left[\mathcal{T}[p^*]\right](k) = \tilde{p}(k). \tag{4.22}$$

That is, p^* is a fixed point in the space of probability distributions, and $\tilde{p}^*$ is a fixed point of the analogous transformation in the space of Fourier representations.

Let p be a distribution which is close to p^*:

$$p(x) = p^*(x) + \epsilon f(x). \tag{4.23}$$

Then, because the Fourier transform is linear,

$$\tilde{p}(k) = \tilde{p}^*(k) + \epsilon\tilde{f}(k). \tag{4.24}$$

What are the eigenfunctions and eigenvalues of $\mathcal{T}$?

$$\mathcal{T}[p^* + \epsilon f_n] = p^* + \lambda_n\epsilon f_n + \mathcal{O}(\epsilon^2), \tag{4.25}$$

$$\tilde{\mathcal{T}}[\tilde{p}^* + \epsilon\tilde{f}_n] = \tilde{p}^* + \lambda_n\epsilon\tilde{f}_n + \mathcal{O}(\epsilon^2). \tag{4.26}$$

Using Eq. (4.21),

$$\tilde{\mathcal{T}}[\tilde{p}^* + \epsilon\tilde{f}_n] = \left[\tilde{p}^*(k/\sqrt{2}) + \epsilon\tilde{f}_n(k/\sqrt{2})\right]^2 \tag{4.27}$$

$$= \left[\tilde{p}^*(k/\sqrt{2})\right]^2 + 2\epsilon\tilde{p}^*(k/\sqrt{2})\tilde{f}_n(k/\sqrt{2}) + \mathcal{O}(\epsilon^2) \tag{4.28}$$

$$= \tilde{p}^*(k) + 2\epsilon\tilde{p}^*(k/\sqrt{2})\tilde{f}_n(k/\sqrt{2}) + \mathcal{O}(\epsilon^2). \tag{4.29}$$

Therefore,

$$\tilde{f}_n(k) = \frac{2}{\lambda_n}\tilde{p}^*\left(\frac{k}{\sqrt{2}}\right)\tilde{f}_n\left(\frac{k}{\sqrt{2}}\right). \tag{4.30}$$

Proposal: let

$$\tilde{f}_n(k) = (ik)^n\tilde{p}^*(k). \tag{4.31}$$

With this ansatz,

$$\tilde{p}^*\left(\frac{k}{\sqrt{2}}\right)\tilde{f}_n\left(\frac{k}{\sqrt{2}}\right) = \left(\frac{ik}{\sqrt{2}}\right)^n\tilde{p}^*\left(\frac{k}{\sqrt{2}}\right) \tag{4.32}$$

$$= \frac{(ik)^n}{(\sqrt{2})^n}\tilde{p}^*(k) \tag{4.33}$$

$$= \frac{1}{(\sqrt{2})^n}\tilde{f}_n(k). \tag{4.34}$$

Consequently,

$$\tilde{f}_n(k) = (\sqrt{2})^n\tilde{p}^*\left(\frac{k}{\sqrt{2}}\right)\tilde{f}_n\left(\frac{k}{\sqrt{2}}\right), \tag{4.35}$$

meaning that the eigenvalue λ_n is given by

$$\frac{2}{\lambda_n} = (\sqrt{2})^n = 2^{n/2} \;\Rightarrow\; \lambda_n = 2^{1-n/2}. \tag{4.36}$$

The first eigenmode, associated with the eigenvalue $\lambda_0 = 2$, is

$$\tilde{f}_0(k) = \tilde{p}^*(k). \tag{4.37}$$

When applied as a perturbation to $\tilde{p}^*$, this yields

$$\tilde{p}(k) = \tilde{p}^*(k) + \epsilon\tilde{p}^*(k), \tag{4.38}$$

meaning that

$$p(x) = (1+\epsilon)p^*(x). \tag{4.39}$$

This is not a normalized probability distribution function, unless $\epsilon = 0$. Therefore, we can neglect the λ_0 mode as meaningless.

The next eigenfunction is

$$\tilde{f}_1(k) = ik\tilde{p}^*(k). \tag{4.40}$$

Applying this perturbation to $\tilde{p}^*$ yields a probability distribution

$$p(x) = \frac{1}{2\pi} \int dk\, (1 + \epsilon ik) e^{ikx} \tilde{p}^*(k). \tag{4.41}$$

Because ϵ is a small number, we can use a Taylor approximation:

$$p(x) = \frac{1}{2\pi} \int dk\, e^{ik\epsilon} e^{ikx} \tilde{p}^*(k) = \frac{1}{2\pi} \int dk\, e^{ik(\epsilon + x)} \tilde{p}^*(k). \tag{4.42}$$

Therefore,

$$p(x) = p^*(x + \epsilon). \tag{4.43}$$

That is, perturbing in the λ_1 eigenmode is equivalent to shifting the mean of $p^*(x)$.

The operation of coarse-graining and then rescaling is our first example of an RG transform, and the cumulants provide our first encounter with the concept that parameters can be relevant or irrelevant under an RG transform. We think of a *flow* in the space of parameters: in this case, the flow induced by the RG transform is the change in the cumulant values. The flow takes us to a *fixed point* which is, in this example, a Gaussian curve. Generally, a quantity is *relevant* if repeated applications of the RG transform cause it to grow, and a quantity is *irrelevant* if it shrinks as we approach the RG fixed point. (Quantities can also be *marginal,* if a linear stability analysis of an RG fixed point fails to indicate whether they will increase or decrease.)

This example also illustrates how fixed points of RG flows relate to *universality.* Gaussian curves show up throughout the sciences, because they arise whenever one considers the sum total effect due to a large number of uncorrelated small influences. Furthermore, once you understand one Gaussian curve, you've pretty much understood them all. This is the essence of a universality class. It is the fact that so many details are irrelevant, in the RG sense, that makes the Central Limit Theorem so powerful.

Using the calculus of variations, one can show that Gaussian curves maximize the Shannon index (also known as the Shannon entropy) for a given variance. That is, if we fix the variance to be σ^2, then the largest value of the Shannon index compatible with this constraint is

$$S_G(\sigma^2) = \frac{1}{2} \log_2(2\pi e \sigma^2). \tag{4.44}$$

As we repeat the RG transform, the variance remains constant, and the probability distribution looks more and more Gaussian, meaning that the Shannon index will approach S_G from below. The Shannon index is an example of a quantity which increases along with the RG flow, attaining a maximum at the flow's fixed point. When RG ideas are applied in the realm of field theory, this idea becomes the c-theorem of Zamolodchikov [44].

Finally, we saw that *exponents* turn out to be important quantities: in this example, they are the eigenvalues $\lambda_n = 2^{1-n/2}$. The importance of exponents that govern scaling behavior holds true across RG theory.

4.2 Extreme Value Distributions

We easily grow accustomed to seeing Gaussian distributions. Whenever we have a random variable whose value is the sum of many uncorrelated small influences, the Central Limit Theorem tells us to describe that variable by a Gaussian curve. But what about when we don't get a Gaussian?

Consider the set $\{H_1, H_2, \ldots, H_K\}$, where the H_i are independently drawn from some probability density $p(H)$. Instead of taking their sum, what if we want their *maximum*? Call this quantity X. What is the probability distribution $p_K(X)$?

We can sneak our way to this by first finding the probability that $X \leq S$, as a function of S. For this to occur, none of the $\{H_i\}$ can be bigger than S. The probability that a particular H_i is bigger than S is the integral of $p(H)$ from S to infinity. Because the distribution is the same for all the $\{H_i\}$, we find that this cumulative probability function is

$$P_K(X \leq S) = \left[1 - \int_S^\infty dH p(H)\right]^K . \tag{4.45}$$

If S is big enough that it is in the tail of the distribution $p(H)$, then the integral will be small, and so

$$P_K(X \leq S) \approx \exp\left[-K \int_S^\infty dH p(H)\right] . \tag{4.46}$$

To make progress, let's suppose that the tail of $p(H)$ decays exponentially. The integral over it is then

$$\int_S^\infty dH p(H) = \int_S^\infty dH a e^{-\lambda H} = \frac{a}{\lambda} e^{-\lambda H} . \tag{4.47}$$

We find the cumulative probability function $P_K(X \leq S)$ by taking the K^{th} power of this, yielding

$$P_K(X \leq S) = \exp\left[-\frac{Ka}{\lambda} e^{-\lambda S}\right] . \tag{4.48}$$

Differentiating this gives the probability density we want, $p_K(X \leq S)$:

$$p_K(X \leq S) = \frac{dP_K(S)}{dS} = ka \exp\left(-\lambda S - \frac{ka}{\lambda} e^{-\lambda S}\right) . \tag{4.49}$$

We can find where this distribution peaks by differentiating the exponential and setting it to zero:

$$\lambda - kae^{-\lambda S^*} = 0 , \tag{4.50}$$

which occurs at an S value of

$$S^* = \frac{1}{\lambda} \log\left(\frac{ka}{\lambda}\right) . \tag{4.51}$$

Knowing this, we can tidy up the probability density by expressing it in terms of the peak position:

$$p_K(X \leq S) = \lambda \exp\left(-\lambda(S - S^*) - e^{-\lambda(S - S^*)}\right) . \tag{4.52}$$

So, we have a two-parameter distribution, whose shape is fixed by specifying S^* and λ. We call this the *Gumbel* or *Fisher–Tippett distribution*; unlike a Gaussian, it is asymmetrical, featuring an exponential tail above S^* and a much more rapid decay below S^*.

We assumed that the underlying distribution decays exponentially in its tail region. Another possibility is that it decays as a power law:

$$p(X) \propto x^{-(1+\alpha)},\ \alpha > 0\,. \tag{4.53}$$

In this case, one finds instead that

$$p_k(X \leq S) = \exp(-S^{-\alpha})\,, \tag{4.54}$$

the so-called *Fréchet* cumulative distribution. A third possibility is that the tail is bounded, that as $x \to a$ the probability density $p(x)$ decays as $(a-x)^{\beta-1}$ for some $\beta \geq 0$. This yields a *Weibull* distribution, whose cumulative curve is

$$p_k(X \leq S) = \exp(-(-S)^{\beta})\,. \tag{4.55}$$

These three distributions provide three more universality classes. The *Fisher–Tippett–Gnedenko theorem* establishes that any distribution of extreme-value statistics will follow one of these three curves, depending upon the tail behavior of the original probability distribution. We can apply RG methods to this, as we did to the Central Limit Theorem. Instead of summing two random variables, we group them into pairs and discard the smaller value in each pair. Once issues of scaling are accounted for properly, the Gumbel, Fréchet and Weibull distributions are each fixed points under this transformation [43]. In the RG derivation of the Central Limit Theorem, we convolve together many random variables, meaning that we take the product of their characteristic functions, whereas in the theory of extreme values, we take the product of the cumulative distributions.

4.3 Isotropic Percolation

In the previous section, we did not put any spatial structure on the set of system components. When we decided to coarse-grain two variables together, any two variables were as good as any other pair. Now, we will move on to a problem where we can apply RG theory to a system that has spatial structure. Our next example of an RG analysis, which will move us closer to eco-evolutionary models, is the topic of *isotropic percolation.*

Percolation is the study of flow through randomized media. In *isotropic* percolation, no spatial orientation is picked out as special. One can also study *directed* percolation, in which there is a preferred direction (*e.g.,* downhill). We will not discuss the directed version in this chapter.

Typically, the first step in defining an isotropic percolation problem is to construct a regular lattice. Each point, or vertex, in the lattice is connected to the same number of other vertices. One way to proceed is to imagine that each edge can be colored with

one of two possible pigments, *e.g.,* white and black. The color of each edge is picked at random, independently of the choices for all other edges. For example, we can say that the probability an edge is marked with black is p, and so an edge chosen at random will be white with probability $1 - p$. We can then ask, for any value of p, what the sizes of the clusters formed by black edges will be. This defines a *bond percolation* problem.

Alternatively, we can make the lattice points our variables of interest. For example, we can say that each lattice point is either *filled* or *empty,* and we fill sites at random with probability p. Depending on the value of p, we will have different statistical distributions for the sizes of the clusters formed by adjacent filled points. This variant is known as *site percolation.*

What constitutes an RG transform in this context? It is easy to imagine what coarse-graining a lattice might mean: we can simply blur out the fine-scale details. This operation will create a new lattice based on the configuration of the original. One way to do this for site percolation is to consider each lattice point together with its immediate neighbors. We "collapse" this set of points down to a single site in the new lattice, and we choose whether the new site is filled or not based on how many of the original sites are filled. Applying this to all the sites in the original lattice creates a new graph in which the small-scale information has been blurred away. It is then convenient to "step back from the picture," rescaling the lengths of the lattice edges so that neighboring points in the derived lattice are separated by the same distance as those in the original. This combination of coarse-graining followed by rescaling is an RG transform.

Suppose that p is small. Then the lattice sites will mostly be empty, though we will have some small islands of filled sites amidst the sea of emptiness. Applying the RG transform will tend to make each of these islands smaller, because filled sites on the edges of the islands will have too many empty neighbors for their coarse-grained images to be filled. So, heuristically speaking, the RG transform will create a new lattice configuration that looks like what we would find *for a smaller value of* p.

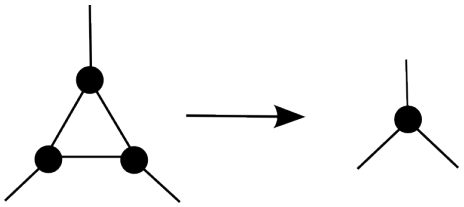
At the opposite extreme, suppose that p is close to 1. Then most of the sites will be filled, with some low density of empty regions. The RG transform will blur away small details, making the smallest empty regions vanish. Empty regions which aren't quite small enough to vanish entirely will be shrunk in the transformed image. So, the result will be a lattice configuration that looks like what we would typically find *for a larger initial choice of* p.

If small values of the filling probability flow to even smaller ones, and large values flow to even larger ones, then there must be a "continental divide" somewhere in the middle, where p is mapped to *itself.* At this point, we expect to find clusters in a continuous range of sizes: the RG transform erases the smallest clusters, so there must be slightly larger ones to replace them, and so forth. When p equals this *critical* value p_c, the distribution of cluster sizes will be *scale-invariant.* Because the lattice will contain clumps of all sizes, when we study the overall properties of the system, the details of how the connections between vertices look at the smallest scale should not matter much. This is the root of why we can group these systems into universality classes.

To make these qualitative considerations concrete, we will work through an example of site percolation. We take a triangular lattice and fill in the lattice sites at random. Let p denote the probability that a lattice site chosen at random is filled. We will see how a simple RG transform can send one value of p to another.

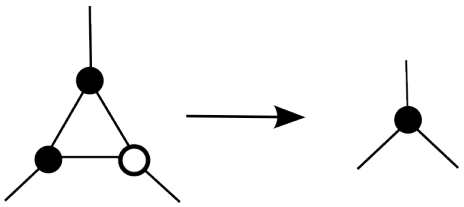
The coarse-graining operation should preserve the connectedness of clusters. This is rather tricky to enforce with a simple mapping, but one way of coming close is to use a *majority rule.* We map a triplet to a filled site if at least two of the original three sites are filled. This turns out to provide a simple approximation that gives remarkably good agreement with the best known results for the triangular lattice.

How does the majority rule create a mapping between values of p? With probability p^3, all three vertices of a triangle will be filled, and when this configuration occurs, we coarse-grain it to a single occupied site:



(4.56)

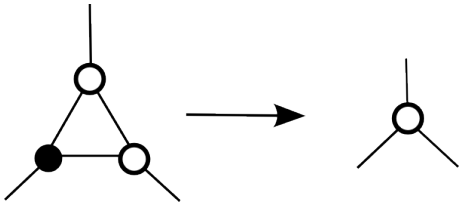
Likewise, if two out of three vertices are occupied, the corresponding site in the coarse-grained lattice is filled:



(4.57)

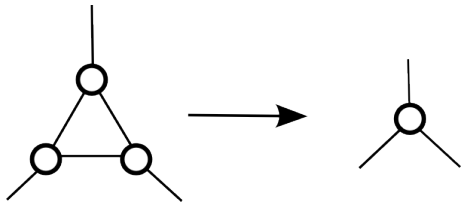
Because there are three ways to choose which vertex is unoccupied, this situation carries a symmetry factor, occuring with probability $3p^2(1-p)$.

We obtain an empty vertex in the coarse-grained lattice if two sites are open:



(4.58)

Or if all three vertices in the original triangle are unfilled:



(4.59)

Let p' denote the probability that a randomly chosen site in the coarse-grained lattice is filled. From the relations above, we see that

$$p' = p^3 + 3p^2(1-p). \tag{4.60}$$

The fixed points of this mapping are given by

$$p_c = p_c^3 + 3p_c^2(1-p_c). \tag{4.61}$$

Trivially, this equation is solved by 0 and by 1, and it has a less obvious solution at

$$p_c = \frac{1}{2}. \tag{4.62}$$

We will call this the *critical* value of the site-filling probability. What happens if p is almost but not quite equal to the critical value? Let us define

$$p = p_c + \delta p,\ p' = p_c + \delta p'. \tag{4.63}$$

We can relate p and p' by the RG flow equation (4.60):

$$p_c + \delta p' = (p_c + \delta p)^3 + 3(p_c + \delta p)^2(1 - p_c - \delta p). \tag{4.64}$$

Expanding out the binomials, multiplying and simplifying, we arrive at

$$p_c + \delta p' = p_c + 6p_c(1-p_c)\delta p. \tag{4.65}$$

Therefore,

$$\delta p' = \frac{3}{2}\delta p. \tag{4.66}$$

We see that a small deviation from the critical value becomes a larger one.

Are there exponents for the percolation problem as there were in the probability scenario we studied in the previous section? In fact, there are. The most important in practice is ν, which relates the characteristic length scale—the typical size of connected clumps—to the distance from the critical threshold:

$$\xi \propto |p - p_c|^{-\nu}. \tag{4.67}$$

Here, we are measuring length in units of the lattice spacing. That the characteristic length ξ varies in this fashion is perhaps most directly appreciated by numerical simulation. We expect that the characteristic length will diverge at the critical threshold, thanks to the scale-invariance we described above.

Coarse-graining a lattice will produce a new grid in which the vertices are further apart. It is traditional to denote the factor by which the vertex separation is increased by the letter b. If $\xi(p)$ is the characteristic cluster size for filling probability p, then $\xi(p')$ is the characteristic cluster size on the lattice produced by coarse-graining, measured in units of the coarse-grained lattice spacing. But this must be related to the original average cluster size by the scaling factor:

$$\xi(p) = b\xi(p'). \tag{4.68}$$

For example, if we projected our original lattice on a wall and measured the typical cluster size to be four centimeters, then coarse-graining will produce a new image, containing fewer pixels per unit area, in which the typical cluster size must still be four centimeters.

If we rewrite our previous expression in terms of the distance from the threshold,

$$\xi(p_c + \delta p) = b\xi(p_c + \delta p'), \tag{4.69}$$

substituting in the scaling form involving the exponent ν yields

$$|\delta p|^{-\nu} = b|\delta p'|^{-\nu}. \tag{4.70}$$

We know from Eq. (4.66) how δp and $\delta p'$ are related, and so we can say that

$$|\delta p|^{-\nu} = b\left(\frac{3}{2}\right)^{-\nu}|\delta p|^{-\nu}. \tag{4.71}$$

Canceling the common factor and solving for ν,

$$\nu = \frac{\log b}{\log(3/2)}. \tag{4.72}$$

We can show geometrically that for this coarse-graining operation on the triangular lattice, $b = \sqrt{3}$. Therefore,

$$\nu = \frac{\log\sqrt{3}}{\log(3/2)} \approx 1.355. \tag{4.73}$$

The *critical exponents* like ν are constant across all members of a universality class. However, other quantities, like the locations of the critical thresholds, are typically model-specific. This means that ν will be the same on a square lattice, for example, while p_c will be different (in fact, it is roughly 0.593).

In order to illustrate the RG concepts, this section has worked through an example where the theory is fairly straightforward to apply. It turns out that had we chosen to start with a square lattice instead of a triangular one, our task would have been significantly more difficult, and we'd have to sweat a bit more even to get an *approximation* for ν and for p_c.

We have seen RG theory at work in two rather different contexts. It can be deployed in many more, and I am not aware of a comprehensive reference that covers them all, or even the ones that are by now well-established. Introductory texts include those by Creswick, Farach and Poole [45] and by McComb [46]. The textbooks by Kardar [28] and by Zee [47] cover some aspects of RG in field theory. Bar-Yam [48] provides a succinct first course on RG that includes its classic use in chaos theory. The admirably slim volume by Cardy [49] includes a study of directed percolation as well.

4.4 Application: Light Scattering

The binary fluid mixture of methanol and cyclohexane [50, 51, 52] is an interesting substance to study, because it provides a convenient example of a phase transition in the 3D Ising universality class [53, 54]. In this section, I will describe an experiment which probes this phase transition using laser light. Measuring the intensity of the light scattered by the methanol–cyclohexane mixture reveals the quantitative character of the transition.

The light source is a 632 nm He–Ne laser. The laser beam passes through the methanol–cyclohexane sample, which is held in an insulating container whose temperature can be controlled electronically. Most of the laser light passes through the methanol–cyclohexane mixture, but some is scattered. Both the methanol and the cyclohexane are transparent by themselves, but their indices of refraction are sufficiently different that a light beam will bend noticeably when passing through an interface between them. Bubbles of one liquid floating within the other will, therefore, scatter light. Laser light scattered from the sample is measured by a photodetector. The average of the recorded intensity (neglecting large spikes due to dust particles floating through the laser beam) and the autocorrelation properties of the signal yield important information about the sample from which the light was scattered.

Ornstein–Zernike scattering theory [50] relates the effectiveness of the fluid mixture at scattering light to the fluid's isothermal compressibility. In turn, the isothermal compressibility is a thermodynamic quantity which obeys a power-law scaling relationship near the critical point. Its divergence as T approaches T_c has the same critical exponent γ as the order-parameter susceptibility $\chi(T)$.

Combining the understanding of critical-point behavior, which tells us how big fluctuations ought to be, with scattering theory, which tells us how fluctuations affect light, implies [50] that the intensity $I_{\text{norm}}(T)$ should vary slowly near T_c and fall off as a power law for T significantly larger than T_c. The exponent of this power law, γ, is the same as the critical exponent for the order-parameter susceptibility $\chi(T)$. In more detail, we can write the following expression for the intensity of the scattered light:

$$I_{\text{norm}}(T) = 1 - \left(1 - \frac{I_{\text{max}}}{I_0}\right) \exp\left[-C\left(\frac{T - T_C}{T_C}\right)^{-\gamma}\right]. \tag{4.74}$$

Here, I_0 is the "baseline" intensity we see when the temperature is far above the critical point and the mixture is a homogeneous, transparent fluid. I_{max} is the maximum

intensity we record. T_c is, as above, the critical temperature, and the constant C indicates the amount of material which the laser beam passes through. I_0, $I_{\max}$ and C depend on the particular experimental setup used, and T_c depends on the fact that the mixture is methanol–cyclohexane instead of something else. However, γ is universal: That is, it should be the same for all systems in the 3D Ising universality class.

The simplest approximation to the dynamics is a *mean-field analysis,* wherein we assume that each bit of the system feels the same average influence from the rest of the system as all the other bits. Here, such a calculation shows that $\gamma = 1$. We can in fact read this off from our analysis of the Ising model in §19.3, where we conclude that

$$\chi(T - T_c) \propto \frac{1}{|T - T_c|} \tag{4.75}$$

on either side of the phase transition. If the data from the light-scattering experiment shows $\gamma \neq 1$, then we can say that the mean-field approximation is not applicable to the methanol–cyclohexane transition.

Figures 4.3 and 4.4 show the intensity $I_{\text{norm}}(T)$ as a function of temperature. First, in Figure 4.3, we see the intensity plotted against the temperature in degrees Celsius. Fitting the parameters I_0, T_c, C and γ in Eq. (4.74) to the experimental observations yields a critical temperature of

$$T_c = 321.5 \pm 0.1\,K, \tag{4.76}$$

and a critical exponent of

$$\gamma = 1.234 \pm 0.005. \tag{4.77}$$

In Figure 4.4, the intensity I_{norm} is plotted against the reduced temperature $(T - T_c)/T_c$. The straight-line decay for larger temperatures makes a power-law dependence an obvious candidate for a functional form.

Because the γ found by experiment is larger than the mean-field value of 1, we can say with some confidence that the mean-field approximation is inadequate for this phase transition.

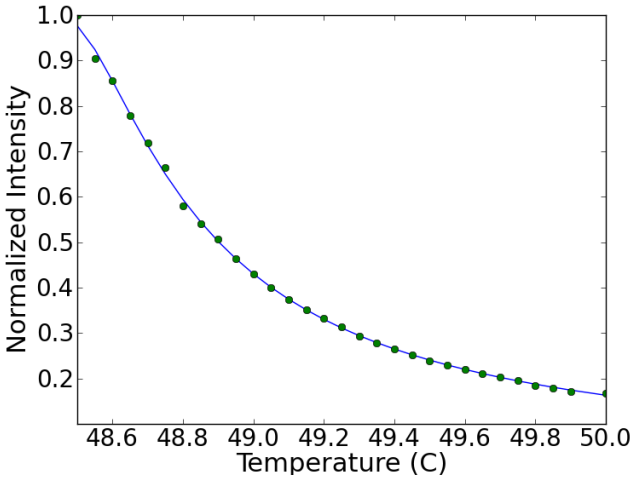


Figure 4.3: Intensity of scattered light as a function of temperature. The amount of light scattered by the sample increases as the temperature approaches the critical point from above. This is an indication of critical opalescence.

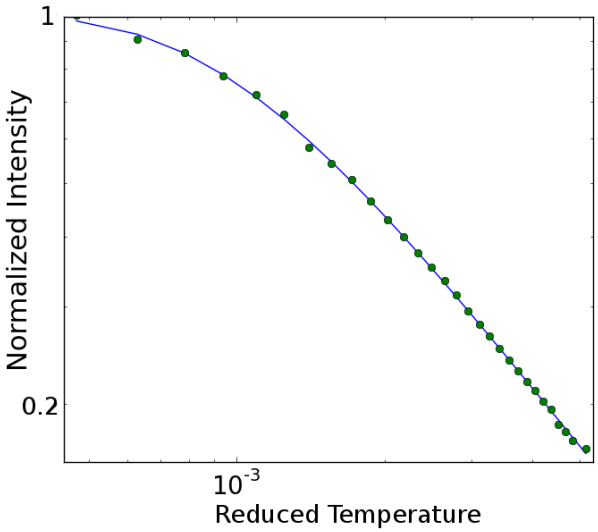


Figure 4.4: Intensity of scattered light as a function of reduced temperature $(T - T_c)/T_c$. The straight-line decay for increasing temperature is indicative of a power-law dependence, as expected from RG calculations combined with Ornstein–Zernike scattering theory [50].

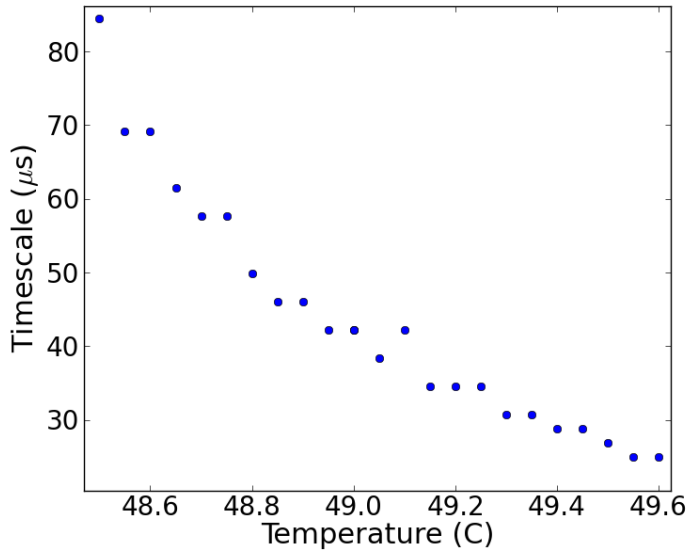


Figure 4.5: Characteristic dynamical timescale, derived from the intensity autocorrelation, as a function of temperature. The increase as the temperature approaches the critical point is an indicator of critical slowing down [28].

5 Laws of Thermodynamics

5.1 Introduction

Every specialty has its own take on thermodynamics. The vocabulary of the mechanical engineer is not the same as that of the chemist, and what is of primary concern to one may be a secondary matter for the other. Physicists, of course, claim to have the deepest understanding of all. I do actually know one very good mechanical engineer who saw the textbook we used for graduate statistical physics, with its opening chapter on thermodynamics, and said that in their department, they spent a whole semester on the topics of that chapter. Of course, taking an entire semester opens up more time to study individual examples, so if you need an answer about a steel pipe of a specific mass in a specific red-brick mill building, you are likely better off asking the engineer.

Presuming that the reader has at least brushed against the subject before, we will speedrun the introductory terminology. A *system* in thermodynamics is any portion of the universe that we conceptually separate from the rest for study. We often speak of the *surroundings* or the *environment* of a system, by which we mean that portion of the rest of the universe that is relevant to the system of interest. The traditional terrain of thermodynamics is macroscopic systems, and indeed systems that are large enough to be divisible into smaller pieces that are themselves still macroscopic. We carry over the concept of *work* from mechanics, and to it we add the idea of changes that do not involve mechanical work, particularly the flow of *heat* but also changes of chemical composition. The *thermodynamic state* of a system is its manner and mode of being at a given time, which we describe numerically using *thermodynamic parameters* (volume, pressure, magnetization and so forth). When a system's thermodynamic state does not change over time, we say that it is in *thermal equilibrium.* We do, of course, want to understand systems that do change over time. The simplest dynamical processes are those in which the system, glimpsed at any single instant, appears to be unchanging: I.e., the change is smooth and slow enough that imagining the system to be in equilibrium is still an effective model. Such processes are called *quasistatic.*

The alert reader may have noticed a certain vagueness in these definitions. How slow is "slow enough" to qualify as quasistatic? And for that matter, how big is big enough to count as macroscopic? These are good questions. One might ask, by analogy, how fine must a paintbrush be that the mark it makes qualifies as a Euclidean line. The answer, which may infuriate those who seek certainty in all the wrong places, is that *it depends.* Invoking a mathematical idealization may be appropriate in one context and inappropriate in another.

Our progress into thermodynamics will involve a sequence of increasingly abstruse

ideas. First, there is *temperature,* and then *energy,* after which we will meet *entropy* and then combine them all into *free energy.*

To begin, let us take up the idea of temperature. In daily experience, without trying to be quantitative about it, we find that subjective sensations of hot and cold correlate with various routine phenomena. It is not immediately obvious that boiling water is "higher" in a meaningful way than snow, for example, or that hotness and coldness can be put on a numerical scale. By analogy, it is common (though not universal) to have sensations of color, but civilization took thousands of years to organize colors by wavelength, and longer still to recognize that visible light is only a thin slice of the electromagnetic spectrum. Likewise, we did a lot of living and dying before we made temperature numerical. And having introduced the idea of a thermometer that provides a quantitative reading, it is still not immediately obvious that all thermometers, with their different "thermometric properties", are really measuring "the same thing". In practice, one must calibrate new thermometers against the old. The conceptual principle that distills down all the empirical results which say we can do this is the *Zeroth Law* of thermodynamics. The Zeroth Law says that if systems A and B are separately each in thermal equilibrium with a third system C, then the two systems A and B will also be in thermal equilibrium with each other. So, thermal equilibrium is an "equivalence relation" that lets us carve up the set of all possible thermal states into nonoverlapping subsets. Temperature is just the label of which equivalence class a thermal state falls into.[1]

Energy is a concept with which a physics student gradually grows more familiar, eventually arriving in a strange place, or at least a place that seems quite foreign from the starting point. What does a moving train have in common with a hefted rock and a bag of sugar? What do all these have in common with sunshine? We can illustrate the oddity by recalling how the concept enters an archetypal physics text, the *Feynman Lectures on Physics.* In the first volume, energy is abstract: the conservation of energy

> states that there is a certain quantity, which we call energy, that does not change in the manifold changes which nature undergoes. That is a most abstract idea, because it is a mathematical principle; it says that there is a numerical quantity which does not change when something happens. It is not a description of a mechanism, or anything concrete; it is just a strange fact that we can calculate some number and when we finish watching nature go through her tricks and calculate the number again, it is the same. (Something like the bishop on a red square, and after a number of moves—details unknown—it is still on some red square. It is a law of this nature.)

The conservation law is then illustrated by analogy, using the parable of a child's toy blocks. The child has a fixed number of blocks, which get lost in increasingly difficult-to-detect ways: the augmented weight of a toy chest, the elevation of the water level

[1] To appreciate the significance of the Zeroth Law, it helps to consider situations where the pairings are *not* "transitive" like this. For example, if Alice and Bob both separately want to date Charlie, it may or may not be the case that Alice and Bob want to date each other. The logistics of such relationships are beyond the scope of this text.

in the bathtub and so forth. Trying to tally the blocks, one "finds a whole series of terms representing ways of calculating how many blocks are in places" that one cannot examine directly, while the sum of that series remains constant. But the analogy is not exact. When it comes to energy, "*there are no blocks*" (emphasis in original). The only concrete term in the series is gone, leaving only the abstract sum.

Thermodynamics is the subject that makes mechanics honest. It is where we learn how to maintain the conservation of energy by accounting for the energy lost as bodies slow down and stop moving. The *First Law of thermodynamics* is the conservation of energy.

In everyday speech, we are not so careful about distinguishing between heat and temperature. Phrases like "bring the heat" and "raise the temperature" might overlap in meaning. But, being more scientific, there is a distinction. If I am baking two loaves of bread, they are at the same temperature in the oven, but it seems quite reasonable that the total heat soaked up within both loaves together is twice that of one loaf alone. So, heat and temperature are related but not identical, and thus the natural next step is to compare them. Dividing heat by temperature gives *entropy,* the key concept at the next level of sophistication. More specifically, entropy was born in the nineteenth-century study of heat engines, when physicists discovered the importance of the ratio between the heat flow in an engine to the temperature at which that heat flow is happening. In a perfect engine, an engine too good for this cruel world, the entropy is constant: The ratio is the same for the input heat that drives the engine and for the waste heat exhaust. But the concept did not stop there, instead gaining a new life in the twentieth century. The modern understanding of it requires probability mathematics and the fusion of thermodynamics with information theory.

Although a pop explanation of entropy tends to invite a grand, cosmic perspective, it is often more practical to set up our equations in terms of the system right in front of us. Energy that is in principle available to do useful actions is called *free energy.* Different formulae are applicable in different circumstances, but they all have the form of the total energy reduced by an amount that depends on temperature and entropy.

5.2 Types of and Notations for Changes

Thermodynamics is a subject that uses at least five different symbols to mean "change in". A Greek capital delta Δ denotes a change of arbitrary size, and a d indicates an infinitesimal change. But as we have seen, some changes depend upon the path we follow, whereas others depend only upon the endpoints. Temperature, energy, pressure, volume and entropy are examples of "state functions", changes in which depend only on the initial and final states, not the details of the process connecting them. Heat and work, on the other hand, are path-dependent quantities. So, when we divide an energy change dE into heat and work contributions, we can incorporate this distinction into our notation:

$$dE = đQ + đW \,. \tag{5.1}$$

Another distinction that is sometimes helpful to include is that between equilibrium and nonequilibrium processes. The latter are sometimes denoted by a lowercase delta, δ.

In addition to Δ, d, đ and δ, we use the symbol ∂ to denote *partial derivatives,* rates of change in a function of multiple variables with respect to one of them. For example, suppose that

$$f(x, y) = 3xy^2 \,. \tag{5.2}$$

Then the partial derivative of f with respect to x at constant y is

$$\left.\frac{\partial f}{\partial x}\right|_y = 3y^2 \,, \tag{5.3}$$

while the other partial derivative, with respect to y at constant x, is

$$\left.\frac{\partial f}{\partial y}\right|_x = 6xy \,. \tag{5.4}$$

The variable that is not being varied may be written as a subscript to a vertical bar, as it is here, or to a closing parenthesis. If it is clear from context, it may be omitted for brevity.

Partial derivatives are not that much more demanding than those familiar from introductory calculus. For example, suppose that

$$z = xy \,. \tag{5.5}$$

Then we can tell that

$$\left.\frac{\partial z}{\partial x}\right|_y = y \,, \tag{5.6}$$

and because $x = z/y$, we know that

$$\left.\frac{\partial x}{\partial z}\right|_y = \frac{1}{y} \,, \tag{5.7}$$

Unsurprisingly, then,

$$\left.\frac{\partial z}{\partial x}\right|_y = \left(\left.\frac{\partial x}{\partial z}\right|_y\right)^{-1} \,. \tag{5.8}$$

But notice what happens when we evaluate

$$\left.\frac{\partial x}{\partial y}\right|_z = -\frac{z}{y^2}, \;\; \left.\frac{\partial y}{\partial z}\right|_x = \frac{1}{x} \,. \tag{5.9}$$

Multiplying these together with our earlier result gives

$$\left.\frac{\partial x}{\partial y}\right|_z \left.\frac{\partial z}{\partial x}\right|_y \left.\frac{\partial y}{\partial z}\right|_x = \left(-\frac{z}{y^2}\right)(y)\left(\frac{1}{x}\right) = -\frac{z}{xy} = -1 \,. \tag{5.10}$$

Both Eq. (5.8) and Eq. (5.10) are examples of general facts.

As physicists, we are accustomed to treating the dx and dy in an ordinary derivative as tiny quantities, like one part in a billionth, that we can manipulate as though they were numbers. Take the chain rule, for instance:

$$\frac{dx}{dy} = \frac{dx}{du}\frac{du}{dy}\,. \tag{5.11}$$

Just cancel the factors of du; what could be simpler? In truth, one doesn't go far wrong thinking this way, as long as one learns the ropes through practice. But we can't treat ∂x the same way we could dx, because ∂x is too squiggly to be meaningful on its own. We can't separate it from the context of what other variables are being held fixed.

5.3 A Microscopic Picture of Thermodynamics

From high-school chemistry on, we have the *ideal gas law* burned into our brains:

$$PV = Nk_BT. \tag{5.12}$$

This relates the pressure P and the volume V of an ideal gas to N, the number of atoms present in that gas, and T, the temperature of the gas. The Boltzmann constant k_B is a quantity we introduce because historically, the concept of temperature predated that of energy, and it is still convenient sometimes to measure them in different units. Thinking of mechanical work, a convenient amount of energy would be that which is present when a standard mass is accelerated to a standard speed. On the other hand, a convenient unit of temperature might be what we get when we split the interval between the freezing and boiling points of water into reasonable increments. The constant k_B intermediates between these two notions of convenience.

One of the earliest bits of wisdom to be codified about kinetic theory is the microscopic picture of how collisions by ideal-gas atoms give rise to pressure. If we imagine a featureless little lump of mass m moving with velocity $\vec{v}$, we know that it carries a momentum $m\vec{v}$. Suppose that our little corpuscle is heading towards a flat wall that is perpendicular to the $\hat{x}$-axis. When the corpuscle hits the wall, it will be reflected back, elastically, so that the $\hat{x}$-component of its momentum will change from mv_x to $-mv_x$. That is, the change in the corpuscle's momentum is $-2mv_x$ along $\hat{x}$, so the wall must be imparted a change of momentum equal in magnitude but opposite in direction, $2mv_x\hat{x}$.

The *force* that the gas applies to the wall is the net effect of such collisions per unit time. Over an area A of wall surface, and supposing that the number density of gas atoms is N/V, the momentum change imparted in a time interval dt is

$$F\,dt = \frac{1}{2}\frac{N}{V}A\,dt\,v_x \cdot 2mv_x. \tag{5.13}$$

The first factor tells us how many collisions are taking place, with the $1/2$ included to make sure we count only those atoms that are moving *towards* the wall.

Now, the velocity of a gas atom is really best thought of as a random variable, so we should be talking about the expected value of the force:

$$\langle F \rangle = \frac{N}{V} Am \langle v_x^2 \rangle. \tag{5.14}$$

But the force per unit area is just the pressure:

$$P = \frac{N}{V} m \langle v_x^2 \rangle. \tag{5.15}$$

It is reasonable to assume that there is no preferred axis of motion, meaning that

$$\langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle. \tag{5.16}$$

By the Pythagorean theorem,

$$\langle v^2 \rangle = \langle v_x^2 + v_y^2 + v_z^2 \rangle, \tag{5.17}$$

and because expectation values with respect to the same probability distribution are additive,

$$\langle v^2 \rangle = \langle v_x^2 \rangle + \langle v_y^2 \rangle + \langle v_z^2 \rangle = 3 \langle v_x^2 \rangle. \tag{5.18}$$

So, putting the pieces together,

$$PV = Nm \left(\frac{1}{3} \langle v^2 \rangle \right). \tag{5.19}$$

We can recognize the expectation value of the kinetic energy:

$$PV = \frac{2}{3} N \left\langle \frac{1}{2} m v^2 \right\rangle. \tag{5.20}$$

Comparing this with the ideal gas law, we find that

$$\langle K \rangle = \left\langle \frac{1}{2} m v^2 \right\rangle = \frac{3}{2} k_B T. \tag{5.21}$$

The 3 in this expression is the dimensionality of space: For each possible axis of motion, we have $\frac{1}{2} k_B T$ of energy.

5.4 Energy Conservation, Work and Heat

One of the simplest things we can do to an ideal gas is to squeeze it. Let us imagine doing so *reversibly,* that is, slowly enough that nothing dramatic happens and we can model its mode of being at each instant as an equilibrium state. What energy is expended in compressing an ideal gas by a given amount? The infinitesimal amount of work done when squeezing a piston of area A against a pressure P and moving it a distance dx is

$$đW = PAdx = -PdV\,. \tag{5.22}$$

The minus sign appears because the volume is being diminished here. Integrating this from an initial volume V_1 to a final V_2, we find

$$W = -\int_{V_1}^{V_2} P dV \,. \tag{5.23}$$

So, we can calculate work by finding the area under a curve on a (P, V) diagram.

Suppose an ideal gas undergoes an isothermal volume change. Then the work done on the gas is

$$W = -Nk_BT \int_{V_1}^{V_2} \frac{1}{V} dV \,. \tag{5.24}$$

Integrating V^{-1} gives the natural logarithm function:

$$W = -Nk_BT \log \frac{V_2}{V_1} \,. \tag{5.25}$$

If the gas is being compressed, $V_1 > V_2$ and so $W > 0$, meaning that positive work is required on the gas to make it shrink, so the sign checks out. We have $W < 0$ if the gas expands.

We can split an infinitesimal energy change into contributions due to heat and due to work:

$$dE = đQ + đW = đQ - PdV \,. \tag{5.26}$$

During an adiabatic process, $đQ = 0$, and so

$$dE = -PdV \,. \tag{5.27}$$

The pressure is therefore minus the derivative of the energy with respect to volume:

$$\frac{dE}{dV} = -P \,. \tag{5.28}$$

But we also know that

$$E = \frac{3}{2} Nk_BT = \frac{3}{2} PV \,, \tag{5.29}$$

which we differentiate to find

$$\frac{dE}{dV} = \frac{3}{2} \frac{dP}{dV} V + \frac{3}{2} P \,. \tag{5.30}$$

These two expressions for the derivative of E with respect to V must equal each other. Combining them and simplifying, we get

$$\frac{dP}{dV} V + \frac{5}{3} P = 0 \,. \tag{5.31}$$

The left-hand side looks almost like something that would come out of the product rule. We have the derivative of pressure, times the volume; and then we have the

pressure, times something that might have been left behind by taking dV/dV. We can exploit this resemblance by multiplying through by $V^{2/3}$:

$$\frac{dP}{dV}V^{5/3} + \frac{5}{3}PV^{2/3} = 0\,. \tag{5.32}$$

Now, the left-hand side is exactly in product-rule form:

$$\frac{d}{dV}\left(PV^{5/3}\right) = 0\,. \tag{5.33}$$

We have deduced that over the course of an adiabatic process, the product $PV^{5/3}$ must be constant!

We can construct an engine cycle out of isotherms and adiabats by making a closed loop in the (P, V) plane. We start at a point, call it 1, and then descend along an isotherm to another point, call it 2. From point 2, we drop along an adiabat (a steeper curve than an isotherm) to point 3, which is at a lower temperature. Then, we move back up along the isotherm through point 3, until we hit point 4, from which we can go up along an adiabat back to point 1. Let us call the temperature along the upper isotherm T_H and that along the lower isotherm T_C. By the ideal gas law, $P = Nk_BT/V$, so along an adiabat, $Nk_BTV^{2/3}$ must be constant. Points 2 and 3 are on the same adiabat, so

$$T_HV_2^{2/3} = T_CV_3^{2/3}\,, \tag{5.34}$$

and likewise,

$$T_CV_4^{2/3} = T_HV_1^{2/3}\,. \tag{5.35}$$

Combining these, we get

$$\frac{V_3}{V_4} = \frac{V_2}{V_1}\,. \tag{5.36}$$

Along an isotherm, $dE = 0$, and so the heat flow into the gas must be balanced exactly by the work going out of it:

$$đQ = PdV\,. \tag{5.37}$$

We can integrate the total heat flow into the gas along the T_H isotherm:

$$Q_H = Nk_BT_H\int_{V_1}^{V_2}\frac{1}{V}dV = Nk_BT_H\log\frac{V_2}{V_1}\,. \tag{5.38}$$

Likewise, the total heat flow *out* along the T_C isotherm is

$$Q_C = Nk_BT_C\log\frac{V_3}{V_4}\,. \tag{5.39}$$

The ratio of the input heat to the output heat is

$$\frac{Q_H}{Q_C} = \frac{T_H\log(V_2/V_1)}{T_C\log(V_3/V_4)} = \frac{T_H}{T_C}\,. \tag{5.40}$$

Another way of expressing this equality is to say that the ratio of the heat that flows to the temperature at which it flows is constant:

$$\frac{Q_H}{T_H} = \frac{Q_C}{T_C}\,. \tag{5.41}$$

5.5 Heat Engines and Entropy

Imagine that we have a thermodynamic system that we can use as a three-stroke engine, with two of the strokes in the cycle being reversible adiabatic paths. Let's say that the system has an internal energy E and two other variables, X and Y (which might be volume and magnetization, for instance). The system starts in state A. One adiabat goes from A to B, and another goes from A to C. The states B and C are assumed to have the same values of X and Y but different energies.

$$E_A > E_C > E_B \,. \tag{5.42}$$

The first stage of the cycle is the adiabatic transition from A to B. This lowers the energy, and since no heat is flowing, the system must be doing work, $E_A - E_B$. The next stage is going from B to C. In this step, X and Y remain constant, so the increase in energy $E_C - E_B$ must be heat absorbed from the environment. Finally, we reset the cycle, going back along an adiabat from C to A. This requires doing work on the system, while no heat flows. The net result of following the cycle is to absorb heat from the environment and turn it into work. *But that is impossible, according to the second law*: We can't turn heat into work with no side effects!

The problem with our hypothetical three-stroke heat engine is the assumption that both the A–B and A–C paths can be adiabatic. If the (X, Y) plane is horizontal and the E axis is vertical, and we consider all the points that can be reached from A by following adiabats, that set can only intersect each vertical line once. So, we can *foliate* the state space, dividing it up into nonintersecting surfaces. The label of a surface is a number we call the *entropy*.

Since our hypothetical three-stroke cycle is a dud, what about a four-stroke engine? We built just such a cycle in the previous section by combining two isotherms with two adiabats.

A *Carnot engine* is an engine that operates between two and only two temperatures, and that can run just as well in reverse. The efficiency of a Carnot engine is the ratio of the work it puts out to the heat that it takes in:

$$\eta = \frac{W}{Q_H} \,. \tag{5.43}$$

Because the total heat in must equal the sum

$$Q_H = W + Q_C \,, \tag{5.44}$$

we can also write the efficiency as

$$\eta = 1 - \frac{Q_C}{Q_H} \,. \tag{5.45}$$

Suppose another engine is more efficient than a Carnot engine. That is, it takes a smaller amount of heat Q'_H at the same temperature T_H to produce the same amount of work that the Carnot engine does. Then we can tie it to a Carnot engine running

in reverse. Take the work output W from the super-engine and feed it into the Carnot refrigerator. The net effect of the two machines operating in tandem is that $Q'_H - Q_H$ of energy comes out of the hot reservoir and is dumped into the cold reservoir. (The only work flow is internal, so by the First Law, the heat out must equal the heat in.) But $Q'_H < Q_H$, so this quantity is negative: Energy is flowing *out* of the *cold* reservoir and into the hot! Indeed, we have ourselves a machine whose only effect is to move heat from low temperature to high, which violates the Second Law.

It follows from the Second Law, then, that the Carnot engine is the most efficient that can operate between two heat reservoirs. Moreover, all Carnot engines operating between the same temperatures are equally efficient. This means that we can plug in our result for the ideal gas example, because any Carnot engine whose internal workings are made of different stuff must still have the same efficiency:

$$\eta = 1 - \frac{T_C}{T_H}\,. \tag{5.46}$$

Imagine that we run a Carnot cycle in the (P, V) plane, with the points on the upper isotherm called A and B, and the corners on the lower isotherm called C and D. The basic "ratio in = ratio out" rule says that

$$\frac{Q_{AB}}{T_H} + \frac{Q_{CD}}{T_C} = 0\,. \tag{5.47}$$

Now, take a smaller cycle which also has B as a vertex. Pick a point E on the same isotherm as A and B, and a point G in between B and C on their adiabat. The heat flows during this cycle must satisfy

$$\frac{Q_{GF}}{T_{FG}} + \frac{Q_{EB}}{T_H} = 0\,. \tag{5.48}$$

The total heat flow along the original A to B curve is

$$Q_{AB} = Q_{AE} + Q_{EB}\,. \tag{5.49}$$

We run the cycle with the B corner cut out, going from A to E, then F, then G, and then through C and D and back to A. Dividing the previous expression by the high temperature, we get

$$\frac{Q_{AB}}{T_H} = \frac{Q_{AE}}{T_H} + \frac{Q_{EB}}{T_H}\,. \tag{5.50}$$

Consequently, substituting this into Eq. (5.47),

$$\frac{Q_{AE}}{T_H} + \frac{Q_{EB}}{T_H} + \frac{Q_{CD}}{T_C} = 0\,. \tag{5.51}$$

And using Eq. (5.48), we find

$$\frac{Q_{AE}}{T_H} - \frac{Q_{GF}}{T_{FG}} + \frac{Q_{CD}}{T_C} = 0\,. \tag{5.52}$$

But the heat flow in going from G to F is just the opposite of that occurring when we go from F to G, so

$$\frac{Q_{AE}}{T_H} + \frac{Q_{FG}}{T_{FG}} + \frac{Q_{CD}}{T_C} = 0\,. \tag{5.53}$$

The sum of each heat flow, divided by the temperature at which it flows, is zero.

We can start with a big Carnot cycle in the (P, V) plane and carve away more and more of it to make any reversible circuit we want, to as close a precision as we want, by taking tiny isotherm and adiabat segments. Going all the way to calculus-tiny traverses, we find that for any reversible cycle,

$$\oint \frac{đQ}{T} = 0\,. \tag{5.54}$$

What if the cycle is not reversible? The equality will become an inequality, and we can see how by comparing a Carnot engine to a not-Carnot engine. Suppose that both engines operate between the same temperatures T_H and T_C and output the same amount of work, W. In order to achive this work output, the not-Carnot engine will need more heat input, since it is less efficient. Call its heat input Q'_H and its heat output Q'_C. By assumption,

$$W = Q_H - Q_C = Q'_H - Q'_C\,. \tag{5.55}$$

From this, we see that

$$Q'_H - Q_H = Q_C - Q'_C\,. \tag{5.56}$$

We want to find the change in the heat/temperature ratio for the not-Carnot engine:

$$\begin{aligned}
\frac{Q'_H}{T_H} - \frac{Q'_C}{T_C} &= \frac{Q'_H}{T_H} - \frac{Q'_C}{T_C} + \frac{Q_H}{T_H} - \frac{Q_C}{T_C} - \frac{Q_H}{T_H} + \frac{Q_C}{T_C} \\
&= \frac{Q'_H - Q_H}{T_H} + \frac{Q_C - Q'_C}{T_C} + \frac{Q_H}{T_H} - \frac{Q_C}{T_C} \\
&= \frac{Q'_H - Q_H}{T_H} + \frac{Q_C - Q'_C}{T_C} \\
&= (Q'_H - Q_H)\left(\frac{1}{T_H} - \frac{1}{T_C}\right)\,.
\end{aligned} \tag{5.57}$$

The first factor is positive, because the not-Carnot engine takes in more heat to do the same work. The second factor is negative, because $T_H > T_C$. Therefore,

$$\frac{Q'_H}{T_H} - \frac{Q'_C}{T_C} < 0\,. \tag{5.58}$$

Following the same trick as before, starting with a big cycle and chopping off corners, we find that for any irreversible cycle,

$$\oint \frac{đQ}{T} < 0\,. \tag{5.59}$$

Putting this together with our previous result, we have that

$$\oint \frac{đQ}{T} \leq 0\,, \tag{5.60}$$

with equality holding if and only if the cycle is reversible. This is *Clausius' theorem,* or the *Clausius inequality.*

Consider a process that goes from state A to state B via one reversible path, and then returns from B to A via another reversible path. Then we can split the integral into pieces:

$$\int_A^B \frac{đQ^{(1)}}{T} + \int_B^A \frac{đQ^{(2)}}{T} = 0\,, \tag{5.61}$$

where the superscripts refer to the path being traversed. If we run the second path in reverse, all the heat flows $đQ^{(2)}$ would be going in the opposite direction. So,

$$\int_A^B \frac{đQ^{(1)}}{T} = \int_A^B \frac{đQ^{(2)}}{T}\,. \tag{5.62}$$

The integral of $đQ/T$ is *independent* of the path we use to get from A to B, as long as it is reversible. This means that all we need to know to compute the integral of $đQ/T$ along any reversible path is where the path starts and stops. We can define a function on the set of states whose difference between B and A gives the integral of $đQ/T$ along any reversible path:

$$\int_A^B \frac{đQ}{T} = S(B) - S(A)\,. \tag{5.63}$$

Our new function S is the entropy.

For example, take an isothermal, reversible expansion of an ideal gas. We have

$$đQ = dE + PdV = PdV\,, \tag{5.64}$$

because the energy E is constant at fixed temperature. An infinitesimal change in our new function S is $đQ$ divided by T, so

$$dS = \frac{P}{T}dV\,. \tag{5.65}$$

By the ideal gas law,

$$dS = \frac{Nk_B}{V}dV\,. \tag{5.66}$$

We integrate to find the change in S between any initial volume V_i and final volume V_f:

$$S_f - S_i = Nk_B \int_{V_i}^{V_f} \frac{1}{V}dV = Nk_B \log\frac{V_f}{V_i}\,. \tag{5.67}$$

This suggests that we could calculate S from the volume by taking $Nk_B \log V$, which is not quite kosher, because V is a number with units and the logarithm function won't

digest it. To get around this, we could define a unit reference volume V_r and say that S is $Nk_B \log(V/V_r)$. Whatever we pick for V_r, it will cancel when we calculate the change in S. So, let us try

$$S = Nk_B \log \frac{V}{V_r} , \tag{5.68}$$

which we reorganize into

$$S = k_B \log \left(\frac{V}{V_r} \right)^N . \tag{5.69}$$

This is starting to look familiar. In an ideal gas of volume V and uniform density, each of the N corpuscles can be in any of the V/V_r different "cells" into which we have divided the total volume. Accordingly, there are $(V/V_r)^N$ different possible "messages" spelling out which corpuscles are in which cells. If all of these messages are equally probable, then the Shannon entropy of that probability distribution is almost exactly the S we have here. The only modifications are the base of the logarithm and the introduction of the constant k_B, which are essentially conventions. This calls for closer investigation.

5.6 The Relation Between Thermodynamic and Statistical Entropy

What entitles us to use the same word, *entropy,* in thermodynamics and in probability theory? We can start to appreciate the answer using a thought-experiment called the *Szilard engine.* This is an imaginary engine whose internal working "fluid" is a single atom [55]. The atom bounces around inside a cylinder that has a partition down the middle, so the atom is either on the left side or the right side. Each end of the cylinder is a movable piston, and the whole apparatus is immersed in a heat bath to maintain a constant temperature.

The operating cycle of a Szilard engine goes as follows. First, suppose that the operator *knows* which side of the partition the atom is on. If she knows that it is on the left, then she slides in the piston on the right until it is halfway into the cylinder. Then she removes the partition and lets the atom bounce around, pushing the piston back out. When the piston is back to its starting position, she restores the partition. The thermal energy of the heat bath becomes useful mechanical work! If the atom is originally on the right, then she does the same procedure but with the left-hand piston, again extracting work. Since it seems that the engine is back to its starting configuration, we appear to have gotten work for nothing — but there is a catch. When the operator restores the partition, the atom could be on either side of it. The atom has been bouncing all around inside the cylinder, and so its final position is effectively randomized. The engineer can get the engine to function *once,* turning one bit of information into mechanical work, but when that is done, she might as well flip a coin to guess where the atom has ended up, and thus which piston she should try using next. If she happens to guess right, she can get another increment of work out of the engine, but if she guesses wrong, then she will be pushing a piston into a region

that is not empty, which requires work to do. Thus, she will be just as likely to put work into the engine as to get any out, and her *expected* work gain will be *zero*. The Szilard engine makes it literally true that knowledge is power, but the corollary is that once used to obtain power, knowledge becomes obsolete.

How much work can she get out of the engine if she guesses correctly? This is just the work done by an isothermal expansion of an ideal gas, with $N = 1$. The final volume is the whole volume of the cylinder, call it V, and the initial volume is $V/2$. Therefore,

$$W_{\text{out}} = k_B T \log 2 \,. \tag{5.70}$$

This is also the work that she has to put *in* if she guesses wrong.

Now, suppose that Szilard engines come from the manufacturer in correlated pairs. Either both cylinders in a pair have their atoms on the right, or both have their atoms on the left. So, a single bit of information will be good for twice the regular amount of work. But what if the operator does not know which way the engines are oriented? If all she can do is try the engines in succession, what is her expected amount of useful work gained? Let's say she guesses that the atoms are on the left, so she tries the right-hand piston on the first cylinder. If she measures an energy gain, she knows that she was correct, and she operates the right-hand piston on the second cylinder too, and she is correct again, for a total work gain of $2k_B T \log 2$. But if she was *wrong* the first time, she measures a net *loss* of $k_B T \log 2$, and the best she can manage is to break even by using the opposite piston on the second cylinder. All together, her expected energy gain is the average

$$\frac{1}{2}(2k_B T \log 2) + \frac{1}{2}(0) = k_B T \log 2 \,. \tag{5.71}$$

This is the work she would expect to obtain per pair of cylinders if she repeated the process for many pairs. Though she might get lucky and win on every pair, the least surprising net amount of work would be $k_B T \log 2$ times the number of trials.

The same logic holds if the cylinders come from the manufacturer in correlated sets of n. The operator might get lucky on her first guess and extract energy from all n cylinders, or she might lose on the first cylinder and only extract energy from $n-1$ of them. In the latter eventuality, she has to input one cylinder's worth of energy, so her net gain is only $(n-2)k_B T \log 2$. Her expected energy gain is thus $(n-1)k_B T \log 2$, the average of these two possibilities.

We can generalize to the case where the initial setup of n cylinders is a string of L's and R's. The goal of the operator is then to infer the entirety of the string given a portion of it. How many guesses can she expect to waste before she narrows down the possibilities and extracts all the remaining energy? The answer must be at least the Shannon entropy of her probability distribution over the possible strings. That, we recall, is her expected number of moves in a guessing game, if she guesses optimally. And here, her only options are to pick a cylinder and a direction in which to try it, which may or may not be what the optimal guessing strategy calls for. If we suppose that learning the original configuration of each cylinder cuts the number of remaining possibilities in half, and that the operator is never partially knowledgeable about a

cylinder (e.g., 75% confident that a particular atom is on the left), then the expected work is

$$\langle W_{\text{out}} \rangle = (n - S_0) k_B T \log 2 \,, \tag{5.72}$$

where S_0 is the Shannon entropy of the operator's original probability distribution. The number n is also the Shannon entropy of the operator's *final* probability distribution, because running all the engines will randomize all the atom positions. So,

$$\langle W_{\text{out}} \rangle = (S_f - S_0) k_B T \log 2 \,. \tag{5.73}$$

This will match the old-school thermodynamic expression if we identify $(S_f - S_0) k_B \log 2$ with the change in thermodynamic entropy.

What if the operator has imperfect knowledge about a cylinder? One way to think about this is to suppose that she has a second physical system, a *descriptor,* that can be correlated with the cylinder. The descriptor might indicate the cylinder's status accurately, or it might not. Let's say that the probability of error is ϵ. Then the mutual information between the descriptor and the cylinder is

$$I = 1 + \epsilon \log_2 \epsilon + (1 - \epsilon) \log_2 \epsilon \,. \tag{5.74}$$

There exists a strategy for manipulating the cylinder that can extract an expected work amount

$$\langle W_{\text{out}} \rangle = (k_B T \log 2) I \,, \tag{5.75}$$

and this gives the upper limit [56].

We could go off in many other directions. What if the string of atom positions is generated by an algorithm [57]? What if we insist that a single atom must be described quantum-mechanically? Pursuing these lines of inquiry at the present stage would be running before we could walk, so for the time being, we may content ourselves with the intuition that *information can be exploited to gain energy.*

5.7 Thermodynamic Potentials

In mechanics, we can find the stable equilibria of a system by minimizing its potential energy. It's worth recalling how this works. Suppose a body of mass m can move along a line, which we take to be the x axis, under the influence of various forces. Its kinetic energy is $\frac{1}{2}mv^2$, and its potential energy is some function of position, $V(x)$. Its total energy is therefore

$$E = \frac{1}{2} m v^2 + V(x) \,. \tag{5.76}$$

If its total energy is *conserved,* then the time derivative of this sum needs to equal zero:

$$\frac{1}{2} m \frac{d}{dt} v^2 + \frac{d}{dt} V = 0. \tag{5.77}$$

The function V has no intrinsic dependence upon time; the potential energy changes over time if the body moves to a new spot where the potential energy takes a different

value. So,

$$mv\frac{dv}{dt} = -\frac{dV}{dx}\frac{dx}{dt}. \tag{5.78}$$

We can divide through by the velocity and thus obtain

$$ma = -\frac{dV}{dx}. \tag{5.79}$$

A body is in stable equilibrium if the net force upon it vanishes, and if a small displacement will bring about a force that pushes back toward the equilibrium point. This means that we want the derivative of V to be zero, and the second derivative of V to be positive, as that implies a small step in the positive x direction will lead to a force in the negative direction and vice versa.

Thermodynamic potentials are functions that generalize this concept to situations involving thermal effects. We can find thermal equilibria by finding extreme points, usually minima, of these functions.

Imagine that we have a cylinder on its side, filled with an ideal gas and in contact with a heat bath to keep its temperature constant. One end of the cylinder is a moveable piston, which the gas pressure will push outward. In order to keep it from getting pushed out all the way, we put a spring in the cylinder. As the piston moves out, the spring stretches, pulling back harder the more it is stretched. What will be the equilibrium position of the piston?

Let's say that the surface area of the piston is A and its distance from the other end is x. The relaxed length of the spring is x_0, and so the force from the spring on the piston is $-k(x - x_0)$. The outward push from the gas is its pressure P times the surface area A. In equilibrium, we have

$$PA = k(x - x_0)\,. \tag{5.80}$$

We assume that the gas is ideal, so $PV = Nk_BT$. The volume is A times the length that the gas occupies, which is x, and so

$$PAx = Nk_BT\,. \tag{5.81}$$

Combining this with the previous expression, we get

$$\frac{Nk_BT}{x} = k(x - x_0)\,. \tag{5.82}$$

From here, we just have to solve a quadratic equation to find x. But let's try thinking about it in another way. If

$$k(x - x_0) - \frac{Nk_BT}{x} = 0\,, \tag{5.83}$$

then a function of which the left-hand side is the derivative will attain an extremum at the equilibrium value of x. The indefinite integral of that combination is

$$\frac{1}{2}k(x - x_0)^2 - Nk_BT(\log x + C)\,. \tag{5.84}$$

It looks like poor form to take the logarithm of a dimensionful quantity like x, but changing the units of x can be absorbed by changing the constant C, so it's really all right. The constant C drops out as usual when we consider the change in this quantity between two positions. When we do that, the difference of logarithms will become the logarithm of a ratio; but the ratio of two position values is also the ratio of *volumes,* thanks to the area A being fixed. And Nk_B times the logarithm of the volume ratio is the *entropy change* of the ideal gas! Consequently, we introduce the combination

$$F = E - TS\,, \tag{5.85}$$

which we call the *Helmholtz free energy.* The change in F between two piston positions is the change in our previous expression, since we are doing everything isothermally, and the gas energy does not change, only the spring energy. Equilibrium occurs when the derivative of F vanishes.

The total volume of our apparatus here is constant: The piston can move within the cylinder, but the cylinder overall stays the same size. The only mechanical work is that which involves the spring. We do not have to account for the cylinder being squeezed from outside, or anything like that. The Helmholtz free energy F turns out to be the valuable thermodynamic potential in situations where changes happen at constant temperature and internal degrees of freedom can vary, while the overall size of the device stays the same.

We can see how the Helmholtz free energy behaves more generally by posing a more abstract scenario. Let us say that we have a system, any system we like, in contact with a reservoir at temperature T. We follow the system through a transformation from an initial state A to a final state B. Using our earlier discussion of the second law and Clausius' theorem, we can say that

$$\int_A^B \frac{đQ}{T} \leq S_B - S_A\,. \tag{5.86}$$

Because this process is an isothermal one, we can pull the temperature T out of the integral and move it to the other side of the inequality:

$$\int_A^B đQ \leq T(S_B - S_A)\,. \tag{5.87}$$

We have found an upper limit for the amount of heat that the system can obtain from its surroundings. In the special case that the trip is reversible, then the inequality becomes an equality.

The first law tells us that the system's total change in energy, $\Delta E = E_B - E_A$, is the work done upon it plus the heat flowing into it:

$$\Delta E = W_{\text{upon}} + Q = -W_{\text{by}} + Q\,. \tag{5.88}$$

We solve for the work done *by* the system and use our upper bound on the heat Q, finding

$$W_{\text{by}} = -\Delta E + Q \leq -\Delta E + T\Delta S\,. \tag{5.89}$$

We recognize the Helmholtz free energy (T is constant, so F can only change as U and S do).

$$W_{\text{by}} \leq -\Delta(E - TS) = -\Delta F\,. \tag{5.90}$$

The *largest amount of work* that the system can do is given by *the drop* in its Helmholtz free energy. This bound is attained if the process is reversible.

It is helpful to investigate what happens if the pressure is held constant as well as the temperature. This is a circumstance one encounters more in beginning chemistry than in beginning physics. Imagine, for example, a reaction happening in a test tube, open to the air and immersed in a large tank of water. The water keeps the temperature effectively fixed, and the pressure throughout the reaction is basically the pressure of the ambient air. We introduce a new quantity, the *Gibbs free energy*:

$$G = E - TS + PV\,. \tag{5.91}$$

Consider the work done by a system as it changes from state A to state B at constant temperature and pressure. As we deduced earlier, this is bounded by the drop in the Helmholtz free energy:

$$W_{\text{by}} = PV_B - PV_A \leq F_A - F_B. \tag{5.92}$$

We rearrange this inequality to put everything pertaining to B on one side and everything pertaining to A on the other:

$$F_B + PV_B \leq F_A + PV_A\,. \tag{5.93}$$

But this is just

$$G_B \leq G_A\,. \tag{5.94}$$

We can treat the thermodynamic potentials in a unified way by introducing the notion of *availability* [58]. To set the stage, we consider a system in contact with an environment, and we imagine that an amount of heat $đQ$ flows into the system. By Clausius,

$$đQ \leq T_e dS\,, \tag{5.95}$$

where T_e is the temperature of the environment. We suppose that the environment can also do mechanical work on the system:

$$đW = -P_e dV\,. \tag{5.96}$$

By the first law, the sum of these must be the change in energy, so

$$dE \leq T_e dS - P_e dV\,, \tag{5.97}$$

and so, moving everything to the same side,

$$dE + P_e dV - T_e dS \leq 0\,. \tag{5.98}$$

We abbreviate this by writing

$$d\alpha \leq 0\,, \tag{5.99}$$

in which we have defined the availability α as the combination

$$\alpha := E + P_e V - T_e S \,. \tag{5.100}$$

Equilibrium with a given environment occurs when $d\alpha = 0$, i.e., when α is minimized.

Let T and P be the temperature and pressure variables for the system itself. Then, for a small reversible change, the energy changes by the amount

$$dE = TdS - PdV \,, \tag{5.101}$$

and so the availability changes by

$$d\alpha = TdS - PdV + P_e dV - T_e dS \tag{5.102}$$

$$= (T - T_e)dS - (P - P_e)dV \,. \tag{5.103}$$

The first term is the work we could extract by running a Carnot engine between the system and the environment, and the second term is the mechanical work that the system can do upon the environment. So, we can keep extracting work until α can decrease no more. We deduce that *the change in availability is the maximum work that we can extract in the given surroundings.*

The availability α reduces to the Helmholtz free energy F when we keep the system and the environment at the same temperature and also hold the total volume fixed. When $T = T_e$ and $dV = 0$, the change $d\alpha$ reduces to

$$d\alpha = dE - TdS \,, \tag{5.104}$$

which is just exactly dF. We likewise obtain the Gibbs free energy by working isothermally ($T = T_e$) and holding the pressure constant, for then

$$d\alpha = dE - TdS + PdV \,, \tag{5.105}$$

recovering dG.

6 Thermodynamics and Multivariable Calculus

6.1 Introduction

In the previous chapter, we made the occasional statement that we were taking the derivative with respect to one quantity while keeping another constant. We now go deeper into that way of thinking. This will pave the way to finding relationships that are often useful when we want to calculate something that is hard to measure from something that is easier to measure.

The total energy of a system is the sum of its free energy and a part that depends upon the entropy:

$$U = F + TS\,. \tag{6.1}$$

So, a tiny change in F is

$$dF = d(U - TS) = dU - TdS - SdT\,, \tag{6.2}$$

which we see by treating the d just like differentiation in calculus. Meanwhile, for a gas the first law says

$$dU = TdS - PdV\,. \tag{6.3}$$

Combining these, we find

$$dF = -PdV - SdT\,, \tag{6.4}$$

and so at constant temperature,

$$P = -\left.\frac{\partial F}{\partial V}\right|_T\,. \tag{6.5}$$

We can likewise find an equation for the entropy S in terms of a derivative of F with respect to something while holding something else constant.

$$S = -\left.\frac{\partial F}{\partial T}\right|_V\,. \tag{6.6}$$

By substituting this into Eq. (6.1), we find an equation for U in terms of F and T:

$$U = F - T\left.\frac{\partial F}{\partial T}\right|_V\,. \tag{6.7}$$

This looks vaguely like something that could come out of differentiating a product. We have F times the derivative of T with respect to T (i.e., 1), and we have T times the derivative of F. Playing around a bit, we see that

$$\frac{\partial}{\partial T}\left(\frac{F}{T}\right) = \frac{T\frac{\partial F}{\partial T} - F}{T^2}\,. \tag{6.8}$$

Multiplying this by T^2 shows that

$$U = -T^2\,\frac{\partial}{\partial T}\bigg|_V \left(\frac{F}{T}\right)\,. \tag{6.9}$$

One of the most fundamental scientific questions is, "If I do *this,* what will happen?" In the context of thermodynamics, we summarize answers to that by finding quantities which express how a bulk property changes when some bulk influence is applied. For example, the *heat capacity at constant volume* captures how energy content changes if heat is allowed to flow while the volume is held constant:

$$C_V = \frac{đQ}{dT}\bigg|_V = \frac{dU - đW}{dT}\bigg|_V = \frac{dU + PdV}{dT}\bigg|_V\,, \tag{6.10}$$

and since constant V means $dV = 0$,

$$C_V = \frac{\partial U}{\partial T}\bigg|_V\,. \tag{6.11}$$

Likewise, we can define a heat capacity at constant *pressure*:

$$C_P = \frac{đQ}{dT}\bigg|_P = \frac{dU - đW}{dT}\bigg|_P = \frac{dU + PdV}{dT}\bigg|_P\,, \tag{6.12}$$

but now, the second term doesn't vanish:

$$C_V = \frac{\partial U}{\partial T}\bigg|_P + P\,\frac{\partial V}{\partial T}\bigg|_P\,. \tag{6.13}$$

The *isothermal compressibility* captures how applying pressure while holding the temperature constant influences the volume:

$$\kappa_T = -\frac{1}{V}\,\frac{\partial V}{\partial P}\bigg|_T\,. \tag{6.14}$$

For an ideal gas, $V = Nk_BT/P$, and so

$$\kappa_T = \frac{Nk_BT}{VP^2} = \frac{1}{P}\,. \tag{6.15}$$

Likewise, the *expansivity* captures how much a change of temperature at constant pressure affects the volume:

$$\alpha_P = \frac{1}{V}\,\frac{\partial V}{\partial T}\bigg|_P\,. \tag{6.16}$$

For an ideal gas,

$$\alpha_P = \frac{Nk_B}{VP} = \frac{1}{T}\,. \tag{6.17}$$

The energy of an ideal gas is a function of N and T, not P or V. So, holding P constant is equivalent to holding V constant: The partial derivative with respect to T will be the same either way. This implies that

$$C_P - C_V = P\left.\frac{\partial V}{\partial T}\right|_P = PV\alpha_P = \frac{PV}{T} = Nk_B\,. \tag{6.18}$$

6.2 Maxwell Relations

We observed above that we can swap the order of partial derivatives with no ill effect:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}\,. \tag{6.19}$$

This mathematical fact has physical weight to it in thermodynamics, because in thermodynamics the first partial derivatives of interesting functions have names of their own. For example, take a thermal system that can be stretched or squeezed in one dimension and in which the amount of matter can fluctuate. The first law reads

$$dE = TdS + Jdx + \mu dN\,. \tag{6.20}$$

The change in energy due to a change only in the entropy must then be the temperature:

$$\left.\frac{\partial E}{\partial S}\right|_{x,N} = T\,, \tag{6.21}$$

and likewise

$$\left.\frac{\partial E}{\partial x}\right|_{S,N} = J\,. \tag{6.22}$$

So, let us differentiate the energy twice:

$$\frac{\partial^2 E}{\partial S \partial x} = \frac{\partial^2 E}{\partial x \partial S}\,, \tag{6.23}$$

where we are holding N constant throughout. The left-hand side is

$$\frac{\partial}{\partial S}\frac{\partial E}{\partial x} = \frac{\partial J}{\partial S}\,, \tag{6.24}$$

and by the same logic,

$$\frac{\partial}{\partial x}\frac{\partial E}{\partial S} = \frac{\partial T}{\partial x}\,. \tag{6.25}$$

Moreover, partial derivatives can be flipped upside down, and so in a rather rabbit-out-of-a-hat way, we have a relation between the change in *entropy* due to *applied force* and the change in *elongation* due to *temperature*:

$$\left.\frac{\partial S}{\partial J}\right|_x = \left.\frac{\partial x}{\partial T}\right|_S . \tag{6.26}$$

Relations found by exchanging the order of partial derivatives are called *Maxwell relations*. We can interpret them in terms of energy conservation. Imagine a cyclic process. The internal energy must be the same before and after the cycle, so

$$\oint (TdS - PdV) = 0 . \tag{6.27}$$

Consequently, the area enclosed by the cycle in the (P, V) plane must be the same as the area we would get if we drew it in the (T, S) plane instead:

$$\oint PdV = \oint TdS . \tag{6.28}$$

This just expresses the First Law: the net work done is the net heat absorbed. Another way to say the same thing is by integrating the area elements $dPdV$ and $dTdS$ over the enclosed regions:

$$\int dPdV = \int dTdS . \tag{6.29}$$

We have changed coordinates while leaving the area unaffected. This means that the Jacobian of the transformation between coordinate systems is 1. Explicitly, the Jacobian is the determinant of the matrix

$$\begin{pmatrix} \left.\frac{\partial T}{\partial P}\right|_V & \left.\frac{\partial T}{\partial V}\right|_P \\ \left.\frac{\partial S}{\partial P}\right|_V & \left.\frac{\partial S}{\partial V}\right|_P \end{pmatrix} \tag{6.30}$$

The Maxwell relations among S, T, P and V can be read out of the condition that the Jacobian is unity [59, 60, 61, 62].

Let A and B be two functions of the variables x and y. Their Jacobian is

$$\frac{\partial(A, B)}{\partial(x, y)} = \det \begin{pmatrix} \frac{\partial A}{\partial x} & \frac{\partial A}{\partial y} \\ \frac{\partial B}{\partial x} & \frac{\partial B}{\partial y} \end{pmatrix} = \frac{\partial A}{\partial x}\frac{\partial B}{\partial y} - \frac{\partial A}{\partial y}\frac{\partial B}{\partial x} . \tag{6.31}$$

From this definition, it is evident that Jacobians are antisymmetric:

$$\frac{\partial(A, B)}{\partial(x, y)} = -\frac{\partial(B, A)}{\partial(x, y)} = \frac{\partial(B, A)}{\partial(y, x)} . \tag{6.32}$$

Moreover, we can quickly verify that

$$\frac{\partial(x, y)}{\partial(x, y)} = 1 . \tag{6.33}$$

Jacobians obeys a chain rule, just like the familiar derivative:

$$\frac{\partial(A,B)}{\partial(w,z)}\frac{\partial(w,z)}{\partial(x,y)} = \frac{\partial(A,B)}{\partial(x,y)}\,. \tag{6.34}$$

We can think of this as splitting a *scaling of area* into two steps. First, we transform from (x,y) to (w,z), which scales a tiny area element $dx\,dy$ by the determinant of that transformation. Then, transforming from (w,z) to (A,B) multiplies that area by another factor.

All partial derivatives are special cases of Jacobians:

$$\frac{\partial(x,B)}{\partial(x,y)} = \left.\frac{\partial B}{\partial y}\right|_x\,, \tag{6.35}$$

everything else canceling.

We saw above that the conservation of energy implies

$$\frac{\partial(T,S)}{\partial(P,V)} = 1\,. \tag{6.36}$$

Let's plug this into the chain rule. For arbitrary thermodynamic state functions A and B, we must have

$$\frac{\partial(T,S)}{\partial(P,V)}\frac{\partial(P,V)}{\partial(A,B)} = \frac{\partial(T,S)}{\partial(A,B)}\,. \tag{6.37}$$

Therefore,

$$\frac{\partial(P,V)}{\partial(A,B)} = \frac{\partial(T,S)}{\partial(A,B)}\,. \tag{6.38}$$

Now, we just pick possibilities for A and B. For example, take $A = T$ and $B = P$. Then

$$\frac{\partial(P,V)}{\partial(T,P)} = \frac{\partial(T,S)}{\partial(T,P)} = -\frac{\partial(V,P)}{\partial(T,P)}\,. \tag{6.39}$$

We read off an equality between partial derivatives:

$$\left.\frac{\partial S}{\partial P}\right|_T = -\left.\frac{\partial V}{\partial T}\right|_P\,. \tag{6.40}$$

Or, try $A = S$ and $B = V$. Then we get

$$\frac{\partial(T,S)}{\partial(S,V)} = \frac{\partial(P,V)}{\partial(S,V)} = -\frac{\partial(S,T)}{\partial(S,V)}\,. \tag{6.41}$$

And again we can read off an equality between partial derivatives:

$$\left.\frac{\partial P}{\partial S}\right|_V = -\left.\frac{\partial T}{\partial V}\right|_S\,, \tag{6.42}$$

which we can flip upside down and present as

$$\left.\frac{\partial S}{\partial P}\right|_V = -\left.\frac{\partial V}{\partial T}\right|_S . \tag{6.43}$$

This is the same result we deduced in Eq. (6.26).

If we take the first law in the form

$$dU = TdS - PdV , \tag{6.44}$$

which applies for reversible infinitesimal transformations, and we divide through by dV while holding temperature constant, we obtain

$$\left.\frac{\partial U}{\partial V}\right|_T = T \left.\frac{\partial S}{\partial V}\right|_T - P . \tag{6.45}$$

Using the Maxwell relation (6.43), this becomes

$$\left.\frac{\partial U}{\partial V}\right|_T = T \left.\frac{\partial P}{\partial T}\right|_V - P . \tag{6.46}$$

For an ideal gas,

$$P = \frac{Nk_BT}{V} , \tag{6.47}$$

and so

$$T \left.\frac{\partial P}{\partial T}\right|_V - P = \frac{Nk_BT}{V} - P = 0 . \tag{6.48}$$

This makes sense, because the energy U of an ideal gas is a function of T, not V.

In a *hard-sphere* or *hard-core gas,* the atoms have a nonzero size, so they can bump into each other, but they don't interact in any other way, either attracting or repelling. This means the ideal gas law is modified, because each atom has less "room to play": not the whole volume V, but V diminished by an amount that depends on how many other atoms there are. The new gas law is

$$P(V - Nb) = Nk_BT , \tag{6.49}$$

where b is some positive constant. What is $\left.\frac{\partial U}{\partial V}\right|_T$ for this gas? It is again zero, as it is for the ideal gas:

$$P = \frac{Nk_BT}{V - Nb} , \tag{6.50}$$

and so

$$T \left.\frac{\partial P}{\partial T}\right|_V - P = T\frac{Nk_B}{V - Nb} - P = 0 . \tag{6.51}$$

In a *van der Waals gas,* the atoms bump into each other, but they can attract each other too. So, there is the "excluded volume" effect we just saw along with an additional modification:

$$\left(P + a\left(\frac{N}{V}\right)^2\right)(V - Nb) = Nk_BT . \tag{6.52}$$

The numbers a and b are constants that depend upon which specific gas one is studying. What is $\frac{\partial U}{\partial V}\big|_T$ for a van der Waals gas? We compute

$$P = \frac{Nk_BT}{V - Nb} - a\left(\frac{N}{V}\right)^2 , \tag{6.53}$$

and so

$$T\,\frac{\partial P}{\partial T}\bigg|_V - P = a\left(\frac{N}{V}\right)^2 . \tag{6.54}$$

The inter-atomic attractions imply that the energy is no longer a function of T alone. However, because

$$\frac{\partial^2 U}{\partial T \partial V} = \frac{\partial^2 U}{\partial V \partial T} , \tag{6.55}$$

and

$$\frac{\partial}{\partial T}\frac{\partial U}{\partial V} = \frac{\partial}{\partial T} a\left(\frac{N}{V}\right)^2 = 0 , \tag{6.56}$$

we see that

$$\frac{\partial}{\partial V}\left(\frac{\partial U}{\partial T}\bigg|_V\right) = \frac{\partial C_V}{\partial V} = 0 . \tag{6.57}$$

6.3 Stability Criteria

Let's consider the isotherms of a van der Waals gas in the pressure–volume plane, as given by Eq. (6.53). For the ideal gas, we had a simple inverse relationship, and each isotherm was (the positive branch of) a hyperbola. But the van der Waals curves are evidently more complicated. To understand their shapes, we differentiate:

$$\frac{\partial P}{\partial V}\bigg|_T = -\frac{Nk_BT}{(V - Nb)^2} + \frac{2aN^2}{V^3} . \tag{6.58}$$

Unlike the ideal-gas isotherms, which are always monotonic, this has the potential for a local extremum. If T is low enough, then the positive term can overpower the negative term, and the pressure would *increase* with volume. In other words, if the gas were to expand a little, the force required to keep it from expanding further would become *larger.* Conversely, if the gas were squeezed, it would become easier to squeeze. This is at odds with the way we have been thinking about gases as homogeneous substances. Any tiny fluctuation will be amplified, running away to lower or higher density. So, there is something unphysical about a portion of each van der Waals isotherm at low temperature, at least under the presumption that the substance described by the equation is a uniform fluid.

Differentiating again,

$$\frac{\partial^2 P}{\partial V^2}\bigg|_T = \frac{2Nk_BT}{(V - Nb)^3} - \frac{6aN^2}{V^4} . \tag{6.59}$$

The edge of the peculiar region is found by setting the first and second derivatives to zero simultaneously. This yields, with some computation,

$$\begin{aligned} P_c &= \frac{a}{27b^2}\,, \\ V_c &= 3Nb\,, \\ k_B T_c &= \frac{8a}{27b}\,. \end{aligned} \tag{6.60}$$

When $T < T_c$, the isotherm contains an unstable interval, where the derivative $\partial P/\partial V$ is positive. A larger interval around this is regarded as unphysical; the endpoints of this interval for each value of T are somewhat difficult to justify on purely thermodynamic grounds.

6.4 Clausius–Clapeyron

In this section, we will derive an expression for how the boiling point of a liquid varies with pressure. We will go about this in three different ways, the first two inspired by Fermi and the third by Kardar.

Let us imagine that we have a container that is partly filled with a liquid. Due to evaporation, there will be vapor from the liquid floating above it. We'll say that the total mass m of our substance is divided into m_l, the mass of the liquid part, and m_g, the mass of the gaseous part:

$$m = m_l + m_g\,. \tag{6.61}$$

Having made this split, we can express the total volume in terms of the *volume per unit mass* (inverse density), and likewise for the total energy:

$$V = m_l v_l + m_g v_g\,, \tag{6.62}$$

$$U = m_l u_l + m_g u_g\,. \tag{6.63}$$

We now consider the isothermal evaporation of an infinitesimal mass dm. In this process, an amount dm of matter leaves the liquid phase and enters the gaseous, thus:

$$\begin{aligned} V + dV &= (m_l - dm)v_l + (m_g + dm)v_g \\ &= V + (v_g - v_l)dm\,. \end{aligned} \tag{6.64}$$

So, the change in volume is

$$dV = (v_g - v_l)dm\,, \tag{6.65}$$

and likewise,

$$dU = (u_g - u_l)dm\,. \tag{6.66}$$

The first law tells us that we can write the infinitesimal heat flow involved as

$$dQ = dU + PdV = dm(u_g - u_l + P(v_g - v_l))\,. \tag{6.67}$$

So, the heat flow per unit mass is

$$\frac{dQ}{dm} = u_g - u_l + P(v_g - v_l)\,. \tag{6.68}$$

This is the *latent heat of vaporization,* and it is the kind of quantity that chemists love to make tables of. So, we give it a symbol, say λ for "latent", and we take the ratio of our energy and volume changes:

$$\begin{aligned}\frac{dU}{dV} &= \left.\frac{\partial U}{\partial V}\right|_T \\ &= \frac{u_g - u_l}{v_g - v_l} \\ &= \frac{\lambda - P(v_g - v_l)}{v_g - v_l} \\ &= \frac{\lambda}{v_g - v_l} - P\,.\end{aligned} \tag{6.69}$$

Comparing this with Eq. (6.46), we see that

$$\left.\frac{\partial P}{\partial T}\right|_V = \frac{\lambda}{T(v_g - v_l)}\,. \tag{6.70}$$

Another way to get to this point is by using the Gibbs free energy, $G = U - TS + PV$. Consider again a system divided into liquid and gas phases, so that the volumes, energies and entropies add:

$$\begin{aligned}V &= V_l + V_g\,, \\ U &= U_l + U_g\,, \\ S &= S_l + S_g\,.\end{aligned} \tag{6.71}$$

Combining these yields

$$G = G_l + G_g\,. \tag{6.72}$$

As before, we work in terms of quantities per unit mass. Along with u_l, u_g and so forth, we have

$$\begin{aligned}G_l &= m_l\gamma_l\,, \\ G_g &= m_g\gamma_g\,.\end{aligned} \tag{6.73}$$

All of these will be functions of T only. If an isothermal process adds a mass dm to the liquid phase, m_g must go down by the same amount, and the Gibbs free energy will be

$$(m_l + dm)\gamma_l + (m_g - dm)\gamma_g = G + dm(\gamma_l - \gamma_g)\,. \tag{6.74}$$

But because the system was originally in equilibrium, G must have originally been at a minimum. So, the variation dG must vanish, and

$$\gamma_l = \gamma_g\,. \tag{6.75}$$

Consequently,

$$(u_g - u_l) - T(s_g - s_l) + P(v_g - v_l) = 0\,. \tag{6.76}$$

Next we differentiate with respect to T:

$$\frac{d}{dT}(u_g - u_l) - T\frac{d}{dT}(s_g - s_l) - (s_g - s_l) + \frac{dP}{dT}(v_g - v_l) + P\frac{d}{dT}(v_g - v_l) = 0\,. \tag{6.77}$$

The first law tells us that

$$du = Tds - Pdv\,, \tag{6.78}$$

from which we see that

$$T\frac{ds}{dT} = \frac{du}{dT} + P\frac{dv}{dT}\,. \tag{6.79}$$

This allows us to cancel some terms in Eq. (6.77), leaving us with

$$-(s_g - s_l) + \frac{dP}{dT}(v_g - v_l) = 0\,. \tag{6.80}$$

The difference $s_g - s_l$ is the change in entropy per unit mass when material is vaporized. In other words, it is the latent heat of vaporization divided by the temperature. So,

$$\frac{dP}{dT} = \frac{\lambda}{T(v_g - v_l)}\,. \tag{6.81}$$

Yet a third way of attaining this equation is to imagine a heat engine running in a cycle where the input heat vaporizes a little material, which condenses back during the output stage. The Carnot efficiency for a high temperature T and an infinitesimally lower temperature $T - dT$ is

$$\eta = 1 - \frac{T - dT}{T} = \frac{dT}{T}\,. \tag{6.82}$$

The efficiency is the ratio of the work done to the heat supplied. And in this case, the heat we supply is a latent heat of vaporization L. The work is the area that the cycle encloses on the P–V diagram, which will approximately be the area of a parallelogram. The height of this parallelogram is the pressure change dP, and the width is the difference in volume between the vapor and liquid phases, $V_g - V_l$. So,

$$\frac{(V_g - V_l)dP}{L} = \frac{dT}{T}\,, \tag{6.83}$$

which we rearrange into

$$\frac{dP}{dT} = \frac{L}{T(V_g - V_l)}\,. \tag{6.84}$$

This is just like what we had before, except phrased in terms of the total latent heat and total volumes instead of the corresponding quantities per unit mass.

Let us suppose that for our substance of interest L is constant with respect to temperature, and that we are working in a situation where the volume of liquid can

be neglected in comparison to the volume of the vapor: $V_l \ll V_g$. Then the Clausius–Clapeyron equation becomes

$$\frac{dP}{dT} = \frac{L}{TV_g} \,. \tag{6.85}$$

Assuming that our vapor is effectively an ideal gas,

$$V_g = \frac{Nk_BT}{P} \,, \tag{6.86}$$

and so

$$\frac{dP}{dT} = \frac{LP}{Nk_BT^2} \,. \tag{6.87}$$

Dividing through by the pressure P, the left-hand side becomes the derivative of the logarithm:

$$\frac{d\log P}{dT} = \frac{L}{Nk_BT^2} \,. \tag{6.88}$$

This integrates easily to

$$\log P = -\frac{L}{Nk_BT} + C \,, \tag{6.89}$$

which upon exponentiation becomes

$$P \propto \exp\left(-\frac{L}{Nk_BT}\right) \,. \tag{6.90}$$

There is yet a fourth way to arrive at the Clausius–Clapeyron relation, at least in this limit where we can neglect the liquid volume ($V_l \ll V_g$). That is to make a *kinetic theory* argument, where we assume that the vapor pressure is proportional to the *number of molecules* in the gaseous phase, which we estimate by calculating the *probability* that a molecule is in the gaseous phase. This gives us a proportionality as in Eq. (6.90), from which we can work backward.

6.5 Chemical Potential

When we derived the Clausius–Clapeyron equation using the Gibbs free energy, we saw that in equilibrium, the free energy per unit mass in the liquid phase must equal the free energy per unit mass in the vapor phase. We can also express this in terms of an energy per particle. The matter in each phase can be understood as a system whose population size can change, and changing the population size is a kind of work:

$$dU = TdS - PdV + \mu dN \,. \tag{6.91}$$

The *chemical potential* is the rate of change of the total energy with respect to the population size N, holding the other variables constant:

$$\mu = \left.\frac{\partial U}{\partial N}\right|_{S,V} \,. \tag{6.92}$$

Goodstein [63] has some clarifying remarks on the chemical potential:

> μ is the change in each of the energy functions when one particle is added in equilibrium. That observation might make it seem plausible that μ would be a positive quantity, but it turns out instead that μ is often negative. Why this is so can be seen [as follows]. To find μ, we ask: How much energy must we add to a system if we are to add one particle while keeping the entropy and volume constant? Suppose that we add a particle with no energy to the system, holding the volume fixed, and wait for it to come to equilibrium. The system now has the same energy it had before, but one extra particle, giving it more ways in which to divide that energy. The entropy has thus increased. In order to bring the entropy back to its previous value, we must extract energy. The chemical potential—that is, the change in energy at constant S and V— is therefore negative. This argument breaks down if it is impossible to add a particle at zero energy, owing to interactions between the particles; the chemical potential will be positive when the average interaction is sufficiently repulsive, and energy is required to add a particle. It will be negative, for example, for low-density gases and for any solid or liquid in equilibrium with (at the same chemical potential as) its own low density vapor.

6.6 Wave Equation for Sound

In this section, we will study the propagation of sound waves through air, write an equation for them and deduce their speed. This is an example of *local equilibrium,* a kind of situation where it is reasonable to think of equilibrium quantities (temperature, pressure, et cetera) being applicable at each point, while their values may vary from place to place. This problem will work in that regime. For simplicity, we will consider plane waves propagating in the x direction: Air moves a little along the x axis, then back again.

Let $\eta(x,t)$ be the displacement of a mass element originally at x. The pressure as a function of position and time is its true equilibrium value plus a perturbation:

$$P(x,t) = P_0 + P_1(x,t)\,, \tag{6.93}$$

and likewise for the density (mass per volume),

$$\mu(x,t) = \mu_0 + \mu_1(x,t)\,. \tag{6.94}$$

Pressure and density are related by some function, the details of which we don't know yet:

$$P = f(\mu)\,. \tag{6.95}$$

For small disturbances, we can apply the Taylor expansion:

$$P(x,t) = f(\mu_0) + \mu_1 f'(\mu_0)\,. \tag{6.96}$$

Therefore, we can write that $P_1 = \kappa\mu_1$, where $\kappa = f'(\mu_0)$.

To find the units of κ, let's go back to $F = ma$. The units of force have to be mass times length per time squared (ML/T^2). Dividing this by area to get pressure, we find that pressure has units of $ML^{-1}T^{-2}$. Density is mass per volume (ML^{-3}), and so κ has units of pressure over density, or

$$\frac{\frac{M}{LT^2}}{\frac{M}{L^3}} = \frac{L^2}{T^2}\,. \tag{6.97}$$

If we are trying to find a speed that is significant for this problem, we should take the square root of κ.

To understand how a wave moves through this medium, we will study the motion of a mass of air with cross-sectional area A and equilibrium density μ_0, between x and $x + \Delta x$. If this air mass is displaced sideways, the surface at x moves to $x + \eta(x,t)$, and the surface at $x + \Delta x$ moves to $x + \Delta x + \eta(x + \Delta x, t)$. Because mass is conserved, we must have

$$\mu_0 A \Delta x = \mu A \left[(x + \Delta x + \eta(x + \Delta x, t)) - (x + \eta(x,t))\right]\,. \tag{6.98}$$

We can simplify this expression using a Taylor series for the spatial dependence of η. We approximate that

$$\eta(x + \Delta x, t) - \eta(x,t) = \frac{\partial \eta}{\partial x}\Delta x\,. \tag{6.99}$$

Substituting this in and cancelling the area factor from both sides, we get

$$\mu_0 \Delta x = \mu\left[\frac{\partial \eta}{\partial x}\Delta x + \Delta x\right]\,, \tag{6.100}$$

and so

$$\mu_0 = \mu\left(\frac{\partial \eta}{\partial x} + 1\right)\,. \tag{6.101}$$

Writing the density as the sum of the equilibrium and the perturbation, this becomes

$$\mu_1 = -(\mu_0 + \mu_1)\frac{\partial \eta}{\partial x}\,. \tag{6.102}$$

Both the perturbation μ_1 and the derivative $\partial\eta/\partial x$ are small, so we make the approximation that their product is zero. Therefore,

$$\mu_1 = -\mu_0 \frac{\partial \eta}{\partial x}\,. \tag{6.103}$$

We now compute the force on our air mass in two different ways. First, we relate it to the acceleration:

$$F = \mu_0 A \Delta x \frac{\partial^2 \eta}{\partial t^2}\,. \tag{6.104}$$

Second, we write the net force in terms of the difference in pressure between x and $x + \Delta x$:

$$\frac{F}{A} = P(x,t) - P(x + \Delta x, t) = -\frac{\partial P}{\partial x}\Delta x\,. \tag{6.105}$$

By setting these two expressions for F/A equal to each other, we deduce that

$$\mu_0 \frac{\partial^2 \eta}{\partial t^2} = -\frac{\partial P}{\partial x} = -\frac{\partial P_1}{\partial x}\,, \tag{6.106}$$

where the second equality follows because P_1 is the only part of the pressure that depends upon position. In turn,

$$-\frac{\partial P_1}{\partial x} = -\frac{\partial}{\partial x}(\kappa \mu_1) = -\kappa \frac{\partial \mu_1}{\partial x} = \mu_0 \kappa \frac{\partial^2 \eta}{\partial x^2}\,. \tag{6.107}$$

So, we end up with the following relation between partial derivatives with respect to time and to space:

$$\frac{\partial^2 \eta}{\partial t^2} = \kappa \frac{\partial^2 \eta}{\partial x^2}\,. \tag{6.108}$$

This we recognize as a *wave equation* for the displacement η. It describes waves that propagate at a speed $\sqrt{\kappa}$.

Finally, we evaluate κ. The important fact is that pressure and temperature changes in this situation happen *adiabatically*, so we can use the relation

$$PV^\gamma = \text{constant}\,. \tag{6.109}$$

The density μ_0 is inversely proportional to the volume:

$$\mu_0 = \frac{Nm}{V}\,, \tag{6.110}$$

so it must be that

$$P_0 = (\text{constant})\mu_0^\gamma\,. \tag{6.111}$$

Applying the idea of local equilibrium, we say that $P \propto \mu^\gamma$ in general. First, we find the derivative $dP/d\mu$ in terms of γ, P and μ. We apply the good old rule for differentating powers:

$$\frac{dP}{d\mu} = (\text{constant})\gamma \mu^{\gamma-1}\,. \tag{6.112}$$

But we can make that unspecified constant drop out with the following clever trick: The multiplicative factor in front of $dP/d\mu$ must be the same as that in front of P itself! So,

$$\frac{dP}{d\mu} = \frac{\gamma P}{\mu}\,. \tag{6.113}$$

Now, we use the ideal gas law to write $dP/d\mu$ in terms of γ, $k_B T$ and m.

$$\frac{dP}{d\mu} = \frac{\gamma PV}{\mu V} = \frac{\gamma k_B T}{m}\,. \tag{6.114}$$

We recall that for each axis of motion, we expect $\frac{1}{2}k_B T$ of kinetic energy:

$$\frac{3}{2}k_B T = \frac{1}{2}m\langle v^2 \rangle\,. \tag{6.115}$$

We can use this to relate the speed of sound to the "root-mean-square" speed of the gas atoms, $\sqrt{\langle v^2 \rangle}$. From

$$k_B T = \frac{1}{3} m \langle v^2 \rangle \,, \tag{6.116}$$

we obtain

$$\frac{\gamma k_B T}{m} = \left(\frac{\gamma}{m}\right) \frac{1}{3} m \langle v^2 \rangle = \frac{\gamma}{3} \langle v^2 \rangle \,. \tag{6.117}$$

So, finally,

$$\sqrt{\kappa} = \left(\frac{\gamma}{3}\right)^{1/2} \sqrt{\langle v^2 \rangle} \,. \tag{6.118}$$

The speed of sound $\sqrt{\kappa}$ is proportional to the characteristic speed of the gas atoms, and the proportionality constant is made from the only unitless numbers we have at hand: γ and the number of dimensions of space.

6.7 Chemical Reactions

Using the tool of free energy, we can also study chemical reactions. We generally express these using an equation of the form

$$n_1 A_1 + n_2 A_2 + \cdots + n_r A_r \leftrightharpoons m_1 B_1 + m_2 B_2 + \cdots + m_s B_s \,. \tag{6.119}$$

Here, the $\{A_i\}$ are the *reactants,* and the $\{B_i\}$ are the *products* — but this designation is somewhat conventional, because the reaction can happen in both directions. What matters is the balance: In the dynamic equilibrium where all sorts of changes are happening, how do they average out?

Following Fermi, we will explore this using the thought-experiment known as a *Van't Hoff reaction box.* This is a big chamber held at a constant temperature T. Inside it is a mixture of gases A_1, A_2, ..., A_r and B_1, B_2, ..., B_s. On the left of the box are attached a set of r cylinders, each of which connects to the box using a semipermiable membrane, so that cylinder number k can exchange gas A_k only. On the right, we do likewise for the B gases. The other end of each cylinder is a moveable piston. We initialize the Van't Hoff box with the B cylinder pistons all the way in, so that those cylinders have zero volume. Meanwhile, the A cylinder pistons are positioned so that cylinder k contains n_k moles of gas A_k. The partial pressures of the gases are equal on both sides of the respective membranes; equilibrium obtains.

Now for the experimental protocol. First, we push the A cylinder pistons in slowly. This means doing work, isothermally. The box is so big by comparison with the cylinders that we can presume the gas concentrations do not change. Next, we slowly slide out the pistons of the B cylinders, letting the gases do isothermal work, until those cylinders contain $m_1, \ldots, m_s$ moles of gases $B_1, \ldots, B_s$ respectively. The total work done by the gas mixture during these two stages is

$$W = N_A k_B T \left(\sum_{j=1}^{s} m_j - \sum_{i=1}^{r} n_i \right) \,, \tag{6.120}$$

where we use the Avogadro constant N_A to turn mole numbers into molecule numbers. This work must equal the change in the free energy. For one mole of gas, the free energy is

$$F = C_V T + U_0 - T(C_V \log T + N_A k_B \log V + a)\,, \tag{6.121}$$

where U_0 is the energy at absolute zero and a is an additive constant that arises in the entropy. These parameterize how the gas deviates from our ideal model at low temperatures. The *concentration* $[A_i]$ of gas A_i is measured in moles per volume, so its reciprocal is a volume per mole. We can thus write the volume term in terms of the concentration:

$$N_A k_B \log V = -N_A k_B \log[A_i]\,. \tag{6.122}$$

For each of the A gases, we compute the molar free energy F_i, and likewise, we find the molar free energy F'_j for each of the B gases. During the experimental protocol, the free energy changes by an amount that must correspond to the work done:

$$N_A k_B T \left(\sum_{j=1}^{s} m_j - \sum_{i=1}^{r} n_i \right) = \sum_{i=1}^{r} n_i F_i - \sum_{j=1}^{s} m_j F'_j\,. \tag{6.123}$$

Expressing the right-hand side in terms of the concentrations and doing some algebra, we find that

$$\frac{[A_1]^{n_1}[A_2]^{n_2}\cdots[A_r]^{n_r}}{[B_1]^{m_1}[B_2]^{m_2}\cdots[B_r]^{m_s}} = K(T)\,, \tag{6.124}$$

where $K(T)$ is a function whose exact form is a little messy. The important thing is that for a given reaction, it depends only upon the temperature, not the volume or the pressure of the reaction box. If ΔU is the change in energy observed when running the reaction in the $A \to B$ direction, then

$$\frac{d\log K(T)}{dT} = -\frac{\Delta U}{N_A k_B T^2}\,. \tag{6.125}$$

If $\Delta U < 0$, the reaction is *exothermal,* and $K(T)$ increases with temperature, whereas an *endothermal* reaction has $\Delta U > 0$, and $K(T)$ decreases with temperature. Suppose a reaction is exothermal. Then increasing the temperature will elevate $K(T)$, which means the chemical equilibrium shifts towards higher concentrations of the A gases. But this means that the system must *absorb* heat, thereby opposing the rise in temperature. This is an example of *Le Chatelier's principle*: If we change the external conditions of an equilibrated system, it will tend to respond by re-equilibrating in such a direction as to oppose that change.

6.8 Dilute Solutions

Dilute solutions behave in many ways like ideal gases. We can appreciate why using a kinetic argument. When we deduced the relation between temperature and average kinetic energy (§5.3), we did not have to care about the shape of the container that

holds our ideal gas. We could give it irregular walls without changing the essence of the problem. Indeed, we could put obstacles inside the container. A dilute solution in thermal equilibrium, like a tiny bit of sugar dissolved in water, is like an "ideal gas" of sugar molecules moving through an obstacle field of water molecules. The obstacles are moving about, but because the whole system is equilibrated, that ultimately won't matter.

The analogue in a solution of the pressure exerted by an ideal gas is the *osmotic pressure* due to collisions of solute molecules. As Fermi says,

> The osmotic pressure of a dilute solution is equal to the pressure exerted by an ideal gas at the same temperature and occupying the same volume as the solution and containing a number of moles equal to the number of moles of the solutes dissolved in the solution.

Suppose we have a vessel filled with a liquid solvent. We separate the vessel into two compartments using a membrane that is permeable to the solvent but not to the solute that we mix into one compartment. The solvent flows through the membrane in both directions, reaching a dynamic equilibrium when the flow rates are equal. The pressures will be different in the two compartments, by an amount equal to the osmotic pressure.

Our next thought-experiment is more elaborate. We take a hollow tube and bend it into a vertical rectangle, so that it forms a closed loop. Into this loop we inject an amount of solvent (e.g., water, alcohol, ...). The solvent collects in the bottom of the loop, and some concentration of solvent vapor will naturally float above it. We separate the bottom segment of the rectangle down the middle with a semipermiable membrane and dissolve a little solute on one side, say the right. Consequently, there will be an osmotic contribution to the pressure.

The liquid level on the side with the solute will be higher, because the osmotic pressure P can support an extra weight of liquid. Call the solvent density ρ_l; because the solution is dilute, this will also be the solution density, practically speaking. The osmotic pressure will hold up an extra height h of liquid, given by

$$P = \rho_l g h \,. \tag{6.126}$$

Meanwhile, the solvent vapor will be pushing down harder upon the left-hand side, the side with the pure solvent, because there is a taller column of it. Calling ρ_g the density of the solvent vapor, this extra pressure must be

$$\Delta P_g = \rho_g g h \,. \tag{6.127}$$

Dividing these two equations, we find that

$$\frac{\Delta P_g}{P} = \frac{\rho_g}{\rho_l} \,. \tag{6.128}$$

Let v_g and v_l be the molar volumes of the solvent in the vapor and liquid phases respectively. Then

$$\Delta P_g = P \frac{v_l}{v_g} \,. \tag{6.129}$$

The osmotic pressure P is given by the counterpart of the ideal gas law:

$$P = \frac{N_s k_B T}{V_l}\,, \tag{6.130}$$

with N_s being the number of molecules of dissolved solute. Inserting this into the previous equation,

$$\Delta P_g = \frac{k_B T}{v_g}\frac{N_s v_l}{V_l}\,. \tag{6.131}$$

The ratio v_l/V_l is the inverse of the number of solvent moles, call it N_l/N_A. Therefore,

$$\Delta P_g = \frac{N_A k_B T}{v_g}\frac{N_s}{N_l}\,. \tag{6.132}$$

In summary, dumping solute into the liquid changes the vapor pressure by a quantifiable amount. This also changes the *boiling point* of the liquid, because that is the temperature at which the vapor pressure equals one atmosphere. This ties back to our discussion of the Clausius–Clapeyron equation (§6.4).

7 Kinetic Theory I: Fundamentals

7.1 The Canonical Probability Distribution

In section 5.3, we got a sense of what the *typical* energy of a particle will be. We'd like to carry this further and be able to say something about the possibility that a particle will have greater or lesser energy than that typical value. In this section, we will derive a general functional form that is quite robustly usable for answering such questions.

We start by considering systems with *discrete* sets of possible energy levels. Our derivation will be based on the idea that if A and B are at the same temperature, a noninteracting composite system AB is also at that temperature. Suppose that E_j is an energy level of system A and E_k is an energy level of B. Then, if there is no interaction between the two systems, AB will have an energy level $E_j + E_k$. If we assume that for all systems prepared at temperature T,

$$P(E_n) = \frac{1}{Z} f(E_n), \tag{7.1}$$

then we have

$$\frac{f(E_j)f(E_k)}{Z_A Z_B} = \frac{f(E_j + E_k)}{Z_{AB}}. \tag{7.2}$$

But we have the freedom to adjust f by an overall multiplicative constant, since the meaningful quantities are the probabilities and any prefactor will cancel when we divide by Z. So, we can declare

$$f(0) = 1, \tag{7.3}$$

which yields

$$Z_A Z_B = Z_{AB}, \tag{7.4}$$

and thus

$$f(E_j + E_k) = f(E_j)f(E_k). \tag{7.5}$$

This is a *functional equation,* one that is solved by a function rather than a value. We can tell that exponential equations are among its solutions, since they are famous for turning sums into products. Looking the topic up in math books, we learn that Eq. (7.5) is *Cauchy's functional equation for the exponential.* Provided that f is continuous at even a single point, then the solution is guaranteed to be of exponential form:

$$f(E) = e^{-\beta E}, \tag{7.6}$$

where the "coolness" β labels the equivalence classes of thermal equilibrium. Showing that β ought to be inversely proportional to the temperature T that we define in

phenomenological thermodynamics takes a little more work. For example, you can do a bit of elementary kinetic theory on an ideal gas and convince yourself that the expectation value of a particle's kinetic energy is proportional to $k_B T$. We did this in deriving Eq. (5.21) above.

This argument is conceptually similar to one given by James Clerk Maxwell. To paraphrase his thoughts into our language, Maxwell argued that the components of a gas atom's velocity — call them v_x, v_y and v_z — are independent random variables. Therefore, the probability distribution must factor cleanly into functions of v_x, v_y and v_z. But if no direction is preferred, then the probability of finding a velocity v can only be a function of the *magnitude* of v, or in other words, a function of v^2. Likewise, we have no reason to prefer "up" over "down" motion along any axis, so the probability densities for each coordinate must be symmetrical about 0. We again obtain a functional equation where a sum becomes a product:

$$p(v^2) = p(v_x^2 + v_y^2 + v_z^2) = p(v_x^2)p(v_y^2)p(v_z^2). \tag{7.7}$$

The probability density is then an exponential function of v^2, and to be well-behaved, it must be a decaying exponential. The rate at which the exponential decays is set by the temperature, since the temperature determines $\left\langle mv^2/2 \right\rangle$.

Calculating probabilities by taking e to the power of minus some energy divided by $k_B T$ is almost an instinctive move in statistical physics. It is the kind of result that is stronger than the argument used to derive it, and that is remembered when the rest of the course is forgotten.

The *canonical* distribution for a system at a temperature T is constructed as follows. Let $E(\mu)$ be the energy of a microstate μ. (The symbol μ is an abbreviation for potentially a lot of information, since if the system is big, we need to describe what each piece of it is doing.) Then the probability of that microstate is

$$p(\mu) = \frac{1}{Z} e^{-\frac{E(\mu)}{k_B T}} . \tag{7.8}$$

The quantity Z is the *partition function,* which we need to include in order to make the probabilities properly normalized:

$$\sum_{\mu} p(\mu) = 1 . \tag{7.9}$$

This tells us what form Z takes:

$$Z = \sum_{\mu} e^{-\frac{E(\mu)}{k_B T}} . \tag{7.10}$$

We often abbreviate $1/(k_B T)$ with the Greek letter beta:

$$\beta = \frac{1}{k_B T} . \tag{7.11}$$

In terms of this variable,

$$p(\mu) = \frac{1}{Z} e^{-\beta E(\mu)} , \tag{7.12}$$

and

$$Z = \sum_{\mu} e^{-\beta E(\mu)} \,. \tag{7.13}$$

Let's solidify our understanding by working with an example. In order to be interesting at all, our system must have at least two different possible microstates. So, let's consider a *two-level system,* where the two microstates are μ_0 and μ_1. Let's say that

$$E(\mu_0) = 0, \;\; E(\mu_1) = \epsilon \,, \tag{7.14}$$

where $\epsilon > 0$. Then the partition function is

$$Z = e^{-\beta E(\mu_0)} + e^{-\beta E(\mu_1)} = e^0 + e^{-\beta\epsilon} = 1 + e^{-\beta\epsilon} \,. \tag{7.15}$$

So, the probabilities for the two microstates are

$$p(\mu_0) = \frac{e^{-\beta E(\mu_0)}}{Z} = \frac{1}{1 + e^{-\beta\epsilon}} \tag{7.16}$$

and

$$p(\mu_1) = \frac{e^{-\beta E(\mu_1)}}{Z} = \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}} \,. \tag{7.17}$$

If we take the ratio of these, the denominators cancel:

$$\frac{p(\mu_1)}{p(\mu_0)} = e^{-\beta\epsilon} \,. \tag{7.18}$$

When the temperature T is low, then the coolness β is very big, and so this ratio is e to a large negative power. This means that the probability of microstate μ_1, the microstate of higher energy, is much smaller than the probability of μ_0. If the temperature is low, it's hard to catch the system in the "excited" state! On the other hand, if the temperature is high, then the coolness β is small, and we can use a Taylor approximation:

$$\frac{p(\mu_1)}{p(\mu_0)} \approx 1 - \beta\epsilon \,. \tag{7.19}$$

The ratio of the probabilities is almost 1, which tells us that the probabilities are almost equal. At high temperature, the difference between the energy levels doesn't matter!

In order to connect the macroscopic picture of thermodynamics with the microstate perspective of the canonical distribution, we equate the energy U from thermodynamics with the expected value of E:

$$U = \langle E \rangle \,. \tag{7.20}$$

By the definition of expected value,

$$U = \sum_{\mu} p(\mu) E(\mu) \,. \tag{7.21}$$

Now, we substitute in the canonical distribution:

$$U = \sum_{\mu} \frac{e^{-\beta E(\mu)}}{Z} E(\mu) \,. \tag{7.22}$$

We can pull the partition function out of the sum, because it is the same for all the terms:

$$U = \frac{1}{Z} \sum_{\mu} e^{-\beta E(\mu)} E(\mu) \,. \tag{7.23}$$

Now comes the sneaky part. If we take the derivative of $e^{-\beta E(\mu)}$ with respect to β, we pull down a factor of $-E(\mu)$:

$$\frac{\partial}{\partial \beta} e^{-\beta E(\mu)} = -E(\mu) e^{-\beta E(\mu)} \,. \tag{7.24}$$

So, we can refashion our sum into

$$U = \frac{1}{Z} \sum_{\mu} \left(-\frac{\partial}{\partial \beta} e^{-\beta E(\mu)} \right) \,. \tag{7.25}$$

Or, pulling the derivative out in front,

$$U = -\frac{1}{Z} \frac{\partial}{\partial \beta} \sum_{\mu} e^{-\beta E(\mu)} \,. \tag{7.26}$$

The sum over μ here just gives us the partition function again!

$$U = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \,. \tag{7.27}$$

We recall from calculus that

$$\frac{d}{dx} \log x = \frac{1}{x} \,, \tag{7.28}$$

and so we recognize that

$$U = -\frac{\partial \log Z}{\partial \beta} \,. \tag{7.29}$$

Derivatives with respect to the coolness β may still be a little unfamiliar, since up until we introduced the canonical distribution, we did everything in terms of the temperature T. How does a derivative with respect to β relate to a derivative with respect to T? We can invoke the trusty chain rule: for an arbitrary function f,

$$\frac{\partial f}{\partial \beta} = \frac{\partial f}{\partial T} \frac{dT}{d\beta} \,. \tag{7.30}$$

We have written the second factor without the partial symbols because T is a function of β alone:

$$T = \frac{1}{k_B \beta} \,. \tag{7.31}$$

Differentiating this with respect to β, we get

$$\frac{dT}{d\beta} = -\frac{1}{k_B\beta^2}\,. \tag{7.32}$$

Let's rewrite the right-hand side in terms of T. One over β^2 is equal to $(k_BT)^2$, and so

$$\frac{dT}{d\beta} = -\frac{(k_BT)^2}{k_B} = -k_BT^2\,. \tag{7.33}$$

Consequently, going back to the chain rule,

$$\frac{\partial f}{\partial \beta} = -k_BT^2\frac{\partial f}{\partial T}\,. \tag{7.34}$$

Now we can use this in our expression for the energy U:

$$U = \frac{\partial(-\log Z)}{\partial\beta} = -k_BT^2\frac{\partial(-\log Z)}{\partial T} = -T^2\frac{\partial(-k_B\log Z)}{\partial T}\,. \tag{7.35}$$

We compare this with Eq. (6.9), and we see that

$$-k_B\log Z = \frac{F}{T}\,. \tag{7.36}$$

Solving for the free energy, we get

$$F = -k_BT\log Z\,. \tag{7.37}$$

This is sometimes called the *bridging equation,* since it connects the microscopic perspective of the canonical distribution with the macroscopic perspective of thermodynamics.

We can get an interesting result if we write $p(\mu)$ in terms of F instead of Z. We start with $\log Z = -\beta F$, from which we get

$$Z = e^{-\beta F}\,. \tag{7.38}$$

Inverting both sides,

$$\frac{1}{Z} = e^{\beta F}\,, \tag{7.39}$$

so we find that

$$p(\mu) = e^{\beta F - \beta E(\mu)}\,. \tag{7.40}$$

Taking the logarithm of both sides gives

$$\log p(\mu) = \beta F - \beta E(\mu)\,. \tag{7.41}$$

Now, we can evaluate the expectation value

$$\langle -\log p(\mu)\rangle = -\sum_\mu p(\mu)\log p(\mu)\,. \tag{7.42}$$

We multiply through by $p(\mu)$ and sum over μ:

$$-\beta F \sum_\mu p(\mu) + \beta \sum_\mu p(\mu) E(\mu) = -\beta F + \beta \langle E \rangle \,. \tag{7.43}$$

Translating back from β to T, we get

$$\langle -\log p(\mu) \rangle k_B T = -F + \langle E \rangle \,. \tag{7.44}$$

So,

$$\langle E \rangle = F + k_B T \langle -\log p(\mu) \rangle \,. \tag{7.45}$$

We see the quantity $k_B \langle -\log p(\mu) \rangle$ plays the role of the thermodynamic entropy S. The thermodynamic and Shannon entropies are simply proportional to each other! We used several assumptions to get here, the most important of which was the basic idea behind our Cauchy–Maxwell derivation of the canonical distribution: The probability $p(\mu)$ depends upon μ only through the energy $E(\mu)$.

7.2 Brownian Motion

We've seen how microscopic collisions can add up to create pressure upon a macroscopic wall. One way to carry this image further is to introduce a *mesoscopic* component, something intermediate between atoms and pistons.

A spherical particle of mass m moves in one dimension through a viscous medium subject to a drag force due to the viscosity and also a randomly fluctuating force caused by random molecular impacts. Newton's second law takes the form

$$m \frac{d^2x}{dt^2} = -(6\pi\eta R) \frac{dx}{dt} + f(t), \tag{7.46}$$

where η is the viscosity and R is the radius of the particle. Because $f(t)$ changes randomly from one instant to the next, we have to work with expectation values. Our goal is to find an equation for the time derivative of $\langle x^2 \rangle$, which we can also write as an expectation value of a derivative:

$$\frac{d}{dt} \langle x^2 \rangle = \left\langle \frac{dx^2}{dt} \right\rangle. \tag{7.47}$$

Because what the force is doing does not depend upon where the particle is, we say they are statistically uncorrelated:

$$\langle x(t)\, f(t) \rangle = 0. \tag{7.48}$$

Because we know something about $\langle xf \rangle$, let's multiply through:

$$\left\langle mx \frac{d^2x}{dt^2} \right\rangle = \left\langle -(6\pi\eta R) x \frac{dx}{dt} \right\rangle + 0. \tag{7.49}$$

The left-hand side looks unfamiliar, so let's rewrite it.

$$\frac{d}{dt}\left(x\frac{dx}{dt}\right) = \left(\frac{dx}{dt}\right)^2 + x\frac{d^2x}{dt^2}. \tag{7.50}$$

So,

$$m\left\langle x\frac{d^2x}{dt^2}\right\rangle = m\left\langle \frac{d}{dt}\left(x\frac{dx}{dt}\right)\right\rangle - m\left\langle \left(\frac{dx}{dt}\right)^2\right\rangle. \tag{7.51}$$

Now it's $x(dx/dt)$ that looks weird. Thinking about it and playing with calculus relations, we eventually recognize that we can write it using the chain rule, because

$$\frac{d}{dt}x^2 = 2x\frac{dx}{dt}. \tag{7.52}$$

Therefore,

$$m\left\langle x\frac{d^2x}{dt^2}\right\rangle = \frac{m}{2}\left\langle \frac{d^2}{dt^2}x^2\right\rangle - m\left\langle \left(\frac{dx}{dt}\right)^2\right\rangle. \tag{7.53}$$

Also,

$$-(6\pi\eta R)x\frac{dx}{dt} = -(3\pi\eta R)\frac{d}{dt}x^2. \tag{7.54}$$

It's helpful to define

$$z := \frac{d}{dt}\langle x^2\rangle. \tag{7.55}$$

This gives

$$\frac{m}{2}\left\langle \frac{d^2}{dt^2}x^2\right\rangle - m\left\langle \left(\frac{dx}{dt}\right)^2\right\rangle = \frac{m}{2}\frac{dz}{dt} - m\langle v^2\rangle, \tag{7.56}$$

$$-(3\pi\eta R)\left\langle \frac{d}{dt}x^2\right\rangle = -(3\pi\eta R)z. \tag{7.57}$$

Putting them together,

$$\frac{m}{2}\frac{dz}{dt} + (3\pi\eta R)z - m\langle v^2\rangle = 0. \tag{7.58}$$

Recalling that k_BT is the kinetic energy,

$$\frac{dz}{dt} + \frac{6\pi\eta R}{m}z = \frac{2k_BT}{m}. \tag{7.59}$$

We can find the long-term behavior by assuming that z isn't changing. Then z will remain fixed at a value z_c given by

$$z_c = \frac{k_BT}{3\pi\eta R}. \tag{7.60}$$

This says that $\langle x^2 \rangle$ grows at a *constant rate.* This is a characteristic feature of Brownian motion. The distance scale of a probability distribution, quantified by $\sqrt{\langle x^2 \rangle}$, grows as the square root of the elapsed time.

Recalling the diffusion equation (2.171) and the discussion that led up to it, we can identify this constant value of z as a diffusion coefficient. In Brownian motion, both the kicks that get the mesoscopic particle moving and the viscosity that slows it back down are due to the same medium. Quantitatively, the *fluctuations,* parameterized by the diffusion coefficient, are tied to the *dissipation,* parameterized by the viscous drag. We have here an example of a *fluctuation-dissipation theorem,* often written as an equality like

$$bD = k_B T \,, \tag{7.61}$$

where in this case we have

$$b = 6\pi\eta R, \;\; D = \frac{1}{2} z_c \,. \tag{7.62}$$

7.3 The Boltzmann Equation

So far, we have mostly discussed systems in thermal equilibrium, while letting slide the question of how systems might come to be equilibrated. We will now take our first real step in this direction by presenting the *Boltzmann equation,* which will help us understand at least small departures from equilibrium. We can think of the Boltzmann equation as a description of the most probable way that the most probable configuration can change. Accordingly, it is not well suited to treating unusual events that require careful coordination to achieve.

When we deduced that the average kinetic energy in an ideal gas is $\frac{3}{2} k_B T$, back in Eq. (5.21), we made use of the fact that the pressure exerted by a gas depends on how many molecules are colliding with the container wall and how fast, but not on which molecules are doing the impacting. One corpuscle is as good as any other for the purpose. The same will hold true in more complicated situations, like a gas that varies in temperature throughout or that grows thinner with altitude. We can, in principle, imagine a joint probability density function for all $N \approx 10^{23}$ molecules in a reasonably-sized portion of gas:

$$\rho(\vec{x}_1, \vec{p}_1, \vec{x}_2, \vec{p}_2, \ldots, \vec{x}_N, \vec{p}_N, t). \tag{7.63}$$

But this is drastic overkill, because molecule 1 striking a wall at position $\vec{x}$ with momentum $\vec{p}$ has the same effect as molecule ten trillion doing so. One way to simplify is to integrate over $N - 1$ of the positions and momenta to obtain a one-particle probability density function:

$$\rho_1(\vec{x}_1, \vec{p}_1, t) := \int d^3\vec{x}_2 d^3\vec{p}_2 \cdots d^3\vec{x}_N d^3\vec{p}_N \, \rho(\vec{x}_1, \vec{p}_1, \vec{x}_2, \vec{p}_2, \ldots, \vec{x}_N, \vec{p}_N, t). \tag{7.64}$$

However, this does not really account for the symmetry of the situation, that there is nothing intrinsically special about particle number 1. In addition, it still expresses a

full probability density function, which is more information than we need at this stage: We are not concerned about unusual events, but commonplace ones.

Accordingly, we introduce a new function, which we will write $f_{\mathrm{I}}(\vec{x}, \vec{p}, t)$. The new style of subscript is to emphasize that this does not refer to a specific particle. Instead, $f_{\mathrm{I}}(\vec{x}, \vec{p}, t)$ gives the *most likely density* of particles in the volume $d^2\vec{x}d^3\vec{p}$ of phase space centered at $(\vec{x}, \vec{p})$.

The Boltzmann equation gives the time evolution of f_{I}, under the assumption that the most likely density changes in the most likely way. This equation contains *streaming terms,* which express how molecules move without affecting one another, and *collision terms,* which incorporate what happens when they do. The former describe a continuous process and are advective in character, whereas the latter introduce the possibility of discontinuous changes of momentum. This is because a collision between two molecules can drastically change the momenta of both. A molecule can "disappear" from $(\vec{x}, \vec{p})$ and "appear" at $(\vec{x}, \vec{p}')$, and so the kind of thinking we used with deriving the diffusion equation doesn't apply, because probability no longer "flows" locally.

Let's first try to understand how f_{I} should change without collisions. First, there should be an advection, since a particle at $(\vec{x}, \vec{p})$ should an instant later have moved to a new position:

$$\vec{x} \to \vec{x} + \frac{\vec{p}}{m} dt. \tag{7.65}$$

This alone would give us a partial differential equation for f_{I} that expresses f_{I} changing because particles both leave from and arrive at the location $\vec{x}$:

$$\frac{\partial f_{\mathrm{I}}(\vec{x}, \vec{p}, t)}{\partial t} = -\frac{\vec{p}}{m} \cdot \frac{\partial}{\partial \vec{x}} f_{\mathrm{I}}(\vec{x}, \vec{p}, t). \tag{7.66}$$

What about changes in particle momenta? Momentum changes when a force is applied, and we can write a force as a gradient of a potential energy. We will use $U(\vec{x})$ to denote a potential applied from outside the system. This gives us another advective term:

$$\frac{\partial f_{\mathrm{I}}(\vec{x}, \vec{p}, t)}{\partial t} = \left[-\frac{\vec{p}}{m} \cdot \frac{\partial}{\partial \vec{x}} + \frac{\partial U}{\partial \vec{x}} \cdot \frac{\partial}{\partial \vec{p}} \right] f_{\mathrm{I}}(\vec{x}, \vec{p}, t). \tag{7.67}$$

The Boltzmann equation takes this expression and adds a collision term that also depends upon f_{I}.

$$\left[\frac{\partial}{\partial t} - \frac{\partial U}{\partial \vec{x}} \cdot \frac{\partial}{\partial \vec{p}} + \frac{\vec{p}}{m} \cdot \frac{\partial}{\partial \vec{x}} \right] f_{\mathrm{I}} = C[f_{\mathrm{I}}]. \tag{7.68}$$

The collision term $C[f_{\mathrm{I}}]$ can get fairly involved, depending on the force by which colliding particles interact. However, we can still deduce the general form of it. First, we make the approximation that interactions happen at a point; rather than considering particles at $(\vec{x}_1, \vec{p}_1)$ and $(\vec{x}_2, \vec{p}_2)$, we take $\vec{x}_1 = \vec{x}_2 = \vec{x}$. This amounts to saying that the time during which two particles are affecting each other is much shorter than the time that particles spend between collisions. And it is in line with our attitude that we are only considering commonplace behaviors, rather than those which require a careful

alignment to achieve — we are simply not keeping track of precise positioning information! So, if two particles with incoming momenta $\vec{p}_1$ and $\vec{p}_2$ collide at $\vec{x}$, we won't try to predict exactly what the outgoing momenta $\vec{p}_1'$ and $\vec{p}_2'$ will be. Instead, we will integrate over the possibilities for those outgoing momenta. As we have argued before, $f_{\rm I}$ will be diminished by particles "leaving" $(\vec{x}, \vec{p}_1)$ and $(\vec{x}, \vec{p}_2)$, and it will be augmented by particles "arriving" at $(\vec{x}, \vec{p}_1')$ and $(\vec{x}, \vec{p}_2')$. To see how often such transitions occur, we imagine ourselves riding along with a particle having velocity $\vec{v}_1 = \vec{p}_1/m$. For any chosen $\vec{v}_2$, we will see a flux of incoming particles that is proportional to $|\vec{v}_1 - \vec{v}_2|$.

Finally, we must consider what a collision can do to particle momenta. When two particles come together and scatter off one another, how much do their momenta change? This will depend upon the force through which the particles interact. But we can encapsulate those details using the idea of a *scattering cross section.* The idea is to think of probabilities in terms of areas, like the bullseye of a target. A scattering cross section is a notional area: We pretend as though a particle which passes within this area scatters, and one that does not will not. A *differential* cross section is a more refined idea, expressing the probability that a scattering event will send a particle going off in a specific direction, which we represent by a differential element of solid angle.

Putting these pieces together, we write the collision term as follows:

$$C[f_{\rm I}] = \int d^3\vec{p}_2 d^2\Omega \left|\frac{d\sigma}{d\Omega}\right| |\vec{v}_1 - \vec{v}_2| [f_{\rm I}(\vec{p}_1')f_{\rm I}(\vec{p}_2') - f_{\rm I}(\vec{p}_1)f_{\rm I}(\vec{p}_2)]. \tag{7.69}$$

It can happen that instead of being short-ranged, interactions between particles reach a long way across the volume of a system. In this case, it is not just nearby particles that affect each other's momenta. We can write a counterpart to the Boltzmann equation by the same heuristic method as before. Suppose that the potential which defines the box in which the system lives is $U(\vec{x})$, as before, and let $V(\vec{x} - \vec{x}')$ be the long-range interaction potential. Then the most likely potential "felt" by the particles at point $\vec{x}$ is the background influence $U(\vec{x})$ plus a contribution given by the most likely density of particles at all other locations:

$$U_{\rm tot}(\vec{x}, t) = U(\vec{x}) + \int d\vec{x}'\, V(\vec{x} - \vec{x}') f_{\rm I}(\vec{x}', t)\,. \tag{7.70}$$

We get the *Vlasov equation* by substituting this into our advection formula in the place of the original background potential:

$$\left[\frac{\partial}{\partial t} - \frac{\partial U_{\rm tot}}{\partial \vec{x}} \cdot \frac{\partial}{\partial \vec{p}} + \frac{\vec{p}}{m} \cdot \frac{\partial}{\partial \vec{x}}\right] f_{\rm I} = 0\,. \tag{7.71}$$

7.4 Application: Charge Pulses in Gases

One fruitful application of these ideas is to *electron swarms* and their motion through neutral gases. This field of inquiry is of major interest for the construction of particle detectors, which often work by the ionization of gas atoms. It is also an excellent

example of the importance of proper approximations, a topic on which later sections will elaborate. To begin with, we note that in the conditions of a typical drift detector, the de Broglie wavelength of the swarming electrons is much smaller than the spacing between atoms (for gas pressures less than about 100 atmospheres). Consequently, the swarm can be modeled as a set of classical particles, each interacting with only one gas atom at a time [64].

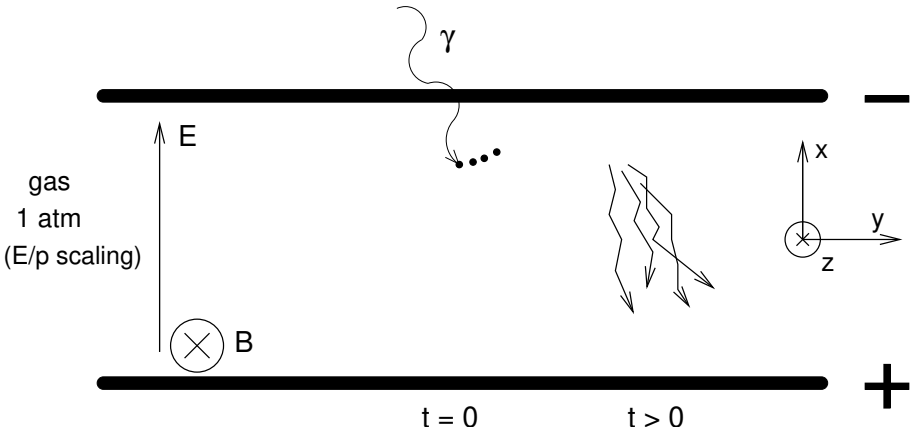


Figure 7.1: Schematic sketch of an electron swarm progressing through a drift chamber.

First, consider a nonrelativistic electron moving through vacuum, subjected to $\vec{E}$ and $\vec{B}$ fields. For our drift chamber geometry, we can choose a coordinate system such that $\vec{B} = B\hat{z}$ and $\vec{E} = E_x\hat{x}$. The electron's acceleration is given by the Lorentz force law,

$$m\frac{d\vec{v}}{dt} = e\vec{E} + e\vec{v} \times \vec{B}. \tag{7.72}$$

In our coordinate system, this simplifies to the three coupled equations

$$\dot{v}_x = \left(\frac{e}{m}\right) E_x + \omega v_y, \tag{7.73}$$

$$\dot{v}_y = -\omega v_x, \tag{7.74}$$

$$\dot{v}_z = 0. \tag{7.75}$$

Decoupling the first two equations gives

$$\ddot{v}_x + \omega^2 v_x = 0, \;\; \ddot{v}_y + \omega^2 v_y + \left(\frac{e}{m}\right) E_x \omega = 0, \tag{7.76}$$

which can be solved to yield

$$\begin{aligned} v_x &= \frac{eE_x}{m\omega}\sin(\omega t) + v_{ox}, \\ v_y &= \frac{eE_x}{m\omega}[\cos(\omega t) - 1] + v_{oy}, \\ v_z &= v_{oz}. \end{aligned} \tag{7.77}$$

Integrating these velocities from an initial time $t = 0$ to a time $t = \Delta t$ later gives expressions for the incremental change in position:

$$\begin{aligned} \Delta x &= \frac{eE_x}{m\omega^2}(1 - \cos\omega\Delta t) + v_{ox}\Delta t, \\ \Delta y &= \frac{eE_x}{m\omega^2}\sin\omega\Delta t - \frac{eE_x}{m\omega}\Delta t + v_{oy}\Delta t, \\ \Delta z &= v_{oz}\Delta t. \end{aligned} \tag{7.78}$$

Except for a change of coordinate axes, these formulas are essentially the same as those given in [65] (in the special case that $\vec{E}$ and $\vec{B}$ are orthogonal). Note that the corresponding equations in [65] lack factors of e/m on several instances of the electric field.

The next phenomenon which we must describe is the interaction of an electron with a gas atom or molecule. These events occur with probabilities dependent upon the electron-atom scattering cross section, the gas density and the applied fields. The simplest case is electron motion through a noble gas, in which all scatterers are monatomic. Since the gas mixtures in many drift detectors contain large proportions of noble gases [66, 67], this is a case of considerable practical interest.

The standard practice is to model the noble-gas atoms as hard spheres whose cross sections σ_m depend upon the energy of the incident electron [68]. This is quantum-mechanically justifiable [69] as long as the electron's energy ϵ is less than that required to excite the atom. Conveniently, helium possesses a first excitation energy of ≈ 19 eV; collisions below this upper limit will be elastic. Newtonian kinematics gives the result that a fraction $\lambda(\theta)$ of the electron's energy will be lost to the atom, where θ is the polar angle between the incident and final velocities:

$$\lambda(\theta) = \frac{2mM}{(M+m)^2}(1 - \cos(\theta)). \tag{7.79}$$

Here, m is the electron mass and M is the mass of the target atom [70]. We shall use λ without an argument to denote the fractional energy loss averaged over all angles θ:

$$\lambda = \frac{2mM}{(M+m)^2}. \tag{7.80}$$

For helium gas, this ratio is roughly 2.7×10^{-4}.

Gases not of the noble variety introduce complications. The presence of asymmetric molecules with translational and vibrational modes allows for a variety of inelastic

collisions, each of which is described by an energy-dependent cross section and a characteristic energy loss. In addition, these molecular gases typically ionize more readily than the noble-gas mixture component. This may be convenient in some experimental situations, where such ionizations are necessary (e.g., the apparatus of [66]), but it makes theoretical work more difficult. Common examples of molecular gases present in drift chamber mixtures include N_2 and hydrocarbons like CH_4, C_2H_6, etc. [64].

These ideas lend themselves fairly readily to a Monte Carlo (MC) implementation. Such an implementation is simple, in principle: Define an electron within the computer's memory by its position and velocity coordinates, propagate that electron according to the Lorentz force law Eq. (7.72), decide stochastically whether or not the electron suffers a collision during that timestep, and in the event of a collision randomize the velocity vector with an appropriate energy loss [70].

The key resource for the MC algorithm is a dependable source of random numbers. Using R to denote a random variable uniformly distributed between 0 and 1, we can write formulae for the kinematic quantities necessary to run the simulation. For isotropic scattering — the case covering hard-sphere noble gases — the polar angle θ is computed by

$$\theta = \cos^{-1}(1 - 2R), \tag{7.81}$$

while the azimuthal angle ϕ is chosen from a uniform distribution,

$$\phi = 2\pi R. \tag{7.82}$$

The most sophisticated part of the MC algorithm is choosing the proper timestep. If computational efficiency is not a primary goal, one can code the MC so that it estimates the mean time between collisions and chooses a timestep significantly smaller. A somewhat more efficient approach is to use the null-collision technique due to Skullerud [71, 72].

Figure 7.2 shows the result of running an MC code that was not designed with much sophistication in mind, but is still serviceable. Drift velocities were computed by measuring the displacement after a given number of collisions and dividing by the amount of time elapsed. Three observations are noteworthy: First, the drift velocity trace as a function of time $v_d(t)$ shows a transient behavior with a timescale of $\approx 3\times10^4$ collisions. MC runs with different starting conditions converge to the same result, after which v_d stays roughly constant for as long as the simulation runs. (In physical time, this transient behavior could exist on the order of a microsecond.) Second, when the MC is run for a range of $\vec{E}$ fields, the results for $\sigma_m = 7\text{Å}^2$ and $\sigma_m = 8\text{Å}^2$ neatly bracket the experimental results for He [73] (using the appropriate λ for He).

Because the relatively simple interactions upon which our MC is built lead to simulation results which agree with experiment, it is tempting to consider an analytic treatment based on the same assumptions. MC researches face difficulties: Psychologically, there is a bit of "equation envy", a worry that the lack of formulae will produce the impression that the treatment lacks understanding. On a more practical level, MC calculations take *time*. A workstation must evaluate $\approx 10^5$ collisions, and even the iterations which do not involve an electron-atom impact require trigonometric manipulations, which evaluate relatively slowly. We turn, therefore, to the highly successful

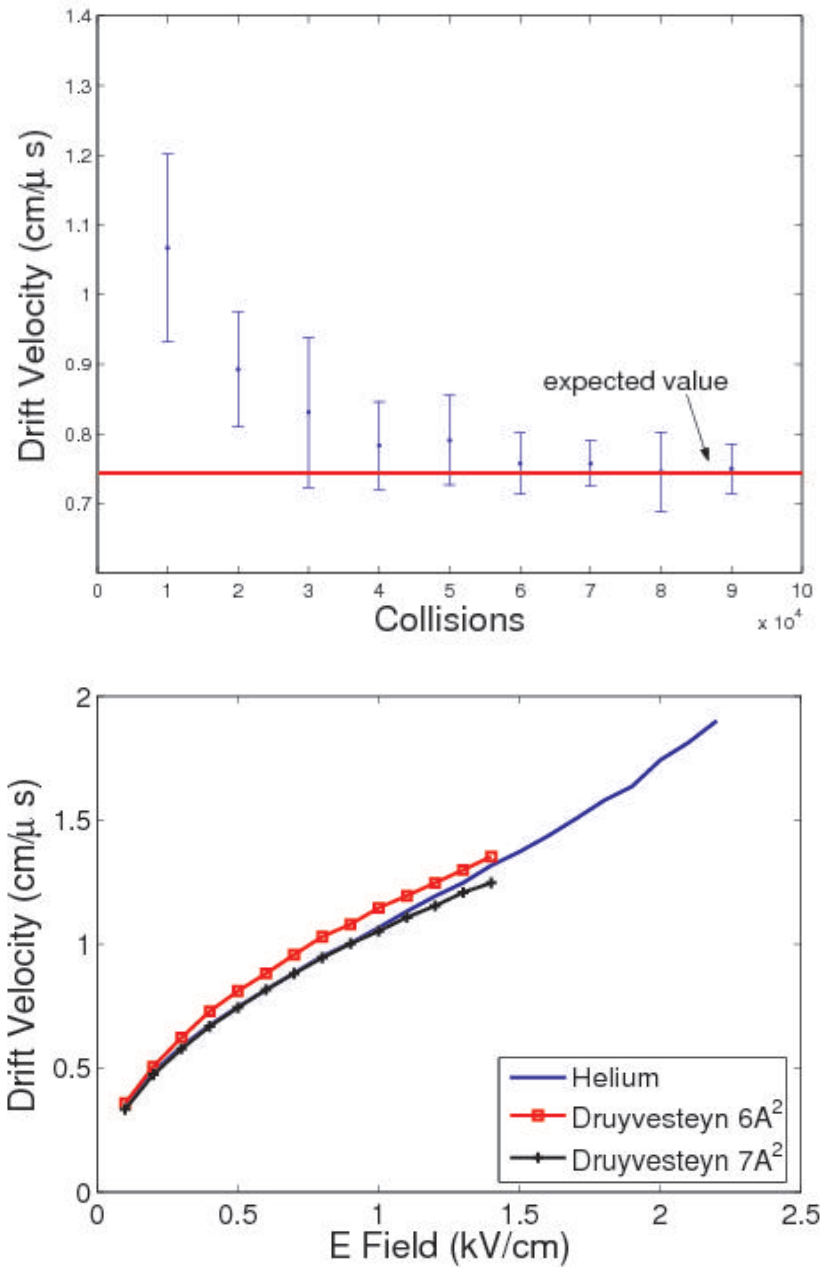


Figure 7.2: Drift velocities computed by Monte Carlo. The upper graph was computed for $E = 0.5$ kV/cm and $\sigma_m = 6\text{Å}^2$.

kinetic theory of gases, in the hope that it will provide expressions for quantities of interest (such as v_d and α) which can be evaluated faster than performing an MC run.

We consider one species of particle, the electrons, moving stochastically amongst another species, the neutral scatterers. For most of this paper, the scatterers will be hard spheres, distributed evenly on a macroscopic scale. This is a reasonable approximation for noble gases, as discussed above. Blum and Rolandi [64] quote several key results of swarm kinetic theory, though with only minimal derivations. Fuller explanations can be found in Shkarofsky *et al.* [74] and the review article [68]. For pedagogical treatments of kinetic theory in general, see *e.g.* Huang [75] and Balescu [76].

Some standard approximations made in neutral-gas kinetic theory also apply here, and have been borne out by Monte Carlo studies. First, the electron-atom interaction is a *local* one, a short-range process whose timescale is much less than the typical time between collisions. Second, the electrons are sufficently far apart that they do not strongly interact with one another. These approximations, valid for dilute and weakly ionized plasmas (the sort found in drift detectors), allow us to apply the kind of thinking we developed earlier in this chapter. The resulting relation is a generalized form of the Boltzmann kinetic equation, written in terms of the electron distribution function f_I and the gas atom distribution, denoted g:

$$\left(\frac{\partial}{\partial t} + \vec{v}\cdot\nabla_q + \frac{e}{m}(\vec{E}+\vec{v}\times\vec{B})\cdot\nabla_v\right) f_\mathrm{I}(\vec{q},\vec{v},t) = C[f_\mathrm{I}, f_\mathrm{I}] + C[f_\mathrm{I}, g]. \tag{7.83}$$

The "streaming terms" on the left express how f_I changes over time due to the applied fields [74]. If the electron swarm propagated through free space — and if the electrons did not interact with one another — then Eq. (7.83) would have only the streaming terms, and it would describe a flow coasting along advectively. However, these electrons are colliding with gas atoms, which intuitively means that an electron may "disappear" from one position in phase space and "reappear" elsewhere with an altered momentum. Therefore, the right-hand side of Eq. (7.83) contains *collision terms,* symbolically denoted with the C operator, which encode how the particle interactions change the swarm distribution f_I.

In the case of a dense swarm — that is, a strongly ionized plasma — we must deal with *long-range* interactions, because the Coulomb potential falls off as $1/r$. Such a case is outside the realm of Eq. (7.83)'s applicability, and must instead be described by a generalized Vlasov equation, analagous to the ones written for the distribution of stars in a galaxy, since gravitational interactions are also $1/r$ potentials. (See, *e.g.,* [77] or [78].) Also, the derivation of Eq. (7.83) depends upon the "assumption of molecular chaos": it is only valid when the two-particle distribution f_II can be factored into independent one-particle distributions. Symbolically,

$$f_\mathrm{II}(\vec{p}_1,\vec{q}_1,\vec{p}_2,\vec{q}_2) \approx f_\mathrm{I}(\vec{p}_1,\vec{q}_1)\cdot f_\mathrm{I}(\vec{p}_2,\vec{q}_2). \tag{7.84}$$

We will return to this in much more depth during the next chapter.

Here, we follow the standard practice in swarm theory, which neglects the first term — electron-electron interactions — in favor of the second term, $C[f_\mathrm{I}, g]$, which expresses the interactions of the light electrons with the heavy, neutral gas atoms [68].

This term can be expected to include a scattering cross-section, representing how likely collisions are to occur, as well as some knowledge of energy losses and momentum transfers, which represent the effect that collisions have on the electrons. A standard result in kinetic theory writes this collision term as an integral over the gas atom's momentum, $\vec{p}_g$. Using $\vec{p}'$ and $\vec{p}_g{}'$ to denote the gas and electron momenta *after* the collision (which we could in principle deduce from the specifics of the interaction potential), we can write the operator C as

$$C[f_\mathrm{I}, g] = -\int d^3\vec{p}_g d\Omega \frac{d\sigma}{d\Omega} |\vec{v} - \vec{v}_g| [f_\mathrm{I}(\vec{p}, \vec{q}) g(\vec{p}_g, \vec{q}) - f_\mathrm{I}(\vec{p}', \vec{q}) g(\vec{p}_g{}', \vec{q})]. \tag{7.85}$$

Note that all functions carry the same position argument, since we have approximated that the collision occurs at the single point $\vec{q}$. (The time arguments have been suppressed for clarity.)

In principle, we would have to provide a similar formalism for the time evolution of $g(\vec{p}, \vec{q})$, the one-particle distribution for gas atoms. This is in fact not necessary for two important reasons: first, that the electrons are a thousandfold lighter than the gas atoms, and second, that we have neglected the issues of ionization or electron attachment which may alter the gas atoms' properties after the swarm passes. As explained above, an electron's fractional energy loss per collision is small, but not entirely negligible; this condition is known as a *pseudo-Lorentz gas.* In this case, the effect of the energy loss is much more pronounced on the electron swarm than it is upon the neutral gas, which remains essentially at the Maxwell–Boltzmann distribution,

$$g(\vec{p}, \vec{q}) = \frac{N}{(2\pi M k_B T)^{3/2}} \exp\left(-\frac{p^2}{2M k_B T}\right), \tag{7.86}$$

where M is the mass of a gas atom, as above, and N denotes the number of atoms per unit volume.

Based on the conditions of the typical experimental setup, swarm theories are typically developed in the *hydrodynamic* regime, where the swarm's evolution is unaffected by boundary conditions and has no memory of its initial configuration. (In traditional kinetic theory, hydrodynamic equations govern a system where all expectation values have relaxed to *local* equilibria; the system is then expected to relax on a much longer timescale to some global equilibrium like the Maxwell–Boltzmann distribution.) Working in this regime allows us to neglect the spatial and temporal dependencies of f_I, regarding it as a function of velocity alone:

$$f_\mathrm{I} = f_\mathrm{I}(\vec{v}). \tag{7.87}$$

For the special case of isotropic scattering, which the Monte Carlo approach treated above, we may expect the distribution to be also isotropic in velocity space. This allows us to write a two-term expression for f_I:

$$f_\mathrm{I}(\vec{v}) = f_0(v) + \frac{\vec{v}}{v} \cdot \vec{f}_1(v). \tag{7.88}$$

This treatment makes intuitive sense for noble gases, and it is supported by the Monte Carlo results derived earlier. However, *it cannot be expected to hold for more complicated scattering molecules.* Recent investigations have focused on the situations when Eq. (7.88) breaks down, necessitating that the angular dependence of $f_1(v)$ be written with an arbitrarily long expansion in spherical harmonics. This can occur in such a common instance as the inclusion of methane (CH_4) in the neutral gas [79].

We are chiefly concerned with the case that $\vec{E}$ and $\vec{B}$ are perpendicular (with $\vec{B}$ possibly zero). Eq. (7.83) can be solved in much the same way as the ordinary Boltzmann equation. Shkarofsky *et al.* give a detailed exposition, while Blum and Rolandi quote the final answer. Again, the electron distribution as a function of energy $f_0(\epsilon)$ is the exponential of minus a quantity, but here the argument is not an energy divided by $k_B T$. Instead, it is an integral over all energies, depending upon the cross section $\sigma_m(\epsilon)$ and the fractional energy loss per collision $\lambda(\epsilon)$. When the $\vec{E}$ and $\vec{B}$ fields are orthogonal, we find the distribution function is

$$f_0(\epsilon) \propto \exp\left[-\frac{3m}{2e^2E^2}\int_0^{\epsilon}\lambda(\epsilon')[\nu^2(\epsilon')+\omega^2]\,d\epsilon'\right] \tag{7.89}$$

where ω is the cyclotron frequency $(e/m)B$ (see reference [80]). $\nu(\epsilon)$ depends upon the number density and the momentum-transfer (or "effective") cross section: $\nu(\epsilon) = \mu\sigma_m(\epsilon)v$, where $v = \sqrt{2\epsilon/m}$.

The constant of proportionality is fixed by normalization. The definition of the probability distribution implies that the integral of $f_0(v)$ over all velocities is related to the spatial density of electrons, n:

$$4\pi\int_0^{\infty} v^2 f_0(v)dv = n. \tag{7.90}$$

Changing Eq. (7.90) to an integral over energy ϵ is only a matter of changing variables. The density n is found to equal

$$n = 2\pi\left(\frac{2}{m}\right)^{3/2}\int_0^{\infty}\epsilon^{1/2}f_0(\epsilon)d\epsilon. \tag{7.91}$$

For convenience's sake, papers may graph $f_0(\epsilon)$ normalized so that the integral of $\epsilon^{1/2}f_0(\epsilon)$ equals 1.

Knowing $f_0(\epsilon)$, the generalized Boltzmann equation can be used to calculate $f_1(\epsilon)$ (see Shkarofsky for details). The result is found in terms of derivatives of f_0:

$$f_1 = -\frac{v}{\nu}\vec{\nabla}_r f_0 - \frac{e\vec{E}}{m\nu}\frac{\partial f_0}{\partial v}. \tag{7.92}$$

As a first approximation, we shall again neglect the term involving spatial derivatives, focusing on f_0's velocity dependence. With probability distributions in hand, we can calculate expectation values by performing the appropriate integrals. For any vector function $\vec{\mathcal{O}} = \mathcal{O}(v)\frac{\vec{v}}{v}$,

$$\left\langle\vec{\mathcal{O}}\right\rangle = \frac{4\pi}{3n}\int\mathcal{O}\vec{f}_1 v^2 dv. \tag{7.93}$$

The drift velocity itself is defined to be

$$\vec{v_d} \equiv \langle \vec{v} \rangle = \frac{4\pi}{3n} \int \vec{f_1} v^3 dv. \tag{7.94}$$

Generally speaking, the electron swarm's diffusion must be defined by a matrix D_{ij} whose indices range over the $\hat{x}$, $\hat{y}$ and $\hat{z}$ axes. In terms of a velocity integral,

$$D_{ij} = \frac{4\pi}{3n} \int_0^\infty f_0 \frac{v^4}{\nu^2 + \omega^2} \begin{pmatrix} \nu & -\omega & 0 \\ \omega & \nu & 0 \\ 0 & 0 & \frac{\nu^2+\omega^2}{\nu} \end{pmatrix}, \tag{7.95}$$

which simplifies in the $\vec{B} = 0$ case to the single relation

$$D_T = \frac{4\pi}{3n} \int_0^\infty f_0 \frac{v^4}{\nu} dv \tag{7.96}$$

$$= \frac{4\pi}{3n} \int_0^\infty f_0 \frac{v^3}{N\sigma(v)} dv. \tag{7.97}$$

Equivalent relations can be derived in "energy space" as well. Differentiating (7.89) gives

$$\frac{df_0}{d\epsilon} = -\frac{3m}{2e^2E^2} \lambda(\epsilon)[\nu^2(\epsilon) + \omega^2] f_0(\epsilon). \tag{7.98}$$

The chain rule gives us a relation between this derivative and the derivative taken with respect to electron velocity.

$$\frac{df_0}{dv} = \frac{d\epsilon}{dv}\frac{df_0}{d\epsilon} = mv\frac{df_0}{d\epsilon}. \tag{7.99}$$

By transforming the integrals to an ϵ dependence, one can find the following expressions for the drift velocity v_d and the transverse diffusion coefficient D_T:

$$v_d = -\frac{1}{3}\left(\frac{2}{m}\right)^{1/2} (eE/\mu) \int_0^\infty [\sigma_m(\epsilon)]^{-1} \left(\frac{df_0}{d\epsilon}\right) \epsilon d\epsilon, \tag{7.100}$$

$$D_T = (1/3\mu)\,(2/m)^{1/2} \int_0^\infty [\sigma_m(\epsilon)]^{-1} f_0(\epsilon) \epsilon d\epsilon. \tag{7.101}$$

Eqs. (7.100) and (7.101) are only valid in the $\vec{B} = 0$ case. Furthermore, when $\vec{B}$ is non-zero, a new quantity becomes of interest, the Lorentz deflection angle. When $\vec{E}$ is perpendicular to $\vec{B}$, the Lorentz angle takes on a computable form. The tangent of the angle α is given by the ratio of two integrals,

$$\tan\alpha = \frac{\int_0^\infty \frac{v^3\omega}{\nu^2+\omega^2}\frac{df_0}{dv}dv}{\int_0^\infty \frac{v^3\nu}{\nu^2+\omega^2}\frac{df_0}{dv}dv}. \tag{7.102}$$

It should be noted that all of these swarm parameters depend upon the electric field $\vec{E}$ and magnetic field $\vec{B}$ only through the ratios $\vec{E}/\mu$ and $\vec{B}/\mu$. A gas containing

Avogadro's number of molecules at the STP volume of 22.4 litres has a number density of roughly 2.69×10^{25} molecules per cubic metre. The experiments listed in [67] report results at a slighly higher temperature, implying a somewhat lower density, $\mu = 2.47 \times 10^{25}$. For a typical drift-chamber electric field on the order of 1 kV/cm, E/μ is on the order of 10^{-21} Vm^2. A convenient unit for E/μ is the townsend (Td), which is defined to be 10^{-17} Vcm^2, or 10^{-21} Vm^2. In an analogous manner, White *et al.* define the huxley (Hx) to be 10^{-27}Tm^3, a convenient scale for the density-reduced magnetic field [81].

The integrals in Eqs. (7.89), (7.100) and (7.101) can be evaluated by hand, in certain special cases described more fully below. In the "worst case scenario", the integration can be performed numerically, using a straightforward quadrature approach. However, when any of our assumptions are relaxed, the analytic expressions for swarm parameters become much more complicated — if they can be expressed at all. Including the effects of anisotropic and inelastic scattering, for example, or modeling $\vec{E}$ and $\vec{B}$ fields crossed at arbitrary angles leads to formalisms of much greater mathematical intricacy. In some models, the temperature $k_B T$ becomes a tensor [82, 81]. These developments are beyond the scope of this paper.

Certain special cases have been studied in which the cross-sectional dependence assumes particularly simple forms. For example, when σ_m is constant over its energy range, we obtain the Druyvesteyn distribution:

$$f_0^D(\epsilon) \propto \exp\left[-\frac{3m\lambda N^2}{2e^2E^2}\left(\frac{\sigma_m^2}{m}\epsilon^2 + \frac{\omega^2}{N^2}\epsilon\right)\right]. \tag{7.103}$$

Also, if we have the cross section decay with increasing energy as $\sigma_m \propto \sigma_0/\sqrt{\epsilon}$, we find that the drift electrons follow a Maxwell–Boltzmann distribution.

$$f_0^M(\epsilon) \propto \exp\left[-\frac{3m\lambda N^2}{2e^2E^2}\left(\frac{2\sigma_0^2}{m} + \frac{\omega^2}{N^2}\right)\epsilon\right]. \tag{7.104}$$

This allows us to define an electron temperature T_e:

$$k_B T_e = \frac{2e^2(E/N)^2}{3m\lambda\left(\frac{2\sigma_0^2}{m} + \frac{\omega^2}{N^2}\right)}. \tag{7.105}$$

Assuming the Maxwellian distribution given by (7.104) and (7.105), substituting $f_0^M(\epsilon) \propto \exp(-\epsilon/k_B T)$ into (7.102) shows that

$$\tan\alpha = \frac{\omega}{\nu_0} = \frac{B}{N}\left(\frac{e}{m}\right)\sqrt{\frac{m}{2}}\sigma_0^{-1}. \tag{7.106}$$

Here we have defined ν_0 to equal the constant collision frequency, which in Maxwell's distribution is independent of the velocity v. This formula makes intuitive sense: The Lorentz angle is a "compromise" between the magnetic field, which pushes the swarm off track, and the scattering effect, which limits its progression. We would thus expect the angle to increase with applied field and decrease with the cross-sectional area, and this is exactly what Equation (7.106) predicts.

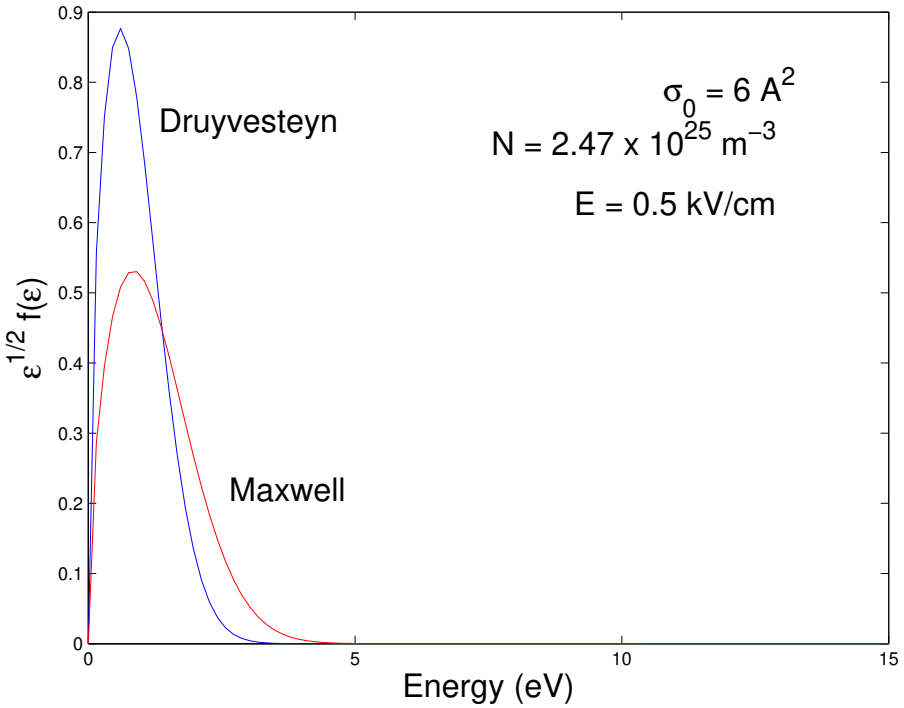


Figure 7.3: Comparison of Maxwellian (7.104) and Druyvesteyn (7.103) distributions, for a typical E-field and gas density.

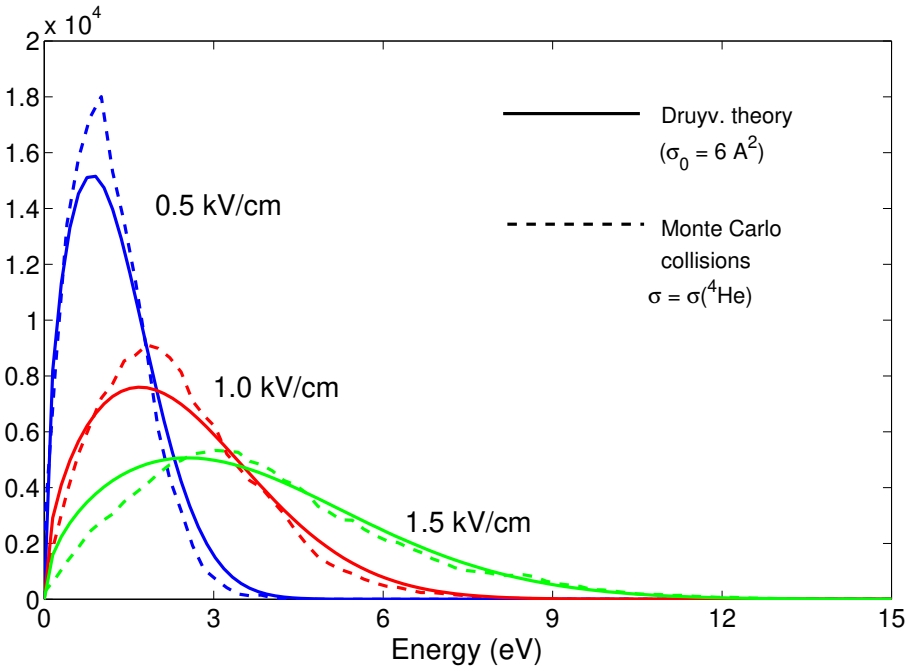


Figure 7.4: Theoretical predictions of the Druyvesteyn model compared with Monte Carlo calculations. The dashed lines indicate a histogram of collision energies, recorded during simulation runs of 5×10^5 collisions. Note that the Druyvesteyn approximation (which is an equilibrium result obtained in the hydrodynamic regime) closely follows the Monte Carlo calculation results.

8 Kinetic Theory II: Hamiltonian Dynamics

8.1 Introduction

One of the persistent themes of a physics education is that it is good to have multiple ways of describing a problem or formulating a theory, because different pictures can make different sets of questions approachable.

The Hamiltonian picture of classical physics is expressed in the language of *phase space.* Phase space is an augmentation of ordinary, geometrical coordinate space: Instead of merely saying where a body is, we introduce additional dimensions to indicate how it is moving. Providing a point within phase space is "informationally complete", in that if we know what phase-space point to associate with a system, we can calculate anything else about that system we'd like to know. The simplest example of a phase space for classical physics is two-dimensional, including one axis for *position* and a second axis for *momentum.* A prototypical application to keep in mind is a pendulum. If we know, for example, that the pendulum bob is exactly vertical, we don't know whether it is going to stay that way. It might be momentarily passing through the vertical position during a rapid swing, or it might be hanging there, placid and still. In order to forecast its future behavior, we need to specify both its location and its manner of motion.

The centerpiece of a Hamiltonian treatment of a physics problem is a function defined on a system's phase space that encodes all of its dynamical possibilities, a function called the system's Hamiltonian:

$$\mathcal{H} : (x_1, \ldots, x_N, p_1, \ldots, p_N, t) \to \mathbb{R}. \tag{8.1}$$

Here, we are using x_i to denote the position coordinates in phase space and p_i to denote the momentum coordinates. A Hamiltonian can also include an explicit dependence upon time t, for example if some external influence like a background magnetic field is varying over the course of our experiment. We will, for the most part, neglect that.

We find how the position and momentum coordinates change over time by seeing how the value of the Hamiltonian *would change* under small displacements in phase space. Hamilton's equations are, using dots to denote time derivatives,

$$\dot{x}_i = \frac{\partial \mathcal{H}}{\partial p_i}, \ \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x_i}. \tag{8.2}$$

We can relate this back to first-year physics by considering an example, a body of

mass m moving in a 1D potential $V(x)$. The Hamiltonian is

$$\mathcal{H}(x,p) = \frac{p^2}{2m} + V(x). \tag{8.3}$$

Taking the partial derivatives as prescribed, we first find that

$$\dot{x} = \frac{\partial \mathcal{H}}{\partial p} = \frac{p}{m}, \tag{8.4}$$

which just states that the momentum is the mass m times the velocity $\dot{x}$, so that makes sense. The rate of change of momentum is

$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial x} = -\frac{dV}{dx}. \tag{8.5}$$

This says that the force is minus the derivative of potential energy with respect to position, which again makes sense. So, we can at least recover the way we did physics before, the "Newtonian" style of classical physics — which, historically speaking, involves a lot of ideas that came after Newton's time, but we call it "Newtonian" in the course catalogue.

One reason we like Hamiltonian mechanics for doing statistical physics is because it's easy to extend to situations involving many particles. We just add terms to the Hamiltonian, and each term is a *scalar,* not a vector. For example, if we have N particles moving about in the same 1D potential well, then we can write a Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{i=1}^{N} V(x_i) + \frac{1}{2}\sum_{i\neq j} U(|x_i - x_j|), \tag{8.6}$$

where the third sum expresses a two-body interaction effect between particles that depends upon the distance between them. We have assumed here that all the particles have the same mass m, which is not at all necessary in principle, but it will be the case for most of the problems we study.

If we do not know exactly where the parts of a system are or how they are moving, we can express what we do know about the system by a probability density function over phase space. Forecasting the system's behavior into the future (or "retrodicting" where it might have been in the past) requires calculating the time evolution of this probability density function.

Suppose that we draw a dot in phase space at (x_a, p_a). Time evolution in accord with Hamilton's equations would carry this dot to new coordinates (x_b, p_b). If the evolution is over a very short time interval δt, then

$$(x_a, p_a) \to (x_b, p_b) = (x_a + \dot{x}|_{(x_a,p_a)}\delta t, p + \dot{p}|_{(x_a,p_a)}\delta t). \tag{8.7}$$

Meanwhile, a dot initially at the slightly different coordinates $(x_a + dx, p_a + dp)$ would go to

$$(x_a + dx, p_a + dp) \to (x_a + dx + \dot{x}|_{(x_a+dx,p_a+dp)}\delta t, p_a + dp + \dot{p}|_{(x_a+dx,p_a+dp)}). \tag{8.8}$$

So, the little intervals dx and dp will potentially be stretched or squeezed, if the time evolution amplifies or diminishes the perturbation to the initial conditions. The new intervals are, to first order in δt,

$$dx' = dx + \frac{\partial \dot{x}}{\partial x} dx \delta t, \tag{8.9}$$

$$dp' = dp + \frac{\partial \dot{p}}{\partial p} dp \delta t. \tag{8.10}$$

Consequently, for sufficiently brief time evolutions,

$$dx'dp' = dxdp \left[1 + \left(\frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{p}}{\partial p} \right) \delta t \right]. \tag{8.11}$$

But using Hamilton's equations, the derivatives within parentheses become

$$\frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{p}}{\partial p} = \frac{\partial^2 \mathcal{H}}{\partial x \partial p} - \frac{\partial^2 \mathcal{H}}{\partial p \partial x} = 0. \tag{8.12}$$

We deduce that *Hamiltonian time evolution preserves areas in phase space.* This is a very general result known as *Liouville's theorem.*

We can write an explicit equation for how a probability density function on phase space — a "Liouville density" — changes as we forecast into the future, by codifying the intuition that the probability we associate to a small region in phase space should increase if there's more chance to enter it than to leave it. The imagery is much like that of classical electromagnetism, where we relate an integral over a volume (the total charge within a Gaussian surface) to an integral over its boundary (the electric flux through that Gaussian surface). In this case, the "flux" is not itself physical; it is a flow of expectation, of anticipation. Hamilton's equations give a directionality to each point in phase space, and the probability density at a point should diminish if those trajectories *diverge* there. Without loss of generality, we can consider a phase space with one position coordinate x and one momentum coordinate p, for which this divergence is

$$\frac{\partial}{\partial x}(\dot{x}\rho) + \frac{\partial}{\partial p}(\dot{p}\rho) = \rho \frac{\partial \dot{x}}{\partial x} + \dot{x} \frac{\partial \rho}{\partial x} + \rho \frac{\partial \dot{p}}{\partial p} + \dot{p} \frac{\partial \rho}{\partial x}. \tag{8.13}$$

The terms proportional to ρ itself vanish by Hamilton's equations. We are left with the relation

$$\frac{\partial \rho}{\partial t} = \frac{\partial \rho}{\partial p} \frac{\partial \mathcal{H}}{\partial x} - \frac{\partial \rho}{\partial x} \frac{\partial \mathcal{H}}{\partial p}. \tag{8.14}$$

This combination of interwoven partial derivatives occurs very commonly, and so we abbreviate it with the *Poisson bracket.* For two functions f, g on phase space, the Poisson bracket of f and g is

$$\{f, g\} := \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial x_i} \right). \tag{8.15}$$

Note that this definition implies

$$\{f, f\} = 0, \tag{8.16}$$

and more generally,

$$\{f, g\} = -\{g, f\}. \tag{8.17}$$

The *Liouville equation* for time evolution is concisely given as

$$\frac{\partial \rho}{\partial t} = -\{\rho, \mathcal{H}\}. \tag{8.18}$$

If we are interested in any other measurable quantity pertaining to a system — say, its total angular momentum around a given axis — this will be a function of its position and momentum coordinates. So, our expectations regarding those coordinates will translate to an expectation value for that new quantity:

$$\langle A(t)\rangle = \int dx_1 \cdots dx_N dp_1 \cdots dp_N \, A(x_1, \ldots, x_N, p_1, \ldots, p_N)\rho(x_1, \ldots, x_N, p_1, \ldots, p_N, t). \tag{8.19}$$

The Liouville equation for time evolution, which follows from Hamilton's equations, implies that

$$\frac{d\langle A(t)\rangle}{dt} = \langle\{A, \mathcal{H}\}\rangle. \tag{8.20}$$

So, if the "observable" A has a vanishing Poisson bracket with the Hamiltonian $\mathcal{H}$, our expectation value for it is conserved. One way this can happen is if A is a function of $\mathcal{H}$, that is, if we can calculate the value of A knowing only the value of $\mathcal{H}$, without having to know the exact phase-space coordinates themselves. Then

$$\frac{\partial A}{\partial x_i} = \frac{dA}{d\mathcal{H}}\frac{\partial \mathcal{H}}{\partial x_i}, \tag{8.21}$$

and likewise for the derivative with respect to p_i, so the Poisson bracket $\{A, \mathcal{H}\}$ will evaluate to zero.

This same calculation shows that if the probability density ρ depends only on the value of the Hamiltonian, then ρ is a *steady-state* solution.

Let's say that Alice and Bob are two physicists contemplating the same physical system. They agree upon the Hamiltonian, but they have different hypotheses about how to locate the system within phase space. Alice writes a Liouville density that we denote ρ_A, and Bob writes a different function ρ_B. If there exists some overlap region for two Liouville densities ρ_A and ρ_B, some subset of phase space where both are nonzero, then Hamiltonian time evolution will carry both ρ_A and ρ_B forward in such a way that Alice and Bob's predictions about the future must overlap to the same extent that their hypotheses about the present do. Hamiltonian time evolution neither creates nor destroys agreement!

One can show that the relative entropy between ρ_A and ρ_B is constant over time:

$$S(\rho_A(t)||\rho_B(t)) = S(\rho_A(0)||\rho_B(0)). \tag{8.22}$$

This fact has interesting consequences. For example, suppose that Alice has a system of interest that she describes with a Liouville density ρ_1. She also has a supply of systems in the laboratory closet, and her description of one of these systems is a Liouville density ρ_0, encapsulating what she knows of how it was calibrated at the factory. Can she build a device that "clones" ρ_1 onto ρ_0, for an arbitrary ρ_1? A "cloning box" would be a machine that, using the rules of classical physics — in other words, using Hamiltonian dynamics — takes in systems described by ρ_1 and ρ_0 and emits systems described by ρ_1 and ρ_1.

If this "cloning box" is to be any good, it had better work for arbitrary inputs. That is, it should take in a pair of systems described by the joint Liouville density $\rho_1\rho_0$ and, by applying some Hamiltonian evolution for some amount of time, leave them described by the joint Liouville density $\rho_1\rho_1$. And it should *also* take $\rho_2\rho_0$ to $\rho_2\rho_2$. But the relative entropy of the initial Liouville densities is

$$S(\rho_1\rho_0||\rho_2\rho_0) = S(\rho_1||\rho_2) + S(\rho_0||\rho_0) = S(\rho_1||\rho_2), \tag{8.23}$$

while the relative entropy of the final Liouville densities is

$$S(\rho_1\rho_1||\rho_2\rho_2) = 2S(\rho_1||\rho_2). \tag{8.24}$$

Because Hamiltonian evolution conserves the relative entropy, no Hamiltonian evolution can effect this kind of transformation for arbitrary ρ_1 and ρ_2! This result is the *no-cloning theorem.* It was first discovered in quantum mechanics, where the analogues of Liouville densities are "quantum states"; only later was it recognized that the same logic applies in classical physics too. One can go into more depth, considering further variations on the idea of information-copying (including processes called "broadcasting"), but the essence of the argument, involving the preservation of an overlap measure, is what we have given here.

8.2 Simple Harmonic Motion

Whenever a potential energy function $V(\vec{x})$ has a local minimum, we can apply a Taylor-series approximation. And so very often, we find ourselves considering situations where *small perturbations* around a *mechanically stable equilibrium* are modeled using quadratic potential wells. The archetype is the simple harmonic oscillator with Hamiltonian

$$\mathcal{H}(x,p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \tag{8.25}$$

Applying Hamilton's equations, we get that

$$\dot{x} = \frac{p}{m},\ \dot{p} = -m\omega^2 x, \tag{8.26}$$

and taking another time derivative, we combine these to find

$$\ddot{x} = \frac{1}{m}\dot{p} = -\omega^2 x. \tag{8.27}$$

This is the simple harmonic oscillator equation of motion, and its solutions are linear combinations of sines and cosines.

Suppose that Alice has an oscillator that is initially at rest, vertically. At time $t = 0$, she applies a kick of unknown but small magnitude, imparting momentum to the oscillator. She writes a Liouville density to express the result,

$$\rho(x, p, t = 0) = f(p)\delta(x). \tag{8.28}$$

The characteristic featue of the simple harmonic oscillator is that all oscillations have the same natural frequency ω: Big swings take the same amount of time as little swings. In phase space, the oscillator executes "orbits" around the origin. Orbits further out have higher energy, but all orbits have the same period, $T = 2\pi/\omega$. So, forecasting the Liouville density forward one quarter-cycle, all of it must lie along the x axis. An initial density that was spread out along p and focused in x becomes focused in p and spread out along x instead. The distribution repeatedly trades off where the uncertainty lies as it cycles around the origin. This implies that the usefulness of any particular measurement depends upon when in the cycle that measurement is made: Initially, Alice could reduce her uncertainty by measuring the momentum, but a momentum measurement at time $T/4$ or $3T/4$ would not be informative.

If the potential well is *anharmonic,* rather than being a nice quadratic curve, then the oscillation frequency will differ between orbits, and so a "blob" in phase space may have some portions cycle around the origin faster than others.

8.3 Chaos and Mixing

Hamiltonian dynamics can exhibit *chaos,* which is roughly defined as "sensitive dependence upon initial conditions". A chaotic system, so the saying goes, has the properties that points initially close to each other in phase space become farther apart exponentially quickly. We have to be a little careful with this "butterfly effect" intuition: The curves $x = e^t$ and $x = 1.001e^t$ grow exponentially further apart, but this kind of divergence is not so very interesting. If we provisionally designate "chaos theory" as the study of those systems featured in books that have "chaos" in the title, we can sharpen our heuristic definition slightly. A chaotic system, for our purposes, is one for which phase-space trajectories diverge rapidly but remain within a bounded volume. For at least some time interval, the divergence is exponential, with a distance measure growing as $e^{\lambda t}$, the parameter λ being known as a *Lyapunov exponent.*

Now, take a Liouville density ρ. How will it evolve with time if the underlying dynamics are chaotic? We can imagine that ρ has a well-defined edge, within which it is nonzero. Focus on two neighboring points on the edge. The chaotic dynamics will carry them farther and farther apart, and so the length of the edge between them should be growing, but by Liouville's theorem, the *area* of the nonzero region of ρ must be constant. So, we have a blob of fixed area, whose edge is stretching, folding, becoming intricately layered.

If we forecast our Liouville density ρ far enough into the future, its image $\rho(T)$ under the chaotic dynamics will have protrusions all across phase space. Any *coarse-grained*

measurement, one that is only roughly sensitive to the phase-space coordinates, will pick up a portion of $\rho(T)$, because every chunk of phase space has a little of $\rho(T)$ in it. For the purposes of predicting the outcomes of such imprecise measurements, we could replace $\rho(T)$ with a "blurred" version thereof. In many cases, this blurring would eliminate correlations between particles and drastically simplify how the Liouville density depends upon the coordinates. The blurred Liouville density is likely to be a function of the Hamiltonian, and thus a steady-state density.

8.4 Reversibility?

Before we get too deep into many-particle dynamics, let us consider a scenario involving only $N = 2$ of them. Alice arranges a *collision* between two bodies of equal mass m. They come in with momenta $\vec{p}_1$ and $\vec{p}_2$, approach one another and interact near the point $\vec{x}$, and then they fly apart again with momenta $\vec{p}\,'_1$ and $\vec{p}\,'_2$. Conservation of momentum implies that

$$\vec{p}_1 + \vec{p}_2 = \vec{p}\,'_1 + \vec{p}\,'_2, \tag{8.29}$$

and if the collision is elastic, then conservation of energy says that

$$p_1^2 + p_2^2 = (p'_1)^2 + (p'_2)^2. \tag{8.30}$$

The exact directions and magnitudes of the outgoing momenta will be some function, perhaps a complicated one, of the incoming momenta and the details of the collision process. We defer consideration of those details for the moment and focus on the conservation laws, which tell us what the *valid* combinations of quantities are.

Alice sets up the collision and films it happening: the particles approaching, interacting and separating again. She calculates the momenta from her recording and verifies that the conservation laws are respected. Then she discovers that her video software has the ability to play the film *backwards.* On a whim, she tries it out, and she calculates the "incoming" and "outgoing" momenta from the reversed film. These vectors *also* satisfy the conservation laws, so they too describe a physically legitimate process. For every valid set of motions, the time-reversal of that set is also valid.

Physicists with a philosophical bent have made much of this. If the microscopic interactions in nature are time-reversible like this, how could it be that they add up to the macroscopic world, with all its irreversibilities — teacups breaking and not reassembling, gases diffusing but never re-concentrating themselves? Surely there is some deep paradox here. Whence time, from laws that are timeless? I myself am never too sure how much to make of these arguments. For one thing, whenever people try to explain what they *mean* by a statement like "the time-reversal of a solution to the equations of motion is also a solution", they tell a story like ours about Alice filming a collision. And in that story, there is definitely a notion of time: Alice goes from less experienced to more so. The moving finger writes, and having writ, moves on — we have tried to explain the supposed timelessness of our physical laws with a parable that already includes the ideas of cause and effect, which seems a shady maneuver. Nor is the mere *mathematical* fact that we can exchange inputs and outputs

too helpful in itself. I carry roughly 25% of Grandpa Stacey's genome, and Grandpa Stacey carried roughly 25% of mine. The probabilities of finding genes in common are time-symmetric, and yet the succession of human generations is quite irreversible!

Physicists like to divide our mathematical descriptions of natural phenomena into *initial* or *boundary conditions* and *dynamical laws.* For example, we have been writing an initial state of knowledge about a system as a Liouville density, and then rolling forward that state of knowledge using a Hamiltonian. The Liouville density and the Hamiltonian are given conceptually different statuses, despite both being the same type of mathematical objects, real-valued functions on phase space. This is a useful way to live one's life when doing weekday physics, but when we get around to the "whence time?!" kind of questions, it is worth questioning that distinction. The astronomers have discovered that the Universe is expanding, and more quickly with each passing day. If we want our laws of physics to be universal, it is hard to find a statement that qualifies better — so, given the confessed incompleteness of modern physics, it is conceivable that we should make room for time-irreversibility as a matter of physical law itself.

Going back to Alice, let us amend our story. Alice runs the collision process forward, and then along comes Bob, who tries to physically reverse it. No trouble, thinks Bob: He'll just place a bumper, an elastically reflecting wall, in the path of each outgoing particle. The particle will bounce back with its momentum inverted, and the collision will then run in reverse.

But wait, Alice realizes! If Bob puts his "mirrors" in the wrong places along the outgoing trajectories, then the bodies will collide with them at *different times.* Particle 1 might hit first and go sailing back with momentum $-\vec{p}_1'$, passing through the original point of collision before particle 2 hits the reflector that Bob put down for it, so the "reversed" particles never actually collide!

Moreover, if Bob does not *align* his reflecting walls properly, then the particles won't go back down their original paths. The reflected momentum could be anything, if Bob is sufficiently sloppy — any vector with the same magnitude. And if Bob hasn't worked to hold his reflectors absolutely still, their motion can contribute energy to the scattered particles, so even the conservation laws won't be helpful. Alice, having no reason to trust that Bob won't muck it all up, has to write a new Liouville density. She calls her original description of her setup $\rho_A(\vec{x}_1, \vec{x}_2, \vec{p}_1, \vec{p}_2, t = 0)$. By ordinary Liouville evolution, she calculates her forecast for a later time $t = T$, well after the collision. This forecast she expresses as a Liouville density $\rho_A(\vec{x}_1, \vec{x}_2, \vec{p}_1, \vec{p}_2, t = T)$. So far, so Hamiltonian. But now along comes Bob, ruining everything. Alice figures that his disruption happened before time T, so by time T, the best she can do for describing the two-particle system is

$$\rho_B(\vec{x}_1, \vec{x}_2, \vec{p}_1, \vec{p}_2, t = T) = \rho_B(\vec{x}_1, \vec{p}_1, t = T)\rho_B(\vec{x}_2, \vec{p}_2, t = T). \tag{8.31}$$

This Liouville density *factors*, because Alice has no reason to think that Bob's intervention has preserved any correlation between the two particles. Knowing the momentum of one no longer helps predict the momentum of the other.

Alice can, computationally, time-evolve in reverse. She can take her $\rho_A(t = T)$ and work backwards to find $\rho_A(t = 0)$, using Liouville's equation. She can do the same

for $\rho_B(t=T)$, finding a $\rho_B(t=0)$ which is the Liouville density that *would evolve* to $\rho_B(t=T)$ given the original dynamical laws. But Liouville's equation preserves the relative entropy:

$$S(\rho_A(t=T)||\rho_B(t=T)) = S(\rho_A(t=0)||\rho_B(t=0)). \tag{8.32}$$

Consequently, we see that the time-reversal of the system after Bob has disrupted it will probably have very little to do with how Alice originally set it up!

We can even see this come into play with a one-particle system. Suppose Bob is trying to reverse the motion of a simple harmonic oscillator by quickly placing a reflecting wall in a good spot and then removing it so as not to interfere further. But Bob measures the oscillator at one of those times when a position observation is not very informative, so he ends up making a pretty wild guess as to where to locate his reflector. The oscillating particle *might* hit Bob's reflector and have its motion reversed, or it might not. Rather than neatly flipping p to $-p$ and reflecting the whole Liouville density across the x axis, he only manages to do so for *part* of it. One blob in phase space becomes two, and a nonzero relative entropy is generated between the probability density that he wished to achieve and the one he actually managed.

8.5 From Liouville to Boltzmann

The state of each particle can be specified by its position $\vec{q}$ and momentum $\vec{p}$, which together make a six-dimensional phase space. Given a total of N particles, we can therefore describe the entire swarm in $6N$ dimensions. Typically, we define a *phase space density*

$$\rho(\vec{p}_1, \vec{q}_1, \vec{p}_2, \vec{q}_2, \ldots, \vec{p}_N, \vec{q}_{N_e}) \tag{8.33}$$

as the probability that the swarm will be found in the small region of phase space around the point labeled by the given p and q coordinates [74]. Using $d\Gamma$ to denote an infinitesimal portion of the $6N$-dimensional phase space Γ, we can define an expectation value for any function $\mathcal{O}$ which depends upon the corpuscle positions and momenta:

$$\langle \mathcal{O} \rangle = \int \rho(\{\vec{p}_i, \vec{q}_i\}) \mathcal{O}(\{\vec{p}_i, \vec{q}_i\}) d\Gamma. \tag{8.34}$$

This description, however, provides too much information. Were we to compute the pressure, for example, that the particle swarm exerts on an object, we would only need to know how likely we are to find *any* of the N particles impacting the object with a particular velocity. In other words, the quantity of physical interest is the *one-particle distribution,* the probability of finding any particle at location $\vec{q}$ with momentum $\vec{p}$. As this quantity may well change over time, we must also consider its dependence upon t. The one-particle distribution, ρ_{I}, is defined in terms of the total probability density ρ:

$$\rho_{\mathrm{I}}(\vec{p}, \vec{q}, t) = \int \prod_{i=2}^{N} d^3\vec{p}_i d^3\vec{q}_i \rho(\vec{p}_1 = \vec{p}, \vec{q}_1 = \vec{q}, \vec{p}_2, \vec{q}_2, \ldots, \vec{p}_{N_e}, \vec{q}_{N_e}, t). \tag{8.35}$$

By integrating over particles numbered 2 through N, we are making the assumption that the probability density ρ is symmetric under exchanges of any two particles. This is the first of several simplifying assumptions that are traditional to kinetic theory, which will get more dramatic as we go along. It is not always applicable. For example, if we have a box divided into left and right halves by a partition, and we put the first 10^{23} atoms on the left and the second 10^{23} atoms on the right, then the probability that atom 1 will impact the left-hand wall is *not the same* as the probability that atom $10^{23}+1$ will do so. Simplifying by symmetrizing forsakes the ability to describe situations where the components of our system come in "communities".

We may similarly define distribution functions involving more particles, ρ_{II},ρ_{III}, and so on. Studying the time evolution of these distributions leads to the *BBGKY hierarchy,* named for Bogoliubov, Born, Green, Kirkwood and Yvon. This hierarchy is an interminably long series of equations in which the time derivative of ρ_{I} depends upon ρ_{II}, the derivative of ρ_{II} depends upon ρ_{III} and so on. The basic idea is that if we follow one particle going along, its momentum may change due to an interaction with another particle, and in order to guess whether such an interaction will happen and change the one-particle distribution, we need to know the two-particle distribution. Likewise, to understand how the two-particle distribution can change, we follow a pair of particles, either of which can interact with any of the particles in the background, and in order to tell when these interactions might happen, we need to know the three-body distribution function, and so on. Simplifying this hierarchy to a manageable equation will be the topic to which we turn next.

The basic idea is that we will again have streaming terms and collision terms. This time, the streaming flow will be expressed by a Poisson bracket, as in the Liouville equation (8.18):

$$\frac{\partial \rho_s}{\partial t} = -\{\rho_s, \mathcal{H}_s\} + (\text{collision terms}). \tag{8.36}$$

Here, we have introduced the notation $\mathcal{H}_s$ to stand for that part of the original many-body Hamiltonian including only the interactions within the first s particles:

$$\mathcal{H} = \mathcal{H}_s + \mathcal{H}_{N-s} + \mathcal{H}'. \tag{8.37}$$

The collision terms will involve an integral for each particle from 1 through s, capturing how that particle might interact with number s+1. These interactions change momenta but not positions, so we will need the *force* between particle n and particle $s+1$. Traditionally, we write this force as minus the gradient of a potential energy $\mathcal{V}$, because the habit of statistical physics is to set things up in terms of energy. The result is

$$\frac{\partial \rho_s}{\partial t} = -\{\rho_s, \mathcal{H}_s\} + (N-s)\sum_{n=1}^{s}\int d^3\vec{p}_{s+1}d^3\vec{q}_{s+1}\frac{\partial \mathcal{V}(\vec{q}_n - \vec{q}_{s+1})}{\partial \vec{q}_n}\cdot\frac{\partial \rho_{s+1}}{\partial \vec{p}_s}. \tag{8.38}$$

The prefactor $N-s$ comes from the symmetry considerations we mentioned earlier. The force between particles enters the same way the external forces did in the Boltzmann equation (7.83), as a dot product with a derivative with respect to momentum. Reaching this point involves taking the Liouville equation for ρ, splitting up the

Hamiltonian as above, integrating over the last $N - s$ particles to get a time-evolution equation for ρ_s and then manipulating that equation through integration by parts.

We can recover an expression with the same form as the Boltzmann equation (7.68) by positing that the joint probability density factors into independent single-particle contributions. Justifying this assumption requires some care. Suppose that the probability density for time t does factor as desired. Then, at a later time t', the particles may have interacted, and measuring one of them can be informative about what another one is doing. So, it appears that the factorization will generally not be maintained under time evolution. Recall that we motivated the Boltzmann equation by considering the *most probable change* of the *most probable number density*. Another way of saying this is that to arrive at the Boltzmann equation, we neglected events of smaller probability. Here, we are formulating that neglect in an alternative way.

9 Proto-Quantum Mechanics

In what follows, we will use bits and pieces of history. We won't concern ourselves with the standard story, or with trying to recapitulate how the development historically unfolded. (The first is dull, and the second is frightfully complicated.) Instead, we will take what physicists had learned by 1924 or so, and play with those ideas to see what *might* have happened.

9.1 Bohr model

Introduced by Niels Bohr in 1913, this model typifies the "old quantum theory", the style of thinking during the 1900–1925 period in which people tried to understand atomic and molecular phenomena by using classical physics and adding some constraints. Indeed, its empirical success makes it one of the high points of that period. Bohr was able to calculate the locations of the spectral lines for atomic hydrogen, in terms of numbers that could be measured from other situations. The essence of the Bohr model is that a lightweight, negatively charged electron orbits a much heavier nucleus whose charge exactly balances that of the electron, and the orbits in which the electron is allowed to exist are those with discretized values of angular momentum:

$$L = n\hbar. \tag{9.1}$$

The constant $\hbar$ could be taken from measurements of the blackbody radiation curve or the ejection of electrons from metal by ultraviolet light. That it is also meaningful for atomic hydrogen says something about nature.

Sometimes we say that in the Bohr model, angular momentum is "quantized"; but this does not mean that the full theory of quantum mechanics is applied to it. The Bohr model is semiclassical, not fully quantum. In it, angular momentum is "quantized" in the same sense that doughnuts are "quantized" because the bakery won't sell half a doughnut.

The force holding the electron in orbit is the inverse-square Coulomb force between two charges of equal magnitude and opposite sign:

$$F = ma = \frac{D}{r^2}, \tag{9.2}$$

where D is q_e^2 in cgs units or $q_e^2/(4\pi\epsilon_0)$ in SI. Since the orbit is circular, the acceleration is v^2/r, and so the kinetic energy is

$$K = \frac{1}{2}mv^2 = \frac{1}{2}mar = \frac{D}{2r}. \tag{9.3}$$

Alternatively, we can write the kinetic energy in rotational form, in terms of the angular momentum and the moment of inertia:

$$K = \frac{L^2}{2I} = \frac{L^2}{2mr^2}. \tag{9.4}$$

In this form, we can plug in Bohr's discretization condition:

$$K = \frac{n^2\hbar^2}{2mr^2}. \tag{9.5}$$

We can equate our expressions for K and solve for the orbit radius, obtaining

$$r = \frac{n^2\hbar^2}{mD}, \tag{9.6}$$

and since the total energy is the sum of kinetic and potential terms, we can readily compute

$$E_n = -\frac{D^2 m}{2\hbar}\frac{1}{n^2}. \tag{9.7}$$

De Broglie found a condition from which the discretization rule can be derived. From Huygens, Young and Maxwell, we learned that light is a wave, but Planck and Einstein taught us that it also particulate aspect to its behavior. So, turning that on its head, why not suppose that electrons, which Thomson found to be particles, also have an associated wave of some kind? And if we want an energy level to be stable, a mode of behavior that can keep doing its own thing, then we should try thinking of it as a *standing wave.* What we want is a wave that wraps evenly around a circular orbit a whole number of times, so that it constructively interferes with itself. If we say that the wavelength is inversely proportional to the momentum,

$$\lambda = \frac{2\pi\hbar}{p}, \tag{9.8}$$

then the orbital circumference must be an integer multiple of this:

$$C = n\lambda \Rightarrow p = \frac{2\pi\hbar}{C/n} = \frac{2\pi n\hbar}{2\pi r} = \frac{n\hbar}{r}. \tag{9.9}$$

We rearrange this into the Bohr condition by noting that $L = pr$.

The empirical success of the Bohr model, when it comes to atomic hydrogen, leads us to take seriously the idea that *energy levels* are *stationary states.* Let's hold on to that concept: Whatever the equations of quantum mechanics are, solutions to them that correspond to energy levels do not change over time, or at most, change in such a way that the statistical predictions we extract from them remain constant.

While the Bohr model predicted the location of spectral lines, it said nothing about their *brightness.* Making progress on this front, it seemed, required a better understanding of transition rates: the faster the rate, the more photons coming out and

the brighter the line. Inamongst the general fumbling came the idea that the electron's orbit ought to be periodic, so we should break that orbit down into a Fourier series and work with the coefficients in the Fourier expansion. By late 1917 or early 1918, Bohr was suggesting that these Fourier coefficients somehow specified transition probabilities: "[T]he amplitude of the harmonic vibrations for a given value of τ will in some way give a measure for the probability of a transition between two states for which $n' - n''$ is equal to τ" [83]. In 1924, Max Born studied perturbations of a classical system, writing a perturbation as a Fourier series with coefficients $\{C_\tau\}$, and sought the corresponding quantum formulae. He decided that the $\{C_\tau\}$ didn't have "a quantum-theoretical meaning", but the squares of their absolute values did [84].

9.2 How to Invent POVMs by Christmas 1925

Born's book *Problems of Atomic Dynamics* collects a series of lectures that he delivered at MIT from 14 November 1925 through 22 January 1926. I did not know that Born had ever lectured at MIT! Nor did I appreciate the facility with which he jumped to self-adjoint matrices. For him, that is the natural extension of the condition that the Fourier transform of a sequence be real-valued ($C_{-r} = C_r^*$). He's also very much in the mindset that the position and the velocity of an electron are not observable. He wants to carry over the form of classical Hamilton (and Hamilton–Jacobi) equations as much as possible, but with the unobservable position and momentum replaced by time-dependent matrices. He talks about taking time derivatives by evaluating the commutator with the Hamiltonian.

"So," an audience member could have asked, "what about the quantum version of Liouville theory? Is there a matrix for the Liouville density whose time derivative is its commutator with the Hamiltonian? Because if position and momentum aren't accessible classically, we can represent the outcomes of the measurements that we can do by conditional probability densities, and calculate expectation values of observable quantities by integrating over the Liouville density function."

"Oh," Born might have said. "That's right. Dirac has just recently told us to replace the classical Poisson bracket with the quantum commutator, so the quantum Liouville matrix would be something that we commute into the Hamiltonian to find its time evolution. And instead of integrating over the Liouville density, we would take the inner product between matrices, because the phase-space integral defined an inner product of density functions and we want to preserve that algebraic structure."

"I guess that means each outcome of a measurement is a conditional probability matrix. Well, not a matrix of conditional probabilities; I mean a Hermitian matrix that plays the role of a conditional probability density on phase space."

"What kind of properties will those matrices have to satisfy?"

"Well... they'd have to have nonnegative inner products with all possible Liouville density matrices."

"And any one of them could be a Liouville density matrix itself, up to scaling, so all the density matrices have to have nonnegative inner products with each other. Oh," Born says as another idea occurs, "if you take the principal axes of any Hamiltonian and

find the matrices that project onto them, they will have nonnegative inner products with the matrices that project onto the principal axes of any perturbation of that Hamiltonian. That feels like it ought to mean something. Wait, wait... that inner product between matrices will be the square of the absolute value of the inner product between the unit vectors along the principal axes. Which fits neatly with the idea that a transition rate is the square of the absolute value of a Fourier coefficient!"

9.3 The Born Rule

In this little dialogue, our imaginary Born and his imaginary interlocutor have speculated their way to the hypothesis that the way to find the probability of an event occurring in quantum mechanics is to take the trace of the product of two matrices. One matrix, which we can write as ρ, is the counterpart of the classical Liouville density function. The other, which we can denote E, plays the role of a conditional probability density. Instead of integrating to find a probability,

$$P(E) = \int dx\, dp\, P(E|x,p)P(x,p)\,, \tag{9.10}$$

we take a different kind of inner product, the *Hilbert–Schmidt* inner product:

$$P(E) = \mathrm{tr}(\rho E)\,. \tag{9.11}$$

Our fictional characters have realized that for this to be consistent, the set of valid E matrices has to have the property that the inner product of any two matrices in it must be nonnegative.

Born notices a special case that strikes him as particularly intriguing. There is a class of matrices that project onto vectors. Given a unit vector $\hat{v}$, he can write the matrix whose entries are

$$[M]_{jk} = v_j^* v_k\,. \tag{9.12}$$

Multiplying any vector by the matrix M takes the part of that vector which is along the direction of $\hat{v}$ and throws away the rest. If ρ and E are both matrices of this type, projecting onto the unit vectors $\hat{u}$ and $\hat{w}$, then the probability $P(E)$ works out to be the square of the absolute value of the inner product between $\hat{u}$ and $\hat{w}$.

$$P(E) = \left| \sum_j u_j^* w_j \right|^2 . \tag{9.13}$$

10 Basic Quantum Mechanics

10.1 Introduction

Before I took statistical physics at the level we are doing in this text, I had rather a lot of quantum mechanics — three whole semesters, in fact. Yet this left me a little unprepared for the quantum physics we use in this subject, and it's worth understanding why. One reason is that fully one third of my undergrad quantum was pretty worthless. They shied away from introducing new mathematics (the story I heard was that the department was afraid they were losing physics majors in the sophomore year), and so they spent a term on "intuition-building" that didn't build much intuition, and "history" that was not very historical. It is common to waste a lot of time in college, though I'd prefer not to bake that into the curriculum itself! Another reason is that my undergrad courses did what in broad strokes we might call "quantum mechanics at zero temperature". They focused almost entirely on situations where the physicist has a level of uncertainty that is the *minimum allowed by quantum theory.* To do statistical physics, we will have to work in more generality. Mathematically, this will mean handling *matrices* filled with complex numbers, whereas undergrad quantum specializes to the case where those matrices can each be written in terms of a single *vector.*

When a student is learning quantum mechanics, sooner or later they must encounter a heap of mathematics. A typical exposition of the algebra might run to dozens of pages, with the physical meaning and the motivation left remarkably unclear. We do not do this out of cruelty — well, not *only* out of cruelty — but in large part because we ourselves do not know any better. It is very easy to pick a statement in the exposition and ask *why,* e.g., "Why does quantum theory use complex numbers?" And very often, the chain of answers to that *why* question runs into ignorance after only a few steps. In another course, we might be able to take the time to work on this, but for present purposes, we will have to take a lot of the mental machinery as given, and we will simply learn to *apply* it.

Now, there are alternatives to putting the "mathematical prerequisites" chapter right at the beginning. We could, for example, try to teach the history of the subject first — names, dates, experiments. There are certainly colorful personalities and character drama in that history, but the algebra is not so neatly strained out. It is often difficult to appreciate why the early pioneers did what they did without having more background knowledge than we can rightfully ask of our students. For example, Pauli found the correct quantum-mechanical energy levels of the hydrogen atom *before* Schrödinger wrote down his equation, by very cannily inventing an analogy to a concept in advanced classical mechanics. This is not the way that any introductory course

ever does the hydrogen atom! So, the "history" in a physics course always involves a measure of myth-making. This is quite a feat for us, because we simultaneously pride ourselves on our solid rationality!

We *could* try another path and attempt to develop the theory using only the mathematics that we presume the students already know. This is what happened to me. I wouldn't recommend it. Life is too short to turn what should be an interesting subject into a semester of plodding through solutions to the same differential equation day after tedious day until they all blur together.

So, we are going to confront the mathematics rather head-on. But first, let us get a sense of what we are aiming for.

10.2 The Parable of the Muffins

Let's try to make a profound statement about reality by thinking hard about baked goods.

I promise this is going somewhere.

A certain bakery has a special deal on muffins. They sell mystery boxes for those who like to live dangerously: mix-and-match sets of three muffins apiece. Each day, Alice, Bob and Charlie buy a mystery box together, and each day, Alice, Bob and Charlie take one muffin apiece back to their respective laboratories for analysis. They each have two testing devices — say, a device that can test whether a muffin is positive for dairy, and another device that can test whether it is positive for tree nuts. We'll call these X tests and Y tests for short. Each day, Alice chooses either to do an X test or a Y test. Bob likewise chooses, independently of Alice, and so does Charlie. Importantly, each muffin can only be tested *once.* Maybe the test destroys the muffin, or maybe it takes so long to do one that they eat their muffins immediately afterward. Whatever the rationale, one test per muffin — that's a rule of the parable.

We can write what they choose to do in a compact way. For example, if all three of them choose to do the X test on their respective muffins, we'll write $A_X B_X C_X$. If Bob and Charlie choose to do the Y test but Alice goes instead with the X test, we'll write $A_X B_Y C_Y$. And so on. We can also write the results compactly, using $+1$ to stand for a positive result and -1 to stand for a negative one. (We could also record the outcomes with zeros and ones, or with trues and falses, greens and blues, etc. Using $+1$ and -1 is just a notation that will turn out to be helpful in a moment.) So, for example, if Alice chooses Y, Bob chooses X and Charlie goes with Y, the results might be $(+1, -1, -1)$. Or they might be $(+1, +1, +1)$, or perhaps $(-1, +1, -1)$.

Over many days of muffin investigation, comparing their notes, they find a dependable pattern. Whenever *two of them* choose to do the Y test, then the *product of their results* is always $+1$. The specific outcome varies randomly from day to day, but there's never only one -1, and they never get all three results being -1. From this pattern, they can draw a couple conclusions. First, once two of them obtain their results, the result of the third is predictable. Let's say their choices are $A_Y B_X C_Y$, as in the previous example, and both Alice and Bob get the result $+1$. Then we can predict that Charlie will get $+1$, because that's the only way the product of the three numbers

can be +1. Or, suppose their choices are $A_Y B_Y C_X$, and both Bob and Charlie get a −1. The two of them report their results and wait for Alice. Knowing Bob and Charlie's results, we can predict that Alice will report a +1 outcome, because that's the only way the product of the three outcomes is +1. A minus times a minus makes a plus, and so a third minus would spoil the plus.

If we had used a different notation for the outcomes, like "green" and "blue" instead of +1 and −1, then we could express this pattern by saying that whenever two of them choose to do the Y test, an even number of the results will be blue.

Now, we deduce something else from the pattern. We can make a prediction about what happens under very different conditions. What about the days when all three choose to measure X?

We've used capital letters to write their choices. Let's use lowercase letters to denote the *properties* of the muffins being measured. An A_X measurement — that is, Alice doing the X test — uncovers the value of a_X. Likewise, Bob doing the Y test reveals the b_Y muffin property, and so on. The pattern so far is that

$$a_X b_Y c_Y = +1\,, \tag{10.1}$$

and

$$a_Y b_X c_Y = +1\,, \tag{10.2}$$

and

$$a_Y b_Y c_X = +1\,. \tag{10.3}$$

Here comes the neat trick. We multiply all three of these facts together.

$$(a_X b_Y c_Y)(a_Y b_X c_Y)(a_Y b_Y c_X) = +1\,. \tag{10.4}$$

There's no meaning to multiplying capital letters in this thought-experiment, because only one choice of test can be made each day. A "product" like $A_X A_Y$ is just nonsense notation. But the lowercase letters, *those* we can multiply, because they stand for properties which are there whether or not anyone measures them. We rearrange the variables, putting like with like:

$$a_X b_X c_X a_Y^2 b_Y^2 c_Y^2 = +1\,. \tag{10.5}$$

On any given day, we don't know the value of a_Y or b_Y or c_Y before the test, and if Alice chooses X instead of Y, then we'll never learn a_Y for that day, and likewise for Bob and Charlie. But the values do have to be waiting there, don't they? Each has to be either +1 or −1, even if we never learn which: That's just a fact determined at the bakery.

The square of −1 is the same as the square of +1, so on every day,

$$a_Y^2 = b_Y^2 = c_Y^2 = +1\,. \tag{10.6}$$

Therefore, we can drop all the Y factors from our previous equation!

$$a_X b_X c_X = +1\,. \tag{10.7}$$

And now we have a prediction: on those days when all three independently choose to do the X test, the product of their answers will be $+1$.

This follows from the dependability of the pattern found on the two-Y-test days and the basic assumption that the tests are testing properties intrinsic to the muffins, properties waiting to be found. The tests don't have to be for dairy and tree nuts — any properties will do. The objects being tested don't have to be muffins, either. Anything that can come in sets of three is suitable. From the dependable pattern in the two-Y-test events, we can draw a conclusion about the triple-X-test events. The product of the X results is going to be $+1$.

How general this result is! From the one pattern, we deduce the other, whether we find that first pattern in muffins, cookies, apples, sand grains...

Electrons? Photons?

Ay, there's the rub.

It is possible to make triplets, not of baked goods but of subatomic particles, that fit the two-Y-test pattern. After doing the Y test on the first two particles, for example, the result of an X test is predictable, using the same rule we described above.

But the prediction about what happens when each particle gets the X test does not hold.

The rules of quantum mechanics imply that there is a way an experimenter can prepare triplets of particles such that they should predict that the product of an X result and two Y results will always be $+1$, while the product of three X results will always be -1.

By rights, this ought to be impossible. But the imagination of nature is more subtle than our own!

Doing the calculation to get the quantum-mechanical answer is not all that hard. It's a spot of matrix algebra that's doable by hand and doesn't require knowing much more than what an "eigenvector" is. (We will develop the details that we need soon.) The much more difficult part is knowing what to make of it!

The predictions of quantum mechanics — which have been checked in the lab, directly and indirectly, and found to work superbly — clash with the basic and seemingly bulletproof calculation we have done here. How can that calculation fail to apply? Where is the gap that means one pattern doesn't have to imply the other? If an argument whose fundamental premise is that *measurements record a property of the thing being measured* is inconsistent with quantum physics, our spectacularly successful guide to living in reality, then what does that say about reality?

Well, ask two rabbis and you'll get three opinions. It's not even clear what making progress on that kind of "what does it all mean?!" question even looks like. If people disagree on the meaning but all use the same math to make the same predictions, on what grounds can we even prefer one proposal about the meaning over another? Gut reaction? Preserving intuitions from classical physics? Remember, classical physics is what turned out to be wrong, so we had to invent quantum mechanics to fix it!

Perhaps one way to keep the question from going stale and vacuous is to find a new way of expressing quantum theory itself. After all, there's no grand guarantee that the formulation of the theory which is good for finding the spectrum of atomic hydrogen (and all the other topics we traditionally assign to undergraduates each semester) is

equally good for all questions. And one feature of the standard presentation is that the really weird stuff, such as what we've confronted today, tends to be buried several chapters in. When the books lay out "axioms" or "postulates" for quantum mechanics, blatant defiance of intuition about intrinsic properties isn't among them. Instead, it is deduced as a consequence, deep into the algebra. But perhaps that is a historical accident, due to the way we got here and not to nature itself!

The parable of the muffins is my attempt at rephrasing a thought-experiment by N. David Mermin, who learned of a *four*-particle scenario devised by Greenberger, Horne and Zeilinger and found that it could be significantly sharpened by going down to three.

10.3 A Single Qubit

Now and then, stories will pop up in the news about the latest hot new thing in quantum computers. If the story makes any attempt to explain why quantum computing is special or interesting, it often recycles a remark along the lines of, "A quantum bit can be both 0 and 1 simultaneously." This is rather like saying that Boston is at both the North Pole and the South Pole simultaneously. Something important has been lost. Our goal in this introductory section is to develop a mental picture for a *qubit,* the basic unit that quantum computers are typically regarded as built out of. To be more precise, we will develop a mental picture for the *mathematics* of a qubit, not for how to implement one in the lab. There are many ways to do so, and getting into the details of any one method would, for our purposes today, be a distraction. Instead, we will be brave and face the issue on a more abstract level.

In the classical study of information, a *bit* is any system that has two unambiguously different modes of being, like "on" and "off", "true" and "false", "0" and "1", "yes" or "no". Digital computers are made of stupendous numbers of switches, each of which can be modeled as a bit. We can speak of the number of bits of information that message conveys, by which we mean the number of yes-or-no questions that one must pose in order to uniquely identify that message out of all the possible messages.

Imagine a light switch in a closed room. The switch can be either on or off. Because I cannot see into the room, I do not know whether it *is* on or off, but I can hazard a guess. If I am mostly confident that it is off, I might (like Alice in the Dutch-book story) put a number to that confidence level and make a gambling commitment, saying that my probability for the light being on is only, e.g., 22%. Then, of course, my probability for the light being off must be 1 minus that, or 78%. The set from which I can draw my value of $p(\text{on})$ is the interval from 0 to 1, inclusive. Picking $p(\text{on}) = 0$ means expressing absolute confidence that the light is off, and picking $p(\text{on}) = 1$ is declaring bet-it-all certainty that the light is on. We can say that my *state space* for this bit is the unit interval $[0, 1]$. I can resolve my uncertainty about this bit by getting an answer to one single question — I poke my head in and check whether the light is on or off.

This example illustrates the salient features of a classical bit: The state space is the interval $[0, 1]$, the uncertainty is due to ignorance about a pre-existing physical

situation, and resolving that uncertainty takes one answer to one binary question.

Now, let's turn from classical bits to qubits. A qubit is a thing that one *prepares* and that one *measures.* The mathematics of quantum theory tells us how to represent these actions algebraically. That is, it describes the set of all possible preparations, the set of all possible measurements, and how to compute the probability of getting a particular result from a chosen measurement given a particular preparation. To do something interesting, one would typically work with multiple qubits together, but we will start with a single one. And we will begin with the simplest kind of measurement, the *binary* ones. A binary test has two possible outcomes, which we can represent as 0 or 1, "plus" or "minus", "ping" and "pong", et cetera. In the lab, this could be sending an atom through a magnetic field and registering whether it swerved up or down; or, it could be sending a blip of light through a polarizing filter turned at a certain angle and registering whether there is or is not a flash. Or any of many other possibilities! The important thing is that there are two outcomes that we can clearly distinguish from each other.

For any physical implementation of a qubit, there are three binary measurements of special interest, which we can call the X test, the Y test and the Z test. Let us denote the possible outcomes of each test by $+1$ and -1, which turns out to be a convenient choice. The *expected value* of the X test is the average of these two possibilities, weighted by the probability of each. If we write $P(+1|X)$ for the probability of getting the $+1$ outcome given that we do the X test, and likewise for $P(-1|X)$, then this expected value is

$$x = P(+1|X) \cdot (+1) + P(-1|X) \cdot (-1)\,. \tag{10.8}$$

Because this is a weighted average of $+1$ and -1, it will always be somewhere in that interval. If for example we are completely confident that an X test will return the outcome $+1$, then $x = 1$. If instead we lay even odds on the two possible outcomes, then $x = 0$. Likewise,

$$y = P(+1|Y) \cdot (+1) + P(-1|Y) \cdot (-1)\,, \tag{10.9}$$

and

$$z = P(+1|Z) \cdot (+1) + P(-1|Z) \cdot (-1)\,. \tag{10.10}$$

To specify the preparation of a single qubit, all we have to do is pick a value for x, a value for y and a value for z. But not all combinations (x, y, z) are physically allowed. The valid preparations are those for which the point (x, y, z) lies on or inside the ball of radius 1 centered at the origin:

$$x^2 + y^2 + z^2 \leq 1\,. \tag{10.11}$$

We call this the *Bloch ball,* after the physicist Felix Bloch (1905–1983). The surface of the Bloch ball, at the distance exactly 1 from the origin, is the *Bloch sphere.* The points where the axes intersect the Bloch sphere — $(1, 0, 0)$, $(-1, 0, 0)$, $(0, 1, 0)$ and so forth — are the preparations where we are perfectly confident in the outcome of one of our three tests. Points in the interior of the ball, not on the surface, imply uncertainty about the outcomes of all three tests. But look what happens: If Alice is perfectly

confident of what will happen should she choose to do an X test, then her expected values y and z must both be zero, meaning that she is *completely uncertain* about what might happen should she choose to do either a Y test or a Z test. There is an inevitable tradeoff between levels of uncertainty, baked into the shape of the theory itself.

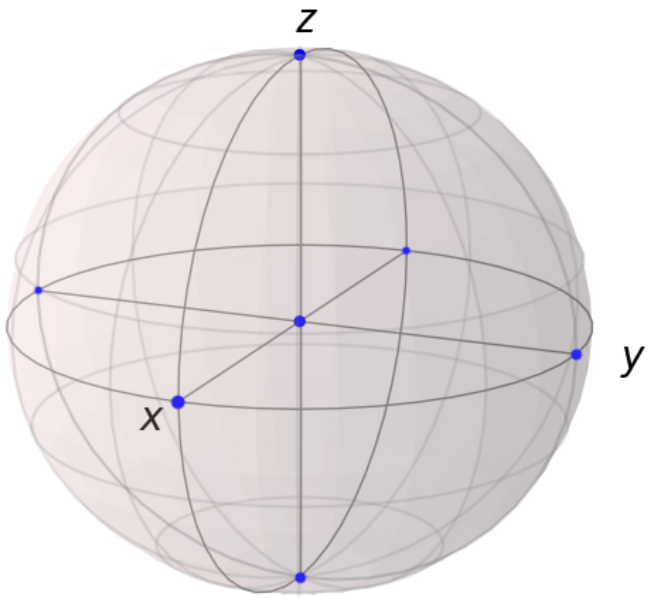


Figure 10.1: Bloch ball, with the center point and the points where the axes intersect the outer sphere marked with dots

We are now well-poised to improve upon the language in the news stories. The point that specifies the preparation of a qubit can be at the North Pole $(0, 0, 1)$, the South Pole $(0, 0, -1)$, or anywhere in the ball between them. We have a whole continuum of ways to be intermediate between completely confident that the Z test will yield $+1$ (all the way north) and completely confident that it will yield -1 (all the way south).

Now, there are other things one can do to a qubit. For starters, there are other binary measurements beyond just the X, Y and Z tests. Any pair of points exactly opposite each other on the Bloch sphere define a test, with each point standing for an outcome. The closer the preparation point is to an outcome point, the more probable that outcome. To be more specific, let's write the preparation point as (x, y, z) and the outcome point as (x', y', z'). Then the probability of getting that outcome given

that preparation is

$$P = \frac{1}{2}(1 + xx' + yy' + zz') \,. \tag{10.12}$$

An interesting conceptual thing has happened here. We have encoded the preparation of a qubit by a set of expected values, i.e., a set of probabilities. Consequently, all those philosophical debates over what probability means will spill over into the arguments about what quantum mechanics means. Moreover, and not unrelatedly, we can ask, "Why *three* probabilities? Why is it the Bloch sphere, instead of the Bloch disc or the Bloch hypersphere?" It would be perfectly legitimate, mathematically, to require probabilities for only two tests in order to specify a preparation point, or to require more than three. That would not be quantum mechanics; the fact that three coordinates are needed to nail down the preparation of the simplest possible system is a structural fact of quantum theory. But is there a deeper truth from which that can be deduced?

One could go in multiple directions from here: What about tests with more than two outcomes? Systems composed of more than one qubit? Very quickly, the structures involved become more difficult to visualize, and familiarity with linear algebra — eigenvectors, eigenvalues and their friends — becomes a prerequisite. People have also tried a variety of approaches to understand what quantum theory might be derivable from. Any of those topics could justify something in between a blog post and a lifetime of study.

For our own purposes, our next step will be to develop the framework in which we can consider multiple qubits together. It might not seem obvious now, but a good way to make progress is to combine our three expected values (x, y, z) into a matrix, like so:

$$\rho = \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix} \,. \tag{10.13}$$

This matrix has some nice properties of the sort that we can generalize to *bigger* matrices. For example, its trace is 1, which feels kind of like how a list of probabilities sums up to 1. Meanwhile, the determinant is the pleasingly Pythagorean quantity

$$\det \rho = \frac{1}{4}(1 - x^2 - y^2 - z^2) \,. \tag{10.14}$$

This will be nonnegative for all the valid preparation points. So, the product of the two eigenvalues of ρ will be positive for every point in the interior; we can only get a zero eigenvalue by picking a point on the surface. Using the trace and the determinant, we can find the eigenvalues thanks to Eq. (22.129):

$$\lambda_\pm = \frac{\mathrm{tr}\rho \pm \sqrt{(\mathrm{tr}\rho)^2 - 4\det\rho}}{2} = \frac{1}{2}(1 \pm \sqrt{x^2 + y^2 + z^2}) \,. \tag{10.15}$$

And indeed, this will always give us positive real numbers, except on the surface of the Bloch ball where the λ_+ solution is 1 while the λ_- solution is 0. Requiring that a matrix's eigenvalues be nonnegative is another property we can generalize.

Another interesting thing happens if we take the square of ρ:

$$\rho^2 = \frac{1}{4}\begin{pmatrix} 1+x^2+y^2+z^2+2z & 2x-2iy \\ 2x+2iy & 1+x^2+y^2+z^2-2z \end{pmatrix}. \tag{10.16}$$

If the point (x, y, z) is on the surface of the sphere, then $\rho^2 = \rho$. This will turn out to be a way to characterize the extreme elements in our set of valid preparations, no matter how big we make our matrices.

We can also take inspiration from our formula for the outcome probability in terms of coordinates, Eq. (10.12). This is hinting to us that, just like we turned a preparation into a matrix, we should turn each measurement outcome into a matrix, too. Furthermore, the rule for combining a preparation and an outcome to get a probability should involve something like a dot product.

10.4 Algebraic Language for Quantum Theory

We can express a workable set of basic postulates for quantum theory in the language and notation that we developed in §22.4, when we got into linear operators.

- A *quantum system* is any part of the physical world that we wish to study, often small, but in principle arbitrarily large.
- To each quantum system is associated a complex Hilbert space.
- A physicist's information about a quantum system — how it was prepared, any hypotheses about its origin and potential behavior — are encapsulated in a *quantum state,* which is a positive semidefinite operator on the system's Hilbert space. Quantum states have trace equal to 1 ("unit trace"). The set of all allowed quantum states that can be ascribed to a system is its *state space.*
- Experiments upon a quantum system are represented mathematically by sets of p.s.d. operators, one for each possible outcome of the experiment, that sum to the identity operator. The *Born rule* says that the probability of an outcome is the trace of the product of the operator representing that outcome and the quantum state.
- State spaces compose according to the tensor product. If $\{|v_i\rangle : i = 1, \ldots, n\}$ is an orthonormal basis for one system's Hilbert space, and $\{|u_j\rangle : j = 1, \ldots, m\}$ is an orthonormal basis for another, then those two systems can be considered jointly using a Hilbert space which has $\{|v_i\rangle \otimes |u_j\rangle\}$ for a basis.
- Transformations like time evolution that send one quantum state to another are *completely positive, trace preserving* (CPTP) maps.

These postulates are not exactly pithy! I spend at least a little time every day wondering why they are that way.

Much of the time, it is mathematically simpler to will work in the quantum theory of finite-dimensional systems, as is often done the study of quantum information and computation [85]. In this theory, each physical system is associated with a Hilbert space $\mathcal{H}_d \simeq \mathbb{C}^d$, where the dimension d can be taken as a physical characteristic of the system. For example, in a quantum computer containing N *qubits,* the dimension is $d = 2^N$. (So, the dimension can be taken to scale with the available budget.) A quantum state, the means of expressing the preparation of a system, is an operator on this Hilbert space that is positive semidefinite and has a trace of unity.

A state ρ is extremal in state space when $\rho^2 = \rho$, which is equivalent to $\mathrm{tr}\rho^2 = 1$. These are the *pure* states, which correspond to the rays in $\mathbb{C}^d$; we can think of them either as rank-1 projection operators or as the rays they project onto. All other states, the *mixed* states, are convex combinations of the pure states. Given two quantum systems for which we have the priors ρ_1 and ρ_2 respectively, the uncorrelated joint prior for the pair is the tensor product $\rho_1 \otimes \rho_2$. We typically introduce an orthonormal basis $\{|n\rangle : n = 0, \ldots, d-1\}$ for $\mathbb{C}^d$ that we call the *computational basis.* Like all orthonormal bases, this defines a POVM, because

$$\sum_n |n\rangle\langle n| = I, \tag{10.17}$$

and so we can speak of "measuring in the computational basis". This is how we represent the act of obtaining a response from a quantum computer at the end of a computation.

The state space of a single qubit is built upon $\mathbb{C}^2$. We take all those 2×2 matrices ρ that are p.s.d. and have unit trace. A convenient way to represent these is to introduce the *Pauli matrices*

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{10.18}$$

Each of these matrices is traceless and has a determinant equal to -1, from which we deduce that their eigenvalues are ± 1. Moreover, each Pauli matrix squares to the identity:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I. \tag{10.19}$$

It is slightly more arduous to check that the trace of the product of any two distinct Pauli matrices is always zero, and that

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \ [\sigma_y, \sigma_z] = 2i\sigma_x, \ [\sigma_z, \sigma_x] = 2i\sigma_y. \tag{10.20}$$

Apart from a factor of 2, these are the same commutator relations we saw for the generators of 3D spatial rotations, back in Eq. (22.209).

Using a notation that we will justify momentarily, let

$$r_x = \mathrm{tr}(\rho\sigma_x), \ r_y = \mathrm{tr}(\rho\sigma_y), \ r_z = \mathrm{tr}(\rho\sigma_z). \tag{10.21}$$

Then we can express ρ as the linear combination

$$\rho = \frac{1}{2}\left(I + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z\right). \tag{10.22}$$

To check the consistency of this expression, multiply both sides by a Pauli matrix and take the trace. Everything involving only one factor of that Pauli matrix will vanish, leaving only the coefficient of that Pauli matrix. Similarly, we find that

$$\mathrm{tr}\rho^2 = \frac{1 + r_x^2 + r_y^2 + r_z^2}{2}. \tag{10.23}$$

This gives us warrant to interpret (r_x, r_y, r_z) as the coordinates of a point within the unit ball, yielding the *Bloch ball* representation of qubit state space. The surface of the ball, the Bloch sphere, consists of all those points that correspond to pure states.

Let $|\psi\rangle$ be a unit vector in $\mathbb{C}^2$. When we form the projector $|\psi\rangle\langle\psi|$, an overall phase will drop out, so without loss of generality we can write $|\psi\rangle$ as

$$|\psi\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{pmatrix}. \tag{10.24}$$

The Pythagorean identity ensures we have the correct normalization. Those factors of $\frac{1}{2}$ look like they came out of nowhere, but they do turn out to be convenient, as we see when we form the projector:

$$|\psi\rangle\langle\psi| = \begin{pmatrix} \cos^2\frac{\theta}{2} & e^{-i\phi}\cos\frac{\theta}{2}\sin\frac{\theta}{2} \\ e^{i\phi}\cos\frac{\theta}{2}\sin\frac{\theta}{2} & \sin^2\frac{\theta}{2} \end{pmatrix}. \tag{10.25}$$

We use the double-angle formulas to simplify this into

$$|\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 1+\cos\theta & (\cos\phi - i\sin\phi)\sin\theta \\ (\cos\phi + i\sin\phi)\sin\theta & 1-\cos\theta \end{pmatrix}. \tag{10.26}$$

This is just our expression for a density matrix ρ in terms of the Pauli matrices, with (r_x, r_y, r_z) being the unit vector specified by the coordinates (θ, ϕ).

Each Pauli matrix has two orthogonal eigenvectors, and the projectors made from these eigenvectors add up to the identity operator. Therefore, each Pauli matrix defines a 2-outcome POVM. Letting j stand for x, y, z, and calling the two outcomes "$j+$" and "$j-$", we have that

$$\sigma_j = (+1)|j+\rangle\langle j+| + (-1)|j-\rangle\langle j-|. \tag{10.27}$$

Using the Born rule to calculate the probabilities of these two outcomes, we find

$$\mathrm{tr}\rho\sigma_j = p(j+) - p(j-), \tag{10.28}$$

which we can interpret as the expectation value of a random variable whose sample space is $\{+1, -1\}$.

This is a very common way of defining POVMs. Indeed, in older books, it is just about the only way ever discussed. Take an orthonormal basis, and associate each vector $|v_i\rangle$ in it with a real number a_i. The projectors defined by these vectors provide a resolution of the identity and thus constitute a POVM. Define a random variable whose sample space is the set of real numbers associated to the vectors, and then the

expectation value of that random variable is given by the Born rule. Interpreting those real numbers as eigenvalues, we can make a self-adjoint matrix

$$A = \sum_i a_i |v_i\rangle\langle v_i|, \tag{10.29}$$

and we can write

$$\langle A \rangle = \mathrm{tr} A\rho. \tag{10.30}$$

We have many ways of characterizing quantum states. We can take $\mathrm{tr}\rho^2$, as we did above. Alternatively, we can compute the *von Neumann entropy*

$$S(\rho) := -\mathrm{tr}\rho \log \rho. \tag{10.31}$$

It may look a little odd at first to be taking the logarithm of a matrix, but we know we can do matrix arithmetic, so we can set up a matrix Taylor series, which means we can define both exponentials and logarithms of matrices. And we can re-express the von Neumann entropy in terms of the eigenvalues of ρ:

$$S(\rho) = -\sum_i \lambda_i \log \lambda_i. \tag{10.32}$$

The eigenvalues of any quantum state ρ are nonnegative, because ρ is always p.s.d., and they add to 1 by the trace-normalization condition, so they qualify formally as a probability distribution. In fact, they provide the Born-rule probabilities for the POVM defined by the eigenvectors of ρ. The von Neumann entropy is the Shannon entropy of this probability distribution. Consequently, it vanishes for pure states. If the dimension of the Hilbert space is finite, then the von Neumann entropy is maximized by the garbage state $\frac{1}{d}I$. It is *additive* over tensor products:

$$S(\rho \otimes \sigma) = S(\rho) + S(\sigma). \tag{10.33}$$

Just as we did in classical probability theory, we can define a *relative* version of the von Neumann entropy,

$$S(\rho||\sigma) := \mathrm{tr}\rho \log \rho - \mathrm{tr}\rho \log \sigma. \tag{10.34}$$

10.5 Dynamics

An operation like time evolution needs to send quantum states to quantum states in a consistent way. If $\rho(t = 0)$ encodes Alice's expectations for the experiments she can perform on her system now, then $\rho(t = \tau)$ encodes what she currently expects to happen in experiments she might perform in the future. Whatever function that she uses to calculate the latter from the former,

$$\rho(t = \tau) = \mathcal{E}(\rho(t = 0)), \tag{10.35}$$

it must produce valid outputs from valid inputs. So, it must send p.s.d. matrices to p.s.d. matrices. Maps that do this are called *positive.* But the maps we use in quantum

theory meet a stronger condition, *complete* positivity. The most straightforward way to characterize the set of CP maps is as follows. Let $\{K_j\}$ be an arbitrary set of $d \times d$ matrices, and define the function

$$\mathcal{E}(\rho) = \sum_j K_j \rho K_j^\dagger \,. \tag{10.36}$$

This is known as the *Kraus operator-sum representation* of a CP map, and the $\{K_j\}$ are *Kraus operators*. The CP maps we will consider most often are *unitary conjugations* of a quantum state:

$$\rho(t = \tau) = U_\tau \rho(t = 0) U_\tau^\dagger. \tag{10.37}$$

By the notation U_τ, we wish to suggest that the operator, when wrapped around a quantum state in this way, implements a time evolution by a duration τ. Unitary conjugations are *trace preserving*: The cyclic property of the trace quickly shows that the trace of the input is the same as that of the output. This makes conjugation by U_τ an example of a *CPTP* map. Moreover, unitary conjugations leave both $\mathrm{tr}\rho^2$ and the von Neumann entropy unchanged.

Where might the Kraus operator-sum representation come from? Well, we know from linear algebra that we can decompose each K_j into the product of a unitary matrix and a p.s.d. matrix:

$$K_j = U_j P_j \,. \tag{10.38}$$

Both factors in this expression, the *polar decomposition* of the Kraus operator, refer to symmetries of the space of quantum states. Unitary matrices implement angle-preserving maps of $\mathbb{C}^d$ to itself. By choosing an appropriate unitary, we can rotate any vector $|\psi\rangle$ into any other vector $|\phi\rangle$. Said another way, unitaries let us send any *pure state* into any other. Meanwhile, p.s.d. matrices provide transformations that send the *interior* of quantum state space to itself. We can think of the set of all $d \times d$ p.s.d. matrices as a cone, with quantum state space being the slice through that cone defined by the condition that the trace be 1. Any p.s.d. matrix has a unique square root that is also p.s.d.; so, for any quantum state ρ, we can define the map

$$\mathcal{E}_\rho(M) = \rho^{1/2} M \rho^{1/2} \,, \tag{10.39}$$

which takes the identity I to the density matrix ρ. Up to normalization, this is the same as sending the garbage state I/d to ρ, of course. These maps preserve positive semidefiniteness, and an inverse map exists when ρ is invertible. The upshot is that the cone of p.s.d. matrices is *homogeneous*: Its symmetries let us map any interior point to any other.

We conclude that the Kraus operator-sum representation is where we naturally end up if we combine (a) the symmetries of pure states, (b) the homogeneity of the p.s.d cone and (c) taking linear combinations. The last consideration can be motivated from the reflection principle we discussed in §2.1. For example, if Alice expects that the time evolution will be described by a unitary chosen at random out of some set $\{U_j\}$, then the reflection principle leads to the map

$$\mathcal{E}(\rho) = \sum_j p_j U_j \rho U_j^\dagger \,, \tag{10.40}$$

for some probabilities $\{p_j\}$.

If the density matrix ρ_0 encodes our expectations for what measurements might give us now, then we can calculate expectations for what measurements might give us in the future, by conjugating with a unitary. And if we label our unitaries with a time index, we can have a whole family of them, with the matrix U_t expressing the operation of letting the system wait in isolation for an interval t.

$$\rho_t = U_t \rho U_t^\dagger \,. \tag{10.41}$$

It is physically reasonable to suppose that this evolution is a continuous change, i.e., that U_{t+dt} is close to U_t for a small time increment dt. And in particular, U_{dt} should be close to the identity:

$$U_{dt} = I + G\,dt \tag{10.42}$$

for some matrix G. What properties might this matrix enjoy? Well, we know that the adjoint of U_{dt} must be its inverse:

$$\begin{aligned} U_{dt}U_{dt}^\dagger &= (I + G\,dt)(I + G\,dt)^\dagger && (10.43)\\ &= (I + G\,dt)(I + G^\dagger\,dt) && (10.44)\\ &= I + (G + G^\dagger)dt + GG^\dagger(dt)^2 && (10.45)\\ &= I\,. && (10.46) \end{aligned}$$

In the limit that our time increment is very small indeed, we can drop the term of order $(dt)^2$, and we find that

$$G^\dagger = -G\,. \tag{10.47}$$

This makes G an *anti*-self-adjoint matrix, which we can write as a self-adjoint times i. Actually, $-i$ works just as well, and is the conventional choice:

$$G = -iA,\ A = A^\dagger\,. \tag{10.48}$$

All told, then,

$$U_{dt} = I - iA\,dt\,. \tag{10.49}$$

If we assume that the effect of time evolution depends only on the interval it lasts, not when that interval starts, then we can make any U_t we want by piling up factors of U_{dt}. We just say that dt is our duration t divided by some huge number N:

$$U_{t/N} = I - \frac{iAt}{N}\,, \tag{10.50}$$

and then we raise this to the Nth power. Going to the appropriate limit,

$$U_t = \lim_{N\to\infty} \left(I - \frac{iAt}{N}\right)^N\,. \tag{10.51}$$

But this is just the matrix exponential:

$$U_t = \exp(-iAt) = I - iAt + \frac{(-iAt)^2}{2!} + \frac{(-iAt)^3}{3!} + \cdots\,. \tag{10.52}$$

The self-adjoint matrix A, which we call the *generator* of the time evolution, defines a POVM by the scheme we described above, and so it has an expectation value:

$$\langle A(t)\rangle = \mathrm{tr}(A\rho_t)\,. \tag{10.53}$$

Because A commutes with every power of itself, it commutes with each term in the infinite series for the exponential, and so it commutes with the unitary U_t. This is a useful fact.

$$\langle A(t)\rangle = \mathrm{tr}(AU_t\rho_0 U_t^\dagger) = \mathrm{tr}(U_t A\rho_0 U_t^\dagger)\,. \tag{10.54}$$

Now, we use the cyclic property of the trace to pull the $U_t^\dagger$ off the end and stick it at the beginning, where it cancels against the U_t to give the identity matrix. Thus,

$$\langle A(t)\rangle = \mathrm{tr}(A\rho_0) = \langle A(0)\rangle\,. \tag{10.55}$$

A self-adjoint matrix generates a time evolution under which the expectation value of that matrix is *conserved!*

We can get a little more granular in our analysis. Suppose that $|\psi\rangle$ is an eigenvector of the generator A:

$$A|\psi\rangle = \lambda|\psi\rangle\,. \tag{10.56}$$

Then, if the state is the density matrix $|\psi\rangle\langle\psi|$, the POVM defined by A will output the $|\psi\rangle\langle\psi|$ result with probability 1. Moreover, because

$$A^n|\psi\rangle = \lambda^n|\psi\rangle\,, \tag{10.57}$$

every term in the matrix exponential for the unitary U_t will just become a *number* when it hits $|\psi\rangle$. Consequently,

$$U_t|\psi\rangle = e^{-i\lambda t}|\psi\rangle\,. \tag{10.58}$$

The vector $|\psi\rangle$ just "spins its wheel", accumulating a phase that varies over time. When we form the density matrix, this phase drops out:

$$U_t|\psi\rangle\langle\psi|U_t^\dagger = e^{-i\lambda t}|\psi\rangle\langle\psi|e^{i\lambda t} = |\psi\rangle\langle\psi|\,. \tag{10.59}$$

In other words, if the state is a projector onto an eigenvector of the generator, the time evolution leaves that projector fixed. We call these states the *eigenstates* of that generator.

What if ρ is an arbitrary state? We can evaluate its time derivative as follows:

$$\dot\rho_t = \lim_{h\to 0}\frac{\rho_{t+h}-\rho_t}{h} \tag{10.60}$$

$$= \lim_{h\to 0}\frac{1}{h}\left(U_h\rho_t U_h^\dagger - \rho_t\right) \tag{10.61}$$

$$= \lim_{h\to 0}\frac{1}{h}\left((I-iAh)\,\rho_t\,(I+iAh) - \rho_t\right)\,. \tag{10.62}$$

Canceling terms and taking the limit,

$$\dot{\rho}_t = [-iA, \rho_t] \,. \tag{10.63}$$

This recalls to mind the Bohr model from the early days of quantum physics — the "old quantum theory" period, when physicists were trying to understand atomic phenomena by taking classical physics and adding constraints of various kinds. In the Bohr model, the electron in a hydrogen atom could sit at one of a discrete set of energy levels and stay at that energy without radiating, despite the prediction of classical electromagnetism that it would. Energy levels were stable. This hints that the generator of time evolution should perhaps be identified with the energy: A state associated with a definite energy value — a state that implies a probability-1 prediction for an energy measurement — would remain fixed over time. Moreover, given any quantum state, the expected value of an energy measurement would stay constant.

This suggests that we should call the generator of a unitary family the quantum Hamiltonian. Further support for this identification comes from an advanced approach to classical mechanics, in which the Liouville equation (8.18) describes the time evolution of a probability density function on phase space. This equation bears an algebraic resemblance to Eq. (10.63), and this resemblance was key to Paul Dirac's method for finding the quantum counterpart to classical concepts. Many students are unfamiliar with this way of doing classical mechanics, since it is often encountered *after* they first see quantum theory. This is one example of how the standard way of teaching quantum mechanics is not historically accurate, even when we pretend to present it in historical order.

We have thought of time evolution as applying to quantum states, i.e., to the density matrices that encode the preparation of a quantum system. We could equivalently think of time evolution in an alternative way: We imagine that the preparation stays fixed, and instead we do a different *measurement,* one whose first stage is to kick back and wait a while. The POVM elements of this measurement will in general be different from, but related to, those that encode the original measurement. Consider our basic formula for getting probabilities, $p = \mathrm{tr}(E\rho)$. We can calculate the time-evolved probability by transforming the state, as we did above:

$$p \to p' = \mathrm{tr}[E(U\rho U^\dagger)] \tag{10.64}$$

but thanks to the cyclic property of the trace, we can put the time dependence into the POVM element instead:

$$p' = \mathrm{tr}[(U^\dagger E U)\rho] \,. \tag{10.65}$$

We have two *pictures* of time dependence. In one picture, we make the preparation time-dependent, and in the other, we do so to the measurement. The two pictures have to give the same answers in the end, but one can be more convenient than the other depending on the scenario being studied. Introductory quantum courses nearly always dwell in the picture where the states are time-dependent. On the other hand, finding the quantum counterpart to a classical theory, and understanding the relation between quantum physics and sophisticated ways of doing classical physics, can be easier in the picture where the time dependence lies in the measurements.

10.6 A Single Qubit, Algebraic Version

The mathematics of a qubit come into play whenever we have two completely distinguishable alternatives. That is, the relevant scenarios are when there is a binary test of some kind, and one preparation of our system can make the "+" outcome completely certain to occur, while another preparation can make the "−" outcome completely certain to occur.

- *Stern–Gerlach experiments.* Atoms are streamed out of a hot oven and sent between the poles of an oddly-shaped magnet. Classical electrodynamics predicts that tiny magnets traversing an asymmetric magnetic field like this will be deflected through a whole range of possible angles, so that the atom beam would come out as a spray, rather like water from a garden hose mostly covered by a person's thumb. But when Stern and Gerlach did the test in 1922, the atom beam split in two. Instead of an "analog" result, they got a "digital" one. (This was in the days when physicists recorded atoms and subatomic particles by firing them at photographic plates and looking for dark spots. In fact, at first Stern and Gerlach saw *nothing.* Their photographic plate refused to darken at all. But then Stern leaned over and breathed on the plate, and the sulfur in his breath from his disgusting cheap cigars provoked the chemical reaction necessary for the dots to appear.) The basic outcomes of a Stern–Gerlach test are "swerve this way" and "swerve that way", and different tests — represented by different axes through the Bloch sphere — can be done by tilting the magnet at different angles.

- *Photon polarization.* Classical E&M tells us that we can have a plane wave in the electric field propagating through empty space. For example,

$$\vec{E}(\vec{r},t) = \begin{pmatrix} E_{x,0}\cos(kz-\omega t+\alpha_x) \\ E_{y,0}\cos(kz-\omega t+\alpha_y) \\ 0 \end{pmatrix} \tag{10.66}$$

  is a wave propagating in the $\hat{z}$ direction with wavenumber $k = 2\pi/\lambda$ and angular frequency ω. The vector $\vec{E}_0$, whose components are $E_{x,0}$ and $E_{y,0}$, specifies the polarization of the light wave. It is horizontally polarized if the y component vanishes and vertically polarized if the x component is zero. At the quantum level, we talk about individual photons being polarized — or, to phrase it more carefully, about the outcomes of polarization tests upon photons. By convention, $|z+\rangle$ and $|z-\rangle$ represent right and left circular polarization. Then $|x+\rangle$ and $|x-\rangle$ represent horizontal and vertical polarization, while $|y+\rangle$ and $|y-\rangle$ correspond to polarization along *diagonal* axes halfway between horizontal and vertical. (The relation between physical geometry and Bloch-sphere geometry can be confusing. We will try to make it clear through examples. Just remember: *antipodal* points on the Bloch sphere correspond to *orthogonal* state vectors.)

- *Two-way interferometers.* You have likely seen optical tables: laboratory setups with mirrors and lenses and light sources and such, all carefully positioned with

respect to each other (and perhaps with a warning sign taped on the side to say *do not touch!*). At the quantum level, we consider individual photons going down one track or another. Or, to phrase it more carefully: We consider the probabilities that a test of the photon-detecting type will have a positive outcome on this path or on the other path. Here, the convention is that a σ_z eigen-POVM represents the act of checking for a photodetector going "click" on one path (say, the "upper") or on the other (the "lower"). Other measurements can be implemented by putting different optical devices in front of the final photodetectors.

- *Asymmetric molecules.* You perhaps have played with molecule models, doing what we might call "ball-and-stick" chemistry. Hydrogen has one plug, carbon has four, and so on. It should come as no surprise to you that a fully quantum treatment of molecular structure is more complicated. However, there is a happy intermediate point between the toys and the full-blown chemistry. This is exemplified by the *ammonia* molecule, NH_3. In the ball-and-stick picture, it's a pyramid of sorts, with the three hydrogen atoms all on the same side of the hydrogen. This gives us something we can model as a qubit, because we can fix an axis and then "ask" (by some laboratory procedure) whether the nitrogen is pointing "up" or "down" along that axis. It so happens — and one can deduce this from experiment, and/or a more detailed study of the molecule — that the pure states which imply a definite prediction for an *energy* measurement are not the same as the pure states which imply a definite prediction for an *orientation* measurement. Instead, the former is a superposition of the latter, and vice versa.

- *Nuclear magnetism.* Big magnets are made of smaller magnets, and so on. A big magnet has a detectable amount of magnetism if the small magnets add up; if they cancel each other out, it won't. Subatomic particles are magnets, too. The energy of a magnet depends upon its orientation with respect to the applied magnetic field. We have a quantum quantity that has a sense of direction: the vector from the origin to a point in the Bloch ball. When we subject subatomic particles to magnetic fields, we treat them as *spin systems.* The quantum counterpart of our classical notion of "energy" will be a 2×2 matrix whose eigenvalues are the possible results of an energy-type measurement, and this matrix will be constructed from a kind of dot product between the magnetic field vector and the Pauli matrices.

- *Neutral kaons.* Back in the day, subatomic particles that were more massive than electrons but less massive than protons were called *mesons.* Among them are the "K-mesons", nowadays called *kaons.* Colliding protons or neutrons against other matter produces neutral kaons (K^0) and, in lesser quantities, neutral anti-kaons ($\bar{K}^0$). Ordinary matter absorbs anti-kaons better than it does kaons. Both K^0's and $\bar{K}^0$'s are radioactive. They can decay either to pairs of pions or to triplets of pions. Decays to pairs of pions happen fast, on the timescale of 9×10^{-11} seconds, while decays to three pions happen much more slowly, on the timescale of 5×10^{-8} seconds. Here, the "computational basis" question is "Are you a K^0

or a $\bar{K}^0$?" A measurement of the decay duration, or of how many pions come out, is a Pauli σ_x measurement. In other words, the eigenvectors of σ_x represent preparations of short-lived (K_S) and long-lived (K_L) kaons.

In this section, we will focus on pure states. So, most of our density matrices will be of the form

$$\rho = |\psi\rangle\langle\psi| \,, \tag{10.67}$$

with

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} . \tag{10.68}$$

In these special cases, we can simplify our formula for the expectation value of an eigen-POVM:

$$\langle A \rangle = \mathrm{tr}(\rho A) = \mathrm{tr}(|\psi\rangle\langle\psi|A) \,. \tag{10.69}$$

The elements of the density matrix are

$$[|\psi\rangle\langle\psi|]_{jk} = \psi_j \psi_k^* \,, \tag{10.70}$$

and so

$$\mathrm{tr}(|\psi\rangle\langle\psi|A) = \sum_{jk} \psi_j \psi_k^* [A]_{kj} = \langle\psi|A|\psi\rangle \,. \tag{10.71}$$

By the same formula, if E is an element in a POVM made from a self-adjoint operator's eigenvectors, say

$$E = |\phi\rangle\langle\phi| \,, \tag{10.72}$$

then the probability of getting outcome E given the state $|\psi\rangle\langle\psi|$ is

$$\mathrm{tr}(|\psi\rangle\langle\psi|E) = \langle\psi|E|\psi\rangle = \langle\psi|\phi\rangle\langle\phi|\psi\rangle = |\langle\psi|\phi\rangle|^2 \,. \tag{10.73}$$

Mathematically, there is no reason to prefer any axis through the Bloch ball over any other. A ball is just symmetrical like that. However, we may for physical reasons care about a particular measurement, like the energy, and it is traditional to associate that measurement with the z axis. The outcomes of this measurement of particular interest will then correspond to the $+1$ and -1 eigenvectors of the Pauli matrix σ_z.

Let us start by considering a case where the Hamiltonian is proportional to σ_z:

$$H = -E_0 \sigma_z = \begin{pmatrix} -E_0 & 0 \\ 0 & E_0 \end{pmatrix} . \tag{10.74}$$

Exponentiating a diagonal matrix is easy. We just take the exponential of each element on the diagonal. Our family of unitaries representing time evolution will thus be given by

$$U_\tau = \begin{pmatrix} e^{i\tau} & 0 \\ 0 & e^{-i\tau} \end{pmatrix} . \tag{10.75}$$

Here, we have introduced the label τ as a dimensionless time parameter. The standard way of making time unitless so that we can take its exponential is to multiply it by an

energy and then divide by the Planck constant $\hbar$. We will be seeing that again and again. In this case, we take

$$\tau = \frac{E_0 t}{\hbar} \,. \tag{10.76}$$

We saw earlier, in Eq. (10.24), that we can specify a pure state in terms of the angle θ from the "north pole" of the Bloch sphere and the azimuthal angle ϕ:

$$|\psi\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{pmatrix} . \tag{10.77}$$

The unitary U_τ introduces opposite phases to the two components of this vector:

$$U_\tau|\psi\rangle = \begin{pmatrix} e^{i\tau}\cos\frac{\theta}{2} \\ e^{-i\tau}e^{i\phi}\sin\frac{\theta}{2} \end{pmatrix} . \tag{10.78}$$

We know that an overall phase doesn't matter, physically, so we factor out an $e^{i\tau}$ and discard it:

$$\begin{pmatrix} \cos\frac{\theta}{2} \\ e^{-2i\tau}e^{i\phi}\sin\frac{\theta}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i(\phi-2\tau)}\sin\frac{\theta}{2} \end{pmatrix} . \tag{10.79}$$

The angle from the vertical remains constant, while the azimuthal angle changes steadily with time. Or, to say it another way, the state rotates around the vertical axis.

We can also do the calculation in a different coordinate system, if we want. The time evolution of a density matrix ρ is

$$\rho \to U_\tau \rho U_\tau^\dagger = \frac{1}{2}(I + xU_\tau\sigma_x U_\tau^\dagger + yU_\tau\sigma_y U_\tau^\dagger + zU_\tau\sigma_z U_\tau^\dagger) \,. \tag{10.80}$$

We figure out what conjugating does to ρ by conjugating each of the Pauli matrices. First,

$$U_\tau\sigma_x U_\tau^\dagger = \begin{pmatrix} 0 & e^{2i\tau} \\ e^{-2i\tau} & 0 \end{pmatrix} . \tag{10.81}$$

Likewise,

$$U_\tau\sigma_y U_\tau^\dagger = \begin{pmatrix} 0 & -ie^{2i\tau} \\ ie^{-2i\tau} & 0 \end{pmatrix} . \tag{10.82}$$

And finally,

$$U_\tau\sigma_z U_\tau^\dagger = \sigma_z \,. \tag{10.83}$$

To find the expectation values of the Pauli matrices, we need to take the trace of their product with ρ. The unitaries U_τ all commute with σ_z, so using the cyclic property of the trace,

$$\mathrm{tr}(U_\tau\sigma_x U_\tau^\dagger\sigma_z) = \mathrm{tr}(\sigma_x U_\tau^\dagger\sigma_z U_\tau) = \mathrm{tr}(\sigma_x\sigma_z) = 0 \,. \tag{10.84}$$

Likewise,

$$\mathrm{tr}(U_\tau\sigma_y U_\tau^\dagger\sigma_z) = \mathrm{tr}(\sigma_y\sigma_z) = 0 \,. \tag{10.85}$$

Therefore, when we multiply σ_z into the density matrix and take the trace, the only contribution can come from the σ_z term:

$$\langle \sigma_z(\tau) \rangle = z \,. \tag{10.86}$$

A few lines of straightforward matrix arithmetic and applying Euler's formula gives

$$\langle \sigma_x(\tau) \rangle = x \cos(2\tau) + y \sin(2\tau) \,. \tag{10.87}$$

Again, we see that the point representing the state circles around the vertical axis at a constant altitude.

Next, we consider what happens if our Hamiltonian is proportional to a different Pauli matrix. Let's pick σ_x.

$$H = E_0 \sigma_x \,. \tag{10.88}$$

We might guess that this will generate rotations around the x axis. Explicitly, this matrix is

$$H = \begin{pmatrix} 0 & E_0 \\ E_0 & 0 \end{pmatrix} \,. \tag{10.89}$$

Restricting our attention to the case of pure states, we can write a pair of differential equations for the individual vector components:

$$i\hbar \frac{d\psi_1}{dt} = E_0 \psi_2 \,, \tag{10.90}$$

$$i\hbar \frac{d\psi_2}{dt} = E_0 \psi_1 \,. \tag{10.91}$$

Adding these two equations gives a differential equation for $\psi_1 + \psi_2$ that we can solve:

$$i\hbar \frac{d}{dt}(\psi_1 + \psi_2) = E_0(\psi_1 + \psi_2) \,. \tag{10.92}$$

When the derivative is just proportional to the original function, we've got ourselves an exponential:

$$\psi_1 + \psi_2 = a e^{-iE_0 t/\hbar} \,. \tag{10.93}$$

Likewise, we can subtract them to get a differential equation for $\psi_1 - \psi_2$:

$$i\hbar \frac{d}{dt}(\psi_1 - \psi_2) = E_0(\psi_2 - \psi_1) \tag{10.94}$$

which we can solve to yield

$$\psi_1 - \psi_2 = b e^{iE_0 t\hbar} \,. \tag{10.95}$$

Adding and subtracting these gives us

$$\psi_1 = \frac{a}{2} e^{-iE_0 t/\hbar} + \frac{b}{2} e^{iE_0 t/\hbar} \,, \tag{10.96}$$

$$\psi_2 = \frac{a}{2} e^{-iE_0 t/\hbar} - \frac{b}{2} e^{iE_0 t/\hbar} \,. \tag{10.97}$$

To illustrate, let's say our initial state is

$$|\psi(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \tag{10.98}$$

Then

$$\frac{a+b}{2} = 1, \ \frac{a-b}{2} = 0 , \tag{10.99}$$

which we solve by setting

$$a = b = 1 . \tag{10.100}$$

Inserting these coefficients into our time-evolution equations and using Euler's formula to simplify, we obtain

$$\psi_1 = \cos \frac{E_0 t}{\hbar} , \tag{10.101}$$

$$\psi_2 = -i \sin \frac{E_0 t}{\hbar} . \tag{10.102}$$

From our expression for $|\psi\rangle$ in spherical coordinates, Eq. (10.24), we recognize this as saying that the angle from the vertical changes smoothly in time,

$$\theta(t) = -\frac{2E_0 t}{\hbar} , \tag{10.103}$$

while the azimuthal angle remains constant. Starting at the "north pole", this Hamiltonian rotates down to the equator, through the south pole and back again. This confirms our intuition that σ_x generates rotations around the x axis.

We now have two hints about how to think of qubit unitaries geometrically. We can explore this further. In the homework, you showed that a unitary evolution of ρ does not affect the purity $\mathrm{tr}\rho^2$, and thus it leaves unchanged the distance $x^2 + y^2 + z^2$ from the center of the Bloch ball. Given any two density matrices, we can show by the same manner that $\mathrm{tr}(\rho_1 \rho_2)$ is unchanged. And if we expand both of these density matrices out into their Bloch-ball coordinates,

$$\rho_1 = \frac{1}{2}(I + x_1\sigma_x + y_1\sigma_y + z_1\sigma_z), \ \rho_2 = \frac{1}{2}(I + x_2\sigma_x + y_2\sigma_y + z_2\sigma_z) , \tag{10.104}$$

we find that every term in the product $\rho_1\rho_2$ that involves two different Pauli matrices vanishes when we take the trace. The only terms that aren't traceless are those proportional to the identity, which come from $I \cdot I$ and the squares of the Pauli matrices. We are left with

$$\mathrm{tr}(\rho_1 \rho_2) = \frac{1}{2}(1 + x_1x_2 + y_1y_2 + z_1z_2) . \tag{10.105}$$

Unitary transformations hold fixed the *dot product* of the Bloch coordinates! This tells us that they implement rotations, and possibly also reflections, of the Bloch ball. Further consideration shows that they cannot actually implement reflections, only rotations.

On the homework, you did the groundwork to show that any 2×2 Hamiltonian can be written as a linear combination of the Pauli matrices and the identity:

$$H = h_I I + h_x \sigma_x + h_y \sigma_y + h_z \sigma_z \,. \tag{10.106}$$

If $|\psi\rangle$ is an eigenvector of this Hamiltonian,

$$H|\psi\rangle = E|\psi\rangle \tag{10.107}$$

then by normalization

$$\langle\psi|H|\psi\rangle = E\langle\psi|\psi\rangle = E\,, \tag{10.108}$$

and

$$\mathrm{tr}(H|\psi\rangle\langle\psi|) = E\,. \tag{10.109}$$

Then, if (x, y, z) are the Bloch coordinates of $|\psi\rangle\langle\psi|$, we must have

$$E = \frac{1}{2}\mathrm{tr}\left((h_I I + h_x \sigma_x + h_y \sigma_y + h_z \sigma_z)(I + x\sigma_x + y\sigma_y + z\sigma_z)\right)\,. \tag{10.110}$$

As before, all of the products that are proportional to a Pauli matrix vanish when we take the trace, and only the terms proportional to the identity survive.

$$E = h_I + x h_x + y h_y + z h_z \,. \tag{10.111}$$

Unless H is proportional to the identity, which is a rather boring Hamiltonian, it will have two distinct eigenvalues and two orthogonal eigenvectors. We can write it as the sum

$$H = E_1|\psi_1\rangle\langle\psi_1| + E_2|\psi_2\rangle\langle\psi_2|\,. \tag{10.112}$$

The two density matrices here are each left invariant under the time evolutions generated by H. They are antipodal to each other on the Bloch sphere, and they define the axis around which those time evolutions rotate the sphere.

From the homework, we know that

$$h_x = \frac{1}{2}\mathrm{tr}(H\sigma_x)\,, \tag{10.113}$$

so if we multiply both sides of the previous expansion by σ_x and take the trace, we get

$$\mathrm{tr}(H\sigma_x) = 2h_x = E_1 x_1 + E_2 x_2\,. \tag{10.114}$$

Meanwhile, because

$$|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| = I\,, \tag{10.115}$$

it must be that

$$x_1 + x_2 = 0\,. \tag{10.116}$$

Therefore,

$$2h_x = (E_1 - E_2)x_1\,. \tag{10.117}$$

The analogous steps work for σ_y and σ_z as well, meaning that we have the Bloch coordinates for the eigenvector $|\psi_1\rangle$:

$$(x_1, y_1, z_1) = \frac{2}{E_1 - E_2}(h_x, h_y, h_z)\,. \tag{10.118}$$

Matters simplify somewhat when H is traceless. Then $h_I = 0$ and

$$H = h_x\sigma_x + h_y\sigma_y + h_z\sigma_z = \begin{pmatrix} h_z & h_x - ih_y \\ h_x + ih_y & -h_z \end{pmatrix}\,. \tag{10.119}$$

The determinant is

$$\det H = -h_z^2 - (h_x - ih_y)(h_x + ih_y) = -h_z^2 - h_x^2 - h_y^2\,. \tag{10.120}$$

This tells us the product of the eigenvalues. The sum of the eigenvalues is 0. Therefore,

$$E_\pm = \pm\sqrt{h_x^2 + h_y^2 + h_z^2}\,. \tag{10.121}$$

A "spin operator" for an arbitrary axis has this form, except that the coefficients are unitless and they are restricted to lie on a sphere of radius 1. We see immediately that the eigenvalues of such an operator are ± 1. Its eigenvectors are the orthogonal vectors represented by antipodal points on the Bloch sphere, along the axis specified by the spin operator's coordinates.

In classical electrodynamics, a body immersed in a magnetic field $\vec{B}$ has a potential energy that depends upon the relative orientation of the applied field and the body's magnetic moment:

$$E = -\vec{\mu}\cdot\vec{B}\,. \tag{10.122}$$

Supposing for example that we have a closed loop of current, we can write the magnetic moment of this loop in terms of the current I and the enclosed area A. In cgs units, using the $\vec{A}$ to denote the "oriented area" — the vector perpendicular to the plane of the loop which has the enclosed area as its length — we have

$$\vec{\mu} = \frac{I}{c}\vec{A}\,. \tag{10.123}$$

For a spinning ring of mass M, carrying a charge Q uniformly distributed, the current is the amount of charge passing a given point per unit time. Let the radius be R and the tangential speed of the spinning ring be v; then the current is

$$I = \frac{Q}{2\pi R}v\,. \tag{10.124}$$

Accordingly, the magnetic moment has a size

$$\mu = \frac{Q}{2\pi Rc}v\pi R^2 = \frac{Q}{2c}Rv\,. \tag{10.125}$$

The angular momentum has magnitude MRv, and so

$$\vec{\mu} = \frac{Q}{2Mc}\vec{L}\,. \tag{10.126}$$

We can devise the quantum counterpart to this formula if we find something that can play the role of the angular momentum $\vec{L}$. Why not use an axis through the Bloch sphere? Any time we describe a qubit by a pure state $|\psi\rangle$, we have a direction in three-dimensional space — the line segment from the origin of the Bloch sphere to the appropriate point on its surface. We distill a large amount of experience from the early years of quantum mechanics by introducing the *spin operators* for the $\hat{x}$, $\hat{y}$ and $\hat{z}$ axes:

$$S_j = \frac{\hbar}{2}\sigma_j\,. \tag{10.127}$$

These have units of angular momentum from the $\hbar$, and the $1/2$ is a convention introduced to make other formulae down the road work out nicely. We will study this much more deeply in later chapters. For now, you may recall from the homework how the commutators for the Pauli matrices resembled those for the generators of spatial rotations. And you might recall from classical mechanics how angular momentum is conserved in situations that are rotationally symmetric (a statement that goes back to Kepler's Second Law). Today, we will content ourselves by positing that the Hamiltonian for a *spin* in a magnetic field is

$$H_S = -\vec{\mu}\cdot\vec{B}\,, \tag{10.128}$$

where

$$\vec{\mu} = \gamma\vec{S} = \frac{\gamma\hbar}{2}\vec{\sigma}\,, \tag{10.129}$$

meaning that

$$H_S = -\gamma\vec{B}\cdot\vec{S} = -\gamma(B_xS_x + B_yS_y + B_zS_z)\,. \tag{10.130}$$

The value of the constant γ (the "gyromagnetic ratio") will depend on what kind of particle we're talking about. A convenient unit for magnetic moments on the atomic scale is the Bohr magneton:

$$\mu_B = \frac{|q_e|\hbar}{2m_ec}\,. \tag{10.131}$$

The nuclear magneton is about two thousand times smaller, because it is defined using the proton mass instead:

$$\mu_N = \frac{|q_e|\hbar}{2m_pc}\,. \tag{10.132}$$

Both the proton and the neutron have magnetic moments on the latter scale:

$$\mu_p \approx 5.586\mu_N,\ \ \mu_n \approx -3.826\mu_N\,. \tag{10.133}$$

The magnetic moment of the electron is very nearly -2 times the Bohr magneton. The sign comes from the electron charge, and the factor of 2 can be deduced from special relativity.

What kind of time evolution will the Hamiltonian H_S generate? The simplest case for which to do the algebra is when $\vec{B} = B\hat{z}$. Then

$$H_S = -\gamma B S_z \,, \tag{10.134}$$

and so

$$U(t,0) = \exp\left(-\frac{iH_S t}{\hbar}\right) . \tag{10.135}$$

If the initial state is

$$|\psi, 0\rangle = \cos\frac{\theta_0}{2}|z+\rangle + \sin\frac{\theta_0}{2}e^{i\phi_0}|-z\rangle \,, \tag{10.136}$$

then our unitary will evolve it to

$$|\psi, t\rangle = \cos\frac{\theta_0}{2}e^{i\gamma Bt/2}|z+\rangle + \sin\frac{\theta_0}{2}e^{i\phi_0}e^{-i\gamma Bt/2}|z-\rangle \,. \tag{10.137}$$

We factor out the phase $e^{i\gamma Bt/2}$ and see that the polar angle θ remains constant in time, while the azimuthal angle ϕ changes linearly:

$$\theta(t) = \theta_0, \;\; \phi(t) = \phi_0 - \gamma Bt \,. \tag{10.138}$$

The Bloch sphere is, after all, a sphere. And as we saw earlier, the Hamiltonian defined by an axis will leave that axis fixed and generate rotations around it. Thus, for $\vec{B}$ in an arbitrary direction, we will have rotation rotation at the *Larmor angular frequency*,

$$\vec{\omega}_L = -\gamma\vec{B} \,. \tag{10.139}$$

For a magnetic field $\vec{B}$ independent of time, $\vec{\omega}_L$ is also constant, and

$$U(t,0) = \exp\left(-\frac{i\vec{\omega}_L \cdot \vec{S}}{\hbar}t\right) . \tag{10.140}$$

What if the applied field is *not* constant over time? We can certainly move the external magnet, or change the amount of current that we supply to an electromagnet. One very important topic, *nuclear magnetic resonance,* arises from combining a constant field combined with a time-varying one:

$$\vec{B}(t) = B_0\hat{z} + B_1(\hat{x}\cos\omega t - \hat{y}\sin\omega t) \,. \tag{10.141}$$

For convenience, we have put the constant field along the $\hat{z}$ axis. We will take $B_0 > B_1$. There are two frequencies involved here. The field in the $\hat{z}$ direction defines a Larmor frequency:

$$\omega_0 = \gamma B_0 \,. \tag{10.142}$$

And then we have the angular frequency ω with which the components in the $\hat{x}$–$\hat{y}$ plane oscillate, the radio-frequency (RF) part of the field. The Hamiltonian for a body of gyromagnetic ratio γ in this field is

$$H_S(t) = -\gamma[B_0 S_z + B_1(\cos\omega t S_x - \sin\omega t S_y)] \,. \tag{10.143}$$

We can throw all sorts of ideas at this Hamiltonian. One route to understanding is to switch to a *rotating frame.* By analogy, we can imagine aliens watching the Earth from a distant spaceship and trying to understand what's going on here. The trajectory of a terrestrial body could be a complicated function in the spaceship's coordinate system, and then become much simpler with respect to coordinate axes that move with the Earth. (Of course, the aliens could also find our behavior confusing for many other reasons.) We take a guess and introduce a family of unitaries generated by S_z:

$$U_\omega(t) = \exp\left(-\frac{i\omega t S_z}{\hbar}\right),\ H_\omega = \omega S_z\,. \tag{10.144}$$

At time $t = 0$, the two frames coincide, because $U_\omega(0)$ is the identity. In the rotating frame, we have a two-entry column vector

$$|\psi_R, t\rangle = U_\omega(t)|\psi_t\rangle\,. \tag{10.145}$$

The time evolution of this vector is what we get by applying the unitary $U_S(t)$ — which we don't yet understand — to the initial state $|\psi, 0\rangle$ and then transforming to the new frame:

$$|\psi_R, t\rangle = U_\omega(t)U_S(t)|\psi, 0\rangle \tag{10.146}$$

$$= U_\omega(t)U_S(t)|\psi_R, 0\rangle\,. \tag{10.147}$$

If we have a family $U(t)$ of unitaries, we can find the Hamiltonian that advances from time t to time $t + \epsilon$:

$$U(t+\epsilon) = e^{-iH(t)\epsilon/\hbar}U(t)\,. \tag{10.148}$$

The difference between $U(t+\epsilon)$ and $U(t)$ is then

$$U(t+\epsilon) - U(t) = [e^{-iH(t)\epsilon/\hbar} - 1]U(t)\,, \tag{10.149}$$

which in the limit of small ϵ becomes

$$U(t+\epsilon) - U(t) = -\frac{i}{\hbar}H(t)\epsilon U(t)\,, \tag{10.150}$$

Dividing through by ϵ and solving for $H(t)$, we obtain

$$H(t) = i\hbar\frac{dU}{dt}U^\dagger\,. \tag{10.151}$$

Applying this to the product $U_\omega(t)U_S(t)$, we compute as follows:

$$H_R(t) = i\hbar\frac{\partial}{\partial t}(U_\omega U_S)U_S^\dagger U_\omega^\dagger \tag{10.152}$$

$$= i\hbar\frac{\partial U_\omega}{\partial t}U_\omega^\dagger + U_\omega i\hbar\frac{\partial U_S}{\partial t}U_S^\dagger U_\omega^\dagger \tag{10.153}$$

$$= H_\omega + U_\omega H_S U_\omega^\dagger\,. \tag{10.154}$$

Switching to the rotating frame does two things to our generator. First, it introduces a contribution that rotates around the $\hat{z}$ axis. Second, it changes the perspective from which we see H_S. We know the formulae for all of these in terms of the magnetic fields and the spin operators:

$$H_R(t) = \omega S_z - \gamma \exp\left(-\frac{i\omega t S_z}{\hbar}\right)(B_0 S_z + B_1(\cos\omega t S_x - \sin\omega t S_y))\exp\left(\frac{i\omega t S_z}{\hbar}\right) \tag{10.155}$$

$$= (-\gamma B_0 + \omega)S_z - \gamma B_1 M(t)\,, \tag{10.156}$$

where we have defined

$$M(t) = \exp\left(-\frac{i\omega t S_z}{\hbar}\right)(\cos\omega t S_x - \sin\omega t S_y)\exp\left(\frac{i\omega t S_z}{\hbar}\right)\,. \tag{10.157}$$

The crafty way to understand this expression is to take its time derivative. This is not so difficult. For arbitrary time-independent A_1 and A_2, we quickly verify that

$$\frac{d}{dt}e^{tA_1}A_2 e^{-tA_1} = e^{tA_1}[A_1, A_2]e^{-tA_1}\,. \tag{10.158}$$

Here, there is a time dependence in the middle contribution, but only through trig functions, which we know how to differentiate. The result is that $M(t)$ is actually *constant* in time:

$$\frac{d}{dt}M(t) = 0\,. \tag{10.159}$$

At $t = 0$, $M(0)$ reduces to S_x. So,

$$H_R = (-\gamma B_0 + \omega)S_z - \gamma B_1 S_x \tag{10.160}$$

$$= -\gamma\left(B_1 S_x + \left(B_0 - \frac{\omega}{\gamma}\right)S_z\right) \tag{10.161}$$

$$= -\gamma\left(B_1 S_x + B_0\left(1 - \frac{\omega}{\omega_0}\right)\right)\,. \tag{10.162}$$

We can express this as the dot product of the spin operator $\vec{S}$ with a modified magnetic field,

$$H_R = -\gamma \vec{B}_R \cdot \vec{S}\,, \tag{10.163}$$

in which

$$\vec{B}_R = B_1\hat{x} + B_0\left(1 - \frac{\omega}{\omega_0}\right)\hat{z}\,. \tag{10.164}$$

The RF signal gives a component that rotates around $\hat{x}$ and changes the rotation rate around $\hat{z}$.

We find the time evolution in the original frame by transforming $|\psi_R, t\rangle$ back with the inverse of $U_\omega(t)$:

$$|\psi, t\rangle = U_\omega^\dagger(t)|\psi_R, t\rangle \tag{10.165}$$

$$= \exp\left(\frac{i\omega t S_z}{\hbar}\right)\exp\left(-i\frac{(-\gamma\vec{B}_R\cdot\vec{S})t}{\hbar}\right)|\psi, 0\rangle\,. \tag{10.166}$$

Suppose the initial spin vector is $|z+\rangle$. If $\omega \ll \omega_0$, then $\vec{B}_R \approx B_0\hat{z} + B_1\hat{x}$. This gives a rapid precession around $\vec{B}_R$, and a slow precession around $\hat{z}$.

On the other hand, if $\omega = \omega_0$, then $\vec{B}_R = B_1\hat{x}$. This gives a spiral motion from $|z+\rangle$ down to the equatorial plane! If we start at $|z+\rangle$, the rotation around the $\hat{x}$ axis will move the state vector a little off the "north pole". The contribution from $U_\omega^\dagger(t)$ will then rotate it a little ways around the $\hat{z}$ axis. By the time it has gone all the way around the $\hat{z}$ axis, it will also have gone some ways around $\hat{x}$. Overall, this makes for a spiral.

If we turn on the RF oscillation for just the right amount of time, a so-called "90° pulse", an initial state $|z+\rangle$ will migrate to the $\hat{x}$–$\hat{y}$ plane, and then the field in the $\hat{z}$ direction will keep it rotating there at the Larmor frequency ω_0.

Classically, a rotating magnetic dipole will radiate electromagnetic waves. This suggests that a spin described by a rotating state vector will emit an electromagnetic signal. Although we have not modeled this in any detail, the intuition holds good; this is how magnetic resonance imaging operates. An MRI machine applies a strong magnetic field in one direction, and this overpowers thermal randomness just enough that a small but nonnegligible fraction of the hydrogen atoms within the field can be described by $|z+\rangle$. A 90° degree pulse is then applied, and the machine "listens" for a signal back. From the recorded signal, the positions of the radiating atoms can be determined. (The clunking noises inside an MRI machine are due to the repositioning of movable magnets that make the $\hat{z}$-axis field strength vary with position, so that the Larmor frequency will depend upon location.)

Further developing the theory of NMR requires appreciating that atoms do not exist in isolation. Two decay processes come into play, on different timescales. *Transverse* relaxation, on a timescale denoted T_2 (in the tens of milliseconds for watery materials), sends pure states into the interior of the Bloch ball, towards the $\hat{z}$ axis. This is due to spins in different atoms interacting with each other. *Longitudinal* relaxation happens on a longer timescale T_1 (hundreds of milliseconds). It is due to thermal equilibrium reasserting itself; we can think of T_1 as the time needed for any energy gained from the RF pulse to be lost again. Because T_1 and T_2 depend on the substance in which our nuclear spins exist, deducing them from data allows for different substances to be identified. Medical applications use this for making useful pictures. The chemists have also been able to wring a few decades of progress out of NMR. The magnetic field strength "felt" by a nucleus depends in part on the surrounding electrons, and so close study of NMR data can tell a chemist about molecular structure.

10.7 Entanglement

Let's make a Hilbert space by putting together two smaller Hilbert spaces. For example, let us say that the space $\mathbb{C}^2$ is associated with one particle, and on the other side of the room, we have a second particle associated with its own $\mathbb{C}^2$. If $|\psi\rangle$ is a state vector that we can ascribe to the first system and $|\phi\rangle$ a state vector that we can ascribe to the second, then the tensor product $|\psi\rangle \otimes |\phi\rangle$ is a valid *joint state* for the two systems considered together. A state for a bipartite system that can be written as

a tensor product is called *separable*. We can make a basis for the joint Hilbert space out of separable states. For example, using $|+z\rangle$ and $|-z\rangle$ to stand for the eigenvectors of the σ_z matrix, we can make the basis

$$\{|+z\rangle \otimes |+z\rangle,\ |+z\rangle \otimes |-z\rangle,\ |-z\rangle \otimes |+z\rangle,\ |-z\rangle \otimes |-z\rangle\}\,. \tag{10.167}$$

The full set of pure states for the bipartite system is made by taking linear combinations of these... But not all states in that set can themselves be written as a tensor product! For instance, consider

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|+z\rangle \otimes |+z\rangle + |-z\rangle \otimes |-z\rangle\right)\,, \tag{10.168}$$

which we can also express as a column vector

$$|\psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}\,. \tag{10.169}$$

It is straightforward to check that there are no vectors in $\mathbb{C}^2$ that we can tensor together to make this $|\psi\rangle$. Any such vector must have the form

$$\begin{pmatrix}a\\b\end{pmatrix} \otimes \begin{pmatrix}c\\d\end{pmatrix} = \begin{pmatrix}ac\\ad\\bc\\bd\end{pmatrix}\,. \tag{10.170}$$

So, if the second row is 0, then either the first row or the fourth row or both must also be zero. Vectors that cannot be written as a tensor product are *entangled.*

Looking again at the four tensor-product vectors in Eq. (10.167), we see that they form an orthonormal basis for $\mathbb{C}^4$. This means that the projectors onto those vectors constitute a quantum measurement. Moreover, we can quickly verify that those vectors are the *eigenvectors* of the tensor-product matrix $\sigma_z \otimes \sigma_z$. Calculating the tensor product is particularly easy here:

$$\sigma_z \otimes \sigma_z = \begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix}\,. \tag{10.171}$$

We can just read the eigenvalues off the diagonal. They are simply the products of the eigenvalues of σ_z. So, by writing

$$\sigma_z \otimes \sigma_z = |z+\rangle \otimes |z+\rangle - |z+\rangle \otimes |z-\rangle - |z-\rangle \otimes |z+\rangle + |z-\rangle \otimes |z-\rangle\,, \tag{10.172}$$

we can construct the "eigen-measurement" of $\sigma_z \otimes \sigma_z$ and give it an expectation value, $\mathrm{tr}(\rho\sigma_z \otimes \sigma_z)$.

This is an example of a more general fact. Suppose that A and B are two arbitrary square matrices, and that

$$A|\psi\rangle = a|\psi\rangle,\ B|\phi\rangle = b|\phi\rangle\,. \tag{10.173}$$

Then we use the interplay of matrix multiplication and the tensor product:

$$(A\otimes B)(|\psi\rangle\langle\psi|\otimes|\phi\rangle\phi) = A|\psi\rangle\langle\psi|\otimes B|\phi\rangle\langle\phi| \tag{10.174}$$
$$= ab(|\psi\rangle\langle\psi|\otimes|\phi\rangle\langle\phi|)\,. \tag{10.175}$$

Thus, the eigenvalues of $A\otimes B$ will be the products of the eigenvalues of the factors A and B.

If we have two quantum systems, and E_A is a measurement outcome matrix for the first while E_B is a measurement outcome matrix for the second, then $E_A\otimes E_B$ represents an outcome of a joint measurement on the two systems together. And if both the density matrix and the outcome matrix split into tensor products, then the *probabilities* factor:

$$p(E_A\otimes E_B) = \mathrm{tr}[(E_A\otimes E_B)(\rho_A\otimes\rho_B)] \tag{10.176}$$
$$= \mathrm{tr}[(E_A\rho_A)\otimes(E_B\rho_B)] \tag{10.177}$$
$$= \mathrm{tr}(E_A\rho_A)\mathrm{tr}(E_B\rho_B) \tag{10.178}$$
$$= p(E_A)p(E_B)\,. \tag{10.179}$$

When the probability for the event "E_A and E_B" factors into the product of the probabilities for E_A and for E_B, those events are independent. In such a situation, learning that E_A has occurred does not affect our odds for whether E_B has occurred, and vice versa. (We have to be a little careful here: This includes the case where both events have probability 1, meaning that we are completely confident that E_A will happen and likewise for E_B. Seeing E_A happen is then completely unsurprising.) So, if we describe our pair of qubits by a tensor-product state $\rho_A\otimes\rho_B$, and we apply a measurement that is also a tensor product like the eigen-measurement of $\sigma_z\otimes\sigma_z$, then the outcome we get on one qubit will be *uncorrelated* with the outcome we get on the other. If we feed the $\sigma_z\otimes\sigma_z$ measurement an eigenstate of that matrix, then we will get the outcome corresponding to that eigenstate with 100% probability. The expectation value of $\sigma_z\otimes\sigma_z$ will be $+1$ or -1, depending on which eigenstate we pick, and this sign tells us the *product of the results* for the σ_z measurements on the individual qubits. In other words, the experiment "do the σ_z test on the left qubit and also do the σ_z test on the right qubit" is represented by $\sigma_z\otimes\sigma_z$.

What if our quantum state for our pair of systems is not a tensor product, i.e., it is entangled? Then we can turn that entanglement into correlations, once we decide what we'd like to correlate! Back in the '90s, Chris Fuchs liked to say that entanglement is the "all-purpose flour" of quantum theory. All-purpose flour is not itself food, but we can use it to make cookies or cakes or béchamel sauce. Entanglement is not itself correlation, but we can use it to create correlations once we pick a recipe, that is, a specific measurement.

The same idea goes for multipartite systems involving more than two pieces. The GHZ–Mermin thought experiment we saw earlier uses the tripartite Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$, i.e., the combination of three qubits. Define the vector

$$|\mu\rangle = \frac{1}{\sqrt{2}} \left(|+z\rangle \otimes |+z\rangle \otimes |+z\rangle - |-z\rangle \otimes |-z\rangle \otimes |-z\rangle \right) . \tag{10.180}$$

The X tests and Y tests we considered in the parable of the muffins will be the σ_x and σ_y measurements applied to the respective qubits. We can verify by direct calculation that $|\mu\rangle$ is an eigenvector of $\sigma_x \otimes \sigma_y \otimes \sigma_y$ with eigenvalue $+1$. Likewise, it is also an eigenvector of $\sigma_y \otimes \sigma_x \otimes \sigma_y$ and of $\sigma_y \otimes \sigma_y \otimes \sigma_x$, also with eigenvalue $+1$. But it is an eigenvector of $\sigma_x \otimes \sigma_x \otimes \sigma_x$ with eigenvalue -1. This is exactly the combination of pluses and minuses that we claimed quantum mechanics could provide.

10.8 Motivating the Form of Thermal States

We teach thermodynamics as a science of bulk matter, each body characterizable using a few parameters held constant or smoothly varying. Energy is useful work or waste heat, without much gradation of utility in between. Manipulations are the application of fire or of the sea [27, 86]. By contrast, quantum physics is a subject where bringing out its riches generally requires precision. The most intriguing and compelling kinds of quantum strangeness tend to wash out when we become sloppy in handling the world [87, 88, 89]. Bringing these subjects together is thus a bit of an adventure. My goal here is to provide a derivation of the canonical, or Boltzmann, class of distributions by starting with thermodynamics and just barely getting its toes wet in quantum theory. This originated with wondering whether a standard bit of pedagogy could be run backward, and it may hold some appeal for those who enjoy re-approaching the familiar from a different starting point. Our purpose will not be to study how a quantum state might "thermalize" in dynamical detail, but rather to understand why a particular class of states should be regarded as the "thermal" ones.

The Zeroth Law of thermodynamics states that thermal equilibrium is transitive, and so we can divide up the set of all possible preparations of thermodynamical systems into equivalence classes. In quantum theory, we encode the preparation of a system by ascribing a density matrix — a positive semidefinite operator of unit trace — to it. Thus, the meeting-of-the-minds for these two subjects, the quantization of thermodynamics, ought to center upon finding a method for generating density matrices given an equivalence class. The First Law tells us that energy is conserved. Quantum theory actively resists the idea that the values a physicist "measures" exist prior to the act of measurement [90]. How, then, can we match the call for constancy with the fact that quantum experiments are active interventions? We cannot with integrity live as though a quantum system has an energy value that an energy measurement simply *reveals,* but if we assume our laboratory abilities are limited, we can minimize the sinfulness of pretending otherwise. Mathematically, this means picking a scheme for building density matrices such that they all commute with one another, and so they can all be diagonalized in a common basis. If all the density matrices are

diagonal in this "energy eigenbasis", then we can come as close as quantum theory will allow to treating them as though they express ignorance about a value of the energy that pre-exists measurement [91]. Moreover, unitary time evolutions generated by a Hamiltonian diagonal in this basis will preserve all these density matrices, giving us the best match we can feasibly get with our classical notion of conservation.

We will at first assume that the density-matrix elements depend *only* upon the energy. Once we understand this case, we will generalize.

To summarize: What we want is a scheme for writing diagonal density matrices where each entry on the diagonal is a function of the eigenvalue of an observable. Requiring that our scheme plays nicely with the tensor product rule constrains what that function can be. So, let's devise a scheme for generating density matrices with energy being the key observable, that is, matrices that are diagonal in the energy eigenbasis. We want to express the idea of equivalence classes, so we will say that when two preparations are equivalent, we will use the same function for mapping energy values to matrix elements.

Let f be a function from the real line to the nonnegative real line. The diagonal entries of a system's density matrix will be found by applying f to each of the energy values we allow for that system. Trace normalization means that we have to scale each $f(E_i)$ by the sum of all the $f(E_j)$.

Now, we take two systems each of which are ascribed a density matrix in this manner. The density matrix of the joint system is the tensor product of the two original matrices. We impose the requirement that this larger matrix also be expressible according to our scheme, using the same function from energies to matrix elements because the joint system is in the same thermodynamic equivalence class as the two halves were separately. Thus, playing nice with the tensor product means that

$$\frac{f(E_1 + E_2)}{Z_{AB}} = \frac{f(E_1)f(E_2)}{Z_A Z_B}, \tag{10.181}$$

We have the freedom to rescale the function $f(E)$ by an arbitrary multiplicative constant, because the meaningful values are the probabilities which we get by normalizing the $f(E)$ by their sum (which is the partition function). Consider the case where $E_1 = E_2 = 0$. In this case, Eq. (10.181) simplifies to

$$\frac{f(0)}{Z_{AB}} = \frac{f(0)^2}{Z_A Z_B}. \tag{10.182}$$

We take advantage of the rescaling freedom to say that $f(0)$ is unity, and so the partition function must be multiplicative. We can therefore drop it from Eq. (10.181), obtaining

$$f(E_1 + E_2) = f(E_1)f(E_2), \tag{10.183}$$

which is the Cauchy functional equation, up to a logarithmic transform [92, §2.1.2]. Its solutions are $f(E) = 0$, which we can rule out by normalizability, and $f(E) = e^{g(E)}$, where g is an arbitrary function satisfying $g(E_1 + E_2) = g(E_1) + g(E_2)$. On physical grounds, we can assume that f is continuous, per the general smoothness we expect

in thermodynamics problems. In fact, we only need to require continuity at a single point in order to fix the form of the function f completely:

$$f(E) = e^{-\beta E}, \tag{10.184}$$

where the "coolness" β labels the equivalence classes and so is an invertible function of the temperature. The choice of inserting a minus sign before β is, at this stage, merely a convention. (The term "coolness" for inverse temperature is due to Baez and Pollard [93].) This transformation of energies is known as *Boltzmann weighting,* and a diagonal density matrix ρ with its entries given by this function is a *Gibbs state.*

So far, we have gotten as far as showing that for any temperature T, there is a number β that characterizes the dependence of probability upon energy for systems equilibrated to the temperature T. We do not yet know the functional dependence of β upon T, but we can make some easy inferences of how it ought to behave. To avoid an ultraviolet catastrophe, it is natural to want $f(E)$ to be a decreasing function of E, or in other words, $\beta > 0$. We also anticipate that higher temperatures should make higher energy levels available, so β should be inversely related to the temperature.

To understand the relation between β and T, we import more knowledge about classical thermodynamics. First, we perform the standard calculation for the expectation value of the energy:

$$\langle E \rangle := \sum_i E_i \, p(E_i) = \sum_i \frac{E_i e^{-\beta E_i}}{Z} = -\frac{1}{Z} \sum_i \frac{\partial}{\partial \beta} e^{-\beta E_i} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta}, \tag{10.185}$$

which we can write compactly as

$$\langle E \rangle = -\frac{\partial \log Z}{\partial \beta}. \tag{10.186}$$

Moreover, since we have a probability distribution $f(E)/Z$, we can take the Shannon entropy of it, which is equal to the von Neumann entropy of the diagonal density matrix specified by those probabilities. (The Shannon entropy can be motivated by information-theoretic arguments that make no reference to thermodynamics [94].) The result is

$$\langle -\log p \rangle := -\sum_i p(E_i) \log p(E_i) = \beta \langle E \rangle + \log Z. \tag{10.187}$$

From thermodynamics, we have the phenomenological relation

$$E = F + TS, \tag{10.188}$$

where F is the Helmholtz free energy (meaning that we have finally gotten around to invoking the Second Law). These two expressions coincide if we make the identifications

$$E = \langle E \rangle, \;\; S = k_B \langle -\log p \rangle, \;\; \beta = \frac{1}{k_B T}, \;\; F = -\frac{1}{\beta} \log Z. \tag{10.189}$$

The first identification is the obvious way to match a quantum observable with the quantity that we understood classically. To further motivate the remaining three, note

that we want the free energy to be *extensive*: Given two independent, identically prepared systems, the free energy of the pair should be double that of either system alone. The partition functions multiply, so we turn to the logarithm to make a quantity that adds. Moreover, in classical thermodynamics, the entropy is the negative derivative of the Helmholtz free energy with respect to temperature. So, using the quotient and chain rules,

$$E = F - T\frac{\partial F}{\partial T} = -T^2\frac{\partial (F/T)}{\partial T} = \frac{\partial (F/T)}{\partial (1/T)}. \tag{10.190}$$

Comparing this with Eq. (10.186), we see the parallel between (F/T) and $-\log Z$, and between $1/T$ and β. And if we make these identifications, we can see what the thermodynamic entropy must correspond to:

$$S = \frac{E-F}{T} = k_B\beta\langle E\rangle + k_B \log Z = -k_B\langle(-\beta E - \log Z)\rangle = -k_B\langle \log p\rangle. \tag{10.191}$$

Mathematically, it appears overall that we can say β is proportional to $1/T$, or we can match the von Neumann entropy with the thermodynamic entropy, and either correspondence implies the other. It is not clear at this stage which identification ought to be taken as the more fundamental. However, if we go back and hit the books, we find that one introductory move is to develop a bit of "baby kinetic theory" for a classical ideal gas, and from this development show that if we have two gases in thermal equilibrium, the mean kinetic energies of their atomic center-of-mass motions are equal. The mean CM kinetic energy is therefore a suitable label for the equivalence classes under the Zeroth Law, and the simple choice is to *declare* that temperature is proportional to mean CM kinetic energy [95]. Making the same declaration here would be no less respectable! From this perspective, identifying the characteristic energy scale β^{-1} with the thermodynamic temperature (up to a choice of units) is the more elementary move, and the correspondence between information-theoretic and thermodynamic entropies is the derived one.

To see how these choices of correspondence work in more detail, we consider an *infinitesimal change* of the information-theoretic entropy, which we imagine to be brought about by a small outside influence gently perturbing the energy levels [88, §9.1]. The change in the expected energy itself is

$$d\langle E\rangle = \sum_i [p(E_i)dE_i + E_i dp(E_i)]. \tag{10.192}$$

The differential form of the first law says that

$$dE = đW + đQ, \tag{10.193}$$

where we have crossed the differentials on the right-hand side to indicate that they individually are path-dependent, while their sum is not. Having already asserted that the thermodynamic E maps onto the quantum-mechanical $\langle E\rangle$, we see that we can

identify

$$đ\langle W\rangle = \sum_i p(E_i)dE_i, \tag{10.194}$$

$$đ\langle Q\rangle = \sum_i E_i dp(E_i). \tag{10.195}$$

The first line tells us how much work is done on the system (provided the perturbation is sufficiently slow). The second line expresses how much heat flows into the system.

The change in the entropy is

$$d\left(\langle -\log p\rangle\right) = -\sum_i d\left(p(E_i)\log p(E_i)\right), \tag{10.196}$$

which we can split into a pair of sums:

$$d\left(\langle -\log p\rangle\right) = -\sum_i dp(E_i)\log p(E_i) - \sum_i p(E_i) d\log p(E_i). \tag{10.197}$$

We insert the Boltzmann weighting into the first sum and simplify the second by differentiating the logarithm:

$$d\left(\langle -\log p\rangle\right) = -\sum_i dp(E_i)(-\beta E_i - \log Z) - \sum_i p(E_i)\frac{dp(E_i)}{p(E_i)}. \tag{10.198}$$

Now, however the energy levels are perturbed, the sum of all the probabilities must remain equal to 1, so the sum of all the $dp(E_i)$ has to vanish. Consequently, nothing will be left except

$$d\left(\langle -\log p\rangle\right) = \beta\sum_i dp(E_i)E_i. \tag{10.199}$$

Comparing this with our expression for the expected heat flow, we see that

$$d\left(\langle -\log p\rangle\right) = \beta đ\langle Q\rangle. \tag{10.200}$$

The coolness β is the integrating factor necessary to turn the path-dependent infinitesimal heat flow $đQ$ into an exact differential. Eq. (10.200) indicates what we said above, that assuming one correspondence implies the other.

Now that we possess the functional form of the Gibbs states, we can make use of them. For example, we can have a system to which we have ascribed a Gibbs state ρ_A interact with a second system to which we have assigned an arbitrary state ρ_B. The simplest such evolution is a joint unitary

$$\rho_A \otimes \rho_B \to \rho' := U(\rho_A \otimes \rho_B)U^\dagger. \tag{10.201}$$

This preserves the von Neumann entropy of the joint state:

$$S^{\mathrm{vN}}(\rho') = S^{\mathrm{vN}}(\rho_A) + S^{\mathrm{vN}}(\rho_B). \tag{10.202}$$

The von Neumann entropy is subadditive, meaning that for the *marginals* of the time-evolved state, we have

$$S^{\mathrm{vN}}(\rho') \leq S^{\mathrm{vN}}(\rho'_A) + S^{\mathrm{vN}}(\rho'_B). \tag{10.203}$$

Therefore, we can write

$$\Delta S_A^{\mathrm{vN}} + \Delta S_B^{\mathrm{vN}} \geq 0, \tag{10.204}$$

whereas the conservation of energy implies

$$\Delta\langle E_A\rangle + \Delta\langle E_B\rangle = 0. \tag{10.205}$$

The interaction will not in general preserve the form of the Gibbs state, but we can still compute the quantity $S^{\mathrm{vN}} - \beta\langle E\rangle$ before and after. Prior to the interaction, this is just $\log Z$. Its change works out to be

$$\Delta(S_A^{\mathrm{vN}} - \beta\langle E_A\rangle) = -\mathrm{tr}\rho'_A(\log\rho'_A - \log\rho_A). \tag{10.206}$$

We recognize the expression for the *relative* von Neumann entropy:

$$S^{\mathrm{vN}}(\sigma'|\sigma) := \mathrm{tr}\sigma'(\log\sigma' - \log\sigma). \tag{10.207}$$

This is provably nonnegative, and so

$$\Delta(S_A^{\mathrm{vN}} - \beta\langle E_A\rangle) = -S^{\mathrm{vN}}(\rho'_A|\rho_A) \leq 0. \tag{10.208}$$

Combining this with the subadditivity and energy-conservation relations, we deduce

$$\Delta(S_B^{\mathrm{vN}} - \beta\langle E_B\rangle) \geq 0. \tag{10.209}$$

Therefore, if a system for which our initial state ρ_B is arbitrary interacts with a system in thermal equilibrium, this quantity must generally increase. Repeatedly interacting with systems at the same temperature, our state for system B tends toward a limit where

$$S_B^{\mathrm{vN}} - \beta\langle E_B\rangle = -\sum_i p(E_i)(\log p(E_i) + \beta E_i) \tag{10.210}$$

is maximized, assuming no selection rules that artificially constrain the accessibility of energy levels. Maximing this quantity with the condition that $p(E_i)$ is properly normalized recovers the Gibbs state at coolness β [88, §9.2].

Now, we return to the question of dependencies on quantities other than the energy. What if there exist other "observables" that commute with each other and with the Hamiltonian, so that we can simultaneously diagonalize the whole lot of them? The above argument shows that interacting with systems that have been ascribed Gibbs states will generally tend to wash out dependencies upon the other "quantum numbers".

We conclude that the canonical mapping from energies to probabilities is just a prior that plays nicely with the tensor product.

Very often, what people do is to go the other way round, starting with the formula for the von Neumann entropy and then looking for states that maximize it given some

constraint. This works! As mentioned above, that yields the Gibbs states. But it's reasonable to ask why one would maximize the von Neumann entropy instead of any other function. The more time I put into convincing myself that, yes, probability theory does make sense in quantum physics after all [96], the less convinced I was that the von Neumann entropy was the "obvious" choice to maximize.

Brandenburger and Steverson also used the Cauchy functional equation to obtain the form of Gibbs states, from a different starting point [97]. They do not invoke quantum concepts, and they do not investigate in depth the question of how the β that arises from solving the Cauchy functional equation relates to the T of thermodynamics. Yunger Halpern and Renes arrive at a special case of the Cauchy functional equation in Lemma 8 of [98], using more details about "resource theory" than we have developed here, and proceed from there to the Gibbs states. (The fact that Gibbs states are all diagonal in the preferred basis makes them poor resources, from a quantum perspective; nonzero off-diagonal elements, or "coherences", are thermodynamically useful. This is consistent with our ethos of obtaining the Gibbs states by merely dipping our toes into the possibilities of quantum theory and resolutely holding out for quasiclassicality. In turn, there is the possibility of pushing still further into the quantum by seeking states that are *robustly* coherent under *changes* of basis [99].) Alterman and coauthors have also studied the origin of Boltzmann weighting and notions of quantum-classical correspondence, in a setting that uses much more detail about specific potential wells than we have here [100].

The fact that canonical partition functions for independent systems are multiplicative is a helpful and easy standard result. Here, we have seen that we can run that bit of pedagogy backwards, starting from a desideratum for how composition behaves and arriving at the Boltzmann weighting $e^{-\beta E}$. It is interesting to step back and consider just what portions of the quantum formalism we used in obtaining this result. To get the Gibbs states themselves, we did not invoke in much detail at all how quantum systems can be correlated, only using the tensor-product rule for compositing state spaces. Further knowledge about the space of possible correlated joint states did not enter until we started talking about thermalization. It is known that we can use more fundamental assumptions, like the impossibility of faster-than-light signaling, to argue that "locally quantum" systems should combine according to the tensor product rule [101, 102]. Whatever the space of joint states is, it at least contains the tensor-product states. Showing that the joint states are specifically positive semidefinite operators on that tensor-product Hilbert space requires further assumptions [103, 104]. However, we did not have to rely upon the positivity of correlated states. We can just say that each system is "locally quantum" and that when we consider them together but uncorrelated (on opposite sides of the lab, perhaps), we use the tensor product. This is enough to power the primary result we obtained above.

11 Quantum Mechanics on a Line

There are at least three ways to go from the study of qubits, as we have developed over the previous chapters, to the quantum physics of a particle on a line. Four, if you count just postulating the answer. One method uses a limiting argument that is a little subtle, and the others rely upon more advanced ways of doing classical physics. We will cover all three options.

11.1 Canonical Quantization

In this approach, we lean upon the Hamiltonian formulation of classical mechanics. Recall that in that formulation, we derived the Liouville equation (8.18) for the time evolution of a probability density on phase space:

$$\frac{\partial \rho}{\partial t} = \{\mathcal{H}, \rho\}. \tag{11.1}$$

We observe a notational similarity between the Liouville equation and the equation that describes the time evolution of a density matrix. Back in Eq. (10.63), we said that given a generator A, the time derivative of the density matrix is

$$\frac{\partial \rho_t}{\partial t} = [-iA, \rho_t]\,. \tag{11.2}$$

This resemblance is more than just a trick of notation. The Poisson bracket and the operator commutator satisfy the same algebraic properties. Both are antisymmetric. And though this is harder to prove, both satisfy the *Jacobi identity*:

$$[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0\,, \tag{11.3}$$

$$\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0\,. \tag{11.4}$$

Dirac had the idea that if we write classical physics using Poisson brackets, we can invent a quantum theory by carrying over the form of the equations and turning variables into operators, in such a way that each Poisson bracket becomes a commutator. This means that each fact about classical physics that we can deduce from the algebraic properties of the brackets alone will have a counterpart in the quantum theory, which should reassure us, because we want to recover the classical predictions in the regime where they have already been validated.

The Poisson bracket of position and momentum variables is unitless:

$$\{x, p\} = 1\,, \tag{11.5}$$

whereas the product and thus the commutator of the quantum operators will have units of angular momentum. Fortunately, from historical considerations, we have a constant with units of angular momentum sitting around: $\hbar$. Can we just declare that the commutator of x and p is $\hbar$? Not quite. Taking the adjoint of the commutator must flip its sign, and $\hbar$ is a real-valued constant unaffected by the adjoint. But we *can* declare

$$[x, p] = i\hbar\,. \tag{11.6}$$

This is known as the *canonical commutator*. It has something of a complicated history. Heisenberg was doing calculations about spectral lines, and Max Born noticed that the operation Heisenberg had invented for combining quantities in his equations was actually matrix multiplication. From this came the idea that in quantum mechanics, position and momentum became matrices with infinitely many rows and columns. The connection with Poisson brackets was found shortly thereafter by Dirac.

We follow the suggestion of turning each Poisson bracket into a commutator divided by $i\hbar$. Then the Liouville equation becomes

$$\dot{\rho}_t = \frac{1}{i\hbar}[H, \rho]\,. \tag{11.7}$$

Now ρ is an operator, and so is the Hamiltonian H. We can take the prefactor into the commutator:

$$\dot{\rho}_t = [-iH/\hbar, \rho]\,. \tag{11.8}$$

This agrees with Eq. (10.63), if we identify the Hamiltonian as the generator of time evolution. Unitary evolution in our theory will then be implemented by operators

$$U_t = e^{-iHt/\hbar}\,, \tag{11.9}$$

where H is something we can write as a function of the position and momentum operators:

$$H = \frac{p^2}{2m} + V(x)\,. \tag{11.10}$$

We have not said anything yet about what these position and momentum operators *are* in terms of anything else, only that they must obey the canonical commutator. Also, we have not specified what kind of operator ρ is. What functions do these operators act upon? We can make progress on this by carrying over more ideas from our qubit studies, with the new proviso that a measurement of the energy has potentially infinitely many possible outcomes. For example, ρ should have trace 1, expectation values should be calculated by $\mathrm{tr}(A\rho)$, and the pure states should be projectors $|\psi\rangle\langle\psi|$. We should still be able to write an arbitrary pure state as

$$|\psi\rangle\langle\psi| = \sum_n \psi_n |n\rangle\langle n|\,, \tag{11.11}$$

where the $|n\rangle$ are eigenvectors of the Hamiltonian:

$$H|n\rangle = E_n|n\rangle\,. \tag{11.12}$$

To summarize: One approach to developing quantum theory is to start with classical mechanics and say, "OK, the variables you knew before are now becoming self-adjoint operators on a vector space of continuous functions, and the position and momentum operators satisfy the canonical commutator." This is a good way to take inspiration from classical calculations for inventing their quantum counterpart. It is a somewhat ambiguous procedure, since the canonical commutator itself implies that the order of the operators must matter. Say that a classical equation involves the product xp. Should the quantum counterpart use xp, px, $(xp + px)/2$ or what?

11.2 Continuum Limit of a Lattice

There is an intriguing plausibility argument for how to go from a particle on a 1D lattice to a particle free to move on a line. In essence, it is a *renormalization group* argument, where the crucial step is adjusting parameters that an experimenter cannot access directly, in such a way that a quantity that is physically relevant in the lab remains fixed. This is a subtle maneuver. (The *Feynman Lectures on Physics* present this in a rather scattered way: In volume 3, chapter 16 refers back to chapter 13, which itself points back to chapter 48 of volume 1.)

Let's warm up by considering a model where we have oscillation of probability between two locations. We set the dimension of the Hilbert space at 2, and we choose the Hamiltonian matrix

$$H = E_0 I - \alpha \sigma_x. \tag{11.13}$$

Explicitly, this is

$$H = \begin{pmatrix} E_0 & -\alpha \\ -\alpha & E_0 \end{pmatrix}. \tag{11.14}$$

Restricting our attention to the case of pure states,

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \tag{11.15}$$

we can write a pair of differential equations for the individual vector components:

$$i\hbar \frac{d\psi_1}{dt} = E_0 \psi_1 - \alpha \psi_2, \tag{11.16}$$

$$i\hbar \frac{d\psi_2}{dt} = -\alpha \psi_1 + E_0 \psi_2. \tag{11.17}$$

Adding these two equations gives a differential equation for $\psi_1 + \psi_2$ that we can solve:

$$\psi_1 + \psi_2 = a e^{-(i/\hbar)(E_0 - \alpha)t}. \tag{11.18}$$

Likewise, we can subtract them to get a differential equation for $\psi_1 - \psi_2$, which we can solve to yield

$$\psi_1 - \psi_2 = b e^{-(i/\hbar)(E_0 + \alpha)t}. \tag{11.19}$$

Adding and subtracting these gives us

$$\psi_1 = \frac{a}{2}e^{-(i/\hbar)(E_0-\alpha)t} + \frac{b}{2}e^{-(i/\hbar)(E_0+\alpha)t}, \tag{11.20}$$

$$\psi_2 = \frac{a}{2}e^{-(i/\hbar)(E_0-\alpha)t} - \frac{b}{2}e^{-(i/\hbar)(E_0+\alpha)t}. \tag{11.21}$$

To illustrate, let's say our initial state is

$$|\psi(0)\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{11.22}$$

Then

$$\frac{a+b}{2} = 1, \ \frac{a-b}{2} = 0, \tag{11.23}$$

which we solve by setting

$$a = b = 1. \tag{11.24}$$

Inserting these coefficients into our time-evolution equations and using Euler's formula to simplify, we obtain

$$\psi_1 = e^{-(i/\hbar)E_0 t}\cos\frac{\alpha t}{\hbar}, \tag{11.25}$$

$$\psi_2 = e^{-(i/\hbar)E_0 t}\sin\frac{\alpha t}{\hbar}. \tag{11.26}$$

Taking the squared magnitude of these gives probabilities that fluctuate like squared trig functions, oscillating between 0 and 1. A measurement near $t = 0$ is likely to find the particle at site 1, whereas a measurement near $t = \pi\hbar/(2\alpha)$ is likely to find it at site 2.

Next, instead of having two sites, what if we had an arbitrarily long chain of them? We suppose that they are ranged along the x axis, a distance b apart from each other, so the coordinate of site number n is

$$x_n = nb. \tag{11.27}$$

The generalization of our two-site Hamiltonian gives a whole host of differential equations, each with a pair of "hopping terms" that connect the probability amplitudes for adjacent sites:

$$i\hbar\frac{d\psi_n}{dt} = E_0\psi_n - \alpha\psi_{n-1} - \alpha\psi_{n+1}. \tag{11.28}$$

When confronted with a differential equation, it is seldom a bad idea to plug in an exponential function and see what happens. Accordingly, we try

$$\psi_n = a(x_n)e^{-iEt/\hbar} \tag{11.29}$$

with

$$a(x_n) = e^{ikx_n}. \tag{11.30}$$

If this works, it will give us a solution with energy E. The time derivative pulls down a $-iE/\hbar$, and all the terms on both sides will be proportional to $e^{-iEt/\hbar}$, so we divide through by that, obtaining

$$Ea(x_n) = E_0 a(x_n) - \alpha a(x_{n+1}) - \alpha a(x_{n-1}). \tag{11.31}$$

Using our guess for the position dependence, this is

$$Ee^{ikx_n} = E_0 e^{ikx_n} - \alpha e^{ik(x_n+b)} - \alpha e^{ik(x_n-b)}. \tag{11.32}$$

We can cancel the dependence on x_n, yielding

$$E = E_0 - \alpha e^{ikb} - \alpha e^{-ikb}, \tag{11.33}$$

which simplifies to

$$E = E_0 - 2\alpha \cos kb. \tag{11.34}$$

A particularly nice special case is when $E_0 = 2\alpha$, which fixes a minimum energy of zero. For small kb, we then have

$$\cos kb \approx 1 - \frac{k^2 b^2}{2}, \tag{11.35}$$

which implies that the energy depends quadratically on k:

$$E = \alpha k^2 b^2. \tag{11.36}$$

We can also relate the energy to the rate at which the phase cycles with time:

$$\omega = \frac{E}{\hbar}. \tag{11.37}$$

If we pile up a lot of plane waves with nearby frequencies to make a lump, what velocity can we say the lump has? The answer is called the *group velocity,* which is calculated by

$$v = \frac{d\omega}{dk}. \tag{11.38}$$

One way to see this emerge is to add two waves:

$$e^{i(\omega_1 t - k_1 x)} + e^{i(\omega_2 t - k_2 x)} = e^{i[(\omega_1+\omega_2)t-(k_1+k_2)x]/2}[e^{i[(\omega_1-\omega_2)t-(k_1-k_2)x]/2} + e^{-i[(\omega_1-\omega_2)t-(k_1-k_2)x]/2}]. \tag{11.39}$$

This is a fast oscillation (a wave of high frequency) modulated by a slow one. If the frequencies ω_1 and ω_2 are nearly equal, then their average is nearly equal to both of them. The speed of the fast oscillation is $(\omega_1 + \omega_2)/(k_1 + k_2)$, but the speed of the modulation is

$$v_{\text{mod}} = \frac{\omega_1 - \omega_2}{k_1 - k_2}. \tag{11.40}$$

In the limit that both differences become very small, this goes over to the derivative. A more elaborate calculation using Taylor series will end in the same place.

Let us calculate the velocity of a lump made by superposing many waves of nearby frequencies:

$$\frac{d\omega}{dk} = \frac{\alpha b^2}{\hbar}\frac{d}{dk}k^2 = \frac{2\alpha b^2}{\hbar}k. \tag{11.41}$$

The velocity is proportional to k, so the energy is proportional to the square of the velocity. This is a kind of relation we are familiar with; we've seen it for years, writing the proportionality constant as half of the mass.

$$E = \frac{1}{2}m_{\text{eff}}v^2, \tag{11.42}$$

with

$$m_{\text{eff}} = \frac{\hbar^2}{2\alpha b^2}. \tag{11.43}$$

The product of the effective mass and the velocity, which ought to be an effective momentum of some kind, nicely gives us the de Broglie relation

$$m_{\text{eff}}v = \hbar k. \tag{11.44}$$

We originally defined the dynamics using a set of coupled differential equations, which in terms of the lattice spacing b are

$$i\hbar\frac{d\psi(x_n)}{dt} = E_0\psi(x_n) - \alpha\psi(x_n + b) - \alpha\psi(x_n - b). \tag{11.45}$$

Let's suppose that a physicist Alice applies this model to some system in her lab, plugging in what she thinks are reasonable figures for α and b. She obtains the effective mass m_{eff}. Then Bob comes along and says, "I buy that value for m_{eff}, but I could have sworn the lattice spacing wasn't your number b. I think it is b' instead." The effective mass is directly relevant to experiment; it is a thing that both of them know how to measure. But the lattice spacing is much, much smaller than any ruler to which they have access.

Alice says, "Fine, then. We will get the same m_{eff} using your value for the lattice spacing, *if* we also change the transition parameter from α to α'." We can absorb the effect of turning one knob in our model by a corresponding turn of another.

We decrease the lattice spacing b, adjusting α so that we keep the effective mass m_{eff} constant. We rewrite the previous differential equation slightly by adding and subtracting $2\alpha\psi(x_n)$:

$$i\hbar\frac{d\psi(x_n)}{dt} = (E_0 - 2\alpha)\psi(x_n) + \alpha[2\psi(x_n) - \psi(x_n + b) - \psi(x_n - b)]\,. \tag{11.46}$$

The quantity in brackets is a discrete approximation to the second derivative:

$$2\psi(x_n) - \psi(x_n + b) - \psi(x_n - b) \approx -b^2\frac{\partial^2\psi}{\partial x^2}\,. \tag{11.47}$$

So, we take a flying leap and write the continuum counterpart of our discrete-space differential equation:

$$i\hbar\frac{\partial\psi}{\partial t} = -\left(\frac{\hbar^2}{2m_{\text{eff}}}\right)\frac{\partial^2\psi}{\partial x^2}\,. \tag{11.48}$$

The function $\psi(x,t)$ is called the *wavefunction,* a term that is often used synonymously with "quantum state" (although it is not as general as our notion of quantum state, since a wavefunction is a pure state). We carry over the idea of squaring the absolute value of a complex number to get a probability, as in Eq. (10.73), but now we express it using a *probability density.* The probability that a position measurement gives a result in the tiny interval between x and $x + dx$ is $|\psi(x,t)|^2 dx$. And so the wavefunction is required to be normalized:

$$\int dx\,|\psi(x,t)|^2 = 1\,, \tag{11.49}$$

where the integral goes over the domain on which $\psi(x,t)$ is defined. Moreover, the expectation value of a self-adjoint operator is

$$(\psi, A\psi) = (A\psi,\psi) = \int dx\,\psi(x,t)^* A\psi(x,t)\,. \tag{11.50}$$

11.3 Hamilton–Jacobi

The path we follow in this section is historically inspired, but not exactly historical. We take ideas that were in the air by 1925 or so, and by picking the right combination ideas in retrospect we can leap our way to the answer.

Which notions from the "old quantum theory" (1900–1925) are we mixing into our brew? First, let's take the principle from the Bohr model that states of fixed energy are stationary: If they evolve over time, it is in a way that means the statistical predictions we derive from them do not change. Second, we adopt Bohr's suggestion for going beyond that model—that transition rates are related to the coefficients in a Fourier expansion. Third, we recognize that quantum predictions ought to reproduce, or reduce to, classical predictions in those regimes where classical physics has already proven itself useful. Combining these latter two ideas leads us to posit that the probability of detecting a particle goes as the squared absolute value of a complex number. This way, when we consider an aggregate of many quanta, our prediction for the most likely outcome should come to resemble the classical equations for wave intensity. Fourth, we go all the way back to Planck and recognize that we have a physical constant with dimensions of angular momentum that we can invoke.

What wave equation should our complex-valued field satisfy? Here we turn for inspiration to the formulation of classical physics that makes particle dynamics look analogous to wave dynamics, *Hamilton–Jacobi theory.* We derive this from the "Newtonian" way of doing classical physics in §24.5. For a single particle, it is not so bad at all. The particle momentum is perpendicular to the lines of constant value in a field, in analogy to how a light ray propagates perpendicularly to its wave fronts. Therefore,

we write the momentum as the gradient of a field:

$$\vec{p} = \vec{\nabla} S \,. \tag{11.51}$$

The field S satisfies a differential equation involving the Hamiltonian function:

$$-\frac{\partial S}{\partial t} = H(\vec{x}, \vec{\nabla} S, t) \,. \tag{11.52}$$

Now, we invent a quantum counterpart to this theory. We zoom in on the idea that particle momentum is along a gradient, i.e., perpendicular to some surfaces of constant value. This is our act of inspiration: We turn the classical surfaces of constant S into the surfaces of constant *phase* in our quantum wave. The more quickly the phase cycles as a function of position, the bigger the momentum (or the *predicted* value of the momentum at the end of the calculation). If in some region the spatial dependence of our complex-valued wave goes as e^{ikx}, the associated momentum should be proportional to k. Moreover, saying that S corresponds to the phase feels right, because a solution will be stationary (like a Bohr energy level) if it changes over time by an overall phase factor. In other words, the change in phase with respect to time corresponds to the classical energy, and the change in phase with respect to position corresponds to the classical momentum, insofar as that is possible.

We put these ideas together into a wave equation for our complex-valued wave ψ:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \,, \tag{11.53}$$

where we have brought in the Planck constant to make the units work, and H is the classical Hamiltonian adapted by replacing every instance of $\vec{p}$ with $-i\hbar\vec{\nabla}$. This gets us to the punchline of the previous section, if we plug in the classical Hamiltonian for free space.

We will see more about how all these expressions hold together as we go.

11.4 Whose Equation?

The continuum counterpart that we have presented above is ubiquitously known as the *Schrödinger equation* for a free particle in one dimension. But is that really the name we should use? Setting aside the fact that Schrödinger's personal life makes him a one-man argument for not naming ideas after people, there is an interesting question of historical priority.

Wikipedia claims that Arthur C. Lunn discovered what we now call the Schrödinger equation some years before Schrödinger. Naturally, I wonder if there is more to say about this than what the references cited there there [105, 106, 107] provide (they have the feel of being faithful recollections, but are light on specifics).

The hearsay is well attested, at least. In a 1964 interview, the physicist Karl Darrow calls the story "impossible to check" [108]. And in another interview, Robert Mulliken (not to be confused with Robert Millikan) shares the story of Lunn having "sent a paper

to the Physical Review which was turned down and which anticipated the quantum mechanics" [109]. Mulliken heard the story from the physical chemist William Draper Harkins. Similarly, Leonard Loeb told Thomas Kuhn that Lunn "was probably a misunderstood genius, and who was completely frustrated, because his one great paper with his one great idea was turned down by a journal" [110].

Lunn did apparently try to present what sounds like a grandiose paper ("Relativity, quantum theory, and the wave theories of light and gravitation") at the American Physical Society meeting in April 1923, but his paper was only "read by title". The abstract ran as follows:

> This paper is a preliminary report on a theory originally sought in order to meet the recognized need for a reconciliation between wave theory and quantum phenomena; its scope of adaptation proves to be quite wide. It includes (1) a wave theory of gravitation in quantitative connection with optical, electronic, and radioactivity data; (2) a related general suggestion of a theory connecting molecular properties with properties of matter in bulk; (3) alternatives for some of the current features in the theories of atomic structure; (4) a new interpretation and deduction of formulas for series and band spectra, using in lieu of the quantum condition a substitute directly related to long familiar physical notions; (5) a modification of Lagrangian dynamics which promises to be of service in the study of complex atomic and molecular structures; (6) a non-quantum theory of specific heat and black radiation. Results so far reached deal mostly with problems approachable by elementary methods or approximate computations. A set of formulas has been obtained which yield computation of the electron constants e, h, m and mass ratios, assuming from observation only the Rydberg constant, velocity of light, gravitation constant, and Faraday constant, with results in each case in practical agreement with measured values.

To a modern eye, this is way, way too much. Gravitation and molecular structure? Pshaw! But, conceivably, at the time these problems could have seemed in closer proximity to each other than they are now.

Darrow says, "I know that in 1924 he wanted to give a twenty or a thirty minute paper before the American Physical Society in Washington, but then authorities of the Society refused him more than ten minutes".

Lunn's abstract in the 1924 proceedings has a similar explain-everything atmosphere:

> **Relativity, the quantum phenomena, and a kinematic geometry of matter and radiation.** A. C. LUNN, University of Chicago. The theory indicated in an earlier paper (Phys. Rev. 21, 711, 1923), has since been developed, extended in scope, and so ordered as to permit of treatment as a deductive space-time geometry. It unites the treatment of the quantum phenomena with the rest of physical theory in a way that yields to illustration by familiar physical images. It resolves into matters of choice a number of hitherto controversial alternatives in the interpretation of phenomena,

> and allows freedom of use of a range of concrete types of representation including many other concepts commonly discarded. Among special topics more recently found to affiliate with the scheme may be mentioned the Stark and Zeeman effects and fine structure, resonance potentials, and the intensity and distribution of general x-radiation. Improvements have been made in the setting of the formulas connecting e, h, and m with pre-electron data. A program has emerged for the foundation of a trial mathematical chemistry by determination of types of atoms, valence, number of isotopes, atomic weights, and spectrum levels.

I can easily imagine a paper with that attempted scope being incomprehensible to whoever had the task of evaluating it, and so any really good morsels within it would have been lost.

I wrote to the *Physical Review* offices on the chance that they had more information and received this reply from Robert Garisto, the Managing Editor of *Physical Review Letters*.

> Thank you for your query. Our records from the early 20th century are fragmentary. I am not sure if we have any from before 1930, much less a complete set that could answer your question.
>
> But I see that Arthur C. Lunn published 7 papers in the Physical Review from 1912–1922. So he was a known author to the editors. Those were different times, and while it is possible that he submitted a paper that was rejected and never published elsewhere, for what it's worth, it strikes me as unlikely.

11.5 Particle in Quarantine

In order to understand the differential equation we have arrived at, let us solve it for the case where $\psi(x,t)$ is restricted to vanish outside an interval. This is known as a *particle in a box,* with the "box" here being one-dimensional. We will take our interval to be the region from $x = 0$ to $x = L$:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}\,,\ 0 \le x \le L\,. \tag{11.54}$$

The time dependence will be nicely simple if t enters in a complex exponential:

$$\psi(x,t) = \psi(x)e^{-iEt/\hbar}\,. \tag{11.55}$$

Then, the partial derivative with respect to t will pull down a factor $-iE/\hbar$, and we will get

$$E\psi(x) = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2}\,. \tag{11.56}$$

So, the spatial part $\psi(x)$ will be a function that, when differentiated twice, is proportional to minus itself. This means that we look to the trigonometric functions. We

use the boundary conditions to select the sine waves

$$\psi(x) = C \sin(kx)\,, \tag{11.57}$$

where

$$k^2 = \frac{2mE}{\hbar^2}\,,\ kL = \pi n\,. \tag{11.58}$$

The eigenvalues of the Hamiltonian — that is, the energy levels — are thus

$$E_n = \frac{\hbar^2 \pi^2 n^2}{2mL^2}\,, \tag{11.59}$$

with $n = 1, 2, 3, \ldots$, since the $n = 0$ solution would be identically zero and thus not normalizable. Integrating $|\psi(x)|^2$ over the allowed interval fixes the normalization:

$$|C| = \sqrt{\frac{2}{L}}\,. \tag{11.60}$$

Significantly, as the energy increases, the number of *nodes* in the ψ-function also increases. This turns out to be a general feature of the theory. We sketch an argument as to why. Suppose that ψ_n and ψ_{n+1} are eigenfunctions of the Hamiltonian defined by a potential $V(x)$, with respective eigenvalues E_n and E_{n+1} satisfying $E_n < E_{n+1}$. Then one can show that

$$\left(\psi_{n+1}\psi_n' - \psi_n\psi_{n+1}'\right)\Big|_a^b = \frac{2m}{\hbar^2}(E_{n+1} - E_n)\int_a^b dx\, \psi_n\psi_{n+1}\,. \tag{11.61}$$

If a and b are the locations of two successive nodes of ψ_n, and $\psi_n(k) > 0$ within this interval (which we can assume without loss of generality), then it follows from the previous equation that ψ_{n+1} must change sign somewhere in the interval $a < x < b$. This in turn implies that ψ_{n+1} must have at least one more node than ψ_n.

11.6 Particle Subject to a Potential

In the previous section, every location in the allowed interval was alike. The next step is to introduce something into our equations that indicates that one place can be different from another. We do this by adding a *potential* term to our differential equation:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V(x)\,. \tag{11.62}$$

The bulk of introductory quantum mechanics involves solving this equation for different choices of $V(x)$. Again, and again.

We can appreciate the physical meaning of this equation and relate it to more familiar ideas by considering expectation values and how they can change over time.

We start with $\langle A\rangle = \langle\psi|A|\psi\rangle$ and differentiate:

$$\frac{d}{dt}\langle A\rangle = \left(\frac{d}{dt}\langle\psi|\right)A|\psi\rangle + \langle\psi|A\left(\frac{d}{dt}|\psi\rangle\right) \tag{11.63}$$

$$= \frac{i}{\hbar}\langle\psi|HA|\psi\rangle - \frac{i}{\hbar}\langle\psi|AH|\psi\rangle \tag{11.64}$$

$$= \frac{i}{\hbar}\langle[H,A]\rangle\,. \tag{11.65}$$

This works just as well in the discrete and the continuous settings. If we use our continuum Hamiltonian

$$H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\,, \tag{11.66}$$

and we plug in the most familiar self-adjoint operator, x, we get

$$\frac{d}{dt}\langle x\rangle = \frac{i}{\hbar}\left\langle\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}, x\right] + [V(x), x]\right\rangle. \tag{11.67}$$

Since $V(x)$ is just a function of x, like a polynomial, then it will commute with x. The other commutator is more interesting. We can evaluate it by applying it to a test function f:

$$\frac{\partial^2}{\partial x^2}(xf) - x\frac{\partial^2}{\partial x^2}f = \frac{\partial}{\partial x}\left(f + x\frac{\partial f}{\partial x}\right) - x\frac{\partial^2 f}{\partial x^2} \tag{11.68}$$

$$= 2\frac{\partial f}{\partial x}\,. \tag{11.69}$$

Therefore,

$$\frac{d}{dt}\langle x\rangle = -\frac{i\hbar}{2m}\left\langle 2\frac{\partial}{\partial x}\right\rangle = \frac{1}{m}\left\langle -i\hbar\frac{\partial}{\partial x}\right\rangle. \tag{11.70}$$

Classically, the derivative of position with respect to time is the velocity, and multiplying that by the mass gives the momentum. This suggests identifying the expression on the right-hand side as the *momentum operator*:

$$p = -i\hbar\frac{\partial}{\partial x}\,. \tag{11.71}$$

Using this definition, our Hamiltonian operator becomes

$$H = \frac{p^2}{2m} + V(x)\,, \tag{11.72}$$

which looks just like the classical sum of kinetic and potential energies.

Recalling the commutator of multiplication and differentiation, Eq. (22.148), we deduce that

$$[x,p] = i\hbar\,. \tag{11.73}$$

We have regained the canonical commutator!

To understand something of the meaning of this p operator, consider a differentiable function f, and another function g that is just f but shifted to the right by an amount dx. In other words, the value of $g(x)$ is the same as evaluating f at a point dx to the left:

$$g(x) = f(x - dx)\,. \tag{11.74}$$

By Taylor series,

$$g(x) = f(x) - dx f'(x)\,. \tag{11.75}$$

Applying the operator $-\frac{d}{dx}$ tells us how to shift a function along the x axis. The operator p takes this and makes it self-adjoint. In other words, p is the generator of translations along x.

Suppose that $\rho(x, p)$ is a classical probability density function on a phase space. The Liouville equation (8.18) describes its time evolution. If we take its Poisson bracket with the classical momentum variable p, we get

$$\{p, \rho\} = \frac{\partial p}{\partial x}\frac{\partial \rho}{\partial p} - \frac{\partial \rho}{\partial x}\frac{\partial p}{\partial p} \tag{11.76}$$

$$= -\frac{\partial \rho}{\partial x}\,. \tag{11.77}$$

Just as bracketing with the Hamiltonian provides a nudge forward in time, bracketing with the momentum provides a nudge in position. This idea carries over into quantum mechanics.

11.7 Harmonic Oscillator

The next simplest Hamiltonian is the harmonic oscillator. Considering all the models for which it is the prototype, it might be the most important one-dimensional Hamiltonian to understand. We can encapsulate the dynamics of the quantum harmonic oscillator by selecting the Hamiltonian operator

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2\,. \tag{11.78}$$

This has the same form as the Hamiltonian for the classical harmonic oscillator, with x as the spatial coordinate and p playing the role of its canonical conjugate momentum. Upon identifying the operator p as $-i\hbar\partial_x$, this becomes

$$H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}m\omega^2 x^2\,. \tag{11.79}$$

To recall a bit about the classical harmonic oscillator, let us use Newton's laws to find the rates of change of position and momentum. The classical potential energy is $\frac{1}{2}m\omega^2 x^2$, and the force is minus the derivative of this with respect to position. The force in turn tells us the time derivative of the momentum:

$$\dot{p} = -m\omega^2 x\,. \tag{11.80}$$

The time derivative of the position is the momentum divided by the mass.

$$\dot{x} = \frac{p}{m}\,. \tag{11.81}$$

Thus, the *second* time derivative of the position must give us the position again, with a minus sign and a constant prefactor:

$$\ddot{x} = \frac{1}{m}\dot{p} = -\omega^2 x\,. \tag{11.82}$$

The solutions of this differential equation are sinusoids of frequency ω, possibly shifted and scaled.

As we did for the particle in a 1D box, let's assume that our $\psi(x,t)$ factors into spatial and a temporal parts. The eigenvalue equation is then

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}m\omega^2 x^2\right)\psi = E\psi\,. \tag{11.83}$$

The potential is symmetric around the origin, so let's look for a solution that is a symmetrical blob of some sort. Exponentials are easy to differentiate, and squaring x is a nice way to make a function that is symmetric around $x = 0$. We pick for an ansatz:

$$\psi(x) = ae^{-bx^2}\,. \tag{11.84}$$

Differentiating this $\psi(x)$ is easy:

$$\psi' = -2bx\psi\,, \tag{11.85}$$

$$\psi'' = -2b\psi + 4b^2x^2\psi\,. \tag{11.86}$$

Plugging this second derivative into the eigenvalue equation gives

$$-\frac{\hbar^2}{2m}\left(4b^2x^2\psi - 2b\psi\right) + \frac{1}{2}m\omega^2x^2\psi = E\psi\,. \tag{11.87}$$

The function ψ now drops out, yielding

$$-\frac{2b^2\hbar^2x^2}{m} + \frac{\hbar^2 b}{m} + \frac{1}{2}m\omega^2x^2 = E\,. \tag{11.88}$$

The right-hand side has no x dependence, so to be consistent, we need the energy E to equal the term on the left that does not include an x:

$$E = \frac{\hbar^2 b}{m}\,. \tag{11.89}$$

And we need all the contributions that do include an x to cancel, which we can achieve by setting

$$\frac{1}{2}m\omega^2 = \frac{2b^2\hbar^2}{m}\,. \tag{11.90}$$

Solving for the parameter b, we get

$$b = \frac{m\omega}{2\hbar}\,. \tag{11.91}$$

And so the energy of this solution is

$$E = \frac{1}{2}\hbar\omega\,. \tag{11.92}$$

This is the *ground state energy* of the harmonic oscillator.

Having figured out b in terms of the oscillator's physical parameters m and ω, we see that

$$\frac{\hbar}{m\omega}\psi' = -x\psi\,. \tag{11.93}$$

So, applying a particular combination of multiplications and differentiations to our ψ will make it vanish:

$$\left(x + \frac{\hbar}{m\omega}\frac{\partial}{\partial x}\right)\psi = 0\,. \tag{11.94}$$

Being annihilated in this way characterizes the ground state. And this provides a clue as to how we can lever our way up and find *all* the eigenvalues and eigenfunctions. If the lowest energy level is $\frac{1}{2}\hbar\omega$, can we find a new way of writing the Hamiltonian so that it is $\frac{1}{2}\hbar\omega$ plus something? The "something" will have to be an operator that, when applied to the ground state, gives 0.

Following this lead, we solve the quantum harmonic oscillator by introducing new operators,

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{ip}{m\omega}\right) \tag{11.95}$$

and

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{ip}{m\omega}\right), \tag{11.96}$$

known as the annihilation and creation operators, respectively. Rewriting the Hamiltonian in terms of these new quantities yields

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right). \tag{11.97}$$

Because x and p canonically fail to commute by an amount i, the commutator of a and $a^\dagger$ is also non-zero:

$$[a, a^\dagger] = 1. \tag{11.98}$$

The spectrum of the harmonic oscillator includes a ground state of energy $\frac{1}{2}$ which is annihilated by the operator a:

$$a|0\rangle = 0. \tag{11.99}$$

Solving this first-order differential equation reveals that $|0\rangle$ is Gaussian in both the position and the momentum bases. The ground-state wavefunction in the position basis is, explicitly,

$$\psi_0(x) \equiv \langle x|0\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega x^2}{2\hbar}\right). \tag{11.100}$$

On top of $|0\rangle$ is built an infinite sequence of excited states, the eigenstates of the Hamiltonian, produced by repeated application of the creation operator $a^\dagger$:

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle. \tag{11.101}$$

If you add a and $a^\dagger$, the terms of opposite signs cancel:

$$a + a^\dagger = 2\sqrt{\frac{m\omega}{2\hbar}}x\,. \tag{11.102}$$

So, we can solve for x:

$$x = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger)\,. \tag{11.103}$$

This expression for x is convenient when we want to calculate expectation values. For example, suppose we ascribe to our harmonic oscillator the energy eigenstate $|n\rangle$, and we want to calculate the expectation value of x^2. First, we expand out:

$$x^2 = \frac{\hbar}{2m\omega}(a^2 + (a^\dagger)^2 + aa^\dagger + a^\dagger a)\,. \tag{11.104}$$

The commutator between a and $a^\dagger$ is the same as saying

$$aa^\dagger = a^\dagger a + 1\,. \tag{11.105}$$

So, we can eliminate the $aa^\dagger$ term in the above:

$$x^2 = \frac{\hbar}{2m\omega}(a^2 + (a^\dagger)^2 + 2a^\dagger a + 1)\,. \tag{11.106}$$

Now, we use the fact that ψ_n is orthogonal to $\psi_{n'}$ for $n \neq n'$. Because a lowers one step on the ladder and $a^\dagger$ raises by one step, therefore,

$$\langle n|a|n\rangle = 0\,, \tag{11.107}$$

and likewise

$$\langle n|a^\dagger|n\rangle = 0\,. \tag{11.108}$$

And this holds true no matter how many powers of a or $a^\dagger$ we include. The combination $a^\dagger a$ is the "number operator" N:

$$N|n\rangle = n|n\rangle\,. \tag{11.109}$$

Putting all the pieces together:

$$\langle x^2 \rangle = \langle n|x^2|n\rangle \tag{11.110}$$

$$= \frac{\hbar}{2m\omega}\langle n|(a^2 + (a^\dagger)^2 + 2a^\dagger a + 1)|n\rangle \tag{11.111}$$

$$= \frac{\hbar}{2m\omega}[\langle n|a^2|n\rangle + \langle n|(a^\dagger)^2|n\rangle + 2\langle n|N|n\rangle + \langle n|n\rangle] \tag{11.112}$$

$$= \frac{\hbar}{2m\omega}(0 + 0 + 2n + 1)\,. \tag{11.113}$$

Simplifying, we obtain

$$\langle x^2 \rangle = \frac{\hbar}{m\omega}\left(n + \frac{1}{2}\right)\,. \tag{11.114}$$

Energy eigenstates are not the only set of states with interesting properties. The states $\{|n\rangle\}$ are eigenstates of the Hamiltonian, but what about eigenstates of other operators, like a and $a^\dagger$? To deduce the form of an eigenstate of the annihilation operator, which we call a *coherent state*, we note that the creation and annihilation operators act something like the operations of multiplication and differentiation.

More precisely, the the harmonic-oscillator commutator, Eq. (11.98), has the same configuration as the commutator of differentiation and multiplication, Eq. (22.148). So, we interpret annihilation as differentiation and creation as multiplication. A superposition of energy eigenstates

$$|\phi\rangle = \sum_n c_n|n\rangle \tag{11.115}$$

is given an alternate representation as a polynomial in the formal variable χ, where the coefficient of χ^n gives, up to a normalization constant of $\sqrt{n!}$, the weighting c_n of the n-th energy eigenstate.

If the annihilation operator a is analogous to differentiation, and we seek a state which a leaves unchanged up to a multiplicative constant, then we should ask what function differentiation leaves unchanged, up to a proportionality factor. The answer, of course, is the exponential, and so we look for a coherent state which is an exponential of the multiplicative analogue, the creation operator. Choosing what turns out to be a convenient normalization,

$$|z\rangle \equiv e^{-\frac{1}{2}|z|^2}e^{za^\dagger}|0\rangle. \tag{11.116}$$

We have here normalized $|z\rangle$ so that $\langle z|z\rangle = 1$. It is a straightforward task to show that $|z\rangle$ is an eigenstate of the annihilation operator a with eigenvalue z. By definition,

$$a|z\rangle = e^{-\frac{1}{2}|z|^2}a\sum_{n=0}^{\infty}\frac{(za^\dagger)^n}{n!}|0\rangle. \tag{11.117}$$

The commutation relation Eq. (11.98) lets us move the annihilation operator to the right of the creation operators, at the cost of a factor n:

$$e^{-\frac{1}{2}|z|^2}a\sum_{n=0}^{\infty}\frac{(za^\dagger)^n}{n!}|0\rangle = e^{-\frac{1}{2}|z|^2}\sum_{n=1}^{\infty}\frac{z^n n(a^\dagger)^{n-1}}{n!}|0\rangle. \tag{11.118}$$

Canceling the extra factors we have introduced and factoring a z out of the sum yields

$$e^{-\frac{1}{2}|z|^2}\sum_{n=1}^{\infty}\frac{z^n n(a^\dagger)^{n-1}}{n!}|0\rangle = e^{-\frac{1}{2}|z|^2}z\sum_{n=1}^{\infty}\frac{(za^\dagger)^{n-1}}{(n-1)!}|0\rangle. \tag{11.119}$$

This is just the original definition of $|z\rangle$, multiplied by z. So, we have verified that

$$a|z\rangle = z|z\rangle. \tag{11.120}$$

At this point, we are more interested in the shape of the equations than in extracting specific numbers, so we adjust our units to make all the constants we don't need to carry around equal to 1. Combining the observation that $\langle z|a|z\rangle = z$ with the definition of a, we deduce that

$$\langle z\rangle = \frac{\langle x\rangle + i\langle p\rangle}{\sqrt{2}}. \tag{11.121}$$

The expectation values for the position and momentum are given by the real and imaginary parts, respectively, of the expectation value $\langle z\rangle$.

This result tells us something valuable about the time evolution of a coherent state. Starting with some $|z_0\rangle$, we would like to find

$$|z_0(t)\rangle = e^{-iHt}|z_0\rangle. \tag{11.122}$$

From Eq. (11.116), we can write $|z_0\rangle$ in terms of the energy basis:

$$|z_0\rangle = e^{-\frac{1}{2}|z_0|^2}\sum_n \frac{(z_0a^\dagger)^n}{n!}|0\rangle, \tag{11.123}$$

or, recalling the normalization from Eq. (11.101),

$$|z_0\rangle = e^{-\frac{1}{2}|z_0|^2}\sum_n \frac{z_0^n}{\sqrt{n!}}|n\rangle. \tag{11.124}$$

The time evolution is then given by

$$\begin{aligned}|z_0(t)\rangle &= e^{-\frac{1}{2}|z_0|^2}\sum_n e^{-int}\frac{z_0^n}{\sqrt{n!}}|n\rangle \\ &= e^{-\frac{1}{2}|z_0|^2}\sum_n \frac{(z_0e^{-it})^n}{\sqrt{n!}}|n\rangle \\ &= e^{-\frac{1}{2}|z_0|^2}\exp\left(z_0e^{-it}a^\dagger\right)|0\rangle,\end{aligned} \tag{11.125}$$

meaning that

$$z_0(t) = e^{-it}z_0. \tag{11.126}$$

The real and imaginary parts of $z_0(t)$ now provide us with the expectation values

$$\langle x(t)\rangle = x_0\cos t - p_0\sin t \tag{11.127}$$

and

$$\langle p(t) \rangle = p_0 \cos t - x_0 \sin t. \tag{11.128}$$

The expectation values for the quantum harmonic oscillator given the coherent state $|z_0\rangle$ obey the classical harmonic oscillator's equations of motion.

11.8 Continuous Random Variables

So far, the image we have mostly kept in mind is of a finite-dimensional vector space. We can let the dimension go to infinity, at the cost of some mathematical subtleties that we will try to avoid as much as possible.

To illustrate, take the real line $\mathbb{R}$, and consider complex-valued functions on it. Adding two functions f and g produces another function, and multiplying f by any constant $\alpha \in \mathbb{C}$ likewise makes another valid function. We can define an inner product on this set of functions by turning the sum that we used for Euclidean dot products into an integral:

$$\langle f|g \rangle := \int_{-\infty}^{\infty} dx\, f(x)^* g(x). \tag{11.129}$$

Note that the first function is complex-conjugated. For this to work as an inner product, it must always yield nonnegative values for $\langle f|f \rangle$:

$$\langle f|f \rangle \geq 0. \tag{11.130}$$

For a completely arbitrary function, it is not always guaranteed that this integral is even evaluable. Therefore, to define a Hilbert space, we impose the requirements that the functions be *square-integrable*: The integral of $|f(x)|^2$ converges. This establishes a Hilbert space whose dimension is infinite. Decomposing a function f into a Fourier series is exactly analogous to writing a vector in $\mathbb{C}^d$ as a sum over an orthonormal basis $\{|n\rangle\}$.

Infinite-dimensional Hilbert spaces are the setting we use to express the quantum theory of position, momentum and other such quantities that we treat as being continuous-valued. This leads to mathematical subtleties that can complicate the learning of the physics. For example, multiplication by x is a linear operator on the Hilbert space we have just defined, because $x(\alpha f) = \alpha(xf)$, and

$$x(f + g) = xf + xg. \tag{11.131}$$

Linear operators on finite-dimensional Hilbert spaces can be written as matrices, and so this should, one would think, correspond to a matrix or matrix-like object with infinitely many rows and columns. But the operator "multiply by x" has no eigenvalues or eigenvectors! If we try to find an "eigenfunction" f_λ of this operator, we'd set

$$xf_\lambda = \lambda f_\lambda, \tag{11.132}$$

and so $(x - \lambda)f_\lambda = 0$, meaning that f_λ must equal 0 for every value of x except λ. This calls for a "delta function" — but delta functions are not square-integrable, and so they

do not lie in the Hilbert space! (They aren't even *functions,* properly speaking, and though one can generalize the notion of "function" to a wider class that encompasses them, this doesn't solve the problem that they lie outside the Hilbert space that was supposed to contain everything.) Physically, we have run into trouble because we have tried to place infinite trust in a theory that could not sustain it — we assumed that everything works out simply even when quantities are infinitely concentrated. Mathematically, we cope with this by admitting that a "position measurement" isn't *really* a measurement of the "observable x", as introductory books would like to pretend; instead, it is a kind of continuum limit of a family of discrete POVMs, much in the vein of how we introduced continuous probability density functions in §2.1.[1]

Much as in §2.1, then, we suppose it is reasonable to ask a question of the form, "Is the measured value of the position less than a given threshold?" Since we are now doing quantum mechanics, we encode this question in an operator. Let us say that $\Pi(x_j)$ is the question "Is the measured value of the position less than x_j?"; these must then satisfy

$$\Pi(x_{\min}) = 0,\ \Pi(x_{\max}) = I\,, \tag{11.133}$$

and have a kind of filtering or seiving property:

$$\Pi(x_m)\Pi(x_n) = \Pi(x_n)\Pi(x_m) = \Pi(\min\{x_m, x_n\})\,. \tag{11.134}$$

Next, we suppose it is reasonable to consider infinitesimally small separations between threshold values. Then we can construct

$$d\Pi(\xi) = \Pi(\xi + d\xi) - \Pi(\xi) \tag{11.135}$$

as the operator encoding the question "Is the measured value of the position in the interval between ξ and $\xi + d\xi$?". The *spectral decomposition* of the position operator x is the integral of $\xi d\Pi(\xi)$. The *spectral theorem* establishes that a self-adjoint operator has a unique decomposition. This is one of those results that we will not need so much explicitly, but it is good to know that it is there if we do.

We can extract insight while keeping these subtleties at bay by working with operators algebraically. This also builds familiarity with the art of inventing quantum analogues of classical systems. To perform canonical quantization, we reinterpret a classical Hamiltonian $\mathcal{H}$ as defining a quantum system by turning the classical variables x and p into operators which satisfy the canonical commutator $[x, p] = i\hbar$. This commutation relation implies that, if we choose to write wavefunctions as functions of position, then the operator p will be implemented as a derivative with respect to the position x. To wit:

In elementary calculus, we learned that the derivative of χ^n with respect to χ is $n\chi^{n-1}$. Differentiation by χ reduces the exponent by one, and multiplication by χ raises it by one, but the effect of differentiation followed by multiplication is not quite the same as the other way around. If we differentiate first and then multiply,

[1] There is also a distinction between "self-adjoint" and "Hermitian" in the infinite-dimensional case, with the latter being a weaker property, because an operator and its adjoint might not be defined on the same domain.

χ^n becomes $n\chi^n$. However, if we multiply first and then differentiate, χ^n becomes $(n+1)\chi^n$. The two results differ by χ^n.

We can write this relationship in commutator notation, if we adopt the convention that the operation immediately to the left of χ^n acts upon it first:

$$[\text{differentiation by } \chi, \text{multiplication by } \chi]\chi^n = \chi^n. \tag{11.136}$$

This relationship holds true if we replace χ^n with any non-zero polynomial function $f(\chi)$, so we can say,

$$[\text{differentiation by } \chi, \text{multiplication by } \chi] = 1. \tag{11.137}$$

If the position operator acts on a wavefunction by multiplying that function by the coordinate x, and if the momentum operator must fail to commute with the position operator by an amount $i\hbar$, then it feels quite natural to write the *position-basis representation of the momentum operator* as

$$p = -i\hbar\partial_x. \tag{11.138}$$

In classical physics, we learn that momentum is that which is conserved when dynamics are invariant under spatial translations. Can we think about invariance under spatial translations in the mathematical language of quantum theory? Let's say that the unitary operator U expresses time evolution for some duration. Another unitary operator, call it T, acts to translate a state spatially. If the dynamics are to be invariant under translation, then it must be that they unfold the same way relative to the starting point regardless of where they are started. So, we can translate spatially either before or after the time evolution, and we must get the same thing:

$$UT|\psi\rangle = TU|\psi\rangle. \tag{11.139}$$

Since this is true for an arbitrary input state $|\psi\rangle$, it must be that T and U commute. But if T commutes with time evolution for *any* duration, then it must commute with time evolution for a very short duration, in which case U is very nearly the identity. More specifically, the difference between U and the identity is a small term proportional to the Hamiltonian H. Consequently,

$$[H, T] = 0. \tag{11.140}$$

This is at a rather abstract level. What can we say about how all this works in terms of coordinates?

We start with the canonical commutator. From this, we can calculate how x interacts with higher powers of p. Let's start with p^2:

$$[x, p^2] = [x, pp] = [x, p]p + p[x, p] = 2i\hbar p. \tag{11.141}$$

Notice that we have an expression in terms of p^2, which works out to an expression involving a factor 2 and a p raised to the 1-th power. This suggests that the commutator of x with p^n will be

$$[x, p^n] = i\hbar n p^{n-1}. \tag{11.142}$$

We can prove this by induction, since we have an initial case already established. Now, since we can talk about powers of p, we can also talk about *functions* of the operator p expanded as Taylor series. Any reasonably well-behaved function $F(p)$ can be expressed as a sum of coefficients f_n multiplied by p^n, and thus we deduce:

$$[x, F(p)] = \sum_n [x, f_n p^n] = \sum_n i\hbar n f_n p^{n-1}. \tag{11.143}$$

Inspecting this result, and pulling the multiplicative constant out in front, we recognize that the final sum gives us just the derivative of $F(p)$ with respect to p.

$$[x, F(p)] = i\hbar \partial_p F(p). \tag{11.144}$$

By interchanging p and x in the above argument, we find that a similar relation holds for functions of the position operator:

$$[p, G(x)] = -i\hbar \partial_x G(x). \tag{11.145}$$

So, the momentum operator p "acting on" a function of x gives $-i$ times Planck's constant times the derivative of the function. This is an interesting result, which we'd like to extend. If this is how p relates to functions of the position *operator,* how does it act on *states* specified in terms of position?

To answer this question, we're going to poke around a little and develop some machinery which we'll use later. We define a new operator, which we'll call a *translation* operator, as the exponential of some constants times p:

$$T(\lambda) = \exp\left(-\frac{i\lambda p}{\hbar}\right). \tag{11.146}$$

Why do we call this a *translation*? First, let's use the equation above to calculate how $T(\lambda)$ commutes with x:

$$[x, T(\lambda)] = i\hbar \left(-\frac{i\lambda}{\hbar}\right) \exp\left(-\frac{i\lambda p}{\hbar}\right) = \lambda T(\lambda). \tag{11.147}$$

We see that the commutator just gives us back $T(\lambda)$ times the parameter λ. Written another way,

$$xT(\lambda) = T(\lambda)(x + \lambda). \tag{11.148}$$

We can explore what this means by doing the very physicist-y thing and cooking up a "position eigenstate". As we saw earlier, such a "state" can't really belong to the Hilbert space of square-integrable functions. But we can attach some extra machinery to that Hilbert space, inventing a "basis" of objects that exist just outside it, in terms of which we can express any element of the Hilbert space itself. We will in fact devise two of these "bases", one for position and one for momentum. Both will satisfy a continuum kind of orthonormality condition:

$$\langle x|x'\rangle = \delta(x - x'), \tag{11.149}$$

$$\langle p|p'\rangle = \delta(p - p'). \tag{11.150}$$

We can expand an arbitrary $|\psi\rangle$ using either of the sets $\{|x\rangle\}$ or $\{|p\rangle\}$. Taking the former, we have

$$|\psi\rangle = \int dx\, |x\rangle\langle x|\psi\rangle := \int dx\, \psi(x)|x\rangle\,. \tag{11.151}$$

The coefficient $\psi(x)$ is what we call the position-space wavefunction for the ket $|\psi\rangle$.

Let's see how the combination $xT(\lambda)$ acts on a "position eigenstate":

$$xT(\lambda)|x\rangle = T(\lambda)(x+\lambda)|x\rangle = (x+\lambda)T(\lambda)|x\rangle\,. \tag{11.152}$$

This result merits close inspection. We started with $xT(\lambda)$ acting on the "eigenstate" $|x\rangle$, which we can just as well call x acting on the state $T(\lambda)|x\rangle$ and we ended with $(x+\lambda)$ — which is just a *number* — times the state $T(\lambda)|x\rangle$. Therefore, $T(\lambda)|x\rangle$ is an "eigenstate" of the operator x with "eigenvalue" $(x+\lambda)$.

Acting with the operator $T(\lambda)$ pushed the "state" $|x\rangle$ over by an amount λ! But any *actual* state in the Hilbert space is a linear combination of the $\{|x\rangle\}$, all of which will be pushed over by the same amount. Hence, T effects *spatial translations* of the states it acts upon.

Since we defined T as an exponential, working it out for small parameters is easy. We just get the identity, plus the small parameter times p (and some constants), plus higher-order terms we lump together because we don't care about them:

$$T(-\epsilon) = e^{i\epsilon p/\hbar} = I + i\frac{\epsilon}{\hbar}p + \mathcal{O}(\epsilon^2) \tag{11.153}$$

Momentum, we note, is (up to some constants) the generator of spatial translations. The quantity which arose in the classical theory as the Noether charge conjugate to spatial translation symmetry becomes, in the quantum theory, the generator of that same symmetry.

To push on this further, we introduce a general state $|\psi\rangle$, which isn't necessarily an eigenstate or anything else special; it's just a state for the system. We then compute the Dirac bracket of this state with the "position eigenstate" $|x\rangle$ and T in the middle:

$$\langle x|T(-\epsilon)|\psi\rangle = \psi(x) + i\frac{\epsilon}{\hbar}\langle x|p|\psi\rangle + \mathcal{O}(\epsilon^2). \tag{11.154}$$

But because

$$\langle x|T(-\epsilon)|\psi\rangle = \langle x+\epsilon|\psi\rangle = \psi(x+\epsilon), \tag{11.155}$$

(we flipped signs because the operator is acting on the *bra* vector, not the ket, which requires taking the adjoint), the Dirac bracket becomes

$$\psi(x+\epsilon) = \psi(x) + i\frac{\epsilon}{\hbar}\langle x|p|\psi\rangle + \mathcal{O}(\epsilon^2). \tag{11.156}$$

We can rearrange this to put the wavefunction terms all one one side, giving us

$$\langle x|p|\psi\rangle = \frac{\hbar}{i}\lim_{\epsilon\to 0}\frac{\psi(x+\epsilon)-\psi(x)}{\epsilon} = -i\hbar\partial_x\psi(x). \tag{11.157}$$

The ket $|\psi\rangle$ was just an arbitrary state, so this relation doesn't depend upon our choice of ket. It's in fact a *general statement* about the operator p as represented in the position basis: $p = -i\hbar\partial_x$, as promised in Eq. (11.138). Again, we see that "acting with p" involves taking a derivative and multiplying by $-i$ and Planck's constant.

I leave as a quick exercise the calculation of the "momentum eigenfunction" in position space. It should not be too difficult to see that

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \exp\left(\frac{ipx}{\hbar}\right). \tag{11.158}$$

Because the "position eigenkets" form a complete basis, we can recover the "momentum eigenstate" by summing over them:

$$|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \, \exp\left(\frac{ipx}{\hbar}\right) |x\rangle. \tag{11.159}$$

A general state $|\psi\rangle$ can of course be projected onto $\langle p|$ just as well as $\langle x|$, giving a wavefunction in *momentum space,* thus:

$$\bar{\psi}(p) = \langle p|\psi\rangle. \tag{11.160}$$

Exercise: what does this say about the Fourier transforms of the position and momentum wavefunctions?

The *Heisenberg equation of motion* for an operator A is

$$\frac{dA}{dt} = i[H, A] + \frac{\partial A}{\partial t}. \tag{11.161}$$

For the position operator,

$$\frac{dx}{dt} = i[H, x], \tag{11.162}$$

and since the only thing in H which does not commute with x is p,

$$\frac{dx}{dt} = i\left(-i\partial_p H\right) = \partial_p H. \tag{11.163}$$

Similarly, we can see that

$$\frac{dp}{dt} = -\partial_x H. \tag{11.164}$$

These expressions, which we derived using the canonical commutator, reproduce the look of the Hamilton equations of motion for classical particle mechanics, but with the classical position and momentum variables replaced by operators.

11.9 Weyl–Schwinger Smoothing

Way back, we saw that we could write any qubit density matrix as a linear combination

$$\rho = \frac{1}{2}\left(I + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z\right). \tag{11.165}$$

In other words, the identity I and the Pauli matrices form an operator basis. We really only need two of the Paulis, since the product of two of them will be proportional to the third. I.e., we can take the elements of our operator basis to be

$$W_{kl} := \sigma_x^k \sigma_z^l \,. \tag{11.166}$$

Allowing complex coefficients, this will serve as a basis not just for self-adjoint matrices, but for unitary ones too.

We can use an idea of Hermann Weyl [111], built upon by Julian Schwinger [112], to run with this. Instead of restricting ourselves to measly 2×2 matrices, let us work with $d \times d$. What are the counterparts of σ_x and σ_z in this setting? Well, interchanging the order of σ_x and σ_z in a product introduces a minus sign, so let us introduce operators X and Z that, when their order is exchanged, gain a *phase*:

$$ZX = \omega XZ \,. \tag{11.167}$$

We take the determinant of both sides and find that most everything drops out:

$$\det(ZX) = \det Z \det X = \omega^d \det XZ = \omega^d \det X \det Z \,. \tag{11.168}$$

and so

$$\omega^d = 1 \,. \tag{11.169}$$

We have deduced that the phase factor ω is a root of unity. We take it to be

$$\omega = e^{2\pi i/d} \,. \tag{11.170}$$

If we pull one factor of X through l factors of Z, we find

$$Z^l X = \omega^l X Z^l \,, \tag{11.171}$$

and likewise,

$$ZX^k = \omega^k X^k Z \,. \tag{11.172}$$

Together,

$$Z^l X^k = \omega^{kl} X^k Z^l \,. \tag{11.173}$$

Any operators that satisfy this *Weyl commutator* will in some basis take the following form. Z is diagonal, with its diagonal entries being powers of the phase ω:

$$Z|n\rangle = \omega^n |n\rangle \,. \tag{11.174}$$

And X will effect a cyclic shift in this basis:

$$X|n\rangle = |n+1\rangle \,, \tag{11.175}$$

interpreting addition modulo d.

Weyl and Schwinger's insight was that, first, the operators

$$W_{kl} := X^k Z^l \tag{11.176}$$

form a basis. Second, X and Z are Fourier conjugates of each other: X cyclically shifts the eigenvectors of Z, and Z cyclically shifts the eigenvectors of X. Third, as the dimension d grows larger and larger, that cyclic shift X starts to look more and more like the *generator of a translation.*

There is a mathematically ill-defined step here, as both Weyl and Schwinger recognized. (Schwinger spells things out in more detail, but my feeling is that what he knew about these operators, Weyl already knew too.) A cyclic shift necessarily brings us back to the starting point eventually. This is fundamentally different from a translation along a line, which can keep going. Schwinger fudges it and says that the $d \to \infty$ limit works with "qualifications", because one is always working in physical situations where the values of position and momentum, "although possibly very large, remain *finite*". He writes X as the exponential of i times a small number times a generator $\hat{p}$, and likewise for Z, and then he shows that the Weyl commutator implies that these generators must satisfy

$$[\hat{x}, \hat{p}] = i\,, \tag{11.177}$$

the canonical commutator (in units with $\hbar = 1$).

We note a certain spiritual commonality with our discussion of going from discrete probability distributions to continuous probability density functions, back in Eq. (2.28) and the following. There, we made a *physical assumption on top of Dutch-book coherence* and said that our expectations for multiple different experiments fit together in a mathematically clean way. Can we not view the Weyl–Schwinger argument similarly? We are positing a sequence $X_2, X_3, \ldots$ of unitaries, along with a conjugate sequence $Z_2, Z_3, \ldots$, and we are adding a physical assumption about how the actions represented by these operators fit together.

There is not a mathematically unique "continuum limit" of the quantum theory of finite-dimensional Hilbert spaces. One must add some physical content, like specifying that we want a theory of a single, nonrelativistic particle in a single spatial dimension. (Consider, for example, the N lowest-lying energy eigenstates of a 3D harmonic oscillator and the N lowest-energy states of a particle in a 1D box; both span vector spaces isomorphic to $\mathbb{C}^N$.) But when we think of this physical addition as telling us how to bind together infinitely many discrete Weyl operators, we go directly to the algebraic formulation of our continuum quantum theory. The continuum theory is that construction which has all the possible finite-dimensional theories as consistent truncations of it, and which embodies the other data properly. All those useful things like Dirac deltas that didn't even fit into our Hilbert space come along for the ride naturally, because the right way to make the transition is not to just contrive an infinite orthonormal basis, but to focus on the algebra.

11.10 Thermalized Harmonic Oscillator

Now that we have the energy levels of the simple harmonic oscillator well in hand, let us see how our concepts of thermal equilibrium apply to it. The energy levels of this system depend upon the oscillator's characteristic frequency ω and a nonnegative

integer n:

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right). \tag{11.178}$$

The definition of the Gibbs states, combined with the Born rule, informs us that if we measure the energy, we will obtain a value E_n with a probability proportional to

$$f(n) = e^{-\beta E_n}, \tag{11.179}$$

where β is the coolness.

In order to get a legitimate probability, we must normalize, so we divide by

$$Z = \sum_{n=0}^{\infty} f(n) = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+1/2)}. \tag{11.180}$$

We can factor out the part involving the 1/2, yielding

$$Z = e^{-\beta\hbar\omega/2}\sum_{n=0}^{\infty} e^{-\beta\hbar\omega n}. \tag{11.181}$$

The sum is a geometric series whose first term is 1 and in which each term is $e^{-\beta\hbar\omega}$ of the one before. Thus,

$$Z = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}}. \tag{11.182}$$

If we differentiate $f(n)$ with respect to β, we pull down the derivative of the argument of the exponential, which is just $-E_n$:

$$\frac{\partial f(n)}{\partial\beta} = -E_n f(n). \tag{11.183}$$

We find the derivative of Z by adding this up:

$$\frac{\partial Z}{\partial\beta} = \sum_{n=0}^{\infty}\frac{\partial f(n)}{\partial\beta} = -\sum_{n} E_n f(n). \tag{11.184}$$

That result looks a lot like an expectation value! We know by definition that

$$\langle E\rangle = \sum_{n=0}^{\infty} E_n p(E = E_n), \tag{11.185}$$

which means that

$$\langle E\rangle = \sum_{n=0}^{\infty}\frac{E_n f(n)}{Z}. \tag{11.186}$$

Basic calculus tells us that

$$\frac{d}{dx}\log x = \frac{1}{x}, \tag{11.187}$$

and using the chain rule,

$$\frac{d}{dx}\log g(x) = \frac{dg/dx}{g(x)}. \tag{11.188}$$

So we apply this to Z, which depends upon the parameter β:

$$\frac{\partial \log Z}{\partial \beta} = \frac{1}{Z}\frac{\partial Z}{\partial \beta}. \tag{11.189}$$

Combining this with the answer to part (b),

$$\frac{1}{Z}\frac{\partial Z}{\partial \beta} = \frac{-\sum_n E_n f(n)}{Z} = -\sum_n E_n \frac{f(n)}{Z}. \tag{11.190}$$

And so, we establish that

$$\langle E\rangle = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\partial \log Z}{\partial \beta}. \tag{11.191}$$

We can evaluate $\log Z$ because we have a formula for Z, Eq. (11.182).

$$\log Z = -\frac{\beta\hbar\omega}{2} - \log(1 - e^{-\beta\hbar\omega}). \tag{11.192}$$

The derivative of this is

$$\frac{\partial \log Z}{\partial \beta} = -\frac{\hbar\omega}{2} - \frac{\hbar\omega e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}. \tag{11.193}$$

And so the expectation value for the energy is

$$\langle E\rangle = \frac{\hbar\omega}{2} + \frac{\hbar\omega e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}. \tag{11.194}$$

It is always good to understand the limiting behavior of an equation like this. When the coolness β is very large, $e^{-\beta\hbar\omega}$ becomes very small, and so $\langle E\rangle$ is approximately $\hbar\omega/2$. At low temperatures, only the ground state contributes anything to how much energy we expect to find! On the other hand, when β is very small — which is just another way of saying, "at high temperature" — the exponential factor $e^{-\beta\hbar\omega}$ will be close to 1, and we can use a Taylor approximation:

$$\langle E\rangle = \frac{\hbar\omega}{2} + \hbar\omega\frac{1 - \beta\hbar\omega}{\beta\hbar\omega} = \frac{\hbar\omega}{2} + k_B T\left(1 - \frac{\hbar\omega}{k_B T}\right). \tag{11.195}$$

So, at higher and higher temperatures, the expectation value for the energy will look more and more like $k_B T$.

12 Perturbation Theory and Scattering

12.1 Prelude

Suppose that we have a Hamiltonian that depends upon a parameter: $H(\lambda)$. The eigenfunctions and eigenvalues of this Hamiltonian will then obey the relation

$$H(\lambda)|\psi_n(\lambda)\rangle = E_n(\lambda)|\psi_n(\lambda)\rangle\,, \tag{12.1}$$

by definition. As usual, the expectation value of the Hamiltonian given that the preparation is an eigenstate is the corresponding energy level:

$$\langle\psi_n(\lambda)|H(\lambda)|\psi_n(\lambda)\rangle = E_n(\lambda)\,. \tag{12.2}$$

Let us differentiate both sides of this equation with respect to the parameter λ:

$$\begin{aligned}\frac{dE_n(\lambda)}{d\lambda} &= \left(\frac{d}{d\lambda}\langle\psi_n(\lambda)|\right) H(\lambda)|\psi_n(\lambda)\rangle + \langle\psi_n(\lambda)|H(\lambda)\left(\frac{d}{d\lambda}|\psi_n(\lambda)\rangle\right) + \langle\psi_n(\lambda)|\frac{dH(\lambda)}{d\lambda}|\psi_n(\lambda)\rangle \\ &= E_n(\lambda)\left(\frac{d}{d\lambda}\langle\psi_n(\lambda)|\right)|\psi_n(\lambda)\rangle + E_n(\lambda)\langle\psi_n(\lambda)|\left(\frac{d}{d\lambda}|\psi_n(\lambda)\rangle\right) + \langle\psi_n(\lambda)|\frac{dH(\lambda)}{d\lambda}|\psi_n(\lambda)\rangle \\ &= E_n(\lambda)\frac{d}{d\lambda}\langle\psi_n(\lambda)|\psi_n(\lambda)\rangle + \langle\psi_n(\lambda)|\frac{dH(\lambda)}{d\lambda}|\psi_n(\lambda)\rangle\,. \end{aligned} \tag{12.3}$$

The inner product in the first term is always 1 by normalization, so its derivative is 0. Thus:

$$\frac{dE_n(\lambda)}{d\lambda} = \langle\psi_n(\lambda)|\frac{dH(\lambda)}{d\lambda}|\psi_n(\lambda)\rangle\,. \tag{12.4}$$

The derivative of the energy is just the expectation value of the derivative of the Hamiltonian. Following Zwiebach, we will call this the *Hellmann–Feynman lemma* [113].

Now, suppose that $H(\lambda)$ has two parts, with the parameter λ controlling the extent to which one of them is enabled:

$$H(\lambda) = H_0 + \lambda G\,. \tag{12.5}$$

We can write a Taylor expansion for the energy E_n in terms of λ:

$$E_n(\lambda) = E_n(0) + \lambda\left.\frac{dE_n(\lambda)}{d\lambda}\right|_{\lambda=0} + \cdots\,. \tag{12.6}$$

The Hellmann–Feynman lemma tells us how to evaluate that derivative: It's just the expectation value of the derivative of $H(\lambda)$. But the only part in $H(\lambda)$ that depends upon λ is the λG term. Therefore, to first order in λ,

$$E_n(\lambda) = E_n(0) + \lambda\langle\psi_n(0)|G|\psi_n(0)\rangle\,. \tag{12.7}$$

This is the most important result in quantum perturbation theory. The first-order correction to an energy level is the expectation value of the perturbation, evaluated using the unperturbed state.

The next most important result is the first-order correction to the state vectors, and after that, the second-order correction to the energy levels. We will derive these next.

12.2 Brillouin–Wigner Series

The standard approach to finding the higher-order corrections involves expanding both the energies and the state vectors in Taylor series and doing a lot of order-by-order matching of terms. It is not particularly opaque, as calculations go, but it is not very memorable either, and handling the details correctly becomes rather persnickety. We will instead follow the method of Brillouin and Wigner, as exposited in Baym's textbook [114].

As before, we will take $H = H_0 + \lambda G$, and to reduce the notational clutter, we will use lowercase letters to label the eigenvectors of the unperturbed Hamiltonian:

$$H_0|n\rangle = \varepsilon_n|n\rangle\,. \tag{12.8}$$

The energy levels of the full Hamiltonian satisfy

$$H|N\rangle = E_n|N\rangle\,. \tag{12.9}$$

It is helpful in perturbation theory to relax the normalization requirement on the perturbed state vectors, reimposing it at the end if necessary. So, we can take the convention that $|N\rangle$ is not itself required to be a unit vector, and

$$\langle n|N\rangle = 1\,. \tag{12.10}$$

Writing out the Hamiltonian H and rearranging a bit, we find

$$(E_n - H_0)|N\rangle = \lambda G|N\rangle\,. \tag{12.11}$$

Taking the inner product of both sides with the unperturbed eigenvector $|m\rangle$,

$$\langle m|(E_n - H_0)|N\rangle = \lambda\langle m|G|N\rangle\,. \tag{12.12}$$

Acting with H_0 to the left, this becomes

$$(E_n - \varepsilon_m)\langle m|N\rangle = \lambda\langle m|G|N\rangle\,. \tag{12.13}$$

If we take $|m\rangle = |n\rangle$, then we get the energy shift

$$E_n - \varepsilon_n = \lambda\langle m|G|N\rangle\,. \tag{12.14}$$

This looks a lot like the formula we derived above, but we have not made any approximations yet, and the two vectors bracketing the G operator are in different bases.

Let's expand $|N\rangle$ in the basis of the original eigenvectors:

$$\begin{aligned} |N\rangle &= \sum_m |m\rangle\langle m|N\rangle \\ &= |n\rangle\langle n|N\rangle + \sum_{m\neq n} |m\rangle\langle m|N\rangle \\ &= |n\rangle + \sum_{m\neq n} |m\rangle \frac{\lambda\langle m|G|N\rangle}{E_n - \varepsilon_m}\,. \end{aligned} \tag{12.15}$$

Here, we have used our normalization convention for $|N\rangle$ to simplify the first term, and we've solved Eq. (12.13) to substitute for $\langle m|N\rangle$. This result might look fairly useless, because we've found $|N\rangle$ in terms of something more complicated that still involves $|N\rangle$. But it actually works out quite nicely, because we can *iterate* it. We take that whole expression for $|N\rangle$ and substitute it back into itself where $|N\rangle$ appears!

Like so:

$$|N\rangle = |n\rangle + \sum_{m\neq n} |m\rangle \frac{\lambda\langle m|}{E_n - \varepsilon_m} \left(|n\rangle + \sum_{j\neq n} |j\rangle \frac{\lambda\langle j|G|N\rangle}{E_n - \varepsilon_j} \right). \tag{12.16}$$

This gives the original vector $|n\rangle$, a term linear in λ that involves G once, and a term quadratic in λ that involves G twice. And we can keep doing this:

$$\begin{aligned} |N\rangle = |n\rangle &+ \lambda \sum_{m\neq n} |m\rangle \frac{\langle m|G|n\rangle}{E_n - \varepsilon_m} \\ &+ \lambda^2 \sum_{m,j\neq n} |j\rangle \frac{\langle j|G|m\rangle}{E_n - \varepsilon_j} \frac{\langle m|G|n\rangle}{E_n - \varepsilon_n} \\ &+ \lambda^3 \sum_{m,j,k\neq n} |k\rangle \frac{\langle k|G|j\rangle}{E_n - \varepsilon_k} \frac{\langle j|G|m\rangle}{E_n - \varepsilon_j} \frac{\langle m|G|n\rangle}{E_n - \varepsilon_n} + \cdots . \end{aligned} \tag{12.17}$$

We don't know what $|N\rangle$ is, so we just keep pushing it further away — and though each new term looks nastier, it will be a smaller correction, because it's higher order in a small number.

The only wrinkle is that we can't calculate the denominators, because they involve the *perturbed* energy levels $\{E_n\}$, and we don't know those! Yet all is not lost. *To lowest order,* the perturbed energy level E_n must be equal to the original energy level ε_n. We know there will be corrections, but those must involve contributions from G and further factors of λ. Consequently, we can truncate the series:

$$|N\rangle \approx |n\rangle + \lambda \sum_{m\neq n} \frac{\langle m|G|n\rangle}{\varepsilon_n - \varepsilon_m} |m\rangle\,. \tag{12.18}$$

This tells us the first-order correction to the eigenstates!

Now, let's try this expression in our formula for the energy shift, Eq. (12.14):

$$E_n - \varepsilon_n = \lambda\langle n|G|n\rangle + \lambda^2\langle n|G\sum_{m\neq n}\frac{\langle m|G|n\rangle}{\varepsilon_n - \varepsilon_m}|m\rangle\,. \tag{12.19}$$

Pulling the prefactor into the sum and recognizing the complex conjugation, we find

$$E_n - \varepsilon_n = \lambda\langle n|G|n\rangle + \lambda^2\sum_{m\neq n}\frac{|\langle m|G|n\rangle|^2}{\varepsilon_n - \varepsilon_m}\,. \tag{12.20}$$

The first term is the result we found earlier using the Hellmann–Feynman lemma, and the second term is the correction to next order.

12.3 Incorporating Time Dependence

Consider a scenario in which the time evolution is generated by the Hamiltonian H_1 from time 0 to time T, and then by H_2 from T until t. The unitary operator implementing this evolution is

$$U(t) = e^{-iH_2(t-T)/\hbar}e^{-iH_1T/\hbar}\,. \tag{12.21}$$

If the operators H_1 and H_2 happen to commute, then we can combine the exponentials just as though they were numbers:

$$e^{-iH_2(t-T)/\hbar}e^{-iH_1T/\hbar} \to e^{-i(H_2(t-T)+H_1T)/\hbar}\,. \tag{12.22}$$

But in general, there will be another factor that depends upon their commutator. In fact, the expression becomes rather involved, incorporating nested commutators of higher and higher order. We will see the gory details in Eq. (14.69). For now, the important thing to know is that if the Hamiltonians are each proportional to some small number λ, then the commutator will be proportional to λ^2, and so on to higher degrees. So, we expect the above expression to emerge an approximation to the correct answer if the Hamiltonians are, in some sense, small enough. And if H depends on time *continously,* we likewise have

$$U(t) = \exp\left(\frac{1}{i\hbar}\int_0^t H(t')dt'\right) \tag{12.23}$$

if all the operators mutually commute. And since we are pinning our hopes that this will approximately hold upon the idea of small perturbations, it is reasonable to look at what happens when the whole integral is small:

$$U(t) \approx I + \frac{1}{i\hbar}\int_0^t H(t)dt'\,. \tag{12.24}$$

Alternatively, let us start with the general time-evolution equation with a time-dependent Hamiltonian, thusly:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle\,. \tag{12.25}$$

We can integrate this formally to obtain

$$|\psi(t)\rangle = |\psi(0)\rangle + \frac{1}{i\hbar}\int_0^t H(t')|\psi(t')\rangle dt' \,. \tag{12.26}$$

As in the previous section, this might look hopeless, because our integral for $|\psi(t)\rangle$ involves $|\psi(t')\rangle$ in the integrand. But we can try iterating, like we did for the Brillouin–Wigner series:

$$|\psi(t)\rangle = |\psi(0)\rangle + \frac{1}{i\hbar}\int_0^t H(t')\left[|\psi(0)\rangle + \frac{1}{i\hbar}\int_0^{t'} H(t'')|\psi(t'')\rangle dt''\right] dt' \,. \tag{12.27}$$

To first order in the Hamiltonian,

$$|\psi(t)\rangle \approx |\psi(0)\rangle + \frac{1}{i\hbar}\int_0^t H(t')dt'|\psi(0)\rangle \,. \tag{12.28}$$

Writing this as a unitary operator applied to $|\psi(0)\rangle$ gives the same approximation as in Eq. (12.24). One way to think about this is that at each instant, $H(t')$ applies a little nudge to the state vector; we are imagining that these touches are small enough that any *later* nudge can be regarded as acting upon the original vector $|\psi(0)\rangle$.

We can apply this way of thinking to perturbation theory by positing that we have a Hamiltonian with a constant part and a time-dependent part:

$$H(t) = H_0 + G(t) \,. \tag{12.29}$$

The sneaky part is that we will work *in a rotating basis.* We change our coordinate axes continually over time in such a way that the evolution generated by H_0 stands still. Then we see what changes G can wring in this basis.

We can absorb the time dependence generated by H_0 thusly:

$$\left|\tilde{\psi}(t)\right\rangle = e^{iH_0t/\hbar}|\psi(t)\rangle \,. \tag{12.30}$$

At time $t = 0$, the two ways of writing the vector coincide. The time-evolution equation for $\left|\tilde{\psi}(t)\right\rangle$ is

$$i\hbar\frac{d}{dt}\left|\tilde{\psi}(t)\right\rangle = \tilde{G}(t)\left|\tilde{\psi}(t)\right\rangle , \tag{12.31}$$

where $\tilde{G}(t)$ is just what $G(t)$ looks like in our rotating basis:

$$\tilde{G}(t) = e^{iH_0t/\hbar}G(t)e^{-iH_0t/\hbar} \,. \tag{12.32}$$

At this juncture, we invoke our first-order approximation:

$$\left|\tilde{\psi}(t)\right\rangle = \left|\tilde{\psi}(0)\right\rangle + \frac{1}{i\hbar}\int_0^t \tilde{G}(t')\left|\tilde{\psi}(0)\right\rangle dt' \,. \tag{12.33}$$

Transforming back to the original basis,

$$|\psi(t)\rangle = e^{-iH_0t/\hbar}\left(|\psi(0)\rangle + \frac{1}{i\hbar}\int_0^t \tilde{G}(t')|\psi(0)\rangle dt'\right) . \tag{12.34}$$

12.4 Scattering as Perturbation

In a *scattering* experiment, we fire particles at a target and see how they bounce off. There are (at least) two ways of relating the theory of scattering to the perturbation theory we have developed over the preceding sections. One way uses an iterative series rather like the Brillouin–Wigner method; this approach requires some familiarity with Green's functions. Another technique is to imagine oneself flying toward the target and seeing its influence as a time-dependent potential. This ends up in the same place and presumes some approximations about "densities of states" along energy or momentum axes. The former is more standard and is the path taken, for example, by Zwiebach [113]; the latter is covered by Schiff [115] and Baym [114].

Either way, the result is a formula for the probability amplitude for scattering in a given direction, as a function of the polar angles θ and ϕ.

12.5 Bragg Scattering

In the first Born approximation (the first term of the Born series), the scattering amplitudes are given by the Fourier transform of the scattering potential:

$$f(\theta,\phi) = -\frac{1}{4\pi}\int d^3\vec{r}\, e^{-i(\vec{k}_s-\vec{k}_i)\cdot\vec{r}} U(\vec{r})\,. \tag{12.35}$$

Here, $\hbar\vec{k}_i$ is the momentum of the incident particle, and $\hbar\vec{k}_s$ is the momentum of the scattered particle, which has the same magnitude but possibly a different direction. The angle between them is θ.

Consider a scattering potential that is translationally invariant. That is, the scattering potential $U(\vec{r})$ can be written as the sum of many copies of a single "fundamental domain":

$$U(\vec{r}) = \sum_{j=1}^{N} u(\vec{r}-\vec{X}_j) \tag{12.36}$$

where the displacements $\vec{X}_j$ are separated by a constant shift $\vec{R}$. We'll prove that in the Born approximation, scattering occurs only in the directions defined by the condition

$$(\vec{k}_s-\vec{k}_i)\cdot\vec{R} = 2\pi n\,, \tag{12.37}$$

where n is an integer.

The scattering amplitude for a momentum transfer $\vec{q}$ is

$$f(\theta,\phi) = -\frac{1}{4\pi}\sum_j \int d^3\vec{r}\, e^{-i\vec{q}\cdot\vec{r}} u(\vec{r}-\vec{X}_j)\,, \tag{12.38}$$

and

$$u(\vec{x}) = \frac{1}{(2\pi)^3}\int d^3\vec{k}\, e^{i\vec{k}\cdot\vec{x}} v_{\vec{k}}\,. \tag{12.39}$$

Using

$$\int d^3\vec{r}\, e^{-i(\vec{q}-\vec{k})\cdot\vec{r}} = \frac{1}{(2\pi)^3}\delta^3(\vec{q}-\vec{k})\,, \tag{12.40}$$

we get

$$f(\theta,\phi) = -\frac{1}{4\pi}u_q \sum_j e^{-i\vec{q}\cdot\vec{X}_j}\,, \tag{12.41}$$

and so the differential cross section is

$$\frac{d\sigma}{d\Omega} = \left(\frac{1}{4\pi}\right)^2 |u_q|^2 \left|\sum_j e^{-i\vec{q}\cdot\vec{X}_j}\right|^2 . \tag{12.42}$$

To have significant scattering, we need

$$e^{-i\vec{q}\cdot\vec{X}_j} = 1\,, \tag{12.43}$$

which means

$$\vec{q}\cdot\vec{X}_j = 2\pi n \qquad \forall j\,, \tag{12.44}$$

and

$$\frac{d\sigma}{d\Omega} \propto N^2\,. \tag{12.45}$$

Expressing the momentum transfer q in terms of the incident and outgoing wave vectors, and recalling the definition of $\vec{X}_j$,

$$(\vec{k}_s - \vec{k}_i)\cdot\vec{R} = 2\pi n\,, \tag{12.46}$$

as desired.

12.6 Scattering from a Charge Distribution

Another scenario we can tackle with the first Born approximation is the scattering of an electron from a *spherically symmetric charge distribution.* Here, the given information is not the potential, but the arrangement of static electrical charge.

The first Born approximation is just the Fourier transform of the scattering potential, and the potential due to a charge distribution can be found by adding up the Coulomb potentials of the pieces making it up. This means that the potential $V(r)$ due to $\rho(r')$ is the convolution of the charge distribution with the Green's function for a point charge. The convolution theorem says that the Fourier transform of a convolution is the product of the Fourier transforms of the functions being convolved. Therefore, the form factor F is given by the Fourier transform of $\rho(r)$.

Scattering off the potential of a point charge, or *Rutherford* scattering, is remarkably difficult to treat directly. The Coulomb force just decreases too slowly with distance!

The resulting complications lead Baym, for example, to resort to confluent hypergeometric functions [114]. But we can get a handle on it by thinking of it as the limit of *Yukawa* scattering, i.e., taking $\mu \to 0$ in

$$U(r) = \beta \frac{e^{-\mu r}}{r} \,. \tag{12.47}$$

The first Born approximation for this potential is

$$f(\theta) = -\frac{2m\beta}{q\hbar^2} \int_0^\infty dr\, e^{-\mu r} \sin(qr) \tag{12.48}$$

$$= -\frac{2m\beta}{\hbar^2(\mu^2 + q^2)} \,. \tag{12.49}$$

Taking the limit $\mu \to 0$, and substituting in the Coulomb form for β, the Rutherford scattering amplitude is

$$f_R(\theta) = -\frac{2mq_e^2}{\hbar^2 q^2} \,. \tag{12.50}$$

Squaring this, the differential Rutherford scattering cross section is

$$\left.\frac{d\sigma}{d\Omega}\right|_R = \left(\frac{2mq_e^2}{\hbar^2}\right)^2 \frac{1}{16k^4 \sin^4(\theta/2)} \,. \tag{12.51}$$

By the convolution theorem,

$$f_I(\theta) = f_R(\theta) F(q) \,, \tag{12.52}$$

where $f_R(\theta)$ is given by Eq. (12.50) and the form factor $F(q)$ is

$$F(q) = \frac{1}{Q} \int d^3\vec{x}\, \rho(\vec{x}) e^{i\vec{q}\cdot\vec{x}} \,. \tag{12.53}$$

The differential cross section is

$$\frac{d\sigma}{d\Omega} = \left.\frac{d\sigma}{d\Omega}\right|_R |F(q)|^2 \tag{12.54}$$

$$= \left(\frac{2mq_e^2}{\hbar^2}\right)^2 \frac{|F(q)|^2}{16k^4 \sin^4(\theta/2)} \,. \tag{12.55}$$

For example, consider a uniformly charged sphere of radius R. Let the charge density of the sphere be ρ_0. The Fourier transform of this charge distribution is

$$F(q) = \frac{2\pi}{Q} \int_{-1}^{+1} d(\cos\theta) \int_0^R dr\, r^2 \rho_0 e^{iqr\cos\theta} \tag{12.56}$$

$$= \frac{2\pi\rho_0}{Q} \int_0^R dr\, r^2 \frac{e^{iqr} - e^{-iqr}}{iqr} \,. \tag{12.57}$$

Evaluating this integral gives

$$F(q) = 3\frac{\sin Rq - Rq\cos Rq}{(Rq)^3}\,, \tag{12.58}$$

where we have used the fact that the total charge Q is $4\rho_0\pi R^3/3$.

Next, let's try the Gaussian distribution

$$\rho(\vec{r}) = \rho_0 \exp\left(-\frac{3r^2}{2R^2}\right)\,. \tag{12.59}$$

The form factor is

$$F(q) = \frac{2\pi\rho_0}{Q}\int_{-1}^{+1} d(\cos\theta)\int_0^\infty dr\, r^2 e^{-3r^2/2R^2} e^{iqr\cos\theta} \tag{12.60}$$

$$= \frac{2\pi\rho_0}{Q}\int_0^\infty dr\, r^2 e^{-3r^2/2R^2}\,\frac{e^{iqr} - e^{-iqr}}{iqr} \tag{12.61}$$

$$= \frac{4\pi\rho_0}{Qq}\int_0^\infty dr\, r e^{-3r^2/2R^2}\sin(qr)\,. \tag{12.62}$$

To evaluate this, we use

$$\int_0^\infty dx\, x^2 e^{-3x^2/2R^2} = \sqrt{\frac{\pi}{6}}\frac{R^3}{3}\,, \tag{12.63}$$

$$\int_0^\infty dx\, x e^{-3x^2/2R^2}\sin(qr) = \sqrt{\frac{\pi}{6}}\frac{R^3}{3} q e^{-q^2R^2/6}\,. \tag{12.64}$$

When we divide by the total charge, everything ends up canceling except the last exponential:

$$F(q) = \exp\left(-\frac{q^2R^2}{6}\right)\,. \tag{12.65}$$

This is the same as the result given in Table I of Hofstadter [116], so I'm happy.

13 Pairs of Qubits

13.1 A Nicely Symmetric Basis

When we introduced the concept of quantum entanglement, we discussed the two-qubit joint state

$$|\psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} . \tag{13.1}$$

Or, using $|+z\rangle$ and $|-z\rangle$ to stand for the eigenvectors of σ_z,

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|+z\rangle \otimes |+z\rangle + |-z\rangle \otimes |-z\rangle \right) . \tag{13.2}$$

As we discussed, this has the property that it cannot be written as a tensor product of two single-qubit state vectors. Another joint state with this property is

$$\frac{1}{\sqrt{2}} \left(|+z\rangle \otimes |-z\rangle + |-z\rangle \otimes |+z\rangle \right) . \tag{13.3}$$

This state has the nice feature that it is *symmetric* with respect to *exchanging* the two qubits. Of course, we can easily think up vectors that aren't entangled and which also have this property:

$$|+z\rangle \otimes |+z\rangle \text{ and } |-z\rangle \otimes |-z\rangle \tag{13.4}$$

spring readily to mind. These three states are all orthogonal to each other. And so, linear combinations of them form a three-dimensional complex vector space. If we could find a fourth vector that is orthogonal to all three of these, then we could span the whole space $\mathbb{C}^4$. Is there such a vector that also has a nice symmetry property with respect to qubit exchange? Let's try to find it, starting by writing

$$|\psi\rangle = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} . \tag{13.5}$$

Being orthogonal to $|+z\rangle \otimes |+z\rangle$ means that a must be 0, and likewise, orthogonality with $|-z\rangle \otimes |-z\rangle$ forces d to be 0. The inner product with our third state is proprotional to $b + c$, so we need $b = -c$. Normalization fixes the magnitude of the nonzero entries,

and we can neglect an overall phase factor because it doesn't matter for calculating probabilities. So, our fourth vector is

$$\frac{1}{\sqrt{2}}\left(|+z\rangle \otimes |-z\rangle - |-z\rangle \otimes |+z\rangle\right) . \tag{13.6}$$

This isn't quite symmetric with respect to qubit exchange. Instead, swapping the two qubits introduces a minus sign. But that's still pretty good! We have a new basis for the vector space $\mathbb{C}^4$, made out of a symmetric *triplet* and an antisymmetric *singlet*.

We can express the idea of exchanging qubits by a linear operator, which we'll call SWAP. Like any linear operator, we can specify it by saying what it does to a basis.

$$\begin{aligned}
\text{SWAP}|+z\rangle \otimes |+z\rangle &= |+z\rangle \otimes |+z\rangle ,\\
\text{SWAP}|+z\rangle \otimes |-z\rangle &= |-z\rangle \otimes |+z\rangle ,\\
\text{SWAP}|-z\rangle \otimes |+z\rangle &= |+z\rangle \otimes |-z\rangle ,\\
\text{SWAP}|-z\rangle \otimes |-z\rangle &= |-z\rangle \otimes |-z\rangle .
\end{aligned} \tag{13.7}$$

As a matrix in this basis,

$$\text{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{13.8}$$

And in Dirac notation,

$$\text{SWAP} = \sum_{lm} |m\rangle\langle l| \otimes |l\rangle\langle m| . \tag{13.9}$$

This last form is helpful when evaluating, for example,

$$\begin{aligned}
\text{tr}(\text{SWAP}(A \otimes B)) &= \sum_{ijlm} \text{SWAP}_{ijlm} A_{li} B_{mj} \\
&= \sum_{ijlm} \delta_{im}\delta_{jl} A_{li} B_{mj} \\
&= \sum_{ij} A_{ji} B_{ij} \\
&= \text{tr}(AB) .
\end{aligned} \tag{13.10}$$

We'll come back to the "swaperator" momentarily.

Where do we go from here? Well, one thing we can do is add more qubits! If we have n qubits, then we can form $n+1$ mutually orthogonal, completely symmetrized pure states. We start with the tensor product of n "up" states, and we progress by steps to the tensor product of n "down" states. At each step in between, we take one state from "up" to "down" and symmetrize. So, inside the set of pure states for a 2^n-dimensional system there lurks a state space for an $(n+1)$-dimensional system, revealed by forming symmetric combinations. This is the starting point for the technique known as the *Majorana representation*.

Another thing we can do is look for problems that our triplet-singlet basis is adapted to solving.

13.2 Pairs of Two-Level Systems

The simplest Hamiltonian we can write that does anything is proportional to the Pauli matrix σ_z:

$$H = A\sigma_z = \begin{pmatrix} A & 0 \\ 0 & -A \end{pmatrix} . \tag{13.11}$$

If we perform an energy measurement upon a system modeled by this Hamiltonian, we get either A or $-A$. What if we have two such systems? We can measure one and ignore the other, which we represent by the operator $A\sigma_z \otimes I$. If instead we want to measure the second and ignore the first, we express that action by the operator $I \otimes A\sigma_z$. The sum of these,

$$H = A\sigma_z \otimes I + I \otimes A\sigma_z = \begin{pmatrix} 2A & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2A \end{pmatrix} , \tag{13.12}$$

models a situation where the time evolution of one qubit has nothing to do with the time evolution of the other. This can be the case, for example, if we are studying two systems that are both immersed in the same external field. That is, we have here the quantum counterpart of the classical situation where, for instance, two electric dipoles are both subject to an electric field in the $\hat{z}$ direction, while being far enough apart that they do not affect each other.

But what is the quantum counterpart of a classical situation where the systems *do* affect each other? If the field felt by one dipole is produced by the other, what then? We can invent a quantum version of this by writing a Hamiltonian that *couples* the two systems, without imposing a preferred choice of axis from outside. We know that we can pick any axis and ask the two qubits to tell us "up" or "down" with respect to that axis. Can we arrange it so that the energy is different depending upon whether these answers are the same or different?

Thinking about symmetric-looking combinations of Pauli matrices, we land upon the following:

$$H = A(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z) = A \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{13.13}$$

And here is where SWAP comes back into the story. We can verify by direct computation that

$$\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z = 2(\text{SWAP}) - I \otimes I . \tag{13.14}$$

This combination of Pauli matrices will naturally act differently upon vectors that are symmetric versus those that are antisymmetric! If SWAP leaves a vector $|\psi\rangle$ unchanged, then

$$\langle\psi|(2(\text{SWAP}) - I \otimes I)|\psi\rangle = 2\langle\psi|\psi\rangle - \langle\psi|\psi\rangle = 1 . \tag{13.15}$$

But if it introduces a minus sign, then

$$\langle\psi|(2(\text{SWAP}) - I \otimes I)|\psi\rangle = -2\langle\psi|\psi\rangle - \langle\psi|\psi\rangle = -3\,. \tag{13.16}$$

In the triplet-singlet basis, SWAP becomes a diagonal matrix. Three of the entries along the diagonal are $+1$, while the other is -1. Consequently, we see that our "dipole interaction" Hamiltonian must be diagonal in that basis, too! We can work this out explicitly by forming the basis-change matrix

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2^{-1/2} & 0 & 2^{-1/2} \\ 0 & 2^{-1/2} & 0 & -2^{-1/2} \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{13.17}$$

The columns of this matrix are what the triplet-singlet basis vectors look like in the orignal basis. So, we apply U to turn the triplet-singlet basis vectors into the original set, apply our Hamiltonian, and then apply $U^{-1} = U^\dagger$ to see what the result looks like in the triplet-singlet basis. The result of calculating $U^\dagger H U$ is a diagonal matrix with three entries equal to A and the last equal to $-3A$.

In fact, this transformation also diagonalizes our *other* Hamiltonian, the one proportional to $\sigma_z \otimes I + I \otimes \sigma_z$. This was already diagonal in our original basis; the conjugation by U just changes which of the entries are nonzero. Very nicely, we can make a Hamiltonian that combines both,

$$H = A(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z) + B(\sigma_z \otimes I + I \otimes \sigma_z)\,, \tag{13.18}$$

and diagonalize it all to make a matrix

$$\begin{pmatrix} A+2B & 0 & 0 & 0 \\ 0 & A & 0 & 0 \\ 0 & 0 & A-2B & 0 \\ 0 & 0 & 0 & -3A \end{pmatrix}. \tag{13.19}$$

If all we have is the A term, then the triplet states will be *degenerate*: The expectation value for the energy is the same given any choice of them. Introducing the B term *lifts* that degeneracy.

13.3 Commutator Relations

The commutator of any two Pauli matrices is proportional to the third. For example,

$$[\sigma_x, \sigma_z] = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} = -2i\sigma_y\,, \tag{13.20}$$

as we can verify by straightforward arithmetic. Making a "two-qubit version" of σ_z led to interesting things, so let's try doing that for all three Pauli matrices:

$$\tau_j = \sigma_j \otimes I + I \otimes \sigma_j,\ j = x, y, z\,. \tag{13.21}$$

By direct computation, we find that

$$[\tau_x, \tau_z] = -2i\tau_y \,, \tag{13.22}$$

just like with the original matrices. And indeed, if we use the Levi-Civita symbol to express all the Pauli commutation equations together,

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l \,, \tag{13.23}$$

we can verify that

$$[\tau_j, \tau_k] = 2i\epsilon_{jkl}\tau_l \,. \tag{13.24}$$

Therefore, anything we deduce just from the commutation relations will apply to the two-qubit operators exactly like it did to the originals.

Because the τ_j matrices do not commute with one another, we cannot find a basis that simultaneously diagonalizes them. So, we get at most one preferred axis in our Hamiltonian this way. But, having picked out the $\hat{z}$ axis as the one we are giving special treatment, we can label the eigenstates of the Hamiltonian by the eigenvalue of τ_z and the eigenvalue of what we can call

$$\tau^2 = \sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z \,. \tag{13.25}$$

If the vector $|\lambda, \mu\rangle$ satisfies

$$\tau^2|\lambda, \mu\rangle = \lambda|\lambda, \mu\rangle, \;\; \tau_z|\lambda, \mu\rangle = \mu|\lambda, \mu\rangle \,, \tag{13.26}$$

then

$$H|\lambda, \mu\rangle = (A\lambda + B\mu)|\lambda, \mu\rangle \,. \tag{13.27}$$

This is an example of calculating an energy in terms of "quantum numbers", eigenvalues of operators that commute with the Hamiltonian and with each other.

13.4 Qubit Induction

What if we wanted to start with *qudits* rather than qubits, that is, with d-dimensional Hilbert spaces where $d > 2$? The tensor product of a d_1-dimensional space with a d_2-dimensional space is a space of dimension $d_1 d_2$:

$$\mathbb{C}_{d_1} \otimes \mathbb{C}_{d_2} \simeq \mathbb{C}_{d_1 d_2} \,. \tag{13.28}$$

Here we are using "$\simeq$" to mean that the two vector spaces are really the same, though we might be describing them in different bases. But by the idea we introduced at the beginning, we can think of each of these spaces as the *fully symmetrized subspace* of pure states for the appropriate numbers of qubits. Let's see where this goes.

We'll use $S(n)$ to denote the space we get by taking n qubits, constructing the $n+1$ orthonormal symmetric vectors as before, and using them as a basis. The basic relation is

$$S(n) \simeq \mathbb{C}_{n+1} \,. \tag{13.29}$$

And so, we immediately have

$$S(n_1) \otimes S(n_2) \simeq \mathbb{C}_{(n_1+1)(n_2+1)} = \mathbb{C}_{n_1 n_2 + n_1 + n_2 + 1} \,. \tag{13.30}$$

We can split this up as a direct sum of vector spaces:

$$\mathbb{C}_{n_1 n_2 + n_1 + n_2 + 1} \simeq \mathbb{C}_{n_1 + n_2 + 1} \oplus \mathbb{C}_{n_1 n_2} \,, \tag{13.31}$$

which we can break down further into

$$\mathbb{C}_{n_1 + n_2 + 1} \oplus \mathbb{C}_{n_1 n_2} \simeq \mathbb{C}_{n_1 + n_2 + 1} \oplus (\mathbb{C}_{n_1} \otimes \mathbb{C}_{n_2}) \,. \tag{13.32}$$

Next, we reinterpret each contribution here in terms of qubits:

$$S(n_1) \otimes S(n_2) \simeq S(n_1 + n_2) \oplus (S(n_1 - 1) \otimes S(n_2 - 1)) \,. \tag{13.33}$$

This is neat, because we can use the formula repeatedly to break down the second term. The process stops when we hit

$$S(n) \otimes S(0) \simeq S(n) \,. \tag{13.34}$$

With each step, we lose two qubits. For example,

$$S(1) \otimes S(1) \simeq S(2) \oplus S(0) \,. \tag{13.35}$$

This is the direct sum of the triplet and the singlet we saw earlier. More generally,

$$S(n_1) \otimes S(n_2) \simeq S(n_1 + n_2) \oplus S(n_1 + n_2 - 2) \oplus \cdots \oplus S(|n_1 - n_2|) \,. \tag{13.36}$$

13.5 Quaternions, Biquaternions and Entanglement

For those with a fondness for vintage geometry, the Pauli matrices are just quaternions in disguise. The fact that the unit-length quaternions constitute the group SU(2) led Baez to coin the slogan, "Qubits are not just quantum — they are quaternionic!" [117]. Poking at this a little more reveals that a pair of qubits considered together are *biquaternionic*. Moreover, a single qubit also has a biquaternionic character, once we consider ascribing to it states of less-than-maximal confidence.

Unitary operations for a single qubit are isomorphic to maximally entangled states for a pair of qubits. In terms of a quantum circuit, we can imagine ascribing a Bell state to a pair of qubits and then applying a unitary to one half of the pair. The details of the unitary are then stored in the correlation — or perhaps better said, the potential correlation — between the qubits.

Just as the complex numbers $\mathbb{C}$ double the dimensionality of the real number line $\mathbb{R}$, the quaternions $\mathbb{H}$ double the dimensionality of $\mathbb{C}$. A quaternion $q \in \mathbb{H}$ is written

$$q = q_0 + q_1 \hat{\imath} + q_2 \hat{\jmath} + q_3 \hat{k}, \tag{13.37}$$

where the three quantities $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ satisfy Hamilton's relations

$$\hat{\imath}^2 = \hat{\jmath}^2 = \hat{k}^2 = \hat{\imath}\hat{\jmath}\hat{k} = -1. \tag{13.38}$$

Since the quaternion units $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ correspond in a natural way to the Pauli matrices, we can map quaternions into the group of qubit unitaries. The only detail we have to be careful about is that each of these square to -1, while the Pauli matrices square to the identity matrix itself. For reasons that will be clear in a moment, let's keep the imaginary unit of the complex numbers, $i \in \mathbb{C}$, symbolically distinct from the $\hat{\imath}$ in the quaternions $\mathbb{H}$. We can then write

$$(q_0, q_1, q_2, q_3) = q_0 + q_1\hat{\imath} + q_2\hat{\jmath} + q_3\hat{k} \in \mathbb{H} \to q_0 I - iq_1\sigma_x - iq_2\sigma_y - iq_3\sigma_z. \tag{13.39}$$

From there, we can also map them to entangled states for pairs of qubits, by the standard trick of acting locally with a unitary on half of a Bell state. This will give us an isomorphism between entangled two-qubit states and certain quaternionic objects, and trusting the isomorphism beyond the case of maximal entanglement that inspired it will provide a new way to think of the standard entanglement measure.

Let's work through this in a little detail. Start with a 2×2 matrix

$$M := \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{13.40}$$

Then we have that

$$\mathrm{tr}(M\sigma_x) = -iq_1 = b + c, \tag{13.41}$$
$$\mathrm{tr}(M\sigma_y) = -iq_2 = i(b - c), \tag{13.42}$$
$$\mathrm{tr}(M\sigma_z) = -iq_3 = a - d. \tag{13.43}$$

So, we can find the quaternion coefficients in terms of the matrix elements thusly:

$$q_0 = a + d,\ q_1 = i(b + c),\ q_2 = c - b,\ q_3 = i(a - d). \tag{13.44}$$

These coefficients will in general be complex numbers themselves, meaning that

$$q := (q_0, q_1, q_2, q_3) \tag{13.45}$$

will be a *biquaternion.*

The fact that the algebra of 2×2 complex matrices is also that of quaternions complexified in this way is known, at least within a niche of quaternion enthusiasts. What this fact implies for quantum theory is less explored.

If M is unitary, then we can apply it to half of a Bell pair to define the state

$$|\psi\rangle = M \otimes I \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}. \tag{13.46}$$

Using Eq. (13.44), we can turn this around and write a biquaternion for any two-qubit state vector $|\psi\rangle$, maximally entangled or not. Is this merely a formal trick, or is it illuminating to do so? To ask the question slightly differently, does this way of writing quantum states make anything easy or natural to calculate? Well, let's find out, starting from the mathematical side: What are the natural things one would do with biquaternions?

There are two ways to take a conjugate of a biquaternion. One is to take the complex conjugate of all the coefficients:

$$q^* := (q_0^*, q_1^*, q_2^*, q_3^*). \tag{13.47}$$

The other is to flip the signs on the $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ components:

$$q^\star := (q_0, -q_1, -q_2, -q_3). \tag{13.48}$$

Let's use this latter method of conjugation to define a norm:

$$|q|_\star := qq^\star = q_0^2 + q_1^2 + q_2^2 + q_3^2, \tag{13.49}$$

all the other cross terms canceling. In terms of the vector components,

$$|q|_\star = 4(ad - bc). \tag{13.50}$$

The *concurrence* [118, 119] of $|\psi\rangle$ is proportional to the magnitude of $|q|_\star$:

$$C(|\psi\rangle) = 2|\psi_0\psi_3 - \psi_1\psi_2| = |ad - bc|. \tag{13.51}$$

What has happened here? It turns out that the four complex components of our biquaternion q are the coefficients of the state vector $|\psi\rangle$ written in a slightly phase-adjusted version of the Bell basis. Each of the four states in this basis is maximally entangled, so the more strongly $|\psi\rangle$ overlaps with any one of these basis vectors, the more entangled it is.

And the other way of conjugating q? Given a biquaternion q that represents a two-qubit pure state $|\psi\rangle$, consider the biquaternion

$$\tilde{q} := -q^* = (-d^* - a^*, i(c^* + b^*), b^* - c^*, i(-d^* + a^*)). \tag{13.52}$$

This biquaternion represents the matrix

$$\tilde{U} = \begin{pmatrix} -d^* & c^* \\ b^* & -a^* \end{pmatrix}, \tag{13.53}$$

which when applied to half of a Bell pair yields the pure state

$$\left|\tilde{\psi}\right\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -d^* \\ c^* \\ b^* \\ -a^* \end{pmatrix}. \tag{13.54}$$

We can quickly check that

$$\left|\tilde{\psi}\right\rangle = \sigma_y \otimes \sigma_y |\psi^*\rangle, \tag{13.55}$$

where $|\psi^*\rangle$ is the state made by complex-conjugating the coefficients of the original state $|\psi\rangle$ in the computational basis. We have that

$$\left\langle\tilde{\psi}\middle|\psi\right\rangle = bc - ad. \tag{13.56}$$

The magnitude of this quantity is the concurrence of $|\psi\rangle$ again.

If the matrix M is Hermitian, then its Hilbert–Schmidt inner products with the Pauli matrices will all be real, so q_0 will be real while q_1, q_2 and q_3 will be purely imaginary. Consequently, we can write a general *mixed state* for a qubit as

$$q_\rho = \frac{1}{2}(1, ix, iy, iz), \tag{13.57}$$

where (x, y, z) are the coordinates of the state ρ within the Bloch ball. To reconstruct ρ from q_ρ, simply take

$$\rho = \frac{1}{2}(I - iq_1\sigma_x - iq_2\sigma_y - iq_3\sigma_z). \tag{13.58}$$

Instead of concurrence, the $\star$-norm now measures mixedness:

$$4|q_\rho|_\star = 1 - (x^2 + y^2 + z^2). \tag{13.59}$$

This vanishes for pure states and is maximized by the "garbage state" $\rho = \frac{1}{2}I$.

By applying the $*$- and $\star$-conjugations together, we can define an adjoint:

$$q^\dagger := (q^*)^\star. \tag{13.60}$$

The product of q with its adjoint $q^\dagger$ is not a scalar, but a biquaternion with generally nonvanishing, yet purely imaginary, $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ components. If q is a two-qubit pure state, then $qq^\dagger$ yields a single-qubit state that is generally a mixed state whose purification is q. In the language of matrices,

$$\mathrm{tr}_B|\psi\rangle\langle\psi| = \begin{pmatrix} |\psi_0|^2 + |\psi_1|^2 & \psi_0\psi_2^* + \psi_1\psi_3^* \\ \psi_2\psi_0^* + \psi_3\psi_1^* & |\psi_2|^2 + |\psi_3|^2 \end{pmatrix} = \frac{1}{2}MM^\dagger. \tag{13.61}$$

This generalizes the familiar fact that when M is unitary, $|\psi\rangle$ is maximally entangled, and so its marginal $\mathrm{tr}_B|\psi\rangle\langle\psi|$ is the garbage state.

There is another way to represent the qubit state space quaternionically, one that treats the four components more symmetrically and which establishes a correspondence between the Bell basis and another set of quantum resources. Let ρ be a single-qubit state, and take its tensor product with the state

$$\Pi_0 := \frac{1}{2}\left(I + \frac{1}{\sqrt{3}}(\sigma_x + \sigma_y + \sigma_z)\right). \tag{13.62}$$

Now, consider measuring the two-qubit system whose preparation is $\rho \otimes \Pi_0$ in the Bell basis. The probability p_n for obtaining the n-th outcome is

$$p_n = \frac{1}{2}\mathrm{tr}(\rho \Pi_n), \tag{13.63}$$

where Π_0 is defined as above and we fill out the set by drawing the rest of a regular tetrahedron in the Bloch sphere:

$$\Pi_1 := \frac{1}{2}\left(I + \frac{1}{\sqrt{3}}(\sigma_x - \sigma_y - \sigma_z)\right), \tag{13.64}$$

$$\Pi_2 := \frac{1}{2}\left(I + \frac{1}{\sqrt{3}}(-\sigma_x + \sigma_y - \sigma_z)\right), \tag{13.65}$$

$$\Pi_3 := \frac{1}{2}\left(I + \frac{1}{\sqrt{3}}(-\sigma_x - \sigma_y + \sigma_z)\right). \tag{13.66}$$

Therefore, the Bell-basis measurement on the two-qubit system is equivalent to measuring the POVM

$$E_n := \frac{1}{2}\Pi_n \tag{13.67}$$

on the original single qubit [120]. This POVM is informationally complete (IC), and so we can replace any qubit state ρ, pure or mixed, with a probability distribution over its outcomes. Its high degree of symmetry makes it optimal, broadly speaking, among qubit IC POVMs [121].

As mentioned earlier, given a two-qubit pure state $|\psi\rangle$, the four components of its biquaternion q are naturally interpreted as coefficients in the Bell basis, up to phase factors that we can neglect when we square those coefficients to obtain probabilities. This inspires the representation

$$q = (\sqrt{p_0}, \sqrt{p_1}, \sqrt{p_2}, \sqrt{p_3}), \tag{13.68}$$

from which we reconstruct a qubit density matrix by

$$\rho = \sum_{n=0}^{3}\left(3p_n - \frac{1}{2}\right)\Pi_n. \tag{13.69}$$

In this representation, all qubit states, whether pure or mixed, are quaternions of length 1. The set of all qubit states is the ball whose boundary is given by

$$\sum_i p_i^2 = \frac{1}{3}. \tag{13.70}$$

Given two quantum states ρ and σ, with their quaternionic representations q_ρ and q_σ defined in this manner, the product $q_\rho q_\sigma^\star$ is the Bhattcharyya coefficient of the probability vectors p_ρ and p_σ:

$$q_\rho q_\sigma^\star = \sum_i \sqrt{(p_\rho)_i (p_\sigma)_i}. \tag{13.71}$$

The four-outcome probability simplex can be given the structure of a Riemannian manifold in such a way that the arccosine of this quantity is the geodesic distance between the points p_ρ and p_σ [122]. Moreover, it provides an upper bound [123] on the mixed-state fidelity between ρ and σ:

$$q_\rho q_\sigma^\star \geq \sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} = \sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}. \tag{13.72}$$

A historical note: The biquaternionic picture came to light during an investigation of a two-qubit SIC, a highly symmetric set of 16 equally but not maximally entangled states [124, 125], which linked up with an inquiry into the symmetries of a regular icosahedron inscribed in the Bloch sphere [126].

14 Fundamentals of Group Theory

14.1 Structure and Transformation

Begin with a set, S, and consider the functions of the set to itself, $f : S \to S$. What can we say about the set of all such maps f? First, we can compose any two maps, $f \circ g$. This composition will obey an associative rule,

$$(f \circ g) \circ h = f \circ (g \circ h), \tag{14.1}$$

and we note the special importance of the *identity map,*

$$e : S \to S, \text{ such that } e(x) = x \ \forall \ x \in S. \tag{14.2}$$

The identity e obeys the rule,

$$e \circ f = f \circ e = f \ \forall \ f. \tag{14.3}$$

If we have $f : S \to S$, under what circumstances will there exist $g : S \to S$ such that $f \circ g = e$? It is not hard to deduce that if f is *one-to-one,* then g exists; we call such a g the *inverse* of f, written f^{-1}.

A *group* is a set G with a binary operation $(\cdot)$ satisfying the following properties:

- Associativity:
$$(a \cdot b) \cdot c = a \cdot (b \cdot c). \tag{14.4}$$
- Identity:
$$\exists e \text{ such that } a \cdot e = e \cdot a = a. \tag{14.5}$$
- Inverses:
$$\forall \ a \in G, \ \exists a^{-1} \text{ such that } a \cdot a^{-1} = a^{-1} \cdot a = e. \tag{14.6}$$

The set $\{f\}$ of all functions from S to itself is a group if the inverse condition is satisfied, i.e., if the functions are one-to-one. In situations where our set of interest satisfies some but not all of the group properties, we have different names. A set with an operation which satisfies associativity and the identity condition but not the inverse condition is called a *monoid,* while a set with an operation satisfying only associativity is called a *semigroup.*

Keeping these ideas in mind, we would like to generalize a bit. Given a set S, let's look at all maps $f : S \to S$ which preserve the "structure" of S. For example, consider a pentagon with vertices labeled 1 through 5. What is the "structure" of this set, and what maps preserve it?

The point of this example is that many valid answers to those questions exist. We can map vertices from one to another, like so:

$$\{1, 2, 3, 4, 5\} \to \{1, 2, 3, 4, 5\}, \tag{14.7}$$

changing the order as we desire,

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \to \begin{pmatrix} 1 \\ 2 \\ 4 \\ 3 \\ 5 \end{pmatrix}. \tag{14.8}$$

These permutations have inverses, clearly enough, and it is easy to conceive the identity operation. Because the permutations are one-to-one, the set of all permutations constitutes a group. For any possible ending configuration — of which there are $5! = 120$ — we can write a one-to-one function taking the initial configuration into that ending state. Therefore, we have a group with 120 elements; we call this particular group the *symmetric group* S_5.

What happens if we mandate that the permutations must preserve the left- and right-hand neighbors of each vertex? When this is our "structure", we arrive at the *cyclic group* of order 5, the group consisting of the five possible rotations. We write this as $\mathbb{Z}_5$.

It is interesting to note that $\mathbb{Z}_5$ is commutative (or *abelian*) while S_5 is not. One way we can think about this is that $\mathbb{Z}_5$ has only one *generator,* the action of rotating by one step in a certain direction. We can build up each member of the group by composing multiple copies of this generator, r, so that our group looks like this:

$$\{r, r^2, r^3, r^4, r^5 = e\}. \tag{14.9}$$

Whenever we have associativity, we can say that

$$r^a r^b = r^{a+b}, \tag{14.10}$$

where we here have the added proviso

$$r^5 = r^0 = e. \tag{14.11}$$

We have blundered our way into modular arithmetic!

At this point, you should convince yourself that S_5 is not commutative.

Let us briefly note what happens when we abandon the rigid requirement that left neighbors become left neighbors — relaxing our concern for chirality and allowing *reflections* as well as rotations. Here we're dealing with *two* generators, one for rotation (which we had before) and one for reflection. Call the former r and the latter s. Note that we can reflect a pentagon in many ways: right-to-left on a page, for example, or taking the mirror image in a verticle line. Geometric intuition tells us that these images are identical *up to a rotation.*

We now introduce a standard notation for group presentation:

$$\langle r, s | r^5 = e, s^2 = e, r^a s = s r^{5-a} \rangle. \tag{14.12}$$

This is one of the more rigorous ways of writing out a group. If we have any "word" made of r and s, we can apply the relations among r, s and e to find equivalent words.

How big is this *dihedral group*? We immediately guess that the answer is the product of the sizes of the two groups, and this is easily checked.

A *subgroup* is a subset that is closed under the group operation.

A subset H of a group G is a *normal* subgroup if it "looks the same" from every vantage point within G. Explicitly,

$$gHg^{-1} = H \ \forall \ g \in G \, . \tag{14.13}$$

We can combine groups to make larger groups. For example, if we have a regular triangle and a regular pentagon, we can apply any symmetry of the triangle and any symmetry of the pentagon, and neither affects the other. The *direct product* of groups G and H is formed in this way, by taking pairs (g_1, h_1) and (g_2, h_2) and using the operations within G and H:

$$(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2) \ \forall \ g_1, g_2 \in G, \ h_1, h_2 \in H \, . \tag{14.14}$$

The *semidirect product* is more complicated. We can motivate it by considering rotations and translations in the Euclidean plane. Let (R, u) denote the transformation consisting of a rotation by R and then a translation by u:

$$f_{(R,u)}(x) = Rx + u \, . \tag{14.15}$$

Now, apply the transformations (R, u) and (R', u') in succession:

$$\begin{aligned} f_{(R,u)} \left(f_{(R',u')}(x) \right) &= f_{(R,u)}(R'x + u') && (14.16) \\ &= RR'x + Ru' + u && (14.17) \\ &= f_{(RR', Ru'+u)}(x) \, . && (14.18) \end{aligned}$$

The composition law is not just the direct product, which would give us $(RR', u + u')$, but instead

$$(R, u)(R', u') = (RR', R(u') + u) \, . \tag{14.19}$$

We apply a function determined by one group element to the other.

Groups can have infinitely many elements. For example, take the *general linear* groups $GL_n(\mathbb{R})$. These consist of the $n \times n$ matrices of real numbers whose determinant is nonzero. The group operation is just matrix multiplication; restricting the determinant makes sure that the matrices are all invertible. We can also define a group $GL_n(\mathbb{C})$, which works the same way but is made out of complex-valued matrices instead. If we require that the determinant is always equal to 1, then we get the *special linear* groups $SL_n(\mathbb{R})$ and $SL_n(\mathbb{C})$ respectively.

For an easy example, start with $GL_2(\mathbb{R})$. The elements of this group are the 2×2 invertible matrices of real numbers. Consider the subset of those defined by

$$M(b) = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} . \tag{14.20}$$

Using ordinary matrix multiplication, we see that

$$M(b')M(b) = \begin{pmatrix} 1 & b'+b \\ 0 & 1 \end{pmatrix} = M(b'+b) . \tag{14.21}$$

The subset that is parameterized by b is closed under the multiplication rule, so it forms a subgroup.

14.2 Representations

A *representation* of a group is a map from group elements to matrices such that the composition of group elements becomes matrix multiplication. One might think of it as "compiling" the high-level language of a group for a computer that calculates with matrices. Because a representation D of a group G must respect its composition law,

$$D(g_1)D(g_2) = D(g_1 g_2) , \tag{14.22}$$

the representation of the identity element is the identity matrix:

$$D(e) = I . \tag{14.23}$$

For example, we can take the cyclic group of order 3, $\mathbb{Z}_3$, and represent it by 1×1 matrices, which are just numbers:

$$D(e) = 1,\ D(r) = e^{2\pi i/3},\ D(r^2) = e^{4\pi i/3} . \tag{14.24}$$

We say that the *dimension* of this representation is 1. We can also construct a representation of dimension 3:

$$D(e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ D(r) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},\ D(r^2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} . \tag{14.25}$$

This is an example of what we call a *regular representation,* which we can construct for any finite group by associating a basis vector with each group element. In this case, our basis vectors are $|e\rangle, |r\rangle, |r^2\rangle$, and we turn the group operation into matrix multiplication by

$$D(g_1)|g_2\rangle = |g_1 g_2\rangle . \tag{14.26}$$

Let D be a representation of a group G, and let M be an invertible matrix. Define the *similarity transformation* of D by M thusly:

$$D'(g) = M^{-1}D(g)M . \tag{14.27}$$

The new function D' from group elements to matrices is also a representation:

$$D'(g_1)D'(g_2) = M^{-1}D(g_1)D(g_2)M = M^{-1}D(g_1g_2)M = D'(g_1g_2)\,. \tag{14.28}$$

Two representations related in this way are said to be equivalent up to similarity.

A representation is *unitary* if all the matrices that it comprises are unitary. All representations of finite groups are equivalent up to similarity with unitary representations.

Because we can turn group elements into matrices, we can do everything we know about from linear algebra to them. A *character* is the trace of the matrix that represents a group element:

$$\chi_D(g) = \mathrm{tr}D(g)\,. \tag{14.29}$$

The cyclic property of the trace implies that characters are invariant under similarity transformations.

A representation is *reducible* if it leaves invariant some proper subspace of the vector space on which it acts. Suppose that the matrices $D(g)$ of a representation act on a vector space V, and let P be a projector onto a subspace of V, i.e., a linear operator that leaves vectors in that subspace untouched and throws away any components orthogonal to it. Then $D(g)$ leaving this subspace invariant means that

$$PD(g)P = D(g)P\,. \tag{14.30}$$

In other words, if we apply $D(g)$ to any vector that we know lives wholly within the subspace, then we get a vector that still lives wholly within the subspace. For example, take the regular representation of $\mathbb{Z}_3$ introduced above. The operator

$$P = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \tag{14.31}$$

projects onto the flat vector, which $D(e)$, $D(r)$ and $D(r^2)$ all leave untouched.

Any representation of a finite group is equivalent to one where all the matrices are block diagonal. To illustrate again with the regular representation of $\mathbb{Z}_3$, we take a cube root of unity $\omega = e^{2\pi i/3}$ and use the similarity transformation defined by

$$S = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2 \end{pmatrix}\,. \tag{14.32}$$

This yields a representation

$$D'(e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ D'(r) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix},\ D'(r^2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega^2 & 0 \\ 0 & 0 & \omega \end{pmatrix}\,. \tag{14.33}$$

For brevity, it is traditional to call irreducible representations "irreps". *Schur's lemma* says that for any irrep that sends a group into complex matrices, if a matrix

M commutes with every $D(g)$, then M is a scalar multiple of the identity matrix. Suppose that G is an abelian group. Then any representation will enjoy the property that

$$D(g)D(h) = D(h)D(g)\,, \tag{14.34}$$

for arbitrary $g, h \in G$. If D is an irrep, then Schur's lemma applies. For any given g, all of the $D(h)$ commute with $D(g)$, and so

$$D(h) = \lambda_h I \tag{14.35}$$

for some complex number λ_h. But the only way that this can actually *be* irreducible is if the identity matrix I leaves no proper subspace invariant — which means that it has to be the 1×1 identity matrix! So, all the irreps of an abelian group have to be one-dimensional.

Let's apply this to a group with infinitely many elements, the rotations of a circle. In the language of matrices and complex numbers, this is the group $U(1)$: all 1×1 complex matrices whose adjoints are their inverses. So, we are thinking of the unit circle in the complex plane, and each point labels the rotation needed to turn the number 1 into that point. This group is abelian, because the multiplication of complex numbers is commutative. Consequently, all its irreps are one-dimensional.

Let D be an irrep of $U(1)$. We can label the elements by angle, so that the periodicity of the circle becomes the condition

$$D(0) = D(2\pi) = 1\,, \tag{14.36}$$

and multiplication in the complex plane becomes

$$D(\theta_1 + \theta_2) = D(\theta_1)D(\theta_2)\,. \tag{14.37}$$

The derivative of the irrep map is, by definition

$$\frac{dD}{d\theta} = \lim_{\Delta\theta\to0} \frac{D(\theta+\Delta\theta) - D(\theta)}{\Delta\theta}\,. \tag{14.38}$$

We use the previous expression to turn the sum into a product:

$$\frac{dD}{d\theta} = \lim_{\Delta\theta\to0} \frac{D(\theta)D(\Delta\theta) - D(\theta)}{\Delta\theta} = D(\theta) \lim_{\Delta\theta\to0} \frac{D(\Delta\theta) - 1}{\Delta\theta}\,. \tag{14.39}$$

We recognize the limit as the derivative evaluated at zero:

$$\frac{dD}{d\theta} = D(\theta)D'(0)\,. \tag{14.40}$$

We have a differential equation for $D(\theta)$ that says it is proportional to its own derivative. Combining this with the condition that $D(0) = 1$, we find that

$$D(\theta) = e^{D'(0)\theta}\,. \tag{14.41}$$

Periodicity means that

$$D(2\pi) = e^{2\pi D'(0)} = 1\,, \tag{14.42}$$

and so $D(0)$ must be ik for some integer k.

This classifies the irreps of $U(1)$. They are all functions of the form $e^{ik\theta}$ for integer k. All of these irreps are unitary!

14.3 Group Theory of a Single Qubit

We have seen that, apart from an overall phase that does not matter for calculating quantum probabilities, we can visualize unit vectors in $\mathbb{C}^2$ as points on the unit sphere in $\mathbb{R}^3$. The unitary group $U(2)$ comprises linear transformations from $\mathbb{C}^2$ to itself that preserve the inner product. When we conjugate any density matrix ρ by any unitary U, we leave $\mathrm{tr}\rho^2$ unchanged, and so the norm of the Bloch vector (r_x, r_y, r_z) must also remain unchanged. In short, the elements of $U(2)$ must act like symmetries of a sphere, possibly with some extra phase information. This means rotations and reflections. Let's think about the rotations first. To specify a rotation of a sphere, we need to give an axis around which to rotate and an angle to rotate through.

Because any unitary matrix preserves the inner product, it will map one orthonormal basis to another. The columns of the unitary matrix will be the image of the original basis, and so they must be orthogonal to each other. Given any $a, b \in \mathbb{C}$ satisfying $|a|^2 + |b|^2 = 1$, the matrix

$$U = \begin{pmatrix} a & -b^* \\ b & a^* \end{pmatrix} \tag{14.43}$$

is guaranteed to be unitary. Furthermore, its determinant will be 1.

Now, we bring the Pauli matrices into the story. To store the axis around which we rotate, we can use a unit vector; call it $\hat{v}$. We'll also need an angle, which we'll call β. We will prove that

$$\exp(i\beta\hat{v}\cdot\vec{\sigma}) = \cos(\beta)I + i(\hat{v}\cdot\vec{\sigma})\sin(\beta), \tag{14.44}$$

where $\hat{v}\cdot\vec{\sigma} = v_i\sigma_i$ in the Einstein convention.

The Pauli matrices satisfy the commutator relation

$$[\sigma_p, \sigma_q] = 2i\epsilon_{pqr}\sigma_r\,, \tag{14.45}$$

and the anticommutator relation

$$\{\sigma_p, \sigma_q\} = 2\delta_{pq}I\,. \tag{14.46}$$

Adding these and simplifying, we obtain

$$\sigma_p\sigma_q = i\epsilon_{pqr}\sigma_r + \delta_{pq}I\,. \tag{14.47}$$

Multiplying both sides by vector components v_p and w_q, we arrive at the following:

$$(\vec{v}\cdot\vec{\sigma})(\vec{w}\cdot\vec{\sigma}) = (\vec{v}\cdot\vec{w})I + i(\vec{v}\times\vec{w})\cdot\vec{\sigma}\,. \tag{14.48}$$

In the case where the vectors $\vec{v}$ and $\vec{w}$ are identical and of unit length, this reduces to

$$(\hat{v}\cdot\vec{\sigma})^2 = I\,. \tag{14.49}$$

By raising both sides repeatedly to higher powers, we can generalize this to

$$(\hat{v}\cdot\vec{\sigma})^{2k} = I\,, \tag{14.50}$$

for any positive integer k.

Now, we consider $\exp(i\beta\hat{v}\cdot\sigma)$. Using the definition of matrix exponentiation,

$$e^{i\beta\hat{v}\cdot\sigma} = \sum_{k=0}^{\infty} \frac{i^k[\beta(\hat{v}\cdot\vec{\sigma})]^k}{k!}\,. \tag{14.51}$$

We separate the real and imaginary terms in this sum:

$$e^{i\beta\hat{v}\cdot\sigma} = \sum_{k=0}^{\infty} \frac{(-1)^k[\beta(\hat{v}\cdot\vec{\sigma})]^{2k}}{(2k)!} + i\sum_{k=0}^{\infty} \frac{(-1)^k[\beta(\hat{v}\cdot\vec{\sigma})]^{2k+1}}{(2k+1)!}\,. \tag{14.52}$$

Using Eq. (14.50), we can simplify both sums:

$$e^{i\beta\hat{v}\cdot\sigma} = \sum_{k=0}^{\infty} \frac{(-1)^k\beta^{2k}}{(2k)!} I + i(\hat{v}\cdot\vec{\sigma})\sum_{k=0}^{\infty} \frac{(-1)^k\beta^{2k+1}}{(2k+1)!}\,. \tag{14.53}$$

Recognizing the Taylor series for cosine and sine, we arrive at

$$e^{i\beta\hat{v}\cdot\sigma} = I\cos\beta + i(\hat{v}\cdot\vec{\sigma})\sin\beta\,. \tag{14.54}$$

With this formula, we can appreciate an important result. An arbitrary unitary for a single qubit can always be written in the form

$$U = e^{i\alpha}\exp(-i\beta\,\hat{v}\cdot\vec{\sigma}/2)\,, \tag{14.55}$$

for some angles α and β and some unit vector $\hat{v}$. To see why this might be true, let's work out the exponential and substitute in the Pauli matrices:

$$\exp(-i\beta\,\hat{v}\cdot\vec{\sigma}/2) = \begin{pmatrix} \cos\frac{\beta}{2} - iv_z\sin\frac{\beta}{2} & -v_y\sin\frac{\beta}{2} - iv_x\sin\frac{\beta}{2} \\ v_y\sin\frac{\beta}{2} - iv_x\sin\frac{\beta}{2} & \cos\frac{\beta}{2} + iv_z\sin\frac{\beta}{2} \end{pmatrix}. \tag{14.56}$$

The norm of the first column is

$$\cos^2\frac{\beta}{2} + (v_x^2 + v_y^2 + v_z^2)\sin^2\frac{\beta}{2} = 1\,. \tag{14.57}$$

And the two columns are orthogonal, meaning that we have a unitary matrix. Indeed, it is a special unitary matrix, i.e., the determinant equals 1.

If the unit vector $\hat{v}$ points straight "up", and we take $\alpha = 0$, then the resulting unitary works out particularly nicely:

$$U = \begin{pmatrix} e^{-i\beta/2} & 0 \\ 0 & e^{i\beta/2} \end{pmatrix}. \tag{14.58}$$

If we apply this to a vector $|\psi\rangle$, then the two components of $|\psi\rangle$ in the σ_z eigenbasis will pick up a relative phase of β with respect to each other. We are just rotating by β around the vertical axis of the Bloch sphere! And indeed, the set of all these unitaries under matrix multiplication gives us a group isomorphic to $U(1)$.

14.4 From Lie Groups to Lie Algebras

To make further progress, we generalize the observation that we can write any element we please of $SU(2)$ as the exponential of a well-chosen weighted sum. A *Lie algebra* takes the idea of commutator equations for matrices and climbs with it to the mountaintops. Lie groups give us Lie algebras when we study what happens in the neighborhood of the group identity element. Like groups, Lie algebras have representations, structure-preserving maps from an algebra into the world of matrices.

First, we lay out what it means to be a Lie algebra. We can define this without referring to a Lie group, actually. A Lie algebra is a vector space equipped with a particular type of binary operation. Because a Lie algebra is a vector space, we can talk about a basis for it, we can multiply elements by scalars, and we can add elements together. The binary operation, the *bracket* for the algebra, takes two elements and outputs a third. To qualify as a bracket, first of all, this operation has to be linear in both its arguments. (This almost goes without saying; we want it to care about the fact that it lives on a vector space.) Second, we require it to be antisymmetric:

$$[u, v] = -[v, u] \tag{14.59}$$

for all u, v in our algebra. This property is familiar from the matrix commutator. Moreover, we require the bracket to satisfy the *Jacobi identity*:

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0\,. \tag{14.60}$$

This is also true of the matrix commutator. If $\{u_j\}$ is a basis for the algebra, then the bracket of any two basis elements has to be expressible as a linear combination of the whole set:

$$[u_j, u_k] = i f_{jkl} u_l\,, \tag{14.61}$$

where the numbers $\{f_{jkl}\}$ are the *structure constants* of the algebra. These tell us which algebra we're talking about. The inclusion of i in this expression is a convention; typically, physicists tend to do it while mathematicians don't.

Time to relate groups and algebras! Let G be a matrix Lie group. The Lie algebra of G, which we will denote $\mathfrak{g}$, is the set of matrices that exponentiate to give elements of G. More carefully put, it is the set of matrices M for which $e^{itM} \in G$ for all real numbers t. Multiplication in the group G relates to the bracket in the algebra $\mathfrak{g}$, in a manner we will now explore.

Let's consider the product $e^A e^B$, where A and B are matrices. If A and B commute, then any power of one will also commute with any power of the other, and we will get e^{A+B} just like we would with exponentials of numbers. But in general, we'll have

$$e^A e^B = e^C\,, \tag{14.62}$$

and we need to find C. We evaluate the left-hand side using the power series for the exponential:

$$\left(I + A + \frac{1}{2}A^2 + \cdots\right)\left(I + B + \frac{1}{2}B^2 + \cdots\right) = I + A + B + AB + \frac{1}{2}A^2 + \frac{1}{2}B^2 + \cdots\,. \tag{14.63}$$

Meanwhile, the right-hand side is

$$e^C = I + C + \frac{1}{2}C^2 + \cdots . \tag{14.64}$$

We will try thinking of C as a series of ever-improving approximations:

$$C = C_1 + C_2 + \cdots , \tag{14.65}$$

and we take our first approximation to be $C_1 = A + B$. From this, we get

$$\begin{aligned} e^C &= I + A + B + \frac{1}{2}(A+B)^2 + C_2 + \cdots \\ &= I + A + B + \frac{1}{2}A^2 + \frac{1}{2}B^2 + \frac{1}{2}AB + \frac{1}{2}BA + C_2 + \cdots . \end{aligned} \tag{14.66}$$

Our series for $e^A e^B$ and for e^C will agree to this order if we set

$$C_2 = \frac{1}{2}AB - \frac{1}{2}BA = \frac{1}{2}[A,B] . \tag{14.67}$$

So, our approximate answer is that

$$e^A e^B = e^{A+B} e^{\frac{1}{2}[A,B]} . \tag{14.68}$$

This is the first iteration of the *Baker–Campbell–Hausdorff formula.* It goes on to involve nested commutators:

$$C = A + B + \frac{1}{2}[A,B] + \frac{1}{12}([A,[A,B]] + [B,[B,A]]) + \cdots . \tag{14.69}$$

A representation of a Lie algebra is a structure-preserving map from the algebra into a set of matrices, such that the role of the bracket can be played by the matrix commutator.

14.5 Highest-Weight Construction

The $\mathfrak{su}(2)$ Lie algebra is

$$[J_a, J_b] = i\epsilon_{abc} J_c , \tag{14.70}$$

where each index runs over the values x, y, z. For example, this is satisfied by taking $J_a = \frac{1}{2}\sigma_a$. We have also seen how we can take $J_a = \frac{1}{2}\sigma_a \otimes I + I \otimes \frac{1}{2}\sigma_a$, and this also furnishes the same algebra. Thinking about "dipole interaction" Hamiltonians led us to construct

$$J^2 = J_a J_a , \tag{14.71}$$

which commutes with any of the generators:

$$[J^2, J_b] = 0 . \tag{14.72}$$

We introduce a pair of operators that will act to raise and lower:

$$J_\pm = J_x \pm iJ_y \,. \tag{14.73}$$

We can quickly check that

$$J_+J_- = J_x^2 + J_y^2 - i[J_x, J_y] = J_x^2 + J_y^2 + J_z \,, \tag{14.74}$$

and if we multiply in the other order,

$$J_-J_+ = J_x^2 + J_y^2 - J_z \,. \tag{14.75}$$

Taking the difference of these gives the commutator

$$[J_+, J_-] = 2J_z \,, \tag{14.76}$$

and we can also easily show that

$$[J_z, J_\pm] = \pm J_\pm \,. \tag{14.77}$$

Moreover,

$$J^2 = J_+J_- + J_z^2 - J_z = J_-J_+ + J_z^2 + J_z \,, \tag{14.78}$$

from which we deduce

$$[J_\pm, J^2] = 0 \,. \tag{14.79}$$

Using what we know so far, we can verify that

$$J_\pm J_\mp = J_x^2 + J_y^2 \pm J_z \tag{14.80}$$

which implies that

$$[J_+, J_-] = 2J_z \,. \tag{14.81}$$

The commutators of $J_\pm$ with J_z are also calculated easily enough:

$$[J_z, J_\pm] = \pm J_\pm \,. \tag{14.82}$$

Because J_z commutes with J^2, we can simultaneously diagonalize them. It is conventional to label the simultaneous eigenvectors of these operators with a pair of numbers, j and m. The convention for J_z is simple enough:

$$J_z|j,m\rangle = m|j,m\rangle \,. \tag{14.83}$$

But the tradition for J^2 is a little funny-looking:

$$J^2|j,m\rangle = j(j+1)|j,m\rangle \,. \tag{14.84}$$

Why might this be the right way to go about it? Well, let's take the example with which we are most familiar, the Pauli matrices (here scaled by 1/2). We can quickly tell that

$$J_z|\pm z\rangle = \pm\frac{1}{2}|\pm z\rangle \,, \tag{14.85}$$

so calling this an $m = \frac{1}{2}$ vector seems reasonable enough. Each Pauli matrix squares to the identity, so each J_a squares to *one-quarter* of the identity, and

$$J^2|\pm z\rangle = \frac{3}{4}|\pm z\rangle = \frac{1}{2}\frac{3}{2}|\pm z\rangle\,. \tag{14.86}$$

We have here factored the constant suggestively, on the hunch that the one-half we saw before should be taken into consideration.

Next, we take the other example we know something about,

$$J_a = \frac{1}{2}\sigma_a \otimes I + I \otimes \frac{1}{2}\sigma_a\,. \tag{14.87}$$

We refer back to the triplet-singlet basis, and we find that the triplet vectors are eigenvectors of this J^2 with eigenvalue 2, while the singlet is an eigenvector with eigenvalue 0. In summary, things seem to hold together if j is $0, \frac{1}{2}, 1, \ldots$, the dimension is $2j+1$ and the eigenvalue of J^2 is $j(j+1)$. We will see as we go that this is a good choice.

Because they commute with J^2, the $J_\pm$ do not affect the eigenvalue j:

$$J^2(J_\pm|j,m\rangle) = J_\pm J^2|j,m\rangle = j(j+1)J_\pm|j,m\rangle\,. \tag{14.88}$$

But on the other hand,

$$\begin{aligned} J_z(J_\pm|j,m\rangle) &= ([J_z,J_\pm] + J_\pm J_z)|j,m\rangle \\ &= (\pm J_\pm + mJ_\pm)|j,m\rangle \\ &= (m\pm 1)J_\pm|j,m\rangle\,. \end{aligned} \tag{14.89}$$

The vector $J_\pm|j,m\rangle$, whatever it is, turns out to be an eigenvector of J_z, with an eigenvalue that we can read off: $m \pm 1$. Consequently, we can think of J_+ as a *raising* operator and J_- as a *lowering* operator.

Accounting for normalization, we have just shown that

$$J_\pm|j,m\rangle = C_\pm(j,m)|j,m\pm 1\rangle\,, \tag{14.90}$$

where $C_\pm(j,m)$ is a function we have yet to figure out. We can grapple with it using the fact that the adjoint of $J_\pm$ is $J_\mp$. So,

$$\langle j,m|J_\mp = \langle j,m\pm 1|C_\pm(j,m)^*\,, \tag{14.91}$$

which combines with our previous result to yield

$$|C_\pm(j,m)|^2 = \langle j,m|J_\mp J_\pm|j,m\rangle\,. \tag{14.92}$$

We want to turn the product of operators into something involving J^2 and J_z so we can evaluate it in terms of eigenvalues. And, fortunately,

$$|C_\pm(j,m)|^2 = \langle j,m|(J^2 - J_z^2 \mp J_z)|j,m\rangle = j(j+1) - m^2 \mp m\,. \tag{14.93}$$

This looks a little more tidy if we write it as

$$|C_\pm(j,m)|^2 = j(j+1) - m(m\pm 1)\,. \tag{14.94}$$

We can take $C_\pm(j,m)$ to be real and nonnegative:

$$C_\pm(j,m) = \sqrt{j(j+1) - m(m\pm 1)}\,. \tag{14.95}$$

This degenerates if we try to raise past $m = j$, since

$$J_+|j,j\rangle = 0\,. \tag{14.96}$$

Likewise, if we try to lower beyond $m = -j$, we discover that

$$J_-|j,-j\rangle = 0\,. \tag{14.97}$$

So, the allowed values of m go in steps of 1 from $-j$ to j. The distance between these endpoints, $2j$, must be an integer, and we recover the pattern hinted at before:

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots\,. \tag{14.98}$$

For each value of j, the vectors $\{|j,m\rangle\}$ provide an orthonormal basis for a space of dimension $2j+1$.

14.6 Example: Ortho-Hydrogen Molecular Beam

Let's get ourselves closer to experiment by doing an example. Here is one that I had to solve for my Qualifyings in graduate school:

> (a) A beam of ortho-hydrogen molecules (proton spin state $S = 1$), traveling along the y axis, passes through a Stern–Gerlach apparatus (SG1) with its magnetic field along the x axis. The emerging molecules with $M_S = 1$ are passed through a second Stern–Gerlach apparatus (SG2) with its magnetic field along the x axis. A constant magnetic field $\vec{B}$ in the z direction acts along the part of the path between the magnets. Show that the fractions of molecules emerging from the three output channels, corresponding to $m_s = 1$, $m_s = 0$ and $m_s = -1$ of the second magnet are
>
> $$\cos^4(\omega t/2),\ 2\sin^2(\omega t/2)\cos^2(\omega t/2),\ \sin^4(\omega t/2), \tag{14.99}$$
>
> where $\omega = 2\mu_p B/\hbar$, t is the time the molecules spend in the field $\vec{B}$, and μ_p is the magnetic moment of the proton.
>
> (b) For a path length in $\vec{B}$ of 20 mm, and molecules with energy 25 keV, it is found that no molecules emerge from the $m_s = 1$ channel of the second magnet when
>
> $$B = 1.80\left(n + \frac{1}{2}\right)\ \mathrm{T}, \tag{14.100}$$

where n is an integer. Deduce the value of μ_p in nuclear magnetons. (The spins of the electrons in ortho-hydrogen are anti-parallel.)

Neglecting all other possible degrees of freedom which the hydrogen molecules might possess, we can treat them as qutrits. We prepare the beam along the x axis, evolve it subject to a magnetic field along the z axis, and measure it along the x axis. We can represent the effect of the preparation procedure by ascribing the eigenstate of S_x with eigenvalue $+1$ to each molecule. The magnetic field along the z axis imposes an energy difference between the eigenstates of S_z. We end by measuring S_x. So, we need to write the $+1$ eigenstate of S_x in terms of the eigenstates of S_z, multiply the components by the appropriate phase factors, and then write the result in terms of the S_x eigenstates again. The coefficients of the resulting components will be the amplitudes which we square to get the Born-rule probabilities we want.

First, we relate the S_z and S_x eigenstates. This all comes back to the representation theory of the Lie algebra $\mathfrak{su}(2)$. The standard approach is to simultaneously diagonalize the operators S^2 and S_z and use their eigenvalues as quantum numbers. Let $|1\rangle$, $|0\rangle$ and $|-1\rangle$ be the eigenstates of S^2 and S_z where the eigenvalue of S^2 is $s = 1$ and the eigenvalues of S_z are 1, 0 and -1. We define the standard raising and lowering operators,

$$S_+ = S_x + iS_y,\ S_- = S_x - iS_y. \tag{14.101}$$

These act on eigenstates by

$$S_+|m\rangle = \sqrt{(s-m)(s+m+1)}|m+1\rangle, \tag{14.102}$$

$$S_-|m\rangle = \sqrt{(s+m)(s-m+1)}|m-1\rangle. \tag{14.103}$$

The results which do not annihilate are

$$S_+|-1\rangle = \sqrt{2}|0\rangle, \tag{14.104}$$

$$S_+|0\rangle = \sqrt{2}|1\rangle, \tag{14.105}$$

$$S_-|0\rangle = \sqrt{2}|-1\rangle, \tag{14.106}$$

$$S_-|1\rangle = \sqrt{2}|0\rangle. \tag{14.107}$$

Consequently, the results of applying S_x to the eigenstates of S_z are

$$S_x|-1\rangle = \frac{\sqrt{2}}{2}|0\rangle, \tag{14.108}$$

$$S_x|0\rangle = \frac{\sqrt{2}}{2}\left(|-1\rangle + |1\rangle\right), \tag{14.109}$$

$$S_x|1\rangle = \frac{\sqrt{2}}{2}|0\rangle. \tag{14.110}$$

In matrix form, we can say that the spin-1 representation of S_x is

$$S_x = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{14.111}$$

This matrix has eigenvectors

$$\frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix},\ \frac{1}{2}\begin{pmatrix} \sqrt{2} \\ 0 \\ -\sqrt{2} \end{pmatrix},\ \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}, \tag{14.112}$$

with eigenvalues -1, 0 and $+1$ respectively.

Our initial state vector is the $+1$ state of S_x,

$$\psi(0) = \frac{1}{2}\left(|1\rangle + \sqrt{2}|0\rangle + |-1\rangle\right). \tag{14.113}$$

Because the magnetic moment is $\mu = 2\mu_p$, the energies of the S_z eigenstates are

$$E_1 = -2\mu_p B,\ E_0 = 0,\ E_1 = 2\mu_p B. \tag{14.114}$$

So,

$$\psi(t) = \frac{1}{2}\left(e^{i\omega t}|1\rangle + \sqrt{2}|0\rangle + e^{-i\omega t}|-1\rangle\right). \tag{14.115}$$

We need this back in terms of the S_x eigenstates to get the probabilities for the final measurement. We must have

$$\psi(t) = c_1|1\rangle_x + c_0|0\rangle_x + c_{-1}|-1\rangle_x, \tag{14.116}$$

for some complex coefficients. In terms of the S_z eigenstates,

$$\psi(t) = \frac{c_1}{2}\left(|1\rangle + \sqrt{2}|0\rangle + |-1\rangle\right) + \frac{c_0}{\sqrt{2}}\left(|1\rangle - |-1\rangle\right) + \frac{c_{-1}}{2}\left(|1\rangle - \sqrt{2}|0\rangle + |-1\rangle\right). \tag{14.117}$$

Now, it's a matter of matching coefficients.

$$e^{i\omega t} = c_1 + c_{-1} + \sqrt{2}c_0, \tag{14.118}$$

$$e^{-i\omega t} = c_1 + c_{-1} - \sqrt{2}c_0, \tag{14.119}$$

$$1 = c_1 - c_{-1}. \tag{14.120}$$

Adding the first two equations gives

$$c_1 + c_{-1} = \cos\omega t, \tag{14.121}$$

and subtracting them gives

$$c_0 = \frac{i}{\sqrt{2}}\sin\omega t. \tag{14.122}$$

Recall the addition formulas for sines and cosines:

$$\cos(a+b) = \cos a\cos b - \sin a\sin b, \tag{14.123}$$

$$\sin(a+b) = \sin a\cos b + \sin b\cos a. \tag{14.124}$$

When the two angles are equal, these simplify to the double-angle formulas,

$$\cos 2a = \cos^2 a - \sin^2 a, \tag{14.125}$$
$$\sin 2a = 2 \sin a \cos a, \tag{14.126}$$

which we can write just as well as half-angle formulas:

$$\cos a = \cos^2(a/2) - \sin^2(a/2), \tag{14.127}$$
$$\sin a = 2 \sin(a/2) \cos(a/2). \tag{14.128}$$

Our results above imply that

$$c_1 = \frac{1}{2} \cos \omega t + \frac{1}{2}, \tag{14.129}$$

which we can rewrite using the half-angle formula and the Pythagorean identity as

$$c_1 = \frac{1}{2} \left(\cos^2(\omega t/2) - \sin^2(\omega t/2) \right) + \frac{1}{2} \left(\cos^2(\omega t/2) + \sin^2(\omega t/2) \right). \tag{14.130}$$

This simplifies to

$$c_1 = \cos^2(\omega t/2). \tag{14.131}$$

For c_{-1}, we get

$$c_{-1} = \frac{1}{2} \cos \omega t - \frac{1}{2}, \tag{14.132}$$

which we rewrite as

$$c_{-1} = \frac{1}{2} \left(\cos^2(\omega t/2) - \sin^2(\omega t/2) \right) - \frac{1}{2} \left(\cos^2(\omega t/2) + \sin^2(\omega t/2) \right), \tag{14.133}$$

yielding after cancelation

$$c_{-1} = -\sin^2(\omega t/2). \tag{14.134}$$

Applying the Born rule to each coefficient gives

$$P_1(t) = \cos^4(\omega t/2), \tag{14.135}$$
$$P_0(t) = 2 \sin^2(\omega t/2) \cos^2(\omega t/2), \tag{14.136}$$
$$P_{-1}(t) = \sin^4(\omega t/2). \tag{14.137}$$

Whew!

Now, for the next step, note that $P_1(t)$ vanishes when

$$\frac{\omega t}{2} = \frac{\pi}{2}, \frac{3\pi}{2}, \ldots \tag{14.138}$$

So, the expected intensity of the $+1$ beam coming out of the second Stern–Gerlach apparatus drops to zero when

$$\frac{2\mu_p B t}{2\hbar} = \pi \left(n + \frac{1}{2} \right). \tag{14.139}$$

If the magnetic field strength at which the measured intensity is zero follows the relationship

$$B = B_0 \left(n + \frac{1}{2} \right), \tag{14.140}$$

then we can say that

$$\mu_p = \frac{\pi \hbar}{B_0 t}. \tag{14.141}$$

The given kinetic energy of 25 keV is so much smaller than the rest mass of a hydrogen molecule (roughly 2 GeV) that we can treat the molecules nonrelativistically. The travel time over a distance L is therefore

$$t = \frac{L}{v} = L\sqrt{\frac{m}{2K}}. \tag{14.142}$$

For m, we can use the mass of a hydrogen molecule, which is roughly twice the mass of a proton ($938\,\mathrm{MeV}/c^2$) plus twice the mass of an electron ($511\,\mathrm{keV}/c^2$), or $1877\,\mathrm{MeV}/c^2$. So,

$$t = 2.0 \times 10^{-2}\,\mathrm{m} \left(\frac{1877\,\mathrm{MeV}/c^2}{0.050\,\mathrm{MeV}} \right)^{1/2} = \frac{2.0 \times 10^{-2}\,\mathrm{m} \cdot 194}{3 \times 10^8\,\mathrm{m/s}} \approx 1.29 \times 10^{-8}\,\mathrm{s}. \tag{14.143}$$

Using $\hbar \approx 1.055 \times 10^{-34}\,\mathrm{Js}$,

$$\mu_p = \frac{\pi \cdot 1.055 \times 10^{-34}\,\mathrm{Js}}{1.8\,\mathrm{T} \cdot 1.29 \times 10^{-8}\,\mathrm{s}} \approx 1.35 \times 10^{-26}\,\mathrm{J/T}. \tag{14.144}$$

The nuclear magneton μ_N is approximately $5.051 \times 10^{-27}\,\mathrm{J/T}$, so

$$\mu_p \approx 2.82 \mu_N. \tag{14.145}$$

The literature value is approximately $2.79\mu_N$.

14.7 $SU(2)$ and $SO(3)$

The group $SU(2)$ is *simply connected.* We can appreciate what this means by using the parameterization of it mentioned earlier: $SU(2)$ comprises all the unitaries of the form

$$U = \begin{pmatrix} a & -b^* \\ b & a^* \end{pmatrix}, \tag{14.146}$$

with the normalization condition

$$|a|^2 + |b|^2 = 1\,. \tag{14.147}$$

In terms of the real and imaginary parts of these variables, we have

$$|\mathrm{Re}\, a|^2 + |\mathrm{Im}\, a|^2 + |\mathrm{Re}\, b|^2 + |\mathrm{Im}\, b|^2 = 1\,. \tag{14.148}$$

The group $SU(2)$ is indeed a manifold, specifically a 3-sphere in $\mathbb{R}^4$. A sphere has the property that any loop drawn upon it can be continuously shrunk to a point. (If we tried to do this for an arbitrary loop on a torus, we'd be out of luck.) Any Lie group of matrices that forms one connected piece and has the loop-contractibility property is simply connected. One can prove that the the representations of a Lie group of matrices G and its Lie algebra $\mathfrak{g}$ are in one-to-one correspondence if and only if G is simply connected.

As it so happens, the spatial rotation group $SO(3)$ is *not* simply connected, even though it "looks just like" $SU(2)$ in the neighborhood of the identity. Instead of being shaped like the sphere S^3, the group $SO(3)$ is shaped like like S^3 with antipodal points identified. This can also be expressed by saying that $SO(3)$ is isomorphic to the real projective space $\mathbb{R}P^3$, the space of lines through the origin of $\mathbb{R}^4$.

One way of getting at this rather important result is to establish that there exists a structure-preserving map from $SU(2)$ to $SO(3)$ that is *two-to-one.* We can find a function

$$\Phi : SU(2) \to SO(3) \tag{14.149}$$

that "hits everything in the target" and sends two distinct points in $SU(2)$ to each point in $SO(3)$. The idea is just to think of what a unitary $U \in SU(2)$ does to the Bloch sphere when it acts by conjugation. It implements a rotation — but both $I \in SU(2)$ and $-I \in SU(2)$ produce the identity in $SO(3)$. For each point in $SU(2)$, there is a point geometrically opposite to it that yields the same rotation. (Remember, $SU(3)$ is a sphere in *four* dimensions, while the Bloch sphere lives in three.)

One big physical consequence of all this is when we are talking about spatial angular momentum, rather than the "intrinsic" angular momentum of particles with spin, we lose the half-integer spin representations. The algebraic theory of $\mathfrak{su}(2)$ gives us j values of $0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$, but the geometry of space restricts j to the nonnegative integers.

15 Spherically Symmetric Potentials

In this chapter, we will develop expertise that is relevant to problems that exhibit spherical symmetry. This includes situations where two particles interact via some effect whose strength depends only upon the distance between them, because we can shift our perspective to a reference frame that fixes their center of mass.

15.1 Relative and Center-of-Mass Variables

Consider two particles, whose position operators shall be $\vec{x}_1$ and $\vec{x}_2$, and whose momentum operators are $\vec{p}_1$ and $\vec{p}_2$. Supposing that the interaction between the particles depends only on the *displacement* between them, we can write a Hamiltonian for the two-body system as follows:

$$H = \frac{\vec{p}_1^2}{2m_1} + \frac{\vec{p}_2^2}{2m_2} + V(\vec{x}_1 - \vec{x}_2). \tag{15.1}$$

We can simplify this problem by dividing up its degrees of freedom between the *center-of-mass motion* and the *relative motion* of the particles. Following classical intuition, we introduce new variables,

$$\vec{x}_{\text{CM}} = \frac{m_1\vec{x}_1 + m_2\vec{x}_2}{m_1 + m_2}, \tag{15.2}$$

$$\vec{x} = \vec{x}_1 - \vec{x}_2, \tag{15.3}$$

$$\vec{p}_{\text{CM}} = \vec{p}_1 + \vec{p}_2, \tag{15.4}$$

$$\vec{p} = \frac{m_2\vec{p}_1 - m_1\vec{p}_2}{m_1 + m_2}. \tag{15.5}$$

Knowing the commutators of the original position and momentum operators, it's not too hard to show that for any components $x_{\text{CM}j}$ and $p_{\text{CM}k}$ of the center-of-mass operators, the commutator is just

$$[x_{\text{CM}j}, p_{\text{CM}k}] = i\hbar\delta_{jk}. \tag{15.6}$$

The same holds for the *relative* degrees of freedom, too:

$$[x_j, p_k] = i\hbar\delta_{jk}. \tag{15.7}$$

In other words, our new operators are just as good position and momentum operators as the old.

By chugging through a little algebra, we can rewrite our original two-body Hamiltonian using the new operators:

$$H = \frac{\vec{p}_{\text{CM}}^2}{2M} + \frac{\vec{p}^2}{2\mu} + V(\vec{x}). \tag{15.8}$$

Here, we've written M for the sum of the masses,

$$M = m_1 + m_2, \tag{15.9}$$

and μ for the *reduced mass*

$$\mu = \left(\frac{1}{m_1} + \frac{1}{m_2}\right)^{-1}. \tag{15.10}$$

When it's written this way, it's easy to see that the Hamiltonian H splits into two parts:

$$H = H_{\text{CM}} + H_r, \tag{15.11}$$

where

$$H_{\text{CM}} = \frac{\vec{p}_{\text{CM}}^2}{2M} \tag{15.12}$$

and

$$H_r = \frac{\vec{p}^2}{2\mu} + V(\vec{x}). \tag{15.13}$$

What's more, these two parts commute with each other,

$$[H_{\text{CM}}, H_r] = 0, \tag{15.14}$$

so we can pick a basis which simultaneously diagonalizes both of them. An eigenstate $|\psi\rangle$ of H satisfies the equation

$$H|\psi\rangle = E|\psi\rangle, \tag{15.15}$$

and the same ket is an eigenket of both H_{CM} and H_r:

$$H_{\text{CM}}|\psi\rangle = E_{\text{CM}}|\psi\rangle, \tag{15.16}$$

$$H_r|\psi\rangle = E_r|\psi\rangle, \tag{15.17}$$

with $E = E_{\text{CM}} + E_r$. Eigenstates of the total Hamiltonian will be tensor products of eigenstates of the CM and relative parts:

$$|\psi\rangle = |\chi\rangle_{\text{CM}} \otimes |\omega\rangle_r, \tag{15.18}$$

let's say. We can rewrite the CM Schrödinger equation above as a statement about position-space wavefunctions, if we take derivatives with respect to CM coordinates:

$$-\frac{\hbar^2}{2M}\partial_{\text{CM}}^2 \chi(\vec{x}_{\text{CM}}) = E_{\text{CM}}\chi(\vec{x}_{\text{CM}}). \tag{15.19}$$

This is just the Schrödinger equation for a free particle. We know this equation rather well! For example, we can say that the momentum eigenstates in position space are plane waves,

$$\chi(\vec{x}_{\rm CM}) = (2\pi\hbar)^{-3/2} \exp\left(\frac{i}{\hbar}\vec{p}_{\rm CM}\cdot\vec{x}_{\rm CM}\right), \tag{15.20}$$

with energies

$$E_{\rm CM} = \frac{\vec{p}^2_{\rm CM}}{2M}. \tag{15.21}$$

We've seen that the system of two coupled particles splits into a center-of-mass part, whose energy is just the translational kinetic energy of a free particle, and a part which embodies the *relative motion* of the two particles. In classical mechanics, we saw the same thing: for example, in the Earth–Moon system, both bodies are orbiting around their common center of mass (which is actually beneath the Earth's surface). For the hydrogen atom, our two particles will be a proton and an electron, bound together by the Coulomb interaction between them. Note that the mass of the proton, m_p, is much larger than that of the electron, m_e, so that the "reduced mass" is very close to the electron mass:

$$\mu = \frac{m_e m_p}{m_e + m_p} = \frac{m_e}{1 + m_e/m_p} \approx m_e\left(1 - \frac{m_e}{m_p}\right). \tag{15.22}$$

The correction term, m_e/m_p, is roughly one part in 1800. This means that the CM frame is *almost* coincident with the rest frame of the proton. We'll be a bit sloppy, but excusably so, if we refer to the relative degrees of freedom as the "electron" part of the system and the CM as the "nucleus."

15.2 Rotation Generators and Angular Momentum

In a decent high-school physics course, one learns that the angular momentum of a point particle can be written

$$\vec{L} = \vec{x}\times\vec{p}. \tag{15.23}$$

When we move from classical to quantum physics, we promote variables to operators whose commutation relations encode the complementarity relationships between observable physical quantities. If we impose the canonical commutation relation

$$[x_j, p_k] = i\hbar\delta_{jk}, \tag{15.24}$$

then it shall come to pass that the components of the quantum angular momentum operator satisfy

$$[L_j, L_k] = i\hbar\epsilon_{jkl}L_l. \tag{15.25}$$

This commutator has an important geometrical meaning. From trigonometry, we can write matrix expressions for rotations in three-dimensional space. For example,

rotating the point (x_1, x_2, x_3) by θ radians around the 3-axis can be represented by the 3×3 matrix

$$R_3(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{15.26}$$

and the corresponding rotation matrices $R_1(\theta)$ and $R_2(\theta)$ for the other axes will likewise have $\sin\theta$ and $\cos\theta$ written in the appropriate slots. If we consider rotations by *infinitesmal* angles θ, such that all terms of higher than linear order in θ can be dropped, then each rotation matrix $R_i(\theta)$ can be written as the identity matrix plus θ times a *generator* which does not depend on θ:

$$R_j(\theta) = I - i\theta J_j. \tag{15.27}$$

Matrix arithmetic will verify that

$$[J_j, J_k] = i\epsilon_{jkl} J_l. \tag{15.28}$$

The generators of 3D rotations are, up to Planck's constant, the components of the angular momentum vector.

In classical mechanics, the rotational kinetic energy is proportional to the square of the angular momentum. Let's work this out using the Einstein summation convention:

$$\begin{aligned} \vec{L} \cdot \vec{L} &= L_c L_c && (15.29)\\ &= \epsilon_{abc} r_a p_b \epsilon_{mnc} r_m p_n && (15.30)\\ &= (\delta_{am}\delta_{bn} - \delta_{an}\delta_{bm}) r_a p_b r_m p_n && (15.31)\\ &= r_a r_a p_b p_b - r_a p_b r_b p_a \,. && (15.32) \end{aligned}$$

Therefore,

$$\vec{L} \cdot \vec{L} = (\vec{r} \cdot \vec{r})(\vec{p} \cdot \vec{p}) - (\vec{r} \cdot \vec{p})^2 \tag{15.33}$$

which we can write in terms of the angle between the position and momentum vectors:

$$L^2 = r^2 p^2 (1 - \cos^2\theta) \,. \tag{15.34}$$

So, the magnitude of the angular momentum L is given by $rp\sin\theta$.

In quantum mechanics, the components of $\vec{r}$ and of $\vec{p}$ all become Hermitian operators that satisfy the canonical commutation relations:

$$[r_a, p_b] = i\hbar\delta_{ab} \,. \tag{15.35}$$

Consequently, we can't just move symbols past each other. The derivation we stepped through above works out in much the same way, except that we get an extra term proportional to $i\hbar$:

$$\vec{L} \cdot \vec{L} = (\vec{r} \cdot \vec{r})(\vec{p} \cdot \vec{p}) - (\vec{r} \cdot \vec{p}) + i\hbar(\vec{r} \cdot \vec{p}) \,. \tag{15.36}$$

We can use this to write the kinetic energy of a point mass m in a way that incorporates the angular momentum:

$$\frac{p^2}{2m} = \frac{1}{2m}\frac{1}{r^2}\left((\vec{r}\cdot\vec{p})^2 - i\hbar(\vec{r}\cdot\vec{p})\right) + \frac{L^2}{2mr^2}\,. \tag{15.37}$$

The operator $\vec{p}$ takes derivatives with respect to position, so the operator $\vec{r}\cdot\vec{p}$ has to include a derivative in the direction of $\vec{r}$:

$$\vec{r}\cdot\vec{p} = -i\hbar r\frac{\partial}{\partial r}\,. \tag{15.38}$$

So, the combination of $(\vec{r}\cdot)\vec{p}$-dependent terms in the kinetic energy becomes

$$(\vec{r}\cdot\vec{p})^2 - i\hbar(\vec{r}\cdot\vec{p}) = -\hbar^2\left(r\frac{\partial}{\partial r}r\frac{\partial}{\partial r} + r\frac{\partial}{\partial r}\right)\,. \tag{15.39}$$

Using the product rule, this simplifies slightly to

$$(\vec{r}\cdot\vec{p})^2 - i\hbar(\vec{r}\cdot\vec{p}) = -\hbar^2\left(r^2\frac{\partial^2}{\partial r^2} + 2r\frac{\partial}{\partial r}\right)\,. \tag{15.40}$$

Dividing through by r^2 gives

$$\frac{1}{r^2}(\vec{r}\cdot\vec{p})^2 - i\hbar(\vec{r}\cdot\vec{p}) = -\hbar^2\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) \tag{15.41}$$

$$= -\hbar^2\frac{1}{r}\frac{\partial^2}{\partial r^2}r\,. \tag{15.42}$$

Therefore, a one-particle Hamiltonian with a potential $V(\vec{r})$ can be written

$$H = \frac{p^2}{2m} + V(\vec{r}) = -\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{L^2}{2mr^2} + V(\vec{r})\,. \tag{15.43}$$

This result is more often approached by way of coordinate transformations, a method to which we will turn in the next section. For now, we observe that a direct computation from the definitions will yield

$$[L_j, r_k] = i\hbar\epsilon_{jkn}x_n\,, \tag{15.44}$$

$$[L_j, p_k] = i\hbar\epsilon_{jkn}p_n\,. \tag{15.45}$$

Moreover, if $\vec{u}$ and $\vec{v}$ are any tuples that transform in this way, then

$$[L_c, \vec{u}\cdot\vec{v}] = [L_c, u_a v_a] \tag{15.46}$$

$$= [L_c, u_a]v_a + u_a[L_c, v_a] \tag{15.47}$$

$$= i\hbar\epsilon_{jca}u_j v_a + i\hbar\epsilon_{jca}u_a v_j \tag{15.48}$$

$$= i\hbar(\epsilon_{jca}u_j v_a + \epsilon_{jca}u_a v_j) \tag{15.49}$$

$$= i\hbar(\epsilon_{jca}u_j v_a + \epsilon_{jac}u_j v_a) \tag{15.50}$$

$$= 0\,. \tag{15.51}$$

This is what we mean when we say that $\vec{r}$ and $\vec{p}$ transform like *vectors,* while their dot product transforms as a *scalar.* Furthermore, it tells us right away that each component of the angular momentum will commute with r^2, with p^2 and with L^2.

15.3 Spherical Coordinates

Our problem of interest has an intrinsic spherical symmetry: the interaction potential does not depend upon direction, only distance. However, we've been working in Cartesian coordinates. Let's shift ourselves into a coordinate system which reflects the symmetry of the problem, so we can get at its essence more easily. How can we turn the Cartesian coordinates x_i into spherical coordinates?

Like this [26]:

$$x_1 = r\sin\theta\cos\phi\,, \tag{15.52}$$
$$x_2 = r\sin\theta\sin\phi\,, \tag{15.53}$$
$$x_3 = r\cos\theta\,. \tag{15.54}$$

Although if we were mathematicians, our θ and ϕ would be defined the other way round.

It follows from these relations that the Cartesian components of the angular momentum can be written in the following way [26]:

$$L_1 = i\hbar\left(\sin\phi\partial_\theta + \frac{\cos\phi}{\tan\theta}\partial_\phi\right)\,. \tag{15.55}$$
$$L_2 = i\hbar\left(-\cos\phi\partial_\theta + \frac{\sin\phi}{\tan\theta}\partial_\phi\right)\,. \tag{15.56}$$
$$L_3 = -i\hbar\partial_\phi\,. \tag{15.57}$$

In addition, we'll be working with the square of the magnitude of the angular momentum,

$$\vec{L}^2 = L_1^2 + L_2^2 + L_3^2, \tag{15.58}$$

which in spherical coordinates works out to be

$$\vec{L}^2 = -\hbar^2\left(\partial_\theta^2 + \frac{1}{\tan\theta}\partial_\theta + \frac{1}{\sin^2\theta}\partial_\phi^2\right)\,. \tag{15.59}$$

15.4 Angular Momentum Eigenfunctions

The next step along this well-trod path is to pick one component of the angular momentum, traditionally L_3, and then simultaneously diagonalize L_3 and $\vec{L}^2$. The eigenfunctions we find are the *spherical harmonics* [26], which are labeled by two integers, the quantum numbers l and m. The quantum number l specifies the eigenvalue of $\vec{L}^2$ and can take the values $0, 1, 2, 3, \ldots$, while for any value of l, the number m ranges from $-l$ to $+l$. We write

$$\vec{L}^2 Y_l^m = l(l+1)\hbar^2 Y_l^m\,, \tag{15.60}$$
$$L_3 Y_l^m = m\hbar Y_l^m\,. \tag{15.61}$$

Spherical harmonics occur in classical electromagnetism as well as in quantum mechanics, since the Laplacian also appears in equations for the electrical potential.

One way to understand the whole shebang, with both the l and the m dependence, is to bring out the representation theory of $\mathfrak{su}(2)$. The angular-momentum operators satisfy the $\mathfrak{su}(2)$ algebra, apart from some factors of $\hbar$ that are really just there for convention, so everything we know about $\mathfrak{su}(2)$, we can apply here. This is, if you like, a fusion of discrete and continuum methods. We are considering functions Y_l^m that are defined to be continuous on the sphere, but we are constructing a discrete set of them, so that they furnish a finite-dimensional representation of $\mathfrak{su}(2)$. For each value of l, we will obtain a finite set of functions that are simultaneous eigenfunctions of both $\vec{L}^2$ and L_z, and these eigenfunctions provide a finite-dimensional vector basis.

Just as we did in the highest-weight construction for $\mathfrak{su}(2)$, we introduce raising and lowering operators by taking linear combinations of our generators:

$$L_\pm = L_x \pm iL_y \,. \tag{15.62}$$

But now we have expressions for L_x and L_y as derivatives in spherical coordinates. With a little work, we obtain

$$L_\pm = \pm\hbar e^{\pm i\phi}\left(\frac{\partial}{\partial\theta} \pm i\cot\theta\frac{\partial}{\partial\phi}\right) . \tag{15.63}$$

Next, we follow the same strategy that we used for the harmonic oscillator: We look for a state that is annihilated by an operator and thus satisfies a differential equation that we can manage. Then, we construct all the other states we need by applying operators whose algebra we understand to that initial state. In this case, we focus on the eigenstate of highest m, which must be annihilated by the raising operator:

$$L_z Y_l^l = \hbar l Y_l^l,\ L_+ Y_l^l = 0 \,. \tag{15.64}$$

From the former condition, we guess that this spherical harmonic will have the form

$$Y_l^l \propto e^{il\phi} f(\theta) \,. \tag{15.65}$$

Differentiating this expression with respect to ϕ just pulls down an il, while differentiating with respect to θ turns the unknown function $f(\theta)$ into $f'(\theta)$. For the raising operator L_+ to annihilate this expression, we must have

$$f'(\theta) + i\cot\theta(il)f(\theta) = 0 \,. \tag{15.66}$$

So, the unknown function satisfies

$$f'(\theta) = l\frac{\cos\theta}{\sin\theta}f(\theta) \,. \tag{15.67}$$

From long experience with derivatives, we know that we can get a prefactor of l by differentating something raised to the l power. An easy function to try would be

$(\sin\theta)^l$, and indeed this fits the bill: We get an l in front, one less factor of $\sin\theta$, and a $\cos\theta$ from the chain rule. Therefore,

$$Y_l^l \propto e^{il\phi}(\sin\theta)^l\,. \tag{15.68}$$

Properly normalized,

$$Y_l^l = \frac{1}{2^l l!}\sqrt{\frac{(2l+1)!}{4\pi}}(-1)^l e^{il\phi}(\sin\theta)^l\,. \tag{15.69}$$

The other spherical harmonics in the l multiplet can be found by repeatedly lowering with L_-.

Spherical harmonics are often written in terms of another set of functions, by splitting off part of the m dependence. For $m \geq 0$, we write

$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}(-1)^m e^{im\phi}P_l^m(\cos\theta)\,. \tag{15.70}$$

And for $m < 0$, we take

$$Y_l^m(\theta,\phi) = (-1)^m[Y_l^{-m}(\theta,\phi)]^*\,. \tag{15.71}$$

The new functions are the *associated Legendre functions.* When $m = 0$, they turn out to be polynomials in $\cos\theta$, known as the *Legendre polynomials* and typically written with the m suppressed for brevity. By plugging this expression for Y_l^m into the eigenvalue equation for the $\vec{L}^2$ operator, we get the differential equation

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{dP_l^m}{d\theta}\right) + (l(l+1)\sin^2\theta - m^2)P_l^m = 0\,. \tag{15.72}$$

Changing variables to $x = \cos\theta$, a little manipulation yields

$$(1-x^2)\frac{d}{dx}\left((1-x^2)\frac{dP_l^m}{dx}\right) + (l(l+1)(1-x^2) - m^2)P_l^m = 0\,. \tag{15.73}$$

We divide through by $1-x^2$ to obtain

$$\frac{d}{dx}\left((1-x^2)\frac{dP_l^m}{dx}\right) + \left(l(l+1) - \frac{m^2}{1-x^2}\right)P_l^m = 0\,. \tag{15.74}$$

It is worthwhile to focus on the Legendre polynomials themselves for a moment. As noted above, they arise in classical physics too, such as electrostatics. Imagine that we are trying to find the electrical potential due to a localized charge distribution. We center the origin of our coordinates within this distribution and then break down the charge density into points, adding up the potential due to each speck of charge. If we are at a distance z up from the origin and the charge speck is at a distance a and tilted from the axis by an angle θ, then the potential it induces will be proportional

to $(z^2 + a^2 - 2za\cos\theta)^{-1/2}$. Writing $x = \cos\theta$ and factoring away the units inspires the *generating function* for the Legendre polynomials:

$$\frac{1}{\sqrt{1 - 2xs + s^2}} = \sum_{l=0}^{\infty} P_l(x) s^l \,. \tag{15.75}$$

Many important properties can be wrung out of the generating function. For example, set $x = 1$. Then

$$\sum_{l=0}^{\infty} P_l(1) s^l = \frac{1}{\sqrt{1 - 2s + s^2}} = \frac{1}{1 - s} \,. \tag{15.76}$$

But this last expression is the sum of a geometric series:

$$\frac{1}{1 - s} = \sum_{l=0}^{\infty} s^l \,. \tag{15.77}$$

Matching these two power series, we find that

$$P_l(1) = 1 \,. \tag{15.78}$$

Taking partial derivatives of the generating function with respect to x and to s yields many relations among the Legendre polynomials, and verifies that they satisfy

$$\frac{d}{dx}\left((1 - x^2)\frac{dP_l}{dx}\right) + l(l+1)P_l = 0 \,. \tag{15.79}$$

This is Eq. (15.74) with $m = 0$, so we really are talking about the same functions.

Eigenfunctions of self-adjoint operators that have distinct eigenvalues must be orthogonal, and functions are orthogonal if they vanish when integrated against one another. So, we can work up a family of polynomials by requiring that each new member be orthogonal to all the previous ones.

We posit that P_l is a polynomial of degree l in $\cos\theta$:

$$P_l(\cos\theta) = \sum_{k=0}^{l} a_k (\cos\theta)^k \,. \tag{15.80}$$

For $l \neq l'$, we must have

$$\int_{-1}^{1} P_l P_{l'} d(\cos\theta) = 0 \,. \tag{15.81}$$

The boundary condition that $P_l(1) = 1$ tells us right away that

$$P_0(\cos\theta) = 1 \,, \tag{15.82}$$

so we're off to a good start. Next, we impose the condition that P_1 be orthogonal to the flat function P_0:

$$\int_{-1}^{1} P_0 P_1 d(\cos\theta) = \int_{-1}^{1} a_1 \cos\theta d(\cos\theta) = 0 \,, \tag{15.83}$$

and so

$$P_1(\cos\theta) = \cos\theta\,. \tag{15.84}$$

The next highest order polynomial is

$$P_2(\cos\theta) = a(\cos\theta)^2 + b(\cos\theta) + c\,, \tag{15.85}$$

with $a + b + c = 1$. Integrating against P_0 gives

$$\int_{-1}^{1} P_2(\cos\theta)d(\cos\theta) = \frac{a}{3} + \frac{b}{2} + c - \left(-\frac{a}{3} + \frac{b}{2} - c\right) = 0\,, \tag{15.86}$$

from which we obtain $c = -a/3$. The analogous inner product with P_1 is

$$\int_{-1}^{1} P_1(\cos\theta)P_2(\cos\theta)d(\cos\theta) = \frac{2b}{3}\,, \tag{15.87}$$

which tells us that $b = 0$. So,

$$P_2(\cos\theta) = \frac{1}{2}(3\cos^2\theta - 1)\,. \tag{15.88}$$

We can likewise compute

$$P_3(\cos\theta) = \frac{1}{2}(6\cos^3\theta - 3\cos\theta)\,, \tag{15.89}$$

$$P_4(\cos\theta) = \frac{1}{8}(35\cos^4\theta - 30\cos^2\theta + 3)\,, \tag{15.90}$$

and so forth.

15.5 Separating the Hamiltonian

So far, we have manipulated our original Schrödinger equation for a two-particle system into the following form, where ψ is the wavefunction which encapsulates the relative degrees of freedom of the two particles; for the hydrogen atom, this is almost, but not exactly, the "wavefunction for the electron."

$$\left(-\frac{\hbar^2}{2\mu}\nabla^2 + V(r)\right)\psi = E\psi. \tag{15.91}$$

The spherical symmetry of our problem leads us to work in spherical coordinates. In that coordinate system, the Laplacian takes the following form [26]:

$$\nabla^2 = \frac{1}{r}\partial_r^2 r + \frac{1}{r^2}\left(\partial_\theta^2 + \frac{1}{\tan\theta}\partial_\theta + \frac{1}{\sin^2\theta}\partial_\phi^2\right). \tag{15.92}$$

Comparing this to the expression for $\vec{L}^2$ whose derivation we sketched above, we can rewrite our Hamiltonian in terms of the "electron's" angular momentum:

$$H = -\frac{\hbar^2}{2\mu}\frac{1}{r}\partial_r^2 r + \frac{1}{2\mu r^2}\vec{L}^2 + V(r). \tag{15.93}$$

Note that the only term which acts on the angular dependency of the wavefunction is the one containing $\vec{L}^2$. If we separate variables and write the wavefunction as the product of a radial function and a spherical harmonic,

$$\psi(\vec{x}) = R_l(r)Y_l^m(\theta,\phi), \tag{15.94}$$

then the total function is an eigenfunction of the Hamiltonian

$$H\psi = E\psi, \tag{15.95}$$

but it is also an eigenfunction of $\vec{L}^2$:

$$\vec{L}^2\psi = l(l+1)\hbar^2\psi. \tag{15.96}$$

For a given value of l, then, our Schrödinger equation can be written

$$\left[-\frac{\hbar^2}{2\mu}\frac{1}{r}\partial_r^2 r + \frac{l(l+1)\hbar^2}{2\mu r^2} + V(r)\right]R_l(r) = ER_l(r). \tag{15.97}$$

We can clean this up a little by scaling by r, thusly:

$$R_l(r) = \frac{1}{r}u_l(r). \tag{15.98}$$

With this transformation, the Schrödinger equation takes the form

$$\left[-\frac{\hbar^2}{2\mu}\partial_r^2 + \frac{l(l+1)\hbar^2}{2\mu r^2} + V(r)\right]u_l(r) = Eu_l(r). \tag{15.99}$$

This is just what we'd get if we wrote a Schrödinger equation for a single particle moving in one dimension with the potential

$$V_{\text{eff.}}(r) = V(r) + \frac{l(l+1)\hbar^2}{2\mu r^2}. \tag{15.100}$$

We call this the *effective potential.* It is just the interaction potential we wrote originally, plus a contribution which depends upon the angular momentum quantum number l, a contribution which is known as the *angular momentum barrier.* This is the quantum analogue of our classical intuition that a particle with large angular momentum spinning around inside a potential well is less likely to be found near the center of the well than is a particle with smaller angular momentum. In other words, the effective potential is where "centrifugal force" sneaks in.

All the work we've done so far applies equally well to any two-body system where the interaction potential depends upon the separation distance. We could, for example, use this machinery to attack the 3D harmonic oscillator (two massive particles connected by a massless spring). Instead of doing that, however, we're going to attack the hydrogenic atoms, where a single electron moves in the Coulomb field of the nucleus. For a nucleus containing Z protons, the interaction potential will be

$$V(r) = -\frac{Ze^2}{r}. \tag{15.101}$$

Therefore, when we have ascribed to the electron an eigenstate of angular momentum quantum number l, we predict that its behavior will obey the Schrödinger equation with the Hamiltonian

$$H_l = -\frac{\hbar^2}{2\mu}\partial_r^2 - \frac{Zq_e^2}{r} + \frac{l(l+1)\hbar^2}{2\mu r^2}. \tag{15.102}$$

Let us get a handle on this by seeing what happens when the angular-momentum term vanishes. Setting $l = 0$, and taking the atomic number $Z = 1$ for simplicity, gives

$$H_0 = -\frac{\hbar^2}{2\mu}\partial_r^2 - \frac{q_e^2}{r}\,. \tag{15.103}$$

An eigenfunction of this Hamiltonian will give us a solution with no angular dependence:

$$-\frac{\hbar^2}{2\mu}\frac{1}{r}\partial_r^2(r\psi) - \frac{q_e^2}{r}\psi = E\psi\,. \tag{15.104}$$

Differentiating $r\psi$ twice with respect to r, we find

$$\partial_r^2(r\psi) = 2\psi' + \psi''\,. \tag{15.105}$$

Substituting this in,

$$-\frac{\hbar^2}{2\mu}\frac{2}{r}\psi' - \frac{\hbar^2}{2\mu}\psi'' - \frac{q_e^2}{r}\psi = E\psi\,. \tag{15.106}$$

Somehow, a first derivative and a second derivative have to resemble the original function enough that they will all cancel out. We make the educated guess that ψ will look (up to normalization) like $e^{-r/a}$ for some constant a. Then each derivative just pulls down a prefactor of $-1/a$, and our equation becomes

$$\frac{\hbar^2}{2\mu}\frac{1}{ar}\psi - \frac{\hbar^2}{2\mu}\frac{1}{a^2}\psi - \frac{q_e^2}{r}\psi = E\psi\,. \tag{15.107}$$

This will all check out if we pick a to be the special value

$$a_0 = \frac{\hbar^2}{\mu q_e^2}\,, \tag{15.108}$$

which is the Bohr radius. The energy of this solution is

$$E = -\frac{\hbar^2}{2\mu}\frac{1}{a_0^2} = -\frac{\mu q_e^4}{2\hbar^2}\,, \tag{15.109}$$

an amount known as one Rydberg, which works out to be about -13.6 electron volts.

15.6 Addition of Angular Momentum

Physics textbooks seem nearly universally committed to using notation that elides tensor products, factors of the identity and so forth. Part of the reason may be that they can be aimed at students who have not yet learned linear algebra as far as tensor products, so the authors of the books bluff their way through by imitating the previous generation of authors who bluffed their way through. This helped to make quantum angular momentum tough going for me, because it obscured which manipulations of the symbols were legitimate and which were not. Symbols can be dropped from notation when they can be recovered from context, but not before.

Like just about everybody I've met in my profession, I had a college course in quantum mechanics that used David J. Griffiths' textbook [127]. When he covers the addition of angular momenta, he gets to the following bit:

> If you think this is starting to sound like mystical numerology, I don't blame you. We will not be using the Clebsch–Gordan tables much in the rest of the book, but I wanted you to know where they fit into the scheme of things, in case you encounter them later on. In a mathematical sense this is all applied **group theory**—what we are talking about is the decomposition of the direct product of two irreducible representations of the rotation group into a direct sum of irreducible representations (you can quote that, to impress your friends).

This has always felt cheap to me. It provides exactly the wrong level of information: enough to sound intimidating, but not enough to be useful. If anything, it makes the subject sound harder and more mysterious than it is. I do not know how many curricula have time for group theory (I have only tried including it in a quantum course at the graduate level), but part of the job of a textbook is to include what the lectures can't. In a modern course with a stronger emphasis on quantum information theory, matrices and qubits and entanglement will already be discussed. The springboard is provided, as it were.

The quoted passage is also, if read at face value, mathematically wrong. I.e., if we take "rotation group" to mean $SO(3)$, the group of rotations in good old 3D space, then we slam into the problem that not all irreps of $SU(2)$ can be irreps of $SO(3)$. The Lie *algebras* of these groups, $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$, are isomorphic, but while the representations of $\mathfrak{su}(2)$ and of $SU(2)$ are in one-to-one correspondence, only the *odd-dimensional* representations of the algebra can come from irreps of $SO(3)$. If we work with "irreducible representations of the rotation group" as the book says, we miss even the base case of two qubits.

We already did the hard work back in §13.4, when we studied the *symmetric subspace* of a collection of qubit state spaces. We wrote $S(n)$ to stand for the vector space we get by taking n qubits, constructing the $n+1$ orthonormal symmetrized vectors (going from "all up" to "all down"), and using these vectors as a basis. We noted that

$$S(n) \simeq \mathbb{C}^{n+1}\,, \tag{15.110}$$

and we deduced that

$$S(n_1) \otimes S(n_2) \simeq S(n_1 + n_2) \oplus S(n_1 + n_2 - 2) \oplus \cdots \oplus S(|n_1 - n_2|)\,. \tag{15.111}$$

There is a natural way for $SU(2)$ to act upon the tensor product of multiple qubits. We just apply the same unitary to each qubit:

$$D(U) = U \otimes U \otimes \cdots \otimes U\,. \tag{15.112}$$

This obviously doesn't prefer one qubit over any of the others or pick out a particular qubit as special; it treats them all symmetrically. To understand what this means for the Lie algebra, we need to look at tiny steps away from the identity. We imagine that $\{U_t\}$ is a whole family of unitaries parameterized by the index t, and we look at what happens near $t = 0$. When we take the derivative of $D(U_t)$ with respect to t, the product rule means that we get a sum of terms, with one term for each qubit. In each term, one factor is the derivative of U_t, and the other factors are all U_t, with everything evaluated at $t = 0$:

$$\begin{aligned}\frac{d}{dt} D(U_t)\bigg|_{t=0} &= U_0' \otimes I \otimes \cdots \otimes I \\ &\quad + I \otimes U_0' \otimes I \cdots \otimes I + \cdots \\ &\quad + I \otimes \cdots \otimes I \otimes U_0'\,.\end{aligned} \tag{15.113}$$

This tells us how to make generators for the many-qubit system out of generators for a single qubit. We can define a multi-qubit version of J_x, of J_y, of everything we'd like. We did this for two qubits in Eq. (14.87), you may recall. The multi-qubit version of the raising operator J_+ is a sum of terms, and each term "hits" one qubit while leaving the others untouched. By starting with the symmetric tensor product state $|0 \cdots 0\rangle$, this produces symmetrized states with more and more 1's.

The orthonormal basis $S(n)$ is therefore a natural home for a representation of $\mathfrak{su}(2)$. And from what we already know, the *tensor product* of two such representations will split into a *direct sum* of representations.

Let's go through an example: what physicists call adding the angular momenta of two spin-1 particles. "Spin-j" means "dimension $2j + 1$". So, $j \in \{0, 1, 2\}$, and for each value of j, m will go from $-j$ to $+j$. In terms of representations,

$$1 \otimes 1 = 0 \oplus 1 \oplus 2. \tag{15.114}$$

The result will have one state which is a scalar under rotations, three states which transform together as a vector (spin-1) and five states which transform together as a tensor (spin-2).

Here, a ket written with two comma-separated numbers, $|j, m\rangle$, stands for an eigenstate of the total angular momentum operators, and a ket written with 0, + or − stands for an eigenstate of single-particle operators. The only way to get $m = 2$ is to have both particles with spins parallel and up:

$$|2, 2\rangle = |{+}{+}\rangle. \tag{15.115}$$

Likewise, the only way to reach the other extreme, $m = -2$, is to have both spins down:

$$|2,-2\rangle = |--\rangle. \tag{15.116}$$

The next lower state in the $j = 2$ representation, $|2,1\rangle$, must be a linear combination of $|+0\rangle$ and $|0+\rangle$. Generally,

$$J_-|j,m\rangle = \sqrt{(j+m)(j-m+1)}|j,m-1\rangle, \tag{15.117}$$

so in this case,

$$J_-|2,2\rangle = \sqrt{4\cdot 1}|2,1\rangle = 2|2,1\rangle. \tag{15.118}$$

The lowering operator for the total system is the sum of the lowering operators for the parts, properly tensored:

$$J_- = J_{1-}\otimes I + I\otimes J_{2-}. \tag{15.119}$$

The first term acts on particle 1, yielding $\sqrt{2}|0+\rangle$, and the second term acts on particle 2, yielding $\sqrt{2}|+0\rangle$. So,

$$|2,1\rangle = \frac{1}{\sqrt{2}}|+0\rangle + \frac{1}{\sqrt{2}}|0+\rangle. \tag{15.120}$$

Coming up from the bottom state, we know in general that

$$J_+|j,m\rangle = \sqrt{(j-m)(j+m+1)}|j,m+1\rangle, \tag{15.121}$$

so in this case,

$$J_+|2,-2\rangle = 2|2,-1\rangle. \tag{15.122}$$

Decomposing J_+ as we did J_-, we get in the same manner as before

$$|2,-1\rangle = \frac{1}{\sqrt{2}}|-0\rangle + \frac{1}{\sqrt{2}}|0-\rangle. \tag{15.123}$$

We've now trapped $|2,0\rangle$ in the middle.

$$J_-|2,1\rangle = \sqrt{(2+1)(2-1+1)}|2,0\rangle = \sqrt{6}|2,0\rangle. \tag{15.124}$$

Now,

$$J_-|2,1\rangle = J_-\left(\frac{1}{\sqrt{2}}|+0\rangle + \frac{1}{\sqrt{2}}|0+\rangle\right) \tag{15.125}$$

$$= \frac{1}{\sqrt{2}}\left(\sqrt{2}|00\rangle + \sqrt{2}|+-\rangle + \sqrt{2}|00\rangle + \sqrt{2}|-+\rangle\right). \tag{15.126}$$

Dividing by $\sqrt{6}$,

$$|2,0\rangle = \frac{1}{\sqrt{6}}|+-\rangle + \sqrt{\frac{2}{3}}|00\rangle + \frac{1}{\sqrt{6}}|-+\rangle. \tag{15.127}$$

That wraps up $j = 2$.

The other state at $m = 1$, $|1,1\rangle$, must be a linear combination of $|+0\rangle$ and $|0+\rangle$ which is orthogonal to $|2,1\rangle$. So,

$$|1,1\rangle = \frac{1}{\sqrt{2}}|+0\rangle - \frac{1}{\sqrt{2}}|0+\rangle. \tag{15.128}$$

Acting with the lowering operator,

$$J_-|1,1\rangle = \sqrt{(1+1)(1-1+1)}|1,0\rangle = \sqrt{2}|1,0\rangle \tag{15.129}$$

$$= J_- \left(\frac{1}{\sqrt{2}}|+0\rangle - \frac{1}{\sqrt{2}}|0+\rangle\right) \tag{15.130}$$

$$= |00\rangle + |+-\rangle - |-+\rangle - |00\rangle. \tag{15.131}$$

Consequently,

$$|1,0\rangle = \frac{1}{\sqrt{2}}|+-\rangle - \frac{1}{\sqrt{2}}|-+\rangle. \tag{15.132}$$

Acting again with the lowering operator,

$$J_-|1,0\rangle = \sqrt{(1+0)(1-0+1)}|1,-1\rangle = \sqrt{2}|1,-1\rangle \tag{15.133}$$

$$= J_- \left(\frac{1}{\sqrt{2}}|+-\rangle - \frac{1}{\sqrt{2}}|-+\rangle\right) \tag{15.134}$$

$$= |0-\rangle - |-0\rangle. \tag{15.135}$$

Therefore,

$$|1,-1\rangle = \frac{1}{\sqrt{2}}|0-\rangle - \frac{1}{\sqrt{2}}|-0\rangle. \tag{15.136}$$

This takes care of $j = 1$.

The only state left is $|0,0\rangle$. It must be a linear combination of $|+-\rangle$, $|00\rangle$ and $|-+\rangle$ which is orthogonal to $|2,0\rangle$ and $|1,0\rangle$. Let's say that

$$|0,0\rangle = a|+-\rangle + b|00\rangle + c|-+\rangle. \tag{15.137}$$

Then

$$\langle 1,0|0,0\rangle = \frac{a}{\sqrt{2}} - \frac{c}{\sqrt{2}}, \tag{15.138}$$

$$\langle 2,0|0,0\rangle = \frac{a}{\sqrt{6}} + b\sqrt{\frac{2}{3}} + \frac{c}{\sqrt{6}}. \tag{15.139}$$

Orthogonality requires $a = c$ and

$$\frac{2a}{\sqrt{6}} = -b\sqrt{\frac{2}{3}} \quad \Rightarrow \quad \frac{a}{\sqrt{6}} = -b\frac{\sqrt{2}}{2}\frac{1}{\sqrt{3}} \quad \Rightarrow \quad a = -b. \tag{15.140}$$

With the Born-rule normalization constraint that $a^2 + b^2 + c^2 = 1$, we get at last that

$$|0,0\rangle = \frac{1}{\sqrt{3}}|+-\rangle - \frac{1}{\sqrt{3}}|00\rangle + \frac{1}{\sqrt{3}}|-+\rangle. \tag{15.141}$$

16 The Nonrelativistic Hydrogen Atom

It seems only fair that, having paid a lot of money for an MIT education, I turn around and give away that knowledge for free. This chapter will be an example, using a topic that I haven't seen treated in much depth by resources that are already available. Therefore, this chapter will explain the solution of the hydrogen atom and similar systems by the factorization method, also known as the method of supersymmetric quantum mechanics.

As an undergraduate, I learned this topic twice, in ways that rather neatly bracket the quality of my college education. The first was a plodding march through differential equations, in a course that was little more than that. I remember nothing of it save that it happened. A semester later, we did hydrogen all over again, in a course that was willing to trust us with a little more mathematics, and the solution stuck with me, leaving an image in my mind that I happily revisited in the later years, correlating it with new things I learned along the way. The lesson is that a solution which is more "elementary" is not always easier in a meaningful sense.

The problem we will address is, give or take, the most complicated item in the standard curriculum to receive an "exact solution". When we say we have an exact solution, we do not mean that any single physical system in the natural world corresponds to our mental idealization in every detail. As the old saying has it, all models are wrong, but some are useful. Rather, *exact* here means that, once we have posed the problem, we can go about solving it without making further approximations.

Physicists often call this problem "solving the hydrogen atom," but this is somewhat misleading. What matters is that we have a positive charge and a negative charge, bound together so that it takes effort to pull them apart. If the positive charge is a proton and the negative charge an electron, then we have the case of hydrogen, but the same solution also applies to many other scenarios: a positron and an electron, a proton and a muon, an electron orbiting a heavier nucleus and so on. Therefore, this chapter will refer to "hydrogenic atoms", that is, physical systems which have a certain plus-and-minus-charge character in common, although they differ in other ways.

The excited states of hydrogenic atoms are our prototype for understanding how the periodic table works, and it's often the first place one runs into the mathematics of angular momentum. Unfortunately, too many standard treatments of introductory QM say that hydrogen has "accidental degeneracies": these states have the same energies as those states for no spectacularly interesting reason. But we are trained to associate degeneracies with *symmetries*—when two sets of eigenstates have the same eigenvalues, we expect some symmetry to be at work. So, is there a symmetry in hydrogenic atoms above and beyond the familiar rotational kind, a symmetry which They haven't been telling us about?

In this chapter, we will explore that question against the background of undergrad-

uate quantum mechanics. First, we will build up some very general machinery [128] for solving problems, and then we will apply those techniques to hydrogenic atoms; by that point, we should have a fair amount of knowledge with which we can move in any one of several interesting directions. The technique we'll employ has a certain charm, because we solve the first part, the angular dependence, using commutator relations, while as we shall see, the *radial* dependence can be solved with *anticommutator* relations. After we see how things work in nonrelativistic quantum mechanics, we'll investigate what changes when we bring special relativity into the story.[1]

I first encountered SUSY QM after a few months of doing quantum mechanics with operators. This chapter targets a hypothetical audience who is roughly as informed now as I was then. For example, I take the quantum harmonic oscillator as something which the reader has already grappled with and seen solved with operator methods (as and $a^\dagger$s). In the next chapter, when we move into relativistic physics, the presentation presumes some prior encounter with special relativity at an introductory level, but we'll develop "four-vector" notation and the Dirac equation ourselves.

To begin, let's familiarize ourselves with the behavior of a *superalgebra.*

16.1 Superalgebra

In fundamental quantum mechanics, we learn that an algebra of operators is defined by commutation relations among those operators. For example, the canonical operators of position and momentum have the commutator $[x, p] = i\hbar$. A more intricate case is the algebra of angular momentum operators, which we encountered when exploring the rotational symmetries of 3D space. To generalize this concept, we define an *anticommutator,* which relates operators in the same way as an ordinary commutator, but with the opposite sign:

$$\{A, B\} \equiv AB + BA. \tag{16.1}$$

If operators are related by anticommutators as well as commutators, we say that they are part of a *superalgebra.* Let's say we have a quantum system described by a Hamiltonian $\mathcal{H}$ and a set of N self-adjoint operators Q_i, each of which commutes with the Hamiltonian. We shall call this system *supersymmetric* if the following anticommutator is valid for all $i, j = 1, 2, \ldots, N$:

$$\{Q_i, Q_j\} = \mathcal{H}\delta_{ij}. \tag{16.2}$$

If this is the case, then we call the operators Q_i the system's *supercharges.* $\mathcal{H}$ will be termed the SUSY Hamiltonian, SUSY being a convenient abbreviation for whichever variation of "supersymmetry" is grammatically appropriate. A SUSY algebra is characterized by its number of supercharges, which we typically denote N.

Because the $N = 2$ case exemplifies many properties of general SUSY theories, it is worthwhile to work it out in some detail. We require two supercharges, $Q_1 = Q_1^\dagger$ and

[1] I wrote much of this material as blog posts, which are still available: http://www.sunclipse.org/?p=323.

$Q_2 = Q_2^\dagger$. The SUSY algebra we defined a moment ago implies the following relations:

$$Q_1 Q_2 = -Q_2 Q_1,\ \mathcal{H} = 2Q_1^2 = 2Q_2^2 = Q_1^2 + Q_2^2. \tag{16.3}$$

It is sometimes more convenient to work with a "complex" supercharge that is not self-adjoint. If we make linear combinations of our supercharges,

$$Q = \frac{1}{\sqrt{2}}(Q_1 + iQ_2),\ Q^\dagger = \frac{1}{\sqrt{2}}(Q_1 - iQ_2), \tag{16.4}$$

then the SUSY algebra implies $\{Q, Q^\dagger\} = \mathcal{H}$. To make this a little more concrete, we can realize a specific incarnation of the superalgebra: let H_1 be some Hamiltonian of interest, and suppose that we can factor H_1 into the product of an operator and its adjoint:

$$H_1 = A^\dagger A. \tag{16.5}$$

Note that this is almost the form of the harmonic oscillator Hamiltonian, except for an energy shift:

$$H_{\text{SHO}} = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right). \tag{16.6}$$

So, this is not an unfamiliar form for a Hamiltonian. Swapping the order of the factors gives another operator, which you can verify is also Hermitian:

$$H_2 = AA^\dagger. \tag{16.7}$$

With A in hand, define the two operators

$$Q = \begin{pmatrix} 0 & 0 \\ A & 0 \end{pmatrix} \tag{16.8}$$

and

$$Q^\dagger = \begin{pmatrix} 0 & A^\dagger \\ 0 & 0 \end{pmatrix}. \tag{16.9}$$

Matrix arithmetic verifies that

$$\{Q, Q^\dagger\} = \begin{pmatrix} A^\dagger A & 0 \\ 0 & AA^\dagger \end{pmatrix}, \tag{16.10}$$

so we can say that the anticommutator of our two charges gives a Hamiltonian $\mathcal{H}$ which is block diagonal,

$$\mathcal{H} = \begin{pmatrix} H_1 & 0 \\ 0 & H_2 \end{pmatrix}. \tag{16.11}$$

H_1 and H_2 can be considered two Hamiltonians acting on subspaces of the original Hilbert space associated with $\mathcal{H}$.

16.2 Partner Hamiltonians

What exactly is so special about operators of the forms $A^\dagger A$ and $AA^\dagger$? Given a Hamiltonian for some system, H_1, if it can be factored into the product of two operators $A^\dagger A$, then we can construct another Hamiltonian $H_2 = AA^\dagger$ which has almost exactly the same energy eigenvalue spectrum. These "isospectral" Hamiltonians may not describe the same physics, and their respective potentials $V_1(x)$ and $V_2(x)$ may look radically different. As usual, a degeneracy in the energy levels corresponds to a symmetry; in this case, the symmetry is the SUSY between our two Hamiltonians.

First, let's take a look at the eigenstates of Hamiltonian number 1. These states satisfy the relationship

$$H_1\left|\psi_n^{(1)}\right\rangle = A^\dagger A\left|\psi_n^{(1)}\right\rangle = E_n^{(1)}\left|\psi_n^{(1)}\right\rangle. \tag{16.12}$$

Now, a surprising thing happens: the operator A maps the eigenstates of Hamiltonian 1 into eigenstates of Hamiltonian 2. Look:

$$H_2 A\left|\psi_n^{(1)}\right\rangle = AA^\dagger A\left|\psi_n^{(1)}\right\rangle, \tag{16.13}$$

but by the equation just above, this means that

$$H_2 A\left|\psi_n^{(1)}\right\rangle = E_n^{(1)} A\left|\psi_n^{(1)}\right\rangle. \tag{16.14}$$

The same logic works in the opposite direction, connecting eigenstates of H_2 with those of H_1. The eigenstates behind door number 2 satisfy

$$H_2\left|\psi_n^{(2)}\right\rangle = AA^\dagger\left|\psi_n^{(2)}\right\rangle = E_n^{(2)}\left|\psi_n^{(2)}\right\rangle, \tag{16.15}$$

so by acting with the operator $A^\dagger$,

$$H_1 A^\dagger\left|\psi_n^{(2)}\right\rangle = A^\dagger AA^\dagger\left|\psi_n^{(2)}\right\rangle = E_n^{(2)} A^\dagger\left|\psi_n^{(2)}\right\rangle. \tag{16.16}$$

We have shown that H_1 and H_2 are *isospectral*. For every eigenstate of one, there lurks an eigenstate of the other *with the same energy.* There is one exception, and it's important: if

$$A\left|\psi_n^{(1)}\right\rangle = 0, \tag{16.17}$$

that is, if H_1 has a zero-energy ground state, the proof does not work, and there is no need for H_2 to have a zero-energy ground state. In fact, as we'll see momentarily, *only* one of H_1 and H_2 may have a zero-energy ground state; for consistency, we usually arrange matters so that H_1 has the extra eigenstate.

16.3 An Aside on Singular Values

There's something worth noting here which I could have discovered for myself when I first saw SUSY QM. I missed it, because at that point in my life, I'd never had

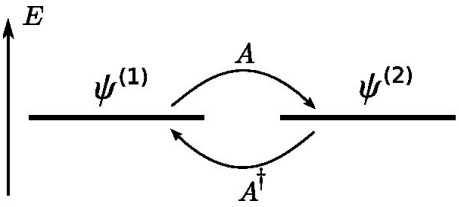


Figure 16.1: Eigenstates of H_1 and H_2 mapped into each other by A and its adjoint $A^\dagger$.

to do serious data analysis on big heaps of experimental results. The isospectrality property we just deduced is another incarnation of an idea which is important in the data analysis context, thanks to the technique of *singular value decomposition,* or SVD for short [26].

SVD works like this: suppose we have a matrix M, not necessarily square. Then we can form the two matrices

$$X = M^\dagger M \tag{16.18}$$

and

$$Y = MM^\dagger. \tag{16.19}$$

If M is an $m \times n$ matrix, then X will be an $n \times n$ one, and Y will be an $m \times m$ one. That is, even if the original matrix is not square, both matrices built from it will be. By the same argument we used before, X and Y must have the same eigenvalues, except possibly for one eigenvalue which equals zero. The square roots of the eigenvalues of X are the *singular values* of the original matrix M. Let $\{|x_i\rangle\}$ be the eigenvectors of X, and $\{|y_i\rangle\}$ the eigenvectors of Y. Then we can work out that

$$M = \sum_{i=1}^{\min(m,n)} s_i |y_i\rangle\langle x_i|, \tag{16.20}$$

where the $\{s_i\}$ are the singular values of M. In this language, we can say that the energy levels of a Hamiltonian which factors into A and $A^\dagger$ are the squares of the singular values of the operator A.

SVD is helpful for understanding when systems of linear equations have solutions, and for finding the best possible approximation of a solution for systems which strictly speaking have none [26]. Furthermore, SVD applied to matrices derived from experimental data is an important tool for finding patterns in large datasets and separating signal from noise.

This isospectrality phenomenon also appears in the study of generalised quantum measurement procedures, known as POVMs, and how one can update statistical information after obtaining new experimental data. In this context, quantum probability is just like classical probability, but with an extra SUSY transformation [101]. Isospectrality even pops up when one tries applying linear algebra to networks, that is, sets of nodes connected by links [129]!

16.4 Superpotentials

Most of the time, we find ourselves dealing with Hamiltonians of the form

$$H = \frac{p^2}{2m} + V(x), \tag{16.21}$$

which, knowing that $p = -i\hbar\partial_x$, we can also write as

$$H = -\frac{\hbar^2}{2m}\partial_x^2 + V(x). \tag{16.22}$$

If we want to factor this H into an operator and its adjoint, we should probably start with an operator which is linear in the derivative of x, thus:

$$A = \frac{\hbar}{\sqrt{2m}}\partial_x + W(x). \tag{16.23}$$

Here, $W(x)$ is some real function of x which we shall call the *superpotential.* Taking the adjoint of A flips the sign on the derivative; you should deduce why by observing that the momentum p is observable and therefore self-adjoint.

$$A^\dagger = -\frac{\hbar}{\sqrt{2m}}\partial_x + W(x). \tag{16.24}$$

Let's say that we have two partner Hamiltonians $H_1 = A^\dagger A$ and $H_2 = AA^\dagger$. In terms of the position variable x, we can write these partner Hamiltonians as

$$\begin{aligned} H_1 &= -\frac{\hbar^2}{2m}\partial_x^2 + V_1(x), \text{ and} \\ H_2 &= -\frac{\hbar^2}{2m}\partial_x^2 + V_2(x). \end{aligned} \tag{16.25}$$

The functions $V_1(x)$ and $V_2(x)$ are *partner potentials.* We can connect the superpotential to H_1's potential function,

$$V_1(x) = W^2(x) - \frac{\hbar}{\sqrt{2m}}\partial_x W(x). \tag{16.26}$$

This relationship is known as the Riccati equation. If we reverse the order of our operators, it turns out that H_2 is a Hamiltonian with a new potential $V_2(x)$, given by

$$V_2(x) = W^2(x) + \frac{\hbar}{\sqrt{2m}}\partial_x W(x). \tag{16.27}$$

We recognize this as the Riccati equation with a change of sign. $V_1(x)$ and $V_2(x)$ are known as *partner potentials,* related through the superpotential $W(x)$.

With one more step, we can relate the superpotential to the ground state wavefunction. Note that the ground state of H_1 is annihilated by A, satisfying the relation

$$A\left|\psi_0^{(1)}\right\rangle = 0. \tag{16.28}$$

Looking back at the form of A, we see that this is a *first-order* differential equation, and we can write its solution as an exponential.

$$\psi_0^{(1)}(x) \propto \exp\left[-\frac{\sqrt{2m}}{\hbar}\int_0^x W(y)dy\right]. \tag{16.29}$$

Note that the zero-energy ground state of H_2 would be annihilated by $A^\dagger$ and would therefore be proportional to

$$\exp\left[\frac{\sqrt{2m}}{\hbar}\int_0^x W(y)dy\right]. \tag{16.30}$$

Only one of these two expressions can give a *normalizable* state: if one behaves nicely, the other will blow up. That's why only one of the two partner Hamiltonians can have a ground state of zero energy.

Often, these equations are shown in "natural units" where $\hbar = 2m = 1$. This can always be done by changing the units of x, and it makes successive steps in the calculations much cleaner. I think it nice to see the equations with all the original units in place at least once; in following sections, however, when units are not illuminating I will set unnecessary constants to unity.

In summary, then: H_1 and H_2 are *isospectral.* That is, while the forms of their eigenfunctions may be different, the eigenvalues associated with those functions are the same; or, in physical terms, the wavefunctions have different shapes, but the energies match. The only exception is the ground state: if H_1 has a zero-energy ground state, then H_2 will not. Furthermore, the operator A maps eigenstates of H_1 into those of H_2, and the operator $A^\dagger$ maps eigenstates in the reverse direction.

We can sketch the full set of H_1 eigenstates as a ladder with, potentially, an infinite number of rungs. The eigenstates of H_2 sit right alongside them; the bottom "rung" is missing, as we mentioned earlier.

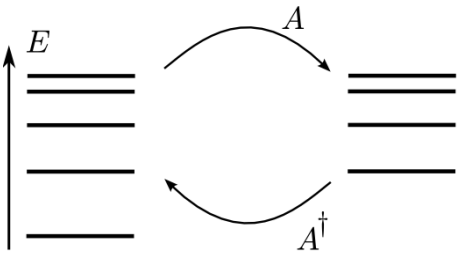


Figure 16.2: Eigenstates of isospectral Hamiltonians.

That "missing rung" has an important implication: instead of mapping the ground state of H_1 into a state of H_2, the operator A annihilates it. We could, therefore, solve for the wavefunction $\psi_0^{(1)}(x)$ by solving a first-order, linear differential equation—a simpler task than tackling the full, second-order Schrödinger equation.

Now, we'll make a qualitative statement which we'll firm up in a moment. If the Hamiltonian H_2 were "roughly similar" to H_1, then we should be able to "pull the same trick again," writing it as

$$H_2 = A_2^\dagger A_2, \tag{16.31}$$

and inventing a third Hamiltonian,

$$H_3 = A_2 A_2^\dagger, \tag{16.32}$$

which will be isospectral with H_2 except for a ground state satisfying

$$A_2 \left| \psi_0^{(2)} \right\rangle = 0. \tag{16.33}$$

Then, why not, we do the same thing a whole bunch of times over again, building a whole hierarchy of Hamiltonians stretching off beyond the margin of our page. We'll make this idea precise in the next section.

For now, note that the "bottom rung" in each "ladder" satisfies a first-order differential equation, and that we can build any state of H_1 by stacking up a sequence of $A^\dagger$ operators. Start with a state annihilated by some A_n and then apply

$$A_{n-1}^\dagger A_{n-2}^\dagger \cdots A_1^\dagger \tag{16.34}$$

to get an eigenstate of H_1 there on the left edge. This is reminiscent of the way we solve the harmonic oscillator using creation and annihilation operators, with the difference that instead of building *up,* we're working in from the *side.*

16.5 Shape Invariance

To see this in action, we need to solidify the notion of rough resemblance we invoked above. Suppose we have a potential V_1 associated with some system of interest. If we compute the superpotential $W(x)$ and find that the partner potential V_2 has a similar shape to V_1, we can calculate the energy states of V_1 in a remarkably straightforward way. In this context, "similar shape" means that if V_1 depends upon a set of parameters, we can reach V_2 by substituting different values of those parameters into V_1. More mathematically, two partner potentials V_1 and V_2 are said to be *shape invariant* if they satisfy

$$V_2(x; a_1) = V_1(x; a_2) + R(a_1) \tag{16.35}$$

where a_1 and a_2 are sets of parameters, and a_2 is uniquely determined by a_1 (that is, $a_2 = f(a_1)$ for some function f). The residual term $R(a_1)$ does not depend upon the coordinate x. We shall see that if H_1 has a shape-invariant partner potential in H_2, the entire spectrum of H_1 can be computed through operator manipulations [130].

In the material which follows, we'll work in "natural units," to save ourselves from carrying around excess notational baggage. Our operators will therefore take the form

$$A = \partial_x + W(x) \tag{16.36}$$

and

$$A^\dagger = -\partial_x + W(x). \tag{16.37}$$

The shape invariance criterion, Eq. (16.35), is the grown-up version of the harmonic oscillator commutation relation for creation and annihilation operators, $[a, a^\dagger] = 1$. To see why, let's look at what happens when we take the superpotential to take the particularly simple form $W(x) = x$. Then our operators are $A = \partial_x + x$ and its conjugate $A^\dagger = -\partial_x + x$, which up to normalization factors of $\sqrt{2}$ are just the familiar old annihilation and creation operators, written in dimensionless form. Moreover, our partner Hamiltonians are

$$H_1 = A^\dagger A = -\partial_x^2 + x^2 - 1, \tag{16.38}$$

and

$$H_2 = AA^\dagger = -\partial_x^2 + x^2 + 1. \tag{16.39}$$

Comparing the two,

$$AA^\dagger = A^\dagger A + 2, \tag{16.40}$$

we see that shape invariance is satisfied without even a transformation of parameters. In addition, scaling the operators each by $\sqrt{2}$ recovers the familiar commutator $[a, a^\dagger] = 1$. We can rewrite the general shape invariance criterion as a commutator itself, if we invent a new unitary operator U which implements the necessary parameter transformation $a_2 = f(a_1)$. In operatorial language, the criterion Eq. (16.35) becomes

$$AA^\dagger = U^\dagger \left(A^\dagger A\right) U + R. \tag{16.41}$$

Noting that $UU^\dagger = I$ by unitarity, we can insert $UU^\dagger$ on the left-hand side:

$$AUU^\dagger A^\dagger = U^\dagger \left(A^\dagger A\right) U + R. \tag{16.42}$$

Or, in commutator form,

$$[AU, (AU)^\dagger] = R. \tag{16.43}$$

It is interesting to ponder how much of the elaborate mathematical machinery built up from the familiar relation $[a, a^\dagger] = 1$ (see *e.g.,* [131, 132, 133]) extends to the more general world of shape invariance.

To a student with some experience in statistical physics, shape invariance bears a certain resemblance to being at the fixed point of a *renormalization group* transformation. In RG theory [134, 43, 135, 46, 136], we apply transformations to a system composed of many parts: for example, we can average over blocks of spins in an Ising model to produce a "coarse-grained" version of that model. Such a transformation clearly affects the state space of the system. We then effect some change of system parameters, such as rescaling our length unit so that the coarse-grained Ising spins have the same apparent size as the original ones. Together, these operations create a *flow* on the space of possible Hamiltonians. Fixed points of this flow are situations where interesting things happen, about which we can deduce a great deal in remarkably slick ways. An analogous fact holds true in SUSY QM: When we live at the "fixed point" defined by the shape invariance criterion, we can answer a great many questions using unexpectedly powerful operator methods.

16.6 Hierarchy of Hamiltonians

We are usually interested in examining the physics of a particular system, for which we have written an H_1. (We are restricted to work in one dimension; however, many higher-dimensional problems can be attacked through separation of variables. When studying the hydrogen atom, the radial dependence may be factored out of the Coulomb Hamiltonian to give a pseudo-one-dimensional potential; when studying the Landau effect, a convenient choice of gauge shows that along one axis, electrons behave as if in a harmonic oscillator potential.) As mentioned above, we can slide the energy scale up and down, so for convenience we fix the $n = 0$ eigenstate of H_1 to have zero energy:

$$E_0^{(1)}(a_1) = 0. \tag{16.44}$$

We can solve for the ground state knowing the superpotential to show that

$$\psi_0^{(1)}(x; a_1) = C \exp\left(-\int^x W(y; a_1) dy\right) \tag{16.45}$$

with C being a constant of normalization. (Having the freedom to choose this constant makes the lower limit of the integral arbitrary.)

To investigate the spectrum of H_1, we construct the series of Hamiltonia $\{H_s\}$, where s ranges over the positive integers:

$$H_s = -\partial_x^2 + V_1(x; a_s) + \sum_{k=1}^{s-1} R(a_k), \tag{16.46}$$

in which we have defined

$$a_s = f^{s-1}(a_1). \tag{16.47}$$

Here, $f^{s-1}(a_1)$ denotes applying the transformation f to the parameter set a_1 a total of $s-1$ times. Each time, we get a new, uniquely determined set of parameters. With this definition, it is easy to see that H_s and H_{s+1} are SUSY partners:

$$H_{s+1} = -\partial_x^2 + V_1(x; a_{s+1}) + \sum_{k=1}^{s} R(a_k), \tag{16.48}$$

which is also equal to

$$-\partial_x^2 + V_2(x; a_s) + \sum_{k=1}^{s-1} R(a_k). \tag{16.49}$$

Because H_s and H_{s+1} are SUSY partners, they share the same eigenenergies, save for the ground state, whose energy is just given by the sum of the residuals,

$$E_0^{(s)} = \sum_{k=1}^{s-1} R(a_k). \tag{16.50}$$

The SUSY between H_1 and H_2 means that they are isospectral; the $n = 1, 2, 3, \ldots$ eigenstates of H_1 have the same energies as the $n = 0, 1, 2 \ldots$ states of H_2. The same holds true for H_2 and H_3, H_3 and H_4, H_{17} with H_{18} and so on until we run out of bound states. (For a problem like the hydrogen atom, we never do: n ranges upward to infinity.) We "lose" one state, the lowest-energy ground state, every time we go from H_s to H_{s+1}. Following the degeneracies through and liberally applying the transitive law, we see that the n-th eigenstate of H_1 is degenerate with the ground state of H_n. Therefore, *we can recover the entire spectrum of H_1 from the equation above.*

We can get the eigenfunctions as well as the eigenvalues by this approach. Up to a normalization constant,

$$\psi_n^{(s+1)} \propto A_s \psi_{n+1}^{(s)}, \tag{16.51}$$

so knowing the eigenfunctions of H_1 would allow us to compute those of H_2. The proportionality works in reverse, too:

$$\psi_n^{(s)} \propto A_{s+1}^\dagger \psi_{n-1}^{(s+1)}. \tag{16.52}$$

Therefore, if we knew the ground-state wavefunction of H_n, we could apply a sequence of operators $A_n^\dagger, A_{n-1}^\dagger, \ldots$ until we reached a state in the spectrum of H_1. But, we *do* know the ground state of H_n: it is that state which is annihilated by the operator A_n. Remember that A_n involves a single first-order derivative ∂_x, so solving $A_n \psi = 0$ only involves solving a first-order differential equation. In this way, for any H_1 which has a shape-invariant potential (SIP), we can *exactly* solve for its eigenvalues and eigenfunctions.

As we noted earlier, this is reminiscent of the operator-based solution of the harmonic oscillator. We had before that we recover the harmonic oscillator, up to overall scaling factors, when $W(x) = x$. In this special case, we can see immediately that the ground state wavefunction is a Gaussian curve:

$$\psi_0^{(1)} \propto \exp\left(-\int^x dy\, y\right) = \exp\left(-\frac{x^2}{2}\right). \tag{16.53}$$

Because the shape invariance relationship in this special case involves only a shift along the energy axis (and not a parameter transformation in addition), we can build up all the eigenstates by repeatedly applying the same operator $A^\dagger$. In addition, because there is no $a_2 = f(a_1)$ complication to worry about, the residual term is the same every time we move from one Hamiltonian in the hierarchy to the next, so the energy shift is the same each time. The spacings between energy eigenvalues will therefore all be the same—a familiar fact for the harmonic oscillator, seen in a new context.

If the potential we wish to study satisfies a certain criterion, which we called "shape invariance," we can construct a hierarchy of Hamiltonians, each missing the lowest-energy eigenstate of the last, and find the complete spectrum of the original Hamiltonian by "working leftward" in the state diagram. We shall see that with the hydrogen atom, each state in the diagram corresponds to a physical eigenstate of the system, but in order to get there, we have to turn the three-dimensional Coulomb potential of the hydrogen atom into the kind of problem we can study with the SUSY QM machinery we've built up so far. Two preliminary steps will be necessary to do this: first,

moving to the center-of-mass reference frame, and second, separating the radial and angular dependencies. The upshot is that our problem cooks down to the problem of understanding a family of Hamiltonians indexed by a nonnegative integer l:

$$H_l = -\frac{\hbar^2}{2\mu}\partial_r^2 - \frac{Ze^2}{r} + \frac{l(l+1)\hbar^2}{2\mu r^2}. \tag{16.54}$$

16.7 The Family of Coulomb Hamiltonians

The problem of the hydrogen atom can be split into a radial part and an angular part. Thanks to spherical symmetry, the angular part can be studied using angular momentum operators and spherical harmonics. We can deduce that the 3D behavior of the electron can be reinterpreted as a 1D wavefunction of a particle in an effective potential which is the two-body interaction potential plus a "barrier" term which depends upon the angular momentum quantum number. Now, we're going to solve the radial part of the problem and thereby find the eigenstates and eigenenergies of the hydrogen atom.

We have a family of Hamiltonians labeled by the angular momentum quantum number:

$$H_l^{\text{Coul}} = -\frac{\hbar^2}{2\mu}\partial_r^2 - \frac{Ze^2}{r} + \frac{l(l+1)\hbar^2}{2\mu r^2}. \tag{16.55}$$

For each eigenstate of any H_l^{Coul}, the hydrogen atom has a set of states labeled by the quantum number m, which takes the values

$$m = 0, \pm 1, \pm 2, \ldots, \pm l. \tag{16.56}$$

However, because the original problem was spherically symmetric, the operator L_3 does not appear in the overall Hamiltonian (there's no preferred axis), and so the energy will not depend upon m. Geometrical symmetry implies that states of different m but the same l are *degenerate*; as we proceed, we'll see an additional example of symmetry implying degeneracy unfold before us.

Carrying around all those constants in H_l^{Coul} would be a pain and not particularly useful, so let's get rid of them. By appropriately rescaling our variables (craftily choosing "better rulers" for distance and energy) we can rewrite H_l^{Coul} in a cleaner form, preserving the essential features of the Hamiltonians without all that dimensionful baggage:

$$H_l^{\text{Coul}} = -\partial_x^2 + \frac{l(l+1)}{x^2} - \frac{1}{x}. \tag{16.57}$$

Now, we have a family of related Hamiltonians, labeled by a nonnegative integer. Why does this situation sound familiar? When we studied supersymmetry and shape invariance, we saw just such a hierarchy of Hamiltonians appear! So, let's see if those ideas are applicable to this problem.

We should note one complication right away: when we explored SUSY QM and shape-invariant partner potentials, we worked with an original Hamiltonian whose ground-state energy was *zero.* Here, however, we're looking at an electron *bound by*

an atom, so the energies of the eigenstates will be *negative.* Shifts of the energy scale aren't going to pose profound difficulties, but we should be on our guard.

16.8 Living up to our Superpotential

Earlier, we took a *superpotential* and from it derived two *partner potentials.* To study H_l^{Coul} in this framework, we'll need a superpotential which, when put through these steps, yields both the $1/x$ Coulomb law and the angular momentum barrier. The superpotential will, therefore, depend upon l. We choose the following:

$$W_l = -\frac{(l+1)}{x} + \frac{1}{2(l+1)}. \tag{16.58}$$

This gives us a pair of operators for each value of l:

$$A_l = \partial_x - \frac{l+1}{x} + \frac{1}{2(l+1)}, \tag{16.59}$$

and

$$A_l^\dagger = -\partial_x - \frac{l+1}{x} + \frac{1}{2(l+1)}. \tag{16.60}$$

The first Hamiltonian is given by one ordering,

$$H_l^{(1)} = A_l^\dagger A_l, \tag{16.61}$$

which works out to be

$$H_l^{(1)} = -\partial_x^2 + \frac{l(l+1)}{x^2} - \frac{1}{x} + \frac{1}{4(l+1)^2}, \tag{16.62}$$

and when we define a Hamiltonian with the other ordering,

$$H_l^{(2)} = A_l A_l^\dagger, \tag{16.63}$$

we get as the partner Hamiltonian

$$H_l^{(2)} = -\partial_x^2 + \frac{(l+1)(l+2)}{x^2} - \frac{1}{x} + \frac{1}{4(l+1)^2}. \tag{16.64}$$

By comparing these two expressions, we get that

$$H_l^{(2)} = H_{l+1}^{(1)} + \frac{1}{4(l+1)^2} - \frac{1}{4(l+2)^2}. \tag{16.65}$$

Here, we see the shape-invariance criterion: $H_l^{(2)}$ is a Hamiltonian which depends upon the parameter l, and we can reach $H_l^{(2)}$ by taking its partner $H_l^{(1)}$, tweaking a

parameter and adding an offset which does not depend upon x. We can see the energy-axis shift mentioned above if we rewrite these operators in terms of the Coulomb Hamiltonian. First, we see that

$$H_l^{(1)} = H_l^{\text{Coul}} + \frac{1}{4(l+1)^2}, \tag{16.66}$$

so the operators $H_l^{(1)}$ and H_l^{Coul} will have the same eigenstates, although their eigenvalues will be separated by a term dependent upon l. We also have

$$H_l^{(2)} = H_{l+1}^{\text{Coul}} + \frac{1}{4(l+1)^2}. \tag{16.67}$$

16.9 Building up the States

We are now in a position to employ the diagrammatic approach we built up earlier. Let's label the eigenstates of a particular Hamiltonian in the hierarchy by ν, such that $\nu = 0$ for the ground state. Each state in our diagram will then be labeled by l and ν, which we write in Dirac notation as $|\nu, l\rangle$. The isospectrality between partner potentials implies that state ν of Hamiltonian $H_l^{(1)}$ has the same energy as state $\nu - 1$ of $H_l^{(2)}$. The state annihilated by A_l is defined by

$$A_l|0, l\rangle = 0, \tag{16.68}$$

which tells us that

$$H_l^{(1)}|0, l\rangle = 0. \tag{16.69}$$

In turn, because of the energy shift, the Coulomb Hamiltonian satisfies

$$H_l^{\text{Coul}}|0, l\rangle = -\frac{1}{4(l+1)^2}|0, l\rangle. \tag{16.70}$$

Likewise, for $l + 1$,

$$H_{l+1}^{(1)}|0, l+1\rangle = 0. \tag{16.71}$$

Using the shape-invariance condition yields

$$H_l^{(2)}|0, l+1\rangle = \left[\frac{1}{4(l+1)^2} - \frac{1}{4(l+2)^2}\right]|0, l+1\rangle. \tag{16.72}$$

By the other key property, isospectrality, we know that $H_l^{(1)}$ has a state of the same energy, one notch higher in ν:

$$H_l^{(1)}|1, l\rangle = \left[\frac{1}{4(l+1)^2} - \frac{1}{4(l+2)^2}\right]|1, l\rangle. \tag{16.73}$$

Again, the physical energy of the hydrogen atom can be found by adding a constant:

$$H_l^{\text{Coul}}|1, l\rangle = -\frac{1}{4(l+2)^2}|1, l\rangle. \tag{16.74}$$

Notice how as we moved from $\nu = 0$ to $\nu = 1$, the value in the denominator went from $l + 1$ to $l + 2$? The rest of the states can be built by working "leftwards" in the state diagram:

$$|\nu, l\rangle = A_l^\dagger A_{l+1}^\dagger \cdots A_{l+\nu-1}^\dagger |0, l + \nu\rangle. \tag{16.75}$$

This just means repeating the process we went through earlier, meaning that the denominator of the eigenvalue will depend on $l + \nu + 1$. Note that the energy of the state labeled by ν and l depends only upon their sum. If we restore the constants we scaled away earlier, we get that

$$E_{\nu l} = -\frac{\mu e^4 Z^2}{2\hbar^2} \frac{1}{(l + \nu + 1)^2}, \tag{16.76}$$

so we can call $l + \nu + 1$ by the new name n and write

$$E_n = -\frac{\mu e^4 Z^2}{2\hbar^2} \frac{1}{n^2}. \tag{16.77}$$

This is the energy of the nth eigenstate of the hydrogenic Hamiltonian. We could work out explicit expressions for states of our choice, but before we plug-and-chug any specific examples, let's take a breather and look at a picture.

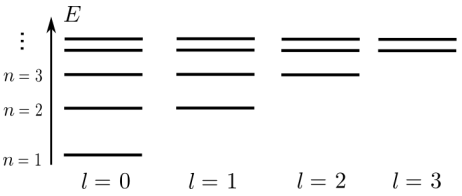


Figure 16.3: The hierarchy of hydrogenic Hamiltonians.

The eigenstate diagram shows *more symmetry* than we had expected on geometrical considerations alone: the only quantum number necessary for determining the energy is n, which corresponds to the energy-level index in Bohr's model of the atom. If our interaction potential were not shape invariant, this degeneracy would be broken.

17 Special Relativity and the Dirac Equation

Or study of the hydrogen atom has been in the grand tradition of textbook quantum mechanics, even though the SUSY machinery is somewhat unusual: We have taken what we know or what we think we know about a situation, and we've rolled it into a set of wavefunctions, which we could use for making statistical predictions about what could happen in various experimental scenarios. After one has worked many problems of this general character, there is a tendency to think of the wavefunctions one writes down not just as calculational artifices, but as *really being out there* — that some kind of ψ-flavored goop is flowing through space, according to its own strange dynamical rules. However, nothing in modern physics is absolutely, ultimately so. The currency of physics is workable approximations. Even our best-confirmed theory of seemingly fundamental constituents of nature — that is, even quantum electrodynamics — is an *effective* theory [135, 136]. Its validity depends not only on the character of the natural world, but on the historical circumstance that we, fragments of nature intervening in the affairs of other fragments, have only been able to intervene in limited ways. Naïvely trusting that the formulas of quantum electrodynamics accurately mirror an ultimate reality is an excellent way to have your calculations blow up in your face and yield meaningless results.

We can tell right off one way in which what we've done so far can only be a limited guide for your experimental grapplings with nature: We have not considered *relativity* at all. We'll now try to remedy this. The majority of problems which involve both relativity and quantum mechanics require the use of *quantum field theory.* However, as luck would have it, we will be able to make progress on this particular problem without developing the full field-theoretic machinery. As Sidney Coleman [137] explained,

> There are very special cases where, because of the details of the dynamics, the kinematic effects of relativity are considerably larger than the effects of pair states. Among them, unfortunately, is the hydrogen atom; and that's why Dirac's theory of the hydrogen atom, which did not include the effects of pairs, was so successful. That was a good thing, because it told people that the basic ideas were right, but it was a bad thing because it led people to spend a lot of time worrying about one-particle and two-particle and three-particle theories, because they didn't realize the hydrogen atom was a very special system.

By studying hydrogenic atoms in considerable detail, we'll be working in a little corner of physics, but an important one. And, as it happens, the mathematics we'll develop while we're busying away in that corner will lead us to unexpected places.

17.1 Lorentz Vectors

Let's ground ourselves with the basic principles of special relativity. First, we have that the laws of physics will appear the same in all inertial frames: If Calvin and Susie are floating past each other in deep space, Calvin can do experiments with springs and whirligigs and beams of light to deduce physical laws, and Susie — who Calvin thinks is moving past with constant velocity — will deduce the same physical laws. Thus, neither Calvin nor Susie can determine who is "really moving" and who is "really standing still."

Second, all observers will measure the same speed of light. Even by testing how fast a ray of light is traveling, Calvin and Susie cannot agree on an absolute standard of rest.[1] In terms of a space-time diagram, where time is conventionally drawn as the vertical axis and space as the horizontal, Calvin and Susie will both represent the progress of a light flash as a diagonal line with the same slope. To make life easy on ourselves, we say that this line has a slope of 1, and is thus drawn at a 45-degree angle from the horizontal. This means we're measuring distance and time *in the same units,* a meter of time being how long it takes light to travel one meter.

Suppose that Calvin sees a burst of light leave a source and arrive a detector a distance x away, taking an amount of time t to do so. We can call the event of the light being emitted A and the event of its absorption B. Calvin writes,

$$x = ct\,. \tag{17.1}$$

Susie, in motion relative to Calvin, watches the same pulse travel from A to B, but he records a different time elapsed, t', and he observes a different distance between the source and detector, x'. The speed of light being the same for Calvin and Susie, we can tell that Susie writes,

$$x' = ct'\,. \tag{17.2}$$

If Calvin chooses his coordinate axes such that the source and detector aren't perfectly along one axis, Calvin will use the Pythagorean theorem to write the distance traveled,

$$d = \sqrt{x^2 + y^2 + z^2}\,, \tag{17.3}$$

and Calvin's equation for the progress of light will be

$$(ct)^2 = x^2 + y^2 + z^2\,. \tag{17.4}$$

If Susie's coordinate axes are similarly tilted, she'll write

$$(ct')^2 = x'^2 + y'^2 + z'^2\,. \tag{17.5}$$

The forms of the equations are the same, but Susie's has primes all over it.

[1] It is interesting that we can express the essence of special relativity so well in terms of physical principles that we can tell conversationally before we start writing equations at all. By contrast, the physics profession lacks a comparably mature understanding of quantum mechanics. People have tried to find one, but no approach has yet been outstandingly successful [138, 139].

Note that if we shuffle all the terms of each equation to the same side, we can see that a quantity remains unchanged by the Lorentz transformation which takes Calvin's coordinates to Susie's:

$$-(ct)^2 + x^2 + y^2 + z^2 = -(ct')^2 + x'^2 + y'^2 + z'^2 = 0\,. \tag{17.6}$$

A massive particle must move slower than light, and so will travel less distance than a light ray in the same amount of time. In that case,

$$(ct)^2 > x^2 + y^2 + z^2\,, \tag{17.7}$$

making the square of the "spacetime interval" negative:

$$s^2 = -(ct)^2 + x^2 + y^2 + z^2 < 0\,, \tag{17.8}$$

so that s is an imaginary number. The essential fact is that Calvin and Susie, while they disagree on the space and time contributions to the interval, *agree* on the interval itself:

$$s^2 = -(ct)^2 + x^2 + y^2 + z^2 = -(ct')^2 + x'^2 + y'^2 + z'^2. \tag{17.9}$$

The interval s is a *Lorentz-invariant* quantity.

We could compress that last equation a little by using summation symbols:

$$s^2 = -(ct)^2 + \sum_{i=1}^{3} x_i^2\,. \tag{17.10}$$

That ct term out in front is irritating. First, if relativity is telling us that Calvin's time is partially Susie's space, then we've no business writing space and time with different units, and having that c cropping up everywhere is just a hangover from a backward way of life. So, let's write,

$$x_0 = ct\,, \tag{17.11}$$

so we can say,

$$s^2 = -x_0^2 + \sum_{i=1}^{3} x_i^2\,. \tag{17.12}$$

Earlier, we said that when we had an expression like $x_i x_i$, most of the time we'd be summing over the repeated index, so that by the "Einstein summation convention" we wouldn't have to write the capital sigma. We'd like to do that here, but we have that time component out in front, with the extra minus sign.

This is where we get sneaky. Quantities will be written as four-component objects called *four-vectors,* in which the components are labeled by an index taking the values 0, 1, 2 and 3. Traditionally, these indices are written with lowercase Greek letters; the position of an index is also significant. The two four-vectors x_μ and x^μ exemplify this difference:

$$x_\mu = (-ct, x_1, x_2, x_3),\ \ x^\mu = (ct, x_1, x_2, x_3)\,. \tag{17.13}$$

The only difference is the sign of the zeroth component. See the trick? If we take the product of these four-vectors, with one index up and the other down, we get what we were looking for:

$$x_\mu x^\mu = -(ct)^2 + x_1^2 + x_2^2 + x_3^2 = s^2 \,. \tag{17.14}$$

Summing over terms with one index up and the other index down yielded us a Lorentz-invariant quantity! In fact, one can *define* the Lorentz transformations as those operations which leave a product of the form $x_\mu x^\mu$ unchanged. As we develop our relativistic quantum mechanics, we'll want to be sure that our equations obey the basic rules of relativity, which means that they'll have to take the same form for Calvin and Susie. That is, we'll want to write all our equations in such a way that we can see their form remains unchanged by a Lorentz transformation, so we'll always be talking about products of four-vectors, with one index up and the other index down.

Rotations and reflections in three dimensions preserve the plain old dot product, and in the language of group theory, they form the Lie group $O(3)$. Lorentz transformations also form a group. It is not $O(4)$, because that minus sign matters; instead we denote it something like $O(3,1)$. The part of the group which is continuously connected to the identity, the analogue of $SO(3)$, is then denoted $SO(3,1)$.

We want to do more than merely describe motion relativistically; we want to *explain* it. In order to advace from kinematics to dynamics, we need relativistic versions of energy and momentum. How do we define these? Well, in pre-Einstein physics, we turned a displacement vector into a velocity vector by dividing by a time interval. This works, i.e., it gives something that transforms like a vector, because spatial rotations leave the time interval unchanged. In special relativity, the situation has changed: Lorentz transformations mix the time and space coordinates together. Suppose we have a displacement four-vector describing the movement of some massive body,

$$dx_\mu = (-c\,dt, dx_1, dx_2, dx_3)\,, \tag{17.15}$$

and we want to make a four-velocity or a four-momentum out of it. We can multiply it by the mass, because the mass of a body in its own inertial frame (its *rest mass*) is a Lorentz invariant. (Inertial observer Alice cannot tell whether she is "really moving" or "really standing still" by measuring the mass of an apple that she holds in her hand.) But to make a velocity, we have to divide by a time. What amount of time can we pick that is itself a Lorentz invariant?

Calvin and Susie, inertial observers in motion relative to each other, will disagree on how much time lapses between two events. But if they are both watching the same object, they can both calculate what the time lapse will be *for an observer moving along with that object.* Tie a wristwatch around an apple; the time shown on the wristwatch when the apple bumps into the wall of the spaceship must be a fact that everyone can agree on. The *proper time,* as we call it, is Lorentz invariant. So, if dx^μ is the four-vector displacement that Calvin uses to describe the apple's motion between two definite events, and $d\tau$ is the time lapse recorded by a clock on the apple itself, then

$$v^\mu = \left(c\frac{dt}{d\tau}, \frac{dx_1}{d\tau}, \frac{dx_2}{d\tau}, \frac{dx_3}{d\tau} \right) \tag{17.16}$$

is a good four-vector velocity. Multiplying by the apple's rest mass m gives a four-vector momentum:

$$p^\mu = m\left(c\frac{dt}{d\tau}, \frac{dx_1}{d\tau}, \frac{dx_2}{d\tau}, \frac{dx_3}{d\tau}\right) . \tag{17.17}$$

Suppose that the apple is moving inertially. In its own rest frame, the spatial components of the four-momentum are all zero, and dt is the same as $d\tau$, so $p_\mu p^\mu = -m^2c^2$. Lorentz transformations preserve this quantity. An observer who sees the apple go zipping past will record a larger time lapse than the apple's proper time — for them, the apple's clock is ticking slow: $dt = \gamma d\tau$. For this observer,

$$p^0 = mc\gamma = \frac{mc}{\sqrt{1 - v^2/c^2}} . \tag{17.18}$$

In the limit of low velocity, $v \ll c$, we can use Taylor expansion to approximate

$$cp^0 = mc^2\left(1 + \frac{1}{2}\frac{v^2}{c^2}\right) = mc^2 + \frac{1}{2}mv^2 . \tag{17.19}$$

The second term here is just the kinetic energy from childhood physics. So, we identify cp^0 as the relativistic energy, with a rest-mass contribution mc^2 and a kinetic contribution. The relativistic relationship among mass, energy and momentum is just the statement that $p_\mu p^\mu$ is Lorentz invariant:

$$p_\mu p^\mu = -m^2c^2 = -\frac{E^2}{c^2} + p^2 . \tag{17.20}$$

Solving this for the energy squared,

$$E^2 = m^2c^4 + p^2c^2 . \tag{17.21}$$

This equation even applies to photons, which have zero rest mass: They possess energy by virtue of always traveling at the speed of light.

When doing electromagnetism, one employs a four-vector version of the vector potential, which incorporates the information for the electric and magnetic fields:

$$\begin{aligned} A_\mu &= \left(-\frac{\phi}{c}, A_1, A_2, A_3\right) , \\ A^\mu &= \left(\frac{\phi}{c}, A_1, A_2, A_3\right) . \end{aligned} \tag{17.22}$$

Recall that in high school, as professors like to say, we learned that the electric field can be written as the gradient of a scalar potential,

$$\vec{E} = -\nabla\phi , \tag{17.23}$$

while the magnetic field is the curl of the vector potential:

$$\vec{B} = \nabla \times \vec{A} . \tag{17.24}$$

In quantum mechanics, it often turns out more convenient to work with these potentials than with the fields we studied in classical electromagnetism.

17.2 Dirac's Very Nice Guess

The first step in developing a relativistic understanding of the quantum theory is to find an equation which does the job of the unitary time evolution we studied in the relativistic regime. As indicated, we'll be focusing on quantities which remain unchanged under Lorentz transformations, *i.e.,* quantities on which all observers resting in inertial frames will agree. We would like to write an equation which not only describes a particle in relativistic motion, but which contains only expressions which all Lorentz observers agree upon.

We can begin by studying the good old unitary wave equation for the time evolution of a state:

$$H|\psi(t)\rangle = i\hbar\partial_t|\psi(t)\rangle\,. \tag{17.25}$$

We would like to write a Hamiltonian which is relativistically invariant. The first guess might be to use Einstein's expression for the energy, $E^2 = m^2c^4 + |\vec{p}|^2c^2$. If we use this energy for our Hamiltonian, we get that the time evolution obeys

$$\sqrt{(mc^2)^2 + \sum_{j=1}^{3}(p_jc)^2}|\psi(t)\rangle = i\hbar\partial_t|\psi(t)\rangle\,. \tag{17.26}$$

But this wasn't what Dirac wanted, because it treats time and space too differently. One observer's time is partly another's space, so singling out a temporal or spatial dependence cannot work in a relativistic equation.

Dirac's solution was to suppose that since the right-hand side of the time-evolution equation contains a first-order time derivative, the left-hand side should contain corresponding first-order derivatives in space. Dirac showed that this could be done by writing the quantity under the radical as the square of an expression linear in p_j. (Recall that in quantum mechanics, $p_j = -i\hbar\partial_j$.) Taking the square root of a perfect square is easy, so Dirac's proposal is algebraically convenient. Writing Dirac's idea out in symbols, we get

$$(mc^2)^2 + \sum_{j=1}^{3}(p_jc)^2 = \left(\alpha_0mc^2 + \sum_{j=1}^{3}\alpha_jp_jc\right)^2\,. \tag{17.27}$$

These conditions cannot be satisfied if the αs are ordinary numbers, but it is easy to find matrices which do the job. We need a total of four, and each will be a 4×4 matrix. Letting the indices μ,ν run from 0 to 3, the conditions on our αs can be written

$$\{\alpha_\mu,\alpha_\nu\} = 2\delta_{\mu\nu}I\,. \tag{17.28}$$

Dirac's choice for solving this equation was the matrices

$$\alpha_0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \tag{17.29}$$

$$\alpha_j = \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix}. \tag{17.30}$$

Many choices of matrices work to satisfy the condition on our matrices, but they lead to equivalent physics, and this choice is a convenient one. Here, I signifies the 2×2 identity matrix, σ_j are the Pauli matrices, and 0 stands for a 2×2 matrix full of zeroes. Now, it is a straightforward matter to write the equation we need:

$$H_D|\psi(t)\rangle = i\hbar\partial_t|\psi(t)\rangle\,, \tag{17.31}$$

where

$$H_D = \alpha_0 mc^2 + \sum_{j=1}^{3} \alpha_j p_j c\,. \tag{17.32}$$

17.3 Antimatter

The relationship defined by our previous two equations is just a *guess* at how the world works. To see if it has any value, we have to work out the consequences of that guess, and then compare those implications to the physical world.

With the Dirac equation in hand, then, we can make a few observations. First, the state $|\psi\rangle$ is now a four-component object, which we call a *spinor.* To understand what this means, we look at the Dirac Hamiltonian's energy eigenvalues. Rewrite the time-independent part of our equation as $H_D|\psi\rangle = E|\psi\rangle$, and try for a plane-wave solution. For convenience, assume that the wave is propagating in the z-direction, giving $\psi(z) = \chi e^{\frac{ipz}{\hbar}}$, where p is the momentum and χ is a constant spinor. Substituting this ansatz into the Dirac equation shows that each value of p has two associated eigenspaces, one containing states of positive energy and the other with negative energy. The α_1 and α_2 terms in the Dirac Hamiltonian will vanish when applied to a state of this form, and the α_3 term will give $p\chi e^{ipz/\hbar}$, so all together the Hamiltonian acts on these states as the matrix

$$\begin{pmatrix} mc^2 & 0 & pc & 0 \\ 0 & mc^2 & 0 & -pc \\ pc & 0 & -mc^2 & 0 \\ 0 & -pc & 0 & -mc^2 \end{pmatrix}. \tag{17.33}$$

This matrix has two distinct eigenvalues, each of multiplicity two:

$$E_\pm(p) = \pm\sqrt{(mc^2)^2 + (pc)^2}\,. \tag{17.34}$$

Dirac viewed the negative-energy solutions as "holes" in the vacuum, which he pictured as a sea of electrons. (This image is analogous to the solid-state physics concept of holes in a solid's conduction band.) In modern times, we associate the negative-energy solution with the electron's antiparticle, the positron: The positron is the *absence* of a negative-energy state, which forms a positive-energy hole. In any case, if we look at the nonrelativistic limit $|E| - mc^2 \approx p^2/2m \ll pc$, the eigenstates which span the two eigenspaces are associated with amplitudes for being in spin-up or spin-down states, with either positive or negative energy. This explains why we need a four-component object to hold all the information associated with a single electron.

By *guessing* that the relativistic motion of the electron should have a certain form, we've found that the electron has to have a "twin" — somewhere, there's an electron with a beard! Four years after Dirac came out with his equation, Carl Anderson [140] published a study of 1300 cosmic-ray tracks which revealed a new, positively charged particle which couldn't be as massive as the proton:

> From an examination of the energy-loss and ionization produced it is concluded that the charge is less than twice, and is probably exactly equal to, that of the proton. If these particles carry unit positive charge the curvatures and ionizations produced require the mass to be less than twenty times the electron mass.

The rest, as they say, is history.

17.4 Incorporating Electromagnetism

The above analysis applies to an electron propagating in free space. We can modify the Dirac Hamiltonian H_D to include electromagnetic interactions fairly easily, by examining the momentum terms. In a nonzero vector potential, the Hamiltonian becomes the following:

$$H_D = \alpha_0 mc^2 + \sum_j \alpha_j \left[p_j - \frac{e}{c} A_j \right] c + e\phi \,. \tag{17.35}$$

Both $\vec{A}$ and the scalar potential ϕ can, of course, be functions of $\vec{x}$ and t. (This is a semiclassical approximation, because it treats the electromagnetic field as obeying Maxwell's laws. A better approximation would involve quantizing the EM field as well, but this leads into quantum field theory, which we've already admitted we won't be getting into.)

From now on, keeping factors of c will be more tiresome than useful, so take $c = 1$, and thus $E = m$. The Einstein summation convention, combined with four-vectors, allows us to give the Dirac Hamiltonian a compact form,

$$H_D = \alpha_0 m + \alpha_j (p_j - eA_j) + e\phi \,, \tag{17.36}$$

which can be rewritten as

$$H_D = \alpha^\mu (p_\mu - eA_\mu) \,, \tag{17.37}$$

where

$$\alpha^\mu = (\alpha_0, \alpha_1, \alpha_2, \alpha_3) \,. \tag{17.38}$$

In fact, ∂_t can also be written ∂_0, since (up to an irrelevant factor c) the quantity x^0 is just the time observed in a particular Lorentz frame. With this idea in mind, we can write yet another four-vector made entirely of derivatives:

$$\partial^\mu = (-\partial_0, \partial_1, \partial_2, \partial_3) \,. \tag{17.39}$$

Another useful object with indices is the flat space metric, denoted $\eta_{\mu\nu}$. It is a matrix whose entries are all zero except along the diagonal, where they are given by $(-1, 1, 1, 1)$. The -1 in the zeroth entry reflects the special status the time dimension has in Einstein's theory. To put our equations in the form most convenient for later work, we define new matrices by

$$\gamma_0 = \alpha_0, \ \gamma_j = \alpha_0 \alpha_j \,, \tag{17.40}$$

which have the anticommutator relation

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu} \,. \tag{17.41}$$

Using the gamma matrices, whose relationship defines a "Clifford algebra," the Dirac equation can be recast in the manifestly covariant form

$$[i\gamma^\mu(\partial_\mu + iA_\mu) - m]\,\psi = 0 \,. \tag{17.42}$$

One can verify this by substituting the definition of γ_μ given above and comparing the result to the original Dirac equation, given by $H_D\psi = i\partial_0\psi$.

17.5 SUSY in the Dirac Hamiltonian

Whew! We spent a considerable amount of wordage developing the Dirac equation. Now, it's time to tie this development back to the supersymmetry material we studied earlier in the nonrelativistic context. The result will be a surprising mapping between relativistic and nonrelativistic quantum mechanics. First, we'll just get the gist of it, and to get started, we'll begin with the final equation we had before,

$$(i\gamma^\mu\partial_\mu - m)\,\psi = 0. \tag{17.43}$$

We can elaborate this to include an electromagnetic field as follows:

$$[i\gamma^\mu(\partial_\mu + iA_\mu) - m]\,\psi = 0. \tag{17.44}$$

The Dirac Hamiltonian H_D has a rich SUSY structure, of which we can catch a glimpse even having pared the problem down to its barest essentials. To take the simplest possible case, consider a Dirac particle living in one spatial dimension, on which there also lives a scalar potential $\phi(x^1)$. (We could call this a "1+1-dimensional" system, to remind ourselves of the difference between time and space.) The SUSY structure can be seen most clearly when we look at the limit of a massless particle; this eliminates the m term we had before.

All that fuss over the matrices γ^μ has its practical benefit now, because writing our formulas in terms of them allows us to write a *new* Dirac equation embodying the analogous physics in *a different number of dimensions.* Just as our first α matrices were defined by the anticommutator relation

$$\{\alpha_\mu, \alpha_\nu\} = 2\delta_{\mu\nu} I, \tag{17.45}$$

we now consider the matrices γ^μ defined by their anticommutator,

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu}. \tag{17.46}$$

The only difference between our original, 3+1-dimensional problem and our current one is that indices like μ and ν will only run over 0 and 1 instead of running up to 3. A useful choice satisfying the anticommutator constraint is

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \gamma^1 &= \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \end{aligned} \tag{17.47}$$

What was a position four-vector is now a "two-vector," but we must remember that moving an index from up to down involves changing the sign on the zeroth component. (The oneth component remains the same: $x_1 = x^1$.) Writing an unadorned x to stand for the entire vector, the Dirac equation is

$$i\gamma^\mu \partial_\mu \psi(x) - \phi(x^0)\psi(x) = 0. \tag{17.48}$$

As usual, we separate out the time dependence, leaving a wavefunction which depends only on x_1:

$$\psi(x) = \exp(-i\omega x^0)\psi(x^1) = \exp(i\omega x_0)\psi(x_1). \tag{17.49}$$

Substituting in this ansatz reduces the 1+1-dimensional Dirac equation to

$$\gamma^0 \omega \psi(x^1) + i\gamma^1 \partial_1 \psi(x^1) - \phi(x^1)\psi(x^1) = 0. \tag{17.50}$$

If we write the two-component wavefunction $\psi(x^1)$ as

$$\psi(x^1) = \begin{pmatrix} \xi(x^1) \\ \chi(x^1) \end{pmatrix}, \tag{17.51}$$

then the reduced Dirac equation becomes the two coupled equations

$$A\xi(x^1) = \omega\chi(x^1), \tag{17.52}$$

$$A^\dagger \chi(x^1) = \omega\xi(x^1), \tag{17.53}$$

where the operators A and $A^\dagger$ are given by

$$A = \partial_1 + \phi(x^1),\ A^\dagger = -\partial_1 + \phi(x^1). \tag{17.54}$$

Decoupling our equations gives a familiar-looking result:

$$A^\dagger A\xi(x^1) = \omega^2 \xi(x^1), \tag{17.55}$$

$$AA^\dagger \chi(x^1) = \omega^2 \chi(x^1). \tag{17.56}$$

Our result identifies readily with the SUSY partner potential concepts developed earlier. $\phi(x^1)$ is just the superpotential of our old, Schrödinger formalism. What's more, ξ and χ are the eigenfunctions of the two Hamiltonians $H_1 = A^\dagger A$ and $H_2 = AA^\dagger$, which we already know are isospectral, except that one has an extra state at zero energy.

Looking back to our results on shape invariance, we can see that every Schrödinger problem with a potential $V_1(x)$ which has a shape-invariant partner $V_2(x)$ can become the scalar potential $\phi(x^1)$ in a Dirac problem!

Of course, physical problems using the Dirac equation exist in more than one dimension, and the electron is not a massless particle. If we soup up our machinery just a little, however, we can apply it to physical situations, such as figuring out what a relativistic charged particle in a Coulomb field will do [141]. This provides the relativistic corrections to our earlier solution of the hydrogenic atom.

17.6 Massive Dirac Electrons in $3+1$ Dimensions

Here, the details of the manipulations do not themselves contain many great insights, and so we will move through the equations with only a little commentary. Choosing $\hbar = c = q = 1$,

$$[i\gamma^\mu(\partial_\mu + iA_\mu) - m]\psi = 0. \tag{17.57}$$

For a central field we can set

$$\vec{A} = 0,\ A_0(\vec{x}, t) = V(r). \tag{17.58}$$

This then implies

$$i\partial_t\psi = H\psi = (\vec{\alpha}\cdot\vec{p} + \beta m + V)\psi, \tag{17.59}$$

where we have defined

$$\alpha^i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}, \tag{17.60}$$

and

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{17.61}$$

We separate in spherical coordinates and focus on the radial equations

$$G'(r) + \frac{kG}{r} - (m + E - V)F = 0, \tag{17.62}$$

$$F'(r) - \frac{kF}{r} - (m - E + V)G = 0. \tag{17.63}$$

For the particular special case of the Coulomb problem, we let

$$V(r) = -\frac{\gamma}{r},\ \text{where } \gamma = Zq^2. \tag{17.64}$$

So,

$$\begin{pmatrix} G_k'(r) \\ F_k'(r) \end{pmatrix} + \frac{1}{r}\begin{pmatrix} k & -\gamma \\ \gamma & -k \end{pmatrix}\begin{pmatrix} G_k \\ F_k \end{pmatrix} = \begin{pmatrix} 0 & m+E \\ m-E & 0 \end{pmatrix}\begin{pmatrix} G_k \\ F_k \end{pmatrix}. \tag{17.65}$$

Here, k is an eigenvalue of the operator $-(\vec{\sigma}\cdot\vec{L}+1)$, whose magnitude is related to the angular momentum by $|k| = J + \frac{1}{2}$ and which takes values in

$$k = \pm 1, \pm 2, \pm 3, \ldots. \tag{17.66}$$

We diagonalize the matrix of coefficients on the left-hand side of our coupled equations by writing

$$D\begin{pmatrix} k & -\gamma \\ \gamma & -k \end{pmatrix} D^{-1}, \tag{17.67}$$

where

$$D = \begin{pmatrix} k+s & -\gamma \\ -\gamma & k+s \end{pmatrix}, \text{ with } s = \sqrt{-\det\begin{pmatrix} k & -\gamma \\ \gamma & -k \end{pmatrix}} = \sqrt{k^2-\gamma^2}. \tag{17.68}$$

Then, introducing the scaled radial variable $\rho = Er$, we can write the transformed radial functions

$$\begin{pmatrix} \tilde{G} \\ \tilde{F} \end{pmatrix} = D\begin{pmatrix} G \\ F \end{pmatrix}. \tag{17.69}$$

Employing the operators

$$A = \partial_\rho - \frac{s}{\rho} + \frac{\gamma}{s}, \tag{17.70}$$

$$A^\dagger = -\partial_\rho - \frac{s}{\rho} + \frac{\gamma}{s}, \tag{17.71}$$

we have

$$A\tilde{F} = \left(\frac{m}{E} - \frac{k}{s}\right)\tilde{G}, \tag{17.72}$$

$$A\tilde{G} = -\left(\frac{m}{E} + \frac{k}{s}\right)\tilde{F}. \tag{17.73}$$

Applying $A^\dagger$ to the first equation and A to the second, we separate the differential equations for the two radial functions, yielding

$$H_-\tilde{F} = A^\dagger A\tilde{F} = \left(\frac{k^2}{s^2} - \frac{m^2}{E^2}\right)\tilde{F}, \tag{17.74}$$

$$H_+\tilde{G} = AA^\dagger\tilde{G} = \left(\frac{k^2}{s^2} - \frac{m^2}{E^2}\right)\tilde{G}. \tag{17.75}$$

Observe that

$$H_+(\rho; s, \gamma) = H_-(\rho; s+1, \gamma) + \frac{\gamma^2}{s^2} - \frac{\gamma^2}{(s+1)^2}. \tag{17.76}$$

We have found that the problem of the Dirac electron in a Coulomb potential satisfies the shape invariance criterion. The transformation of the parameters upon which the potential depends is just a simple shift from s to $s+1$, and the position-independent residual term is

$$R(a_2) = \frac{\gamma^2}{a_1^2} - \frac{\gamma^2}{a_2^2}. \tag{17.77}$$

As in the nonrelativistic problem, we can find the spectrum of energy levels by summing up the residual terms through successive transformations.

$$\left(\frac{k^2}{s^2} - \frac{m^2}{E_n^2}\right) = \sum_{k=2}^{n+1} R(a_k) = \gamma^2 \left(\frac{1}{s^2} - \frac{1}{(s+n)^2}\right). \tag{17.78}$$

At last, we have

$$E_n = m\left(1 + \frac{\gamma^2}{(s+n)^2}\right)^{-1/2}, \tag{17.79}$$

where $n = 0, 1, 2, \ldots$. This includes the rest mass of the electron; a more direct comparison to the nonrelativistic problem can be made by restoring the constant c and subtracting off an mc^2. The result is

$$E = -\frac{m\gamma^2}{2(n+|k|)^2}\,, \tag{17.80}$$

which can also be written using the "principal quantum number"

$$N = n + |k| \tag{17.81}$$

and the orbital and total angular momentum numbers l and j. That is, we can carry over the nonrelativistic labels for the eigenstates. Using the Bohr radius a_0 and the fine-structure constant

$$\alpha = \frac{q_e^2}{\hbar c} \simeq \frac{1}{137}\,, \tag{17.82}$$

one finds

$$E_{Nlj} = -\frac{q_e^2}{2a_0}\frac{1}{N^2}\left[1 + \frac{\alpha^2}{N^2}\left(\frac{N}{j+\frac{1}{2}} - \frac{3}{4}\right)\right]. \tag{17.83}$$

This formula gives what we call the *fine structure* of the hydrogen spectrum, i.e., one stage of understanding finer than the Bohr energy levels. The Bohr energy levels have a characteristic energy scale of $\alpha^2 mc^2$, and the fine-structure corrections are of order $\alpha^4 mc^2$. They lift some, but not all, of the degeneracies we found in the nonrelativistic treatment. For example, in the nonrelativistic model, the $n = 2, l = 1$ states were degenerate; incorporating the fine structure, we get a $j = 1/2$ state that has the same energy as the $n = 2, l = 0$ level, and a $j = 3/2$ level that is slightly higher. All the energy shifts are downwards from the nonrelativistic values, with lower j getting pushed down more.

This is as far as we can trust the Dirac equation to tell us energy levels. The *hyperfine* correction further amends the fine structure. It is due to the coupling between the

magnetic moments of the electron and the proton, and so it is smaller than the fine structure corrections by the ratio of the electron and proton masses. And full-blown quantum electrodynamics, a theory where particle number is not fixed, gives us the *Lamb shift* of order $\alpha^5 mc^2$ that further lifts degeneracies beyond what Dirac's theory can account for.

18 Semiclassical Methods

18.1 Continuum Wigner Function

Back in the 1930s, Eugene Wigner wanted to do quantum statistical physics. More specifically, he wanted a way to incorporate Boltzmann-style probabilities, those proportional to $e^{-\beta E}$, into quantum mechanics. But quantum mechanics had no analogue for a probability density on phase space, because it doesn't admit the idea of a phase space. Classically, we could speak of the position and momentum of a corpuscle as quantities tied to it and independent of how we chose to measure it, but quantum physics disallows that. So, there was no way to write a probability density $P(x, p)$, because the physics doesn't allow x and p to be simultaneously meaningful. Wigner had done some work with Szilard, however, and he realized that an expression they had devised could be repurposed into a function that was *almost* a probability density for a quantum particle: It was normalized, integrating to 1; taking the marginal over x or p worked as desired; yet unlike a probability density, it went *negative.* Wigner published a solo-author paper in 1932 [142], and ever since, the expression that he told everyone he co-invented with Szilard has been called the *Wigner function.*

Wigner did not present an argument for where the formula came from, instead noting that there were many expressions that satisfied the conditions he had in mind, but the one he settled on "seems to be the simplest" [142]. The tradition since then has been to present the definition without explanation. We can do a little better by thinking about properties we'd like it to satisfy. We'll work in the context of quantum mechanics on the line, so our pure states will be wavefunctions $\psi(x)$, and our mixed states will be density operators. First, an average over density operators should be equivalent to an average over quasiprobability distributions. I.e., the map from ρ to $W(x, p)$ should respect convex combinations. We'll take it to be a *linear* function of ρ at each point (x, p). We could impose the weaker condition that it be an affine function, i.e., a linear one plus a constant offset; this would also respect weighted averages, but it would play less well with more general linear combinations. We'd like our map to handle operators that aren't necessarily normalized, like the elements of POVMs, and at this stage, we may as well take the path that is easiest for us. So, let the map be linear. This means that we can write it as an inner product of the density operator with something else:

$$W(x, p) = \mathrm{tr}(\rho \hat{\delta}(x, p)) \,. \tag{18.1}$$

We don't yet know what the $\hat{\delta}$ operators are, just that there is one for each point in our quasi-phase space, and they ought to act something like delta functions, each selecting one point out of the whole space. We recall that the Fourier transform of the ordinary

delta function is a flat line, or equivalently, the delta function can be represented as an integral over a continuum of frequency modes. So, the $\hat{\delta}$ operators should involve an integral over a complex exponential.

We know that it is a little awkward to write a "position eigenstate" $|x\rangle$ (as we discussed in §11.8). But we can introduce them as calculational aids anyway, to help ourselves at intermediate stages. Let's do this and express our $\hat{\delta}$ operators that way:

$$\hat{\delta}(x,p) = \int dx'\, dx''\, f_{(x,p)}(x',x'')|x'\rangle\langle x''|\,. \tag{18.2}$$

We need a complex exponential in there somewhere, and the typical argument for those in quantum theory on a line looks like i times momentum times position over $\hbar$. We don't want to single out a special point on the line, so the dependence of $f_{(x,p)}(x',x'')$ on x should only involve the *difference* of the arguments from x. We also know that when ρ is a pure state,

$$\langle x''|\rho|x'\rangle = \langle x''|\psi\rangle\langle\psi|x'\rangle = \psi(x'')^*\psi(x')\,. \tag{18.3}$$

In this form, it's easy to see that boosting ψ along the momentum axis by an amount p' introduces a factor of $e^{ip'(x'-x'')/\hbar}$. Inserting this factor has to be equivalent to changing the value of p by a shift p'. Let's consider operators of the form

$$\hat{\delta}(x,p) = a\int dx'\, dx''\, |x'\rangle\langle x''|e^{ip(x'-x'')/\hbar}\,, \tag{18.4}$$

where a is a constant to be worried about later, and the relation among x, x' and x'' is something which we need to figure out. Then our quasiprobability will be of the form

$$a\int dx'\, dx''\, \langle x''|\rho|x'\rangle e^{ip(x'-x'')/\hbar}\,. \tag{18.5}$$

This integral needs to evaluate to a real number in all circumstances. So, we see how it behaves under complex conjugation. The first factor becomes $\langle x'|\rho|x''\rangle$, and the second becomes $e^{-ip(x'-x'')/\hbar}$. Exchanging the labels x' and x'' cancels the effect of switching the sign on the exponent, which is good. Moreover, integrating the whole thing over p, the exponential factor gives us a delta function peaked at $x' = x''$, and in this case, the whole expression needs to reduce to $\langle x|\rho|x\rangle$. Putting all these considerations together, we find they are satisfied if one label looks like $x+y$ and the other like $x-y$.

Wigner's (and Szilard's) expression amounts to taking

$$\hat{\delta}(x,p) = \frac{1}{\pi\hbar}\int dy\, |x+y\rangle\langle x-y|e^{2ipy/\hbar}\,, \tag{18.6}$$

so that

$$W(x,p) = \frac{1}{\pi\hbar}\int dy\, \langle x-y|\rho|x+y\rangle e^{2ipy/\hbar}\,. \tag{18.7}$$

This is the Wigner function of ρ. As is usual in anything related to Fourier transforms, factors of 2 can migrate around the formula depending on who writes it.

The Wigner function $W(x,p)$ is real-valued and enjoys the property that its marginals over x and p give probability densities for momentum and position measurements respectively. Indeed, if $f(x)$ is a function of the position operator alone, and $g(p)$ is a function of the momentum operator alone, then

$$\langle f(x) + g(p) \rangle = \int dx\, dp\, [f(x) + g(p)] W(x,p)\,. \tag{18.8}$$

Moreover, if W_ρ and W_σ are the Wigner function representations of the density operators ρ and σ, then

$$\mathrm{tr}(\rho\sigma) = 2\pi\hbar \int dx\, dp\, W_\rho(x,p) W_\sigma(x,p)\,. \tag{18.9}$$

This implies that

$$\int dx\, dp\, [W(x,p)]^2 \leq \frac{1}{2\pi\hbar}\,, \tag{18.10}$$

with equality obtaining if and only if the quantum state is pure. Furthermore, if we *convolve* any Wigner function with any other, the result is everywhere nonnegative, because its value at each point is, up to a positive prefactor, the inner product between two density operators.

The coherent states $|z\rangle$ we defined in Eq. (11.116) turn out to have Wigner functions that look like *Gaussian* distributions. Working in "natural units" where $\hbar = 1$ and writing

$$z = \frac{1}{\sqrt{2}}(X + iP)\,, \tag{18.11}$$

one finds that

$$W(x,p) = 2e^{-(x-X)^2 - (p-P)^2}\,. \tag{18.12}$$

Hudson's theorem says that the only states whose Wigner functions are everywhere nonnegative are mixtures of Gaussians, i.e., coherent states and "squeezed" modifications thereof:

$$|z,\eta\rangle = \exp(\eta(a^\dagger)^2 - \eta^* a^2)|z\rangle,\ \ \eta \in \mathbb{C}\,. \tag{18.13}$$

This raises an intriguing possibility. If we stuck to using only Gaussian states and quantum operations that stay within the realm of Gaussian states, could we find a portion of quantum mechanics that works just like classical physics? Is there a way of starting with Liouville dynamics, like we studied in classical kinetic theory, and imposing a restriction that the Liouville densities never become too sharp, resulting in a theory that is *equivalent to a subtheory of quantum mechanics*?

Yes, it is [143]! The idea is to introduce a new parameter into Liouville mechanics that has units of action and plays the role of the Planck constant. One then imposes the restriction that the position and momentum variables satisfy an uncertainty relation — the observer cannot become confident of the values of conjugate variables simultaneously — and that the Liouville density maximizes the continuum Shannon entropy. The resulting "epistemically restricted" Liouville mechanics has a set of allowed states, measurements and transformations that coincide with Gaussian quantum

theory. A particularly noteworthy aspect of this theory is that it includes entangled states. So, entanglement by itself is not as strange as often advertised: We can mimic it using vanilla uncertainty about classical degrees of freedom! True strangeness requires not just entanglement, but a richer set of measurements than the Gaussian subtheory provides.

Models of this type can even reproduce the interference phenomena that have so frequently been touted as revealing the essence of quantum mechanics [144]. The reason is rather cute: If all states are probability distributions with a substantial spread over the underlying phase space, then so too is the *vacuum*. A state labeled by having no photon in a given path of an interferometer, for example, must actually be a distribution. So, even if the photon number is zero, the underlying dynamics can still carry a signal — like hearing the pitch at which a rest is played!

18.2 Single-Particle Partition Function

Let's suppose that we have a quantum particle whose dynamics we describe by the Hamiltonian H. Its eigenvalues are the energy levels $\{E_n\}$, and the canonical probability distribution is

$$Z = \sum_n e^{-\beta E_n} = \sum_n \langle n|e^{-\beta H}|n\rangle \,. \tag{18.14}$$

We can write this in a basis-independent way:

$$Z = \sum_n \mathrm{tr}(e^{-\beta H}|n\rangle\langle n|) = \mathrm{tr}\left(e^{-\beta H}\sum_n |n\rangle\langle n|\right) = \mathrm{tr}e^{-\beta H} \,, \tag{18.15}$$

because the energy eigenstates provide a resolution of the identity.

For a particle on a line, the Hamiltonian will look like

$$H = \frac{p^2}{2m} + V(x) \,, \tag{18.16}$$

and exponentiating that will get messy because of what happens when one takes the exponential of things that don't commute. As we saw leading up to Eq. (14.69),

$$e^A e^B = e^{A+B} e^{\frac{1}{2}[A,B]} \,, \tag{18.17}$$

provided that A and B both commute with $[A, B]$. This condition is satisfied by the position and momentum operators. So,

$$e^{-\beta H} = e^{-\beta p^2/2m} e^{-\beta V(x)} + \mathcal{O}(\hbar) \,, \tag{18.18}$$

meaning that we get a classical-looking expression plus a contribution that would vanish if $\hbar$ were negligible.

We can use this to evaluate the partition function, neglecting contributions that grow with $\hbar$, using position and momentum "eigenstates". Because position and momentum are Fourier conjugates,

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \,. \tag{18.19}$$

We evaluate the partition function by taking the trace, via the position basis:

$$Z = \int dx \, \langle x|e^{-\beta p^2/2m} e^{-\beta V(x)}|x\rangle \, . \tag{18.20}$$

The $e^{-\beta V(x)}$ factor acts upon the ket and becomes a number:

$$Z = \int dx \, e^{-\beta V(x)} \langle x|e^{-\beta p^2/2m}|x\rangle \, . \tag{18.21}$$

We know what $e^{-\beta p^2/2m}$ does in the momentum basis, so we insert two resolutions of the identity:

$$Z = \int dx \, dp' \, dp'' e^{-\beta V(x)} \langle x|p'\rangle \langle p'|e^{-\beta p^2/2m}|p''\rangle \langle p''|q\rangle \, . \tag{18.22}$$

We act with $e^{-\beta p^2/2m}$ to the right (or to the left), employ the orthogonality of the momentum kets and use the Fourier overlap condition between bases. The result is

$$Z = \frac{1}{2\pi\hbar} \int dx \, dp \, e^{-\beta H(x,p)} \, , \tag{18.23}$$

which looks just like applying Boltzmann weighting to a classical Hamiltonian, apart from that prefactor. We haven't made $\hbar$ go away completely, but we have reduced it to doing a boring job: It looks like it's just canceling the units acquired by integrating over x and p. In other words, this is what we'd get if we tried a wholly classical calculation in which we approximated the discrete sum in the partition function by an integral, dividing up the phase space into "cells" of area $2\pi\hbar$.

Doing this in three dimensions just means going from scalars to vectors and changing the prefactor to $(2\pi\hbar)^{-3}$. The partition function then becomes a Gaussian integral that evaluates to

$$Z = V \left(\frac{m k_B T}{2\pi\hbar^2} \right)^{3/2} \, , \tag{18.24}$$

where V is the volume of the three-dimensional box.

Because we know the partition function, we can find the free energy using the bridging equation, and thus we can calculate the pressure:

$$P = -\frac{\partial F}{\partial V} = \frac{\partial k_B T \log Z}{\partial V} = \frac{k_B T}{V} \, . \tag{18.25}$$

This is just the ideal gas law for $N = 1$.

18.3 Decoherence

Alice holds a quantum system before her and plans to perform a sequence of binary measurements, one for each tick of a clock. She represents each measurement by a POVM, and she factors each POVM element into Kraus operators. She assumes that

the Kraus operators describing a measurement outcome at any time t do not depend upon which outcomes she might obtain prior to t. Her Kraus operator representing a positive outcome (a "click") during the interval $[t, t+dt]$ can for convenience be written

$$M_1(dt) = L\sqrt{dt}\,, \tag{18.26}$$

where L is the *Lindblad operator* for that time interval. (For brevity, we suppress the t dependence of the Lindblad operators in our notation.) A standard calculation then gives the probability of a click during that interval. If $\rho(t)$ is the density matrix encoding Alice's expectations pertaining to time t, then the Born rule says to compute the probability by

$$dp(t) = \mathrm{tr}(M_1^\dagger(dt) M_1(dt)\rho(t)) = \mathrm{tr}(L^\dagger L\rho(t))\,dt\,. \tag{18.27}$$

Because Alice has a factorization of her POVM elements into Kraus operators, she can compute the post-measurement state to which she expects to update after receiving a click:

$$\rho(t+dt) = \frac{L\rho(t)L^\dagger}{\mathrm{tr}(L^\dagger L\rho(t))}\,. \tag{18.28}$$

Alice also has a set of Kraus operators $\{M_0(dt)\}$ representing the events of *not* detecting a click. Because the elements of any POVM must sum to the identity, these satisfy

$$M_0^\dagger(dt)M_0(dt) = I - M_1^\dagger(dt)M_1(dt) = I - L^\dagger L dt\,. \tag{18.29}$$

We can think of this as fixing the "magnitude" of $M_0(dt)$. Invoking the polar decomposition, we can write $M_0(dt)$ as $\sqrt{I - L^\dagger L dt}$ times a unitary, which in turn we can treat as generated by some Hamiltonian:

$$M_0(dt) = \sqrt{I = L^\dagger L dt}\exp\left(-\frac{i}{\hbar}H dt\right)\,. \tag{18.30}$$

To first order in the time increment dt, then,

$$M_0(dt) \approx \left(I - \frac{1}{2}L^\dagger L dt\right)\left(I - \frac{i}{\hbar}H dt\right) \approx I - \frac{1}{2}L^\dagger L dt - \frac{i}{\hbar}H dt\,. \tag{18.31}$$

Alice's probability for *not* receiving a "click" during the time interval $[t, t+dt]$ is

$$1 - dp(t) = \mathrm{tr}(M_0(dt)^\dagger M_0(dt)\rho(t))\,. \tag{18.32}$$

Her post-measurement state in the absence of a "click" event is

$$\rho(t+dt) = \frac{M_0(dt)\rho(t)M_0^\dagger(dt)}{1 - dp(t)}\,, \tag{18.33}$$

which to first order in dt is

$$\rho(t+dt) = \frac{\left(I - \frac{1}{2}L^\dagger L dt - \frac{i}{\hbar}H dt\right)\rho(t)\left(I - \frac{1}{2}L^\dagger L dt + \frac{i}{\hbar}H dt\right)}{1 - dp(t)}\,. \tag{18.34}$$

Alice can consider all possible sequences of clicks and not-clicks, weighted by their respective probabilities. Using the reflection principle (§2.2), she writes a density matrix $\rho(T)$ that is her forecast *now* of what her expectations *will* be at time T. This will be the same density matrix that she would obtain by solving the differential equation

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H,\rho] - \frac{1}{2}(L^\dagger L\rho + \rho L^\dagger L) + L\rho L^\dagger \,. \tag{18.35}$$

Now, the radical step: Alice decides that it doesn't actually matter to her forecast whether she switches on the clicky detectors or not. She figures that, honestly, it's irrelevant for her purposes, and so *she adopts the reflection state* $\rho(T)$ as her forecast about her future expectations, *regardless* of whether she plans to switch on the detectors. The density matrix $\rho(T)$ is a *decohered* quantum state [145].

19 Canonical Distributions

19.1 Introduction

In an earlier chapter, we introduced the Gibbs states, which depend upon a Hamiltonian $\mathcal{H}$ and an inverse temperature β:

$$\rho = \frac{1}{Z} e^{-\beta \mathcal{H}}. \tag{19.1}$$

We noted that for any such quantum state, $[\rho, \mathcal{H}] = 0$, and so ρ will be constant over time, establishing one property that we would like states of thermal equilibrium to have. The normalization constant Z is surprisingly important for a number that almost seems to be introduced just to keep the equations tidy. It is known as the *partition function*, and we can evaluate it in terms of the eigenvalues E_μ of the Hamiltonian:

$$Z = \sum_\mu e^{-\beta E_\mu}. \tag{19.2}$$

As we saw in the last chapter, we can obtain the expectation value for the energy by the relation

$$\langle E \rangle = -\frac{\partial \log Z}{\partial \beta}. \tag{19.3}$$

It is always worth checking the consistency of units. The partition function Z is unitless, because it is just the sum of bare numbers. β has units of inverse energy, and so differentiating with respect to β brings units of energy, which is just what we need for $\langle E \rangle$.

In thermodynamics, we had the relation

$$E = F + TS, \tag{19.4}$$

which expressed the fact that the total energy of a system is divided into the amount we can extract to do useful work and the part that is tied up in waste heat. We can write the thermodynamic entropy S as

$$S = -\frac{\partial F}{\partial T}, \tag{19.5}$$

and so the total energy E is

$$E = F - T\frac{\partial F}{\partial T}, \tag{19.6}$$

which with a little calculus is the same as saying that

$$E = -T^2 \frac{\partial}{\partial T}\left(\frac{F}{T}\right). \tag{19.7}$$

In the previous chapter, we argued that the parameter β in the definition of the Gibbs state is really $1/(k_B T)$, and so with a little more calculus we can express the total energy E in terms of the β parameter:

$$E = \frac{\partial(\beta F)}{\partial \beta}. \tag{19.8}$$

Comparing this way of writing the E we use in thermodynamics with the expectation value $\langle E \rangle$ we got out of quantum mechanics, we are led to identifying

$$\beta F = -\log Z. \tag{19.9}$$

Or, in terms of the temperature,

$$F = -k_B T \log Z. \tag{19.10}$$

This equation bridges the microscopic worldview, captured in Z, and the macroscopic ideas of thermodynamics, represented by the free energy F. It also says that the free energy is the *cumulant generating function* in disguise, an idea to which we will return later.

Much of what we will do from now on will involve calculating or estimating a partition function Z. In other words, we will be calculating probabilities using the distribution

$$p(\mu) = \frac{1}{Z} e^{-\beta E_\mu}, \tag{19.11}$$

often called *Boltzmann weighting.* This is a formula that can be applied in classical physics, simply by taking it as given, though it does feel more proper to find an "ultimate justification" for it in quantum mechanics, which we generally consider a more fundamental theory.

To see how these calculations go, we start with an example: a *two-level system* with N sites. Each site can either be "on", in which case it has an energy ε, or "off", in which case it has energy 0. The total energy is therefore

$$E = \varepsilon \sum_{i=1}^{N} n_i, \tag{19.12}$$

with each n_i either 1 or 0. We label each possible configuration of the system by a string of 1's and 0's. The partition function is a sum over all such strings:

$$Z = \sum_{\{n_i\}} \exp\left(-\beta\varepsilon \sum_{i=1}^{N} n_i\right). \tag{19.13}$$

Exponentials turn sums into products, so

$$Z = \sum_{\{n_i\}} e^{-\beta\varepsilon n_1} e^{-\beta\varepsilon n_2} \cdots e^{-\beta\varepsilon n_N}. \tag{19.14}$$

Each factor only depends upon one of the n_i, so we can distribute the summation, obtaining

$$Z = (1 + e^{-\beta\varepsilon})^N. \tag{19.15}$$

This is an example of a very general and useful fact: The partition function of a system made out of *independent pieces* is just the *product* of the partition functions of the pieces themselves. Here, each site is independent of all the others, "deciding" to be on or off without regard for what the others are doing, and each site has the same value ε of the energy gap between on and off. Therefore, the total partition function is the partition function for an individual site, raised to the power N.

We find the free energy straightforwardly:

$$F = -Nk_BT \log(1 + e^{-\varepsilon/k_BT}). \tag{19.16}$$

And the total energy is

$$E = -\frac{\partial \log Z}{\partial \beta} = \frac{N\varepsilon e^{-\beta\varepsilon}}{1 + e^{-\beta\varepsilon}}. \tag{19.17}$$

This has units of energy and scales linearly with N, as it should.

Many calculations dating from the early years of quantum physics have this flavor: We consider systems with discrete energy levels, establish a probability distribution that depends upon those energies and deduce the consequences. This just barely taps into the possibilities of quantum theory! For example, a fully quantum treatment of a set of two-level systems would model it as a set of N qubits. For the remainder of this chapter, we will go as far as we can using discretized energy levels and Boltzmann weighting. The results of such calculations can often be taken as "semiclassical" approximations of a fully quantum description, in one limit or another.

19.2 Blackbody Radiation

A problem of considerable practical interest and of great import for the history of physics is the situation of *electromagnetic radiation in thermal equilibrium with its container.* We imagine a box with heated walls that can emit and absorb electromagnetic waves. What is the EM field inside the box doing? We get at the answer by establishing a canonical distribution over the possible ways that the EM field can oscillate. A *mode* of the EM field is labeled by its polarization α and a direction vector

$$\vec{k} = (k_x, k_y, k_z). \tag{19.18}$$

Putting the field in a box of length L on each side discretizes the possible values of $\vec{k}$:

$$\vec{k} = \frac{2\pi}{L}(n_x, n_y, n_z). \tag{19.19}$$

In the limit of large box size L, we can think of a density of allowed vectors,

$$\frac{V}{(2\pi)^3} d^3\vec{k}. \tag{19.20}$$

For each $\vec{k}$, there are two modes with different polarizations, which we could call $\alpha = +$ and $\alpha = -$. Each mode, labeled by $(\vec{k}, \alpha)$, acts like a harmonic oscillator with a characteristic frequency given by

$$\omega(\vec{k}) = ck. \tag{19.21}$$

Thus, we can model the whole system by piling up harmonic oscillators, which we already understand individually.

The expectation value for the total energy is found by summing up that for all the oscillators:

$$\langle E \rangle = \sum_{\vec{k},\alpha} \hbar ck \left(\frac{1}{2} + \frac{e^{-\beta\hbar ck}}{1 - e^{-\beta\hbar ck}} \right). \tag{19.22}$$

The sum over the second term is mathematically trickier, while the sum over the first term is more conceptually problematic. Taking it naïvely, we are adding up a contribution $\hbar ck/2$ for each oscillator, over *infinitely many* oscillators, yielding an infinite result! This is a common problem that arises when we put infinite trust in a theory. Physically, it is not so reasonable to insist that our description of a phenomenon — of the electromagnetic field in a box, let's say — remains good in arbitrarily extreme conditions. For example, very high-frequency vibrations of the field will have wavelengths that are smaller than the spacing between the atoms in the box walls. Should we really use the same boundary conditions for a wavelength of 1 meter and for a wavelength of 1 femtometer? Probably not!

For the time being, we handle this by *dropping* the ground-state contributions. The $\hbar ck/2$ terms do not depend upon the temperature, which is the dependence we really want to get a grasp on, so we bypass all the troubles of the "zero-point energy" by defining

$$E^* = \sum_{\vec{k},\alpha} \hbar ck \frac{e^{-\beta\hbar ck}}{1 - e^{-\beta\hbar ck}} = \sum_{\vec{k},\alpha} \frac{\hbar ck}{e^{\beta\hbar ck} - 1}. \tag{19.23}$$

In the limit of large box size, the step between one valid $\vec{k}$ and another will be very small, so we can turn this sum into an integral:

$$E^* = \frac{2V}{(2\pi)^3} \int d^3\vec{k}\, \frac{\hbar ck}{e^{\beta\hbar ck} - 1}. \tag{19.24}$$

We have picked up a factor of 2 because both polarizations contribute equally. To evaluate the integral, we'd really like to change variables,

$$\xi = \beta\hbar ck. \tag{19.25}$$

But we are integrating over a three-dimensional space, so we need to consider the volume element,

$$d^3\vec{k} = \frac{4\pi\xi^2 d\xi}{(\beta\hbar c)^3}. \tag{19.26}$$

This lets us write our "excitation energy" as a heap of prefactors times an integral that seems like one we can look up:

$$E^* = V\frac{\hbar c}{\pi^2}\left(\frac{k_B T}{\hbar c}\right)^4 \int_0^\infty d\xi \frac{\xi^3}{e^\xi - 1}. \tag{19.27}$$

The integral over ξ can in fact be evaluated in terms of the Riemann zeta function:

$$\int_0^\infty d\xi \frac{\xi^3}{e^\xi - 1} = 6\zeta(4) = \frac{\pi^4}{15}. \tag{19.28}$$

Consequently, the thermal excitation energy per unit volume in the EM field is

$$\frac{E^*}{V} = \frac{\pi^2}{15}\left(\frac{k_B T}{\hbar c}\right)^3 k_B T. \tag{19.29}$$

If we open a little hole in the side of our box, what colors of light will come out? We can calculate this by finding the probability that a mode will be thermally excited as a function of its frequency ω. The result is that the *power* per unit area radiated in a frequency window $d\omega$ is $W(\omega)d\omega$, where

$$W(\omega) = \frac{\hbar}{4\pi^2 c^2}\frac{\omega^3}{e^{\beta\hbar\omega} - 1}. \tag{19.30}$$

This is the *Planck blackbody radiation* law. The curve of $W(\omega)$ has a peak whose location depends upon the Boltzmann energy $k_B T$. Integrating the Planck curve over all frequencies ω gives

$$W_{\text{tot}} = \int_0^\infty d\omega\, W(\omega) = \sigma T^4, \text{ where } \sigma = \frac{\pi^2}{60}\frac{k_B^4}{\hbar^3 c^2}. \tag{19.31}$$

This is the *Stefan–Boltzmann law*.

19.3 The Ising Model

The Ising model is defined by the Hamiltonian

$$\mathcal{H} = -J\sum_{\langle ij\rangle}\sigma_i\sigma_j - h\sum_i \sigma_i, \tag{19.32}$$

where the σ_i are spin variables which can take the values ± 1, J is a coupling constant with units of energy and h is a magnetic field applied from outside. This expresses two effects: First, it is energetically favorable for a spin to be aligned with the external

field, and second, it is energetically favorable for two neighboring spins to be aligned with each other. The sum over $\langle ij \rangle$ is over adjacent spins, and the spins can be arranged with many different patterns of adjacency — a 1D chain, a 2D square grid, anything that is topologically conceivable. Many choices of topology will imply that the system has a *phase transition* between ordered and disordered regimes, a transition which provides a model for how magnets lose their magnetism when heated. In this section, we will obtain a first taste of how one characterizes such a phase transition, by applying a *mean-field approximation* and extracting *critical exponents*. It is a well-nigh ubiquitous theme in the study of phase transitions that quantities of interest vary as some power of the distance to the transition point (where "distance" might be along the temperature axis, for example). A mean-field treatment of the Ising model is a good starting point to appreciate this way of thinking.

First, consider the extreme case in which the spins are independent, that is, $J = 0$. This is an approximation to what happens when the temperature is very high, and the thermal energy can overwhelm the interaction energy between spins. The partition function in this case is

$$Z(J=0) = \prod_{i=1}^{N} \left(e^{\beta h} + e^{-\beta h} \right) = \left[2 \cosh \left(\frac{h}{k_B T} \right) \right]^N . \tag{19.33}$$

The free energy is found by

$$F(J=0) = -k_B T \log Z(J=0), \tag{19.34}$$

and the magnetization is determined by differentiating the free energy density:

$$m = -\frac{1}{N} \frac{\partial F}{\partial h} = \tanh \left(\frac{h}{k_B T} \right) . \tag{19.35}$$

The mean-field approximation consists of pretending that each spin feels the same average magnetic field due to all the other spins. Let us suppose that the spins are arranged in a D-dimensional lattice, so that each spin has $2D$ neighbors: front-back, left-right, up-down, hyper-up and hyper-down, etc. The field experienced by an individual spin is then

$$B = h + 2DJm. \tag{19.36}$$

Self-consistency requires that

$$m = \tanh \left(\frac{h + 2DJm}{k_B T} \right) . \tag{19.37}$$

The *spontaneous* magnetization, that amount of magnetization we can expect without any applied external field h, is

$$m = \tanh \left(\frac{2DJm}{k_B T} \right) . \tag{19.38}$$

If $k_B T > 2DJ$, then the hyperbolic tangent curve rises sublinearly, and the only valid solution to this equation is $m = 0$. When $k_B T < 2DJ$, there are two valid solutions of

equal magnitude, one positive and one negative. The critical temperature is therefore $T_c = 2DJ$. Algebraic manipulations on our formula for m give us

$$m = \tanh\left(\beta h + m\left(\frac{T_c}{T}\right)\right) \tag{19.39}$$

$$= \frac{\tanh(\beta h) + \tanh\left(m\frac{T_c}{T}\right)}{1 + \tanh(\beta h)\tanh\left(m\frac{T_c}{T}\right)}. \tag{19.40}$$

If both h and m are small, which we expect to be a valid approximation when T is near T_c, then we can apply Taylor expansion.

$$\beta h = m\left(1 - \frac{T_c}{T}\right) + m^3\left(\frac{T_c}{T} - \left(\frac{T_c}{T}\right)^2 + \frac{1}{3}\left(\frac{T_c}{T}\right)^3 + \dots\right) + \dots \tag{19.41}$$

The critical exponent which we are most interested in is that which governs the isothermal susceptibility χ:

$$\chi = \left.\frac{\partial m}{\partial h}\right|_T. \tag{19.42}$$

Differentiating the Taylor expansion gives

$$\beta = \chi\left(1 - \frac{T_c}{T}\right) + 3m^2\chi\left(\frac{T_c}{T} - \left(\frac{T_c}{T}\right)^2 + \frac{1}{3}\left(\frac{T_c}{T}\right)^3\right). \tag{19.43}$$

When T is greater than the critical temperature T_c, the system is in the disordered phase, so $m = 0$, and if m vanishes, then

$$\chi = \frac{1}{k_B}\frac{1}{T - T_c} + \dots \tag{19.44}$$

The critical exponent γ is therefore 1. Below the critical temperature, in the ordered phase, we have

$$m = \sqrt{3}\left(\frac{T_c - T}{T}\right)^{1/2}, \tag{19.45}$$

and so it turns out that

$$\chi = \frac{1}{2k_B}\frac{1}{T - T_c} + \dots, \tag{19.46}$$

meaning that the exponent controlling the divergence on the other side of the critical point is also 1.

Exact solutions for systems of interacting components are hard to come by and should be appreciated when they can be found. To that end, we take the time here to solve the one-dimensional Ising model, simplifying matters by working at zero external field ($h = 0$) and putting the spins in a loop ($\sigma_{i+N} = \sigma_i$). Our Hamiltonian becomes

$$\mathcal{H} = -J\sum_{i=1}^{N}\sigma_i\sigma_{i+1}. \tag{19.47}$$

The partition function is a sum over all possible strings $\{\sigma_i\}$:

$$Z = \sum_{\{\sigma_i\}} e^{-\beta \mathcal{H}(\{\sigma_i\})}. \tag{19.48}$$

For brevity, let us write $K = J\beta$ for the ratio of the interaction energy to the Boltzmann thermal energy. Then

$$Z = \sum_{\{\sigma_i\}} \prod_{i=1}^{N} e^{K\sigma_i\sigma_{i+1}}. \tag{19.49}$$

We can evaluate this by turning it into a bit of matrix arithmetic. We introduce the *transfer matrix* T by defining

$$T = \begin{pmatrix} e^K & e^{-K} \\ e^{-K} & e^K \end{pmatrix}. \tag{19.50}$$

Then the partition function becomes simply

$$Z = \mathrm{tr} T^N. \tag{19.51}$$

We find this trace by calculating the eigenvalues of T, which are

$$\lambda_\pm = e^K \pm e^{-K}. \tag{19.52}$$

In terms of the hyperbolic trig functions,

$$\lambda_+ = 2\cosh K, \tag{19.53}$$
$$\lambda_- = 2\sinh K, \tag{19.54}$$

with the former being larger. Consequently,

$$Z = \lambda_+^N + \lambda_-^N, \tag{19.55}$$

which becomes better and better approximated by the first term alone as $N \to \infty$. The free energy per spin, in the limit where we have many spins, is thus

$$\lim_{N\to\infty} \frac{F}{N} = \lim_{N\to\infty} \frac{1}{N}(-k_B T \log \lambda_+^N) = -k_B T \log(2\cosh K). \tag{19.56}$$

This is a slightly exotic-looking function, but it has no singularities or other weird behavior for any $K \geq 0$. Nothing divides the free energy into different regimes of behavior at different temperatures. We deduce that *the one-dimensional Ising model at zero applied field has no phase transition.* This definitely contrasts with the result of the mean-field approximation! What is going on? Generally speaking, mean-field approximations work better *in larger dimensions,* where each spin can "talk" to more other spins, sampling the collection more widely and making the immediate neighborhood more plausibly representative of the whole. The absence of a phase transition is a peculiarity of 1D.

The Ising model in a given dimension is an exemplar of a universality class. Many systems other than magnets have phase transitions that agree quantitatively with the behavior Ising model. The necessary conditions are that the system have two possible choices at each point in space, and that the interactions within the system are short-ranged.

19.4 Statistical Mechanics of a Polymer

A polymer is a big molecule made by sticking together little molecules. In the simplest case, the individual units — the monomers — are the same kind of molecule. We can make a simplified but still decent model for keratin fibers, for example, by saying that we have N monomers, all laid end-to-end. Each monomer can be in either of two configurations, call them α and β. If a monomer is in configuration α, it has a length a and an energy ϵ_a. Likewise, if a monomer is in configuration β, it has length b and energy ϵ_b. We can use the canonical probability distribution to find how the length of this polymer depends upon the temperature.

First, we invoke the canonical distribution:

$$P(\text{configuration}, T) = \frac{1}{Z} \exp\left(-\frac{E(\text{configuration})}{k_B T}\right). \tag{19.57}$$

The partition function for a single monomer is

$$Z_1 = e^{-\epsilon_a/k_B T} + e^{-\epsilon_b/k_B T}, \tag{19.58}$$

and so,

$$P(\alpha, T) = \frac{e^{-\epsilon_a/k_B T}}{e^{-\epsilon_a/k_B T} + e^{-\epsilon_b/k_B T}}, \tag{19.59}$$

$$P(\beta, T) = \frac{e^{-\epsilon_b/k_B T}}{e^{-\epsilon_a/k_B T} + e^{-\epsilon_b/k_B T}}. \tag{19.60}$$

On average, we expect $N_\alpha = P(\alpha, T)N$ monomers to be in state α and $N_\beta = P(\beta, T)N$ monomers to be in state β. The length contributed by all the monomers in state α is aN_α, and likewise for state β. So,

$$l(T) = \left[aP(\alpha, T) + bP(\beta, T)\right] N, \tag{19.61}$$

or

$$l(T) = \left(\frac{ae^{-\epsilon_a/k_B T} + be^{-\epsilon_b/k_B T}}{e^{-\epsilon_a/k_B T} + e^{-\epsilon_b/k_B T}}\right) N. \tag{19.62}$$

At low temperatures, the lower-energy conformation α should be favored. To understand this in more detail, factor out the Boltzmann weight of state α:

$$l(T) = \left(\frac{a + be^{-(\epsilon_b - \epsilon_a)/k_B T}}{1 + e^{-(\epsilon_b - \epsilon_a)/k_B T}}\right) N. \tag{19.63}$$

When $k_B T$ is small compared to the energy difference $\epsilon_b - \epsilon_a$, the argument of the exponentials is strongly negative, and so

$$l(T \to 0) \approx aN. \tag{19.64}$$

By contrast, when $k_B T$ is large, we can use Taylor expansion to write

$$l(T) = \left(\frac{a + b\,(1 - (\epsilon_b - \epsilon_a)/k_B T)}{2 - (\epsilon_b - \epsilon_a)/k_B T} \right) N, \tag{19.65}$$

which as the temperature increases approaches

$$l(T \to \infty) \approx \left(\frac{a+b}{2} \right) N. \tag{19.66}$$

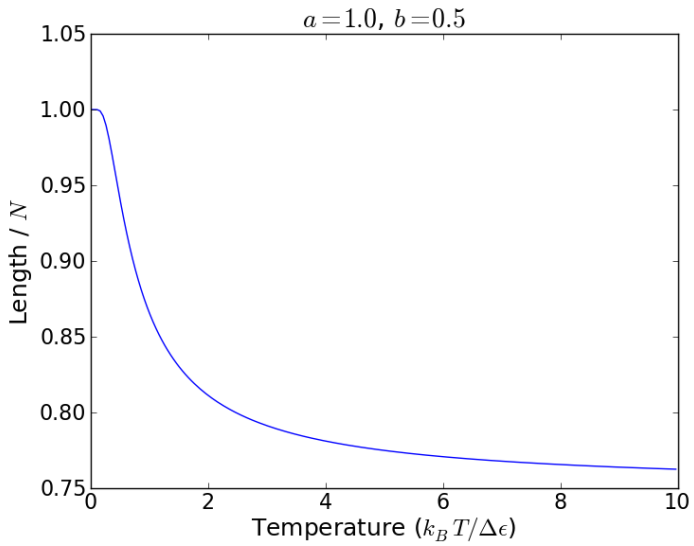


Figure 19.1: Normalized polymer chain length $l(T)/N$ as a function of temperature. Here, $a = 1.0$ and $b = 0.5$. Temperature is measured relative to the energy difference $\Delta\epsilon = \epsilon_b - \epsilon_a$.

Our intuition for thermal equilibrium is a battle between entropy and energy. At high temperatures, entropy wins, and because there are many more ways to have an even mix of α and β than to be all in one conformation, the chain is evenly mixed. At low temperatures, energy wins, and we minimize the free energy by putting all the monomers into the lower-energy state, which is α.

19.5 Biophysics

In this section, we will introduce how statistical physics can intersect with biology at the molecular scale.

To really understand molecules, one has to get into types of bonds, which requires quantum mechanics. Covalent bonds involve the sharing of electrons between atoms, and hydrogen bonds are (comparatively weak, but significant) attractions between a hydrogen on one molecule and a nonhydrogen atom on another. The calculations that show these bonds are possible are quantum-mechanical, and we will not be going into too much depth on those topics. Instead, we will first take up some situations where the key physics is the electrostatic interactions of charged particles, which can be treated classically.

The first question we will address is why ionic compounds dissociate in solution. By themselves, oppositely-charged particles will minimize their potential energy by squeezing together. In thermal equilibrium with a surrounding fluid, however, what matters is the *free* energy. An increase ΔS of the entropy can overpower the change ΔE of the energy, leading to the dissociated state being preferred. So, let us calculate the change in free energy, which will have a term from the binding energy, counterbalanced against a term involving the entropy. We can estimate the latter by calculating the number of positions available for a set of N particles distributed throughout a volume V:

$$\Delta F = \Delta E - T\Delta S = -\mathcal{E}_b + Tk_b \log\left(\frac{V}{N\lambda^3}\right). \tag{19.67}$$

Here, λ denotes the characteristic "thermal wavelength", which characterizes the available volume per particle. We don't yet have a means of calculating this, but as we did in Eq. (5.69), we shall introduce a volume and hope for the best. In equilibrium, we set $\Delta F = 0$ and find the concentration to be

$$\frac{N}{V} = \frac{1}{\lambda^3} e^{-\beta\mathcal{E}_b}. \tag{19.68}$$

We can use Coulomb's law to compute the electrostatic contribution to the binding energy, $\mathcal{E}_c$, given that the dielectric constant of water is ϵ. The electrostatic contribution to the binding energy for a pair of charges is proportional to their product and inversely proportional to their separation:

$$\mathcal{E}_c = \frac{q_1 q_2}{\epsilon r} = \frac{e^2 z^2}{\epsilon r}. \tag{19.69}$$

As always, what matters is how this energy scale compares to the thermal energy $k_B T$:

$$\beta\mathcal{E}_c = \frac{\beta e^2 z^2}{\epsilon r} = z^2 \frac{l_B}{r}, \tag{19.70}$$

where we have defined the *Bjerrum length*

$$l_B := \frac{e^2}{k_B T \epsilon}. \tag{19.71}$$

For water, the dielectric constant ϵ is ≈ 81, and so the Bjerrum length works out to be roughly 7.1 Å. The Bjerrum length gives an indication of how tightly the Coulomb interactions are "screened": When the separation between two unit charges is greater than l_B, we can expect that thermal energy will dominate over electrostatic interactions.

Substances which dissolve to produce ions are called *electrolytes.* We can distinguish between strong electrolytes, like NaCl, NaOH and HCl, which dissociate almost totally, and weak electrolytes which produce few ions. Water itself naturally dissociates into H^+ and OH^- ions; perhaps surprisingly, the fact that water is H_2O means that some tiny fraction of it *won't* be H_2O. At room temperature, the dissociation of water molecules produces about 10^{-7} hydrogen ions per mole, which is why we say that water has a pH of 7.

What sort of electrical environment can we expect to find inside a living cell? First of all, the "A" in "DNA" stands for *acid*: In solution, DNA dissociates to release H^+ ions, meaning that its backbone will be negatively charged. Proteins are *polyampholytes,* being able to develop positively and negatively charged regions both, because some of the amino acids from which proteins are made have side groups that dissociate to release OH^-. Molecules like the DNA backbone, which carry a uniform charge, are known as *polyelectrolytes.* Together, both sorts of macromolecules are referred to as *macroions,* and the small charged particles they give up into the cytoplasm are called *counterions.*

Macroions interact with one another. Proteins bind to DNA, and to each other, and to other stuff floating about within the cell. Can we capture any of this with a partition function? For simplicity, we can begin by holding the macroions at some fixed separation. Because we are not expressly concerned with the counterions, we integrate over the counterion locations:

$$e^{-\beta F} = Z = \int \prod_i d^3 r_i \, e^{-\beta \mathcal{H}_c}. \tag{19.72}$$

Here, the Hamiltonian $\mathcal{H}_c$ must reflect the Coulomb interactions among *all* charges — a horrendously complicated quantity for any real biological system. We will step through this in the manner of a homework problem.

(a) Simplify the problem using a mean-field approximation, wherein the local density of counterion species α is given by an average density, modulated by a Boltzmann weighting factor which depends on the temperature. If the average densities are $\bar{n}_\alpha$ and the effective potential at position $\vec{r}$ is $\phi(\vec{r})$, what is $n_\alpha(\vec{r})$?

(b) Write the Poisson equation for the electric potential in the system, given that the dielectric constant of water is ϵ.

(c) Find the charge density at any point, in terms of the macroions and the counterions.

(d) Impose self-consistency between the solutions to parts (a), (b) and (c) to derive a differential equation for the electric potential in terms of the macroion distribution, the average counterion densities $\bar{n}_\alpha$ and the thermal energy $k_B T = 1/\beta$.

The local density of counterion species α is given by an average density, modulated according to a Boltzmann weighting factor.

$$n_\alpha(\vec{r}) = \bar{n}_\alpha \exp\left(-\beta\phi(\vec{r})z_\alpha e\right). \tag{19.73}$$

The charges, of course, determine the electric potential, which we can compute using the Poisson equation.

$$\nabla^2\phi = -\frac{4\pi}{\epsilon}\rho(\vec{r}). \tag{19.74}$$

The charge density at any point has a contribution from the macroions and from the fluctuating counterion density:

$$\rho(\vec{r}) = \rho_{\text{macroions}}(\vec{r}) + \sum_\alpha z_\alpha e\bar{n}_\alpha e^{-\beta\phi z_\alpha e}. \tag{19.75}$$

Self-consistency requires that

$$\nabla^2\phi = -\frac{4\pi}{\epsilon}\left[\rho_{\text{macroions}}(\vec{r}) + \sum_\alpha z_\alpha e\bar{n}_\alpha e^{-\beta\phi z_\alpha e}\right], \tag{19.76}$$

a relation known as the *Poisson–Boltzmann equation.* Like most partial differential equations, it is only soluble in certain special cases or given some simplifying approximations.

Suppose we have a 2D charged membrane, extending along the yz-plane, which carries a negative charge density $\sigma = -e/d^2$. This is a rough model of a macromolecule bearing unit negative charges a distance d apart. The liquid medium will then carry counterions with $z = 1$. Our next homework problem is the following:

(a) Write the Poisson–Boltzmann equation for this arrangement.

(b) Find how the counterion density depends on the distance from the membrane.

By symmetry, the only significant variations will be along the x direction, so we can simplify Eq. (19.76) to read

$$\frac{d^2\phi}{dx^2} = -\frac{4\pi}{\epsilon}e\bar{n}e^{-\beta e\phi(x)}. \tag{19.77}$$

To make progress, we recall how to solve differential equations. The first standard method is *to recognize the equation as one we have seen before.* In this case, that doesn't work; the equation is unfamiliar, not an exponential decay or a harmonic oscillation. The second standard method is *to guess and check.* Since the equation is not of a familiar form that yields an exponential or a trigonometric function, let us try substituting in a logarithm. To avoid carrying around a lot of constants, let's consider the equation

$$\frac{d^2\phi}{dx^2} = ae^{b\phi(x)}. \tag{19.78}$$

We insert the ansatz

$$\phi(x) = c \log W(x), \tag{19.79}$$

whose derivatives are

$$\phi' = \frac{cW'}{W}, \tag{19.80}$$

$$\phi'' = c\frac{WW'' - (W')^2}{W^2}. \tag{19.81}$$

Substituting and rearranging a little, we find,

$$WW'' - (W')^2 = \frac{a}{c}W^{2+bc}, \tag{19.82}$$

and so we set $c = -2/b$ to simplify the right-hand side. If $W(x)$ is a linear function, then W'' will vanish, and

$$(W')^2 = \frac{ab}{2}. \tag{19.83}$$

So, the original differential equation can be solved by

$$W(x) = W(0) + x\sqrt{\frac{ab}{2}}. \tag{19.84}$$

If we establish the boundary condition that the potential ϕ vanishes at $x = 0$, then we will want $W(0) = 1$. We can re-express the slope of W as a characteristic distance scale x_0:

$$W(x) = 1 + \frac{x}{x_0}. \tag{19.85}$$

Restoring the original constants and solving for $\phi(x)$, we obtain

$$\phi(x) = \frac{2}{\beta e}\log\left[1 + \frac{x}{x_0}\right]. \tag{19.86}$$

The parameter x_0 is a length scale at which the counterion density changes significantly. To investigate what this parameter means, we can go to the $x \ll x_0$ limit, in which Eq. (19.86) becomes

$$\phi(x) \approx \frac{2}{\beta e}\cdot\frac{x}{x_0}. \tag{19.87}$$

We can understand this in the following way: Near the surface, at a distance close enough where screening is unimportant, we expect the electric field will be (using a Gaussian pillbox argument)

$$E = \frac{2\pi\sigma}{\epsilon}, \tag{19.88}$$

implying that the potential is

$$\phi = -\frac{2\pi\sigma}{\epsilon}|x|. \tag{19.89}$$

Now we can check back with Eq. (19.87) and solve for the length scale x_0:

$$x_0 = \frac{\epsilon}{e\sigma\pi\beta}. \tag{19.90}$$

Using the Bjerrum length defined in Eq. (19.71), we can rewrite Eq. (19.90) as

$$x_0 = \frac{d^2}{\pi l_B}. \tag{19.91}$$

This characteristic distance scale is called the *Guoy–Chapman length.* It expresses the thickness of the "diffusive boundary layer" of ions insulating the membrane.

Now, we can calculate how the counterion density falls as a function of distance.

$$n(x) \propto \frac{1}{W^2} \propto \frac{1}{\left(1 + \frac{x}{x_0}\right)^2}, \tag{19.92}$$

which for $x \gg x_0$ behaves as $\frac{1}{x^2}$. This contrasts with the exponential decay we would see in the absence of dielectric: if we had a uniformly charged plate in vacuum, the potential would be a *linear* function of distance. A linear potential inserted into the Boltzmann-weighting function gives an exponentially decaying $n(x)$.

In the previous example, all the screening effects were due to the counterions necessary to neutralize the overall macromolecule charge. Now we would like to consider a situation where the screening is dominated by *salt ions.* That is, we dump in a quantity of salt and let it dissolve, so the dominant contribution to the screening effect is due to the dissociated charges we have introduced. Our procedure is now as follows:

(a) What condition does the sum $\sum_\alpha z_\alpha \bar{n}_\alpha$ satisfy for salt ions in solution?

(b) Linearize the Poisson–Boltzmann equation, and show that

$$\nabla^2\phi = -\frac{4\pi}{\epsilon}\rho_{\text{macro}}(\vec{r}) + \frac{\phi}{\lambda^2}. \tag{19.93}$$

Provide an expression for the characteristic length λ.

(c) What does the potential ϕ become in the limit that the macromolecule can be approximated by a point?

(d) In this limit, what is the total interaction energy of all the macromolecules?

Again, we attempt to linearize the Poisson–Boltzmann equation (19.76).

$$\sum_\alpha z_\alpha \bar{n}_\alpha e^{-\beta z_\alpha e\phi} \approx \sum_\alpha z_\alpha \bar{n}_\alpha \left[1 - \beta z_\alpha e\phi + \cdots\right]. \tag{19.94}$$

For salt ions, $\sum_\alpha z_\alpha \bar{n}_\alpha = 0$. Therefore,

$$\nabla^2\phi = -\frac{4\pi}{\epsilon}\rho_{\text{macro}}(\vec{r}) + \frac{\phi}{\lambda^2}, \tag{19.95}$$

where

$$\lambda^{-2} = \sum_\alpha \frac{4\pi}{\epsilon} \beta e^2 z_\alpha^2 \bar{n}_\alpha = 4\pi l_B \sum_\alpha z_\alpha^2 \bar{n}_\alpha. \tag{19.96}$$

Eq. (19.95) is the *Debye–Hückel* equation, and the parameter λ is the Debye *screening length.*

If our macromolecule can be approximated by a point on this distance scale,

$$\rho_{\text{macro}}(\vec{r}) = Q\delta^3(r)) \tag{19.97}$$

$$\Rightarrow \phi(\vec{r}) = k_B T \frac{l_B}{|\vec{r}|} e^{-\frac{|\vec{r}|}{\lambda}}. \tag{19.98}$$

The total interaction energy of all the macromolecules is therefore given by

$$\beta E_{\text{int}} = \sum_{m<n} \frac{l_B}{|\vec{r}_{mn}|} e^{-\frac{|\vec{r}_{mn}|}{\lambda}}. \tag{19.99}$$

What is the energy cost of bending a gene? There will be some "mechanical" resistance to bending, due to the covalent bonds in the nucleotide chain; in addition, we must consider the electrostatic forces between the charges carried on the DNA helices. The DNA backbone carries unit negative charges separated by $b \approx 0.17$ nm. Bending reduces the distance between these charges, and since the charges all carry the same sign, we have to exert additional energy.

In this example, consider a DNA backbone bent into a curve whose radius of curvature is R. Denote the distance between charges at positions m and n on the *unbent* DNA by $s = b(m - n)$. The distance between charges m and n on the *bent* DNA is d_{mn}.

(a) By how much does bending the polymer reduce the distance between two charges at positions m and n?

(b) How does this change of distance affect the Coulomb energy of the m–n pair? Remember that the electrostatic interaction is screened by the ions in solution.

(c) What is the total energy cost of bending the DNA?

(d) Define the *bending rigidity* κ:

$$E_{\text{bending}} = \frac{\kappa}{2} \int_0^L ds \frac{1}{R(s)^2} = \frac{\kappa}{2} \cdot \frac{L}{R^2}. \tag{19.100}$$

What is the electrostatic contribution to the bending rigidity?

Bending the polymer reduces the distance between charges m and n from the separation $s \equiv b(m - n)$ to the new value

$$d_{mn} = 2R \sin\left(\frac{s}{2R}\right) = 2R\left[\frac{s}{2R} - \frac{1}{6}\left(\frac{s}{2R}\right)^3 + \cdots\right]. \tag{19.101}$$

The change in separation from the unbent value is

$$\delta s = -\frac{1}{24} \cdot \frac{s^3}{R^2}, \tag{19.102}$$

and so the increased Coulomb energy of one pair is

$$\begin{aligned} k_B T l_B \delta \left(\frac{e^{-\frac{s}{\lambda}}}{s} \right) &= k_B T l_B \frac{s\delta e^{-\frac{s}{\lambda}} - e^{-\frac{s}{\lambda}} \delta s}{s^2} \\ &= k_B T l_B \left(\frac{e^{-\frac{s}{\lambda}}}{\lambda} \frac{s^2}{24R^2} + \frac{e^{-\frac{s}{\lambda}} s}{24R^2} \right) \\ &= \frac{1}{24} k_B T l_B \cdot \frac{e^{-\frac{s}{\lambda}}}{R^2} \left(\frac{s^2}{\lambda} + s \right). \end{aligned}$$

Therefore, the *overall* cost of bending can be calculated as

$$\delta E = \frac{1}{24} k_B T l_b \cdot \frac{1}{R^2} \cdot \frac{L}{b} \cdot \int_b^L \frac{ds}{b} e^{-\frac{s}{\lambda}} \left(s + \frac{s^2}{\lambda} \right), \tag{19.103}$$

where we can approximate the integral on the right as

$$\int_b^L \frac{ds}{b} e^{-\frac{s}{\lambda}} \left(s + \frac{s^2}{\lambda} \right) \approx \frac{1}{b} \lambda^2 (1 + s). \tag{19.104}$$

Using this approximation gives a total energy cost of

$$\delta E = \frac{k_B T}{8} \cdot \frac{L l_b \lambda^2}{R^2 b^2}. \tag{19.105}$$

Our work indicates that the electrostatic contribution to the bending rigidity is

$$\kappa_e = \frac{k_B T}{4} \cdot \frac{l_B \lambda^2}{b^2}. \tag{19.106}$$

The overall effective bending rigidity will be a combination of this electrostatic effect and a contribution from the molecular structure,

$$\kappa_{\text{eff}} = \kappa_{\text{base}} + \kappa_e. \tag{19.107}$$

The higher the salt concentration, the shorter the screening length and the lower the electrostatic rigidity — so the more the DNA can bend!

19.6 Grand Canonical Distributions

In thermodynamics, we sometimes treat a system as having energy simply because it contains a certain number of particles. This is helpful when, for example, we are considering reactions between different chemical compounds, where the population size

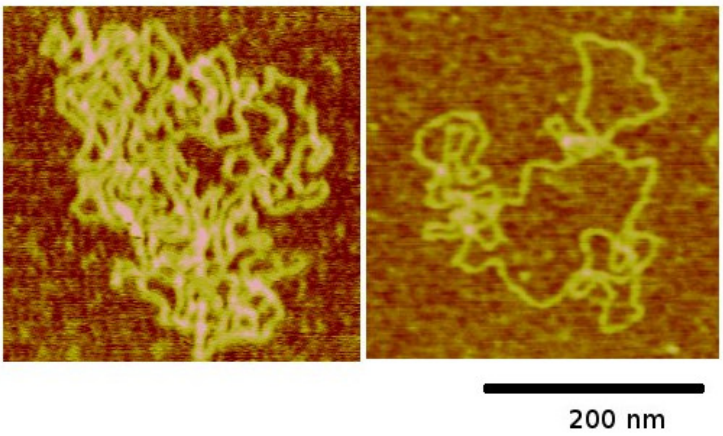


Figure 19.2: Atomic Force Microscope images of DNA deposited on a flat surface. On the left, the ion concentration in the solution was higher, and so the DNA molecules bunched more tightly. Pictures by Davide Venturelli, Cendrine Moskalenko and the author.

of each chemical species can go up or down as that chemical is produced or consumed. Alternatively, we might only have one chemical present in our reaction vessel, but with some of the substance in a different phase than the rest — a cube of ice floating in liquid water, or liquid ethane in the bottom while ethane vapor floats above. We address these situations by introducing the *chemical potential* of a species, which is the amount of energy μ required to introduce one extra unit of that species to the population. This energy counts as part of the total, but it is not easily liberated to do useful work, so we write the free energy as $E-TS-\mu N$. In our computations involving temperature, we should replace E with $E - \mu N$; so, the Boltzmann weighting factor $e^{-\beta E}$ becomes $e^{-\beta E+\beta\mu N}$. This leads us to construct the *grand canonical* partition function

$$Z(\beta, \mu) = \sum_s e^{\beta E_s + \beta \mu N_s}. \tag{19.108}$$

We can rearrange this into a sum over the N-particle partition functions, thusly:

$$Z(\beta, \mu) = \sum_{N=0}^{\infty} e^{\beta \mu N} Z_N(\beta). \tag{19.109}$$

The grand canonical distribution is built by Boltzmann weighting, just like the canonical distribution was, but it has the grand canonical partition function as its normalization constant. It is applicable when particles (or "effective particles") can be created or destroyed. Mathematically speaking, it is useful when we need to eliminate an awkward constraint in a sum. The grand canonical distribution gives us an expectation value for the number of particles, just like the canonical gave us an expectation value

for the energy:

$$\langle N \rangle = \frac{\partial \log Z}{\partial(\beta\mu)} \,. \tag{19.110}$$

Often, one defines the *fugacity*

$$z = e^{\beta\mu} \,, \tag{19.111}$$

in terms of which the grand canonical partition function becomes

$$Z = \sum_{N=0}^{\infty} z^N Z_N \,. \tag{19.112}$$

Recalling §2.7, we see that this has the form of a *generating function* over the Z_N. Because the free energy is $-k_B T \log Z$, the grand canonical partition function is analogous to a generating function for moments, while the free energy is analogous to a generating function for cumulants. This becomes very important when doing perturbation theory.

One example where we can apply the grand canonical distribution is to a *gas of photons*. Take a cubic box of side length L, and suppose that an individual particle has an energy that depends upon its wavenumber:

$$E = E(|\vec{k}|) \,. \tag{19.113}$$

To describe photons, we should say that for a given $\vec{k}$, a particle can come in two different polarizations, or spins. Moreover, the chemical potential μ is zero: We can create photons with arbitrarily small energy.

The free energy is a function of the box size and the temperature:

$$F(L,T) = \int d^3\vec{k}\, f(|\vec{k}|) \,. \tag{19.114}$$

To find this explicitly, we begin by calculating the partition function. This is just the product of single-particle partition functions, with two factors for each $\vec{k}$:

$$Z = \prod_k \left[\frac{1}{1 - e^{-\beta E(k)}} \right]^2 \,. \tag{19.115}$$

The free energy is given by

$$\begin{aligned} F &= -k_B T \log Z && (19.116)\\ &= 2k_B T \sum_{\vec{k}} \log\left(1 - e^{-\beta E(k)}\right) && (19.117)\\ &= 2k_B T \int \left(\frac{L}{2\pi}\right)^3 d^3\vec{k} \log\left(1 - e^{-\beta E(k)}\right) \,. && (19.118) \end{aligned}$$

Therefore,

$$f = \frac{k_B T L^3}{4\pi^3} \log\left(1 - e^{-\beta E(k)}\right) \,. \tag{19.119}$$

We know from thermodynamics that we can find the entropy from the relation

$$S = -\frac{\partial F}{\partial T}\,. \tag{19.120}$$

The average energy is given by a sum over states,

$$\langle E\rangle = \sum_i \frac{E e^{-\beta E_i}}{Z} = -\frac{\partial \log Z}{\partial \beta}, \tag{19.121}$$

which we can rewrite as

$$\langle E\rangle = -\frac{\partial(-\beta F)}{\partial\partial\beta}\,. \tag{19.122}$$

Employing the product rule gives

$$\langle E\rangle = F + \beta\frac{\partial F}{\partial \beta} = F - T\frac{\partial F}{\partial T} = F + TS\,. \tag{19.123}$$

which checks out.

From the early days of quantum theory, we know that we can write a photon's energy as

$$E = \hbar c|\vec{k}|\,. \tag{19.124}$$

With this in hand, the free energy becomes

$$F(L,T) = -\frac{\pi^2}{45}\frac{(k_B T)^4}{(\hbar c)^3}L^3\,. \tag{19.125}$$

Now, let's compress the box, so that its side length changes from L to αL. We do this slowly, and we make sure that the walls are thermally insulating. If the initial temperature is T_i, what is the final temperature T_f?

Because the compression is adiabatic and reversible, $\Delta S = 0$. The initial entropy obeys the proportionality

$$S_i \propto T^3 L^3\,, \tag{19.126}$$

and since the entropy does not change during the compression,

$$T_f^3(\alpha L)^3 = T_i^3 L^3\,. \tag{19.127}$$

Solving for T_f,

$$T_f = \frac{T_i}{\alpha}\,. \tag{19.128}$$

Next, let us suppose that the container is filled with electrons and positrons. We can again begin by calculating the partition function.

$$Z(L,T) = \prod_{\vec{k}}\left[1 + e^{-\beta E_e(|\vec{k}|)}\right]^4\,, \tag{19.129}$$

which implies a free energy of

$$F_e(L,T) = -4k_BT \int \left(\frac{L}{2\pi}\right)^3 d^3\vec{k} \log\left[1 + e^{-\beta E_e(|\vec{k}|)}\right] . \tag{19.130}$$

The form for f_e can be read off:

$$f_e = -\frac{k_BTL^3}{2\pi^3} \log\left[1 + e^{-\beta E_e(|\vec{k}|)}\right] . \tag{19.131}$$

If the thermal energy is very big, $k_BT \gg m_ec^2$, the rest energy can be ignored. In that case,

$$F_e(L,T) = -\frac{7\pi^2}{180}\frac{(k_BT)^4}{(\hbar c)^3}L^3 . \tag{19.132}$$

Let's consider a situation where to begin, the temperature is low ($k_BT_i \ll m_ec^2$), but then the box is compressed until $k_BT_f \gg m_ec^2$. The entropy is still conserved. Initially,

$$S_i = S_\gamma = -\frac{\partial F_\gamma}{\partial T} . \tag{19.133}$$

After the compression,

$$S_f = S_\gamma + S_e = -\frac{\partial}{\partial T}(F_\gamma + F_e) . \tag{19.134}$$

By differentiating the photon-gas free energy, we have that

$$S_i = \frac{4\pi^2}{45}\frac{k_B^4}{(\hbar c)^3}T_i^3L^3 \tag{19.135}$$

and, for the final state,

$$S_f = \frac{4\pi^2}{45}\frac{k_B^4}{(\hbar c)^3}T_f^3(\alpha L)^3 + \frac{28\pi^2}{180}\frac{k_B^4}{(\hbar c)^3}T_f^3(\alpha L)^3 . \tag{19.136}$$

Solving for T_f gives an expression involving lots of fractions:

$$T_f = \frac{T_i}{\alpha}\left(\frac{\frac{4}{45}}{\frac{4}{45} + \frac{28}{180}}\right)^{1/3} . \tag{19.137}$$

We multiply the numerator and denominator by 45/4 to simplify this, obtaining

$$T_f = \frac{T_i}{\alpha}\left(\frac{1}{1 + \frac{7}{4}}\right)^{1/3} = \frac{T_i}{\alpha}\left(\frac{4}{11}\right)^{1/3} . \tag{19.138}$$

20 Nonequilibrium Processes

In this chapter, we will further explore processes that go outside the regime of thermal equilibrium.

20.1 Rényi Entropy and Free Energy

When we developed the idea of the Helmholtz free energy F, we considered slow isothermal processes. What happens if, instead, the changes applied to the system are abrupt?

Imagine two independent random variables described by the same probability distribution. What is the probability that they both yield the same value? The probability that both result in the value i is p_i^2, and so the total probability of "rolling doubles" regardless of the specific value is $\sum_i p_i^2$. Turning this "collision" or "coincidence" probability into a surprisal gives an entropy-like expression:

$$R_2(p) = -\log \sum_i p_i^2 \,. \tag{20.1}$$

This is the *Rényi entropy* of order 2, which is the prototype for the general expression for the Rényi entropies:

$$R_q(p) = \frac{1}{1-q} \log \sum_i p_i^q \,. \tag{20.2}$$

The prefactor diverges as $q \to 1$, but using l'Hôpital's rule, we can show that R_1 is really just the Shannon entropy:

$$R_1(p) = -\sum_i p_i \log p_i \,. \tag{20.3}$$

We can find a thermodynamic meaning for the Rényi entropies following an argument due, in various portions, to Beck, Schlögl, Baez, Polettini and Downes [146].

Let p_T be the canonical probability distribution at temperature T for a system with some set of energy levels:

$$p_T(i) = \frac{1}{Z} e^{-E_i/k_B T}, \;\; Z = \sum_i e^{-E_i/k_B T} \,. \tag{20.4}$$

To make the Rényi entropy of order q thermodynamic, we multiply it by Boltzmann's constant:

$$S_q(T) = k_B R_q(p_T) = \frac{k_B}{1-q} \log \sum_i p_T(i)^q \,. \tag{20.5}$$

Plugging in the canonical form,

$$S_q(T) = \frac{k_B}{1-q} \log \sum_i \frac{e^{-qE_i/k_BT}}{Z^q} \,. \tag{20.6}$$

We can pull the Z^q out of the sum, which leaves an expression that is just the partition function for the same system at a different temperature, T/q:

$$S_q(T) = \frac{k_B}{1-q} \log \frac{Z(T/q)}{Z(T)^q} \,. \tag{20.7}$$

Let's call the new temperature T', so that

$$q = \frac{T}{T'} \,, \tag{20.8}$$

and then

$$S_{T/T'}(T) = \frac{k_B}{1-(T/T')} [\log Z(T') - (T/T') \log Z(T)] \,. \tag{20.9}$$

Multiplying by 1 in the form T'/T', we obtain

$$S_{T/T'}(T) = k_B \frac{T' \log Z(T') - T \log Z(T)}{T' - T} \,. \tag{20.10}$$

But k_B times the temperature times the log of the partition function is minus the free energy. So,

$$S_{T/T'}(T) = -\frac{F' - F}{T' - T} \,. \tag{20.11}$$

The Rényi entropy of order q tells us the maximum amount of work that a system can do if we *quench* its temperature by a factor of q.

Suppose $\{p_i\}$ is a canonical distribution at temperature T:

$$p_i = \frac{1}{Z} e^{-E_i/k_BT} \,, \tag{20.12}$$

and let $\{p'_i\}$ be any other probability distribution over the same set of microstates. The expectation value for the energy given the latter distribution is

$$E' = \sum_i p'_i E_i \,, \tag{20.13}$$

and if we define a free energy using the Shannon entropy, we find

$$F' = E' - TS' = -k_BT \sum_i p'_i \log(Zp_i) + k_BT \sum_i p'_i \log p'_i \,. \tag{20.14}$$

This simplifies to

$$F' = -k_BT \log Z + k_BT \sum_i p'_i \log \frac{p'_i}{p_i} \,. \tag{20.15}$$

The first term is the equilibrium free energy of the original canonical distribution, and the second term is proportional to the relative entropy between $\{p_i\}$ and $\{p'_i\}$.

20.2 Relaxation from Impulsive Work

What if it is not the temperature of a system that we change abruptly, but instead the mechanical condition of it? For example, let us imagine a system in contact with a reservoir at coolness β. We suddenly jolt the system in some way, or we snap off some constraint upon its motion. The system generates or receives *mechanical work,* but very quickly: so quickly that heat does not have time to flow. Then, gradually, the reservoir does its job, and the system re-equilibrates to the coolness β, but having changed in the interval. How can we characterize this kind of process, and in particular, what can we say about the work flow?

Before the re-equilibration, the concepts of equilibrium thermodynamics won't all be applicable, so we turn to the statistical perspective and start thinking about energy levels. If our system is an Ising magnet, for instance, we know that we can do work to it by applying a magnetic field. Changing the strength of the external field changes the energy value associated with each microstate. On the other hand, for any fixed value of the field, pouring heat into the magnet doesn't change the energy levels, only which ones are likely to be occupied. So, we have a provisional definition of *work* in microscopic terms: Work is the energy change attributable to shifting the energy levels themselves by altering macroscopic parameters. Let us emphasize that this is a *provisional* definition, since we cannot be guaranteed that macroscopic notions map to microscopic ones perfectly cleanly. But it seems a good enough idea to adopt for the time being.

At the very beginning and the very end of our experimental protocol, we can give an equilibrium description. The coolness will be β at both times, the initial free energy will be some value F and the final free energy some value F'. Everything we can calculate we get from the partitition function, which initially is Z and at the end is Z'. Let's compare these by dividing them:

$$\frac{Z'}{Z} = e^{-\beta(F'-F)} \,. \tag{20.16}$$

The initial probability distribution has the canonical form

$$p(i) = \frac{1}{Z} e^{-\beta E_i} \,. \tag{20.17}$$

Writing the final energy levels as $\{E'_j\}$, we have

$$e^{-\beta(F'-F)} = \frac{1}{Z} \sum_j e^{-\beta E'_j} \,. \tag{20.18}$$

Now we do a sneaky thing: Let $p(j)$ denote, *not* the final canonical distribution, but the distribution over the microstates just when the work stage of the protocol is done. The joint probability $p(i,j)$ encodes the dynamics that happen during the work stage. We know on general grounds that

$$p(i,j) = p(j|i)p(i) \,, \tag{20.19}$$

and by normalization,

$$\sum_i p(j|i) = 1\,. \tag{20.20}$$

We have an expression that equals 1, so we can insert it:

$$e^{-\beta(F'-F)} = \frac{1}{Z}\sum_j e^{-\beta E'_j}\sum_i p(j|i)\,. \tag{20.21}$$

Next, we turn the conditional probability into a joint probability:

$$e^{-\beta(F'-F)} = \frac{1}{Z}\sum_j e^{-\beta E'_j}\sum_i \frac{p(i,j)}{p(i)}\,. \tag{20.22}$$

The partition function lurking out front can come into the sum and cancel against that $p(i)$, leaving us with

$$e^{-\beta(F'-F)} = \sum_{i,j} e^{-\beta E'_j} e^{\beta E_i} p(i,j)\,. \tag{20.23}$$

The expression on the right-hand side now has the form of an expectation value:

$$e^{-\beta(F'-F)} = \left\langle e^{-\beta(E'_j - E_i)} \right\rangle, \tag{20.24}$$

where the average is taken over the initial microstates and the microstates at the end of the work stage. But by assumption, *no heat flows* during that part of the protocol, so the energy change $E'_j - E_i$ must be the *work* required to go from i to j. We thus have a relationship between the change in free energy after thermalization and the work flow before thermalization:

$$e^{-\beta(F'-F)} = \left\langle e^{-\beta W} \right\rangle, \tag{20.25}$$

This is the *Jarzynski equality*. Our path to it was inspired by an argument of Page and Sorkin, though arranged differently to clarify the assumptions that in many discussions of this topic have been quite opaque to me.

Because the Jarzynski equality involves the expected value of an exponential function of the work W, we can apply the linked-cluster theorem, Eq. (2.146):

$$\left\langle e^{-\beta W} \right\rangle = \sum_n \frac{\langle(-\beta W)^n\rangle}{n!} = \sum_n \frac{(-\beta)^n}{n!}\langle W^n\rangle = \exp\left(\sum_n \frac{(-\beta)^n}{n!}\,\langle W^n\rangle_c\right). \tag{20.26}$$

We can relate this to the problem of Brownian motion (§7.2) by considering a body floating in viscous goop. We imagine dragging the body with a force f for a time Δt. This requires an amount of work $W = f\Delta x$, which is a random variable because Δx is not fixed. Supposing that the medium is sufficiently viscous that the body reaches its terminal velocity f/b pretty much immediately, then the *expected* displacement will be

$$\langle \Delta x\rangle = \frac{f}{b}\Delta t\,. \tag{20.27}$$

But the actual displacement could be a little more or a little less than this, because of random jitter. Even in the absence of an applied force, the body will migrate diffusively, meaning that the variance of the displacement will grow as the elapsed time:

$$\left\langle (\Delta x)^2 \right\rangle_c = 2D\Delta t \,. \tag{20.28}$$

So, the mean work is

$$\langle W \rangle = \frac{f^2}{b} \Delta t \,, \tag{20.29}$$

while the variance of W is

$$\left\langle W^2 \right\rangle_c = f^2 \left\langle (\Delta x)^2 \right\rangle_c = 2Df^2 \Delta t \,. \tag{20.30}$$

In this case, the free energy does not change, and moreover it is reasonable to characterize the probability distribution for W using its first two cumulants. Together, these observations imply that

$$\exp\left(-\beta \left\langle W \right\rangle_c + \frac{\beta^2}{2} \left\langle W^2 \right\rangle_c \right) = 1 \,, \tag{20.31}$$

and consequently,

$$-\beta \langle W \rangle + \frac{\beta^2}{2} \left\langle W^2 \right\rangle_c = 0 \,. \tag{20.32}$$

We therefore have a relation between the mean and the variance:

$$\left\langle W^2 \right\rangle_c = \frac{2}{\beta} \langle W \rangle = 2k_B T \langle W \rangle \,. \tag{20.33}$$

Applying what we said about these a moment ago,

$$2Df^2 \Delta t = \frac{2k_B T f^2}{b} \Delta t \,. \tag{20.34}$$

The details of the experimental protocol — the magnitude and duration of the applied force — drop out, leaving us with

$$bD = k_B T \,, \tag{20.35}$$

the fluctuation-dissipation relation we saw from another perspective in §7.2.

If we express our mean-variance relation in terms of the displacement, rather than the work, we get

$$\langle \Delta x \rangle = \frac{f}{2k_B T} \left\langle (\Delta x)^2 \right\rangle_c \,. \tag{20.36}$$

Or,

$$\langle \Delta x \rangle = \frac{f}{2k_B T} [\left\langle (\Delta x)^2 \right\rangle - \langle \Delta x \rangle^2] \,. \tag{20.37}$$

This points to a more general statement about fluctuation-dissipation theorems: They relate a *response function* of a system to a measure of its correlation. Further developing this line of thought leads into the subject of *linear response theory* [147].

20.3 Comparing Forward and Reverse Processes

The next level of sophistication is to compare two different out-of-equilibrium processes. The first, which we'll call the *forward* protocol, is what we studied in the previous section. The second, the *reverse* protocol, begins with the system nicely thermalized in whatever equilibrium it eventually settles to after the forward protocol, and then applies the opposite sudden jolt — suddenly turning all the knobs the same amount as in the forward protocol, but in the other direction. If the forward protocol requires some amount of work input, we figure that the reverse protocol will generate an amount of work output. How might these quantities be related?

The event that the forward protocol requires a work input of W can in principle happen in multiple ways, i.e., by any transition from an E_i to an E'_j for which the difference equals W. So, we get the total probability by adding up these contributions:

$$p_f(\text{require } W) = \sum_{i,j} p_f(i,j)(E'_j - E_i)\delta(E'_j - E_i - W)\,. \tag{20.38}$$

In the reverse protocol, the same logic applies, only the initial microstate is some j and the post-jolt one is some i, and we have to flip a couple signs to agree with our changed perspective:

$$p_r(\text{generate } W) = \sum_{i,j} p_r(i,j)(-(E_i - E'_j))\delta(-(E_i - E'_j) - W)\,. \tag{20.39}$$

Simplifying,

$$p_r(\text{generate } W) = \sum_{i,j} p_r(i,j)(E'_j - E_i)\delta(E'_j - E_i - W)\,. \tag{20.40}$$

By the general rules of conditional probability and the assumption that the forward protocol begins with a thermalized system,

$$p_f(i,j) = p_f(j|i)p_f(i) = p_f(j|i)\frac{e^{-\beta E_i}}{Z}\,. \tag{20.41}$$

Likewise, the reverse protocol assumes a thermalized system to start with:

$$p_r(i,j) = p_r(i|j)p_r(j) = p_r(i|j)\frac{e^{-\beta E'_j}}{Z'}\,. \tag{20.42}$$

We can connect these expressions by positing that the probability of going from i to j under the forward protocol is the same as the probability of going from j to i under the reverse protocol. Maybe some weird situation wouldn't satisfy this, but it does sound like what reversing the dynamics ought to mean. We therefore presume that

$$p_f(j|i) = p_r(i|j)\,. \tag{20.43}$$

So, we have

$$p_f(i,j)Ze^{\beta E_i} = p_r(i,j)Z'e^{\beta E'_j}\,, \tag{20.44}$$

which we can solve to yield

$$p_f(i,j) = p_r(i,j)e^{\beta(E'_j - E_i)}\frac{Z'}{Z} \tag{20.45}$$

$$= p_r(i,j)e^{\beta(E'_j - E_i)}e^{-\beta(F'-F)}\,. \tag{20.46}$$

Substituting this back into our expression for the probability of the forward protocol requiring a particular amount of work, we find

$$p_f(\text{require } W) = \sum_{i,j} p_r(i,j)(E'_j - E_i)e^{\beta(E'_j - E_i)}e^{-\beta(F'-F)}\delta(E'_j - E_i - W)\,. \tag{20.47}$$

The free energy is a constant, so we can pull it out of the sum, and the delta function constrains $E'_j - E_i$ to equal W, meaning that

$$p_f(\text{require } W) = p_r(\text{generate } W)e^{\beta W}e^{-\beta\Delta F}\,. \tag{20.48}$$

This is the *Crooks fluctuation theorem.*

Moving the $e^{\beta W}$ factor to the other side and integrating over W brings us back to the Jarzynski equality:

$$\int dW\, p_f(\text{require } W)e^{-\beta W} = e^{-\beta\Delta F}\,, \tag{20.49}$$

since the integral over the other probability distribution must equal unity by normalization.

20.4 Driven Diffusive Flow

We can get further out of the equilibrium mindset by considering a scenario in which there is an ongoing stochastic flow into our system [148, 149, 150, 151]. Imagine, for example, a nutrient spreading through a population of bacteria. The nutrient may arrive irregularly over time and be concentrated more in some places than others; it spreads diffusively and is consumed at a rate that depends upon the local population density of bacteria. In this section, we consider diffusion with sources and sinks. After employing Fourier transform techniques to solve the case of diffusion driven by deterministic deposition, we will turn to a scenario in which the input is stochastic.

We begin with an inhomogeneous diffusion equation:

$$\partial_x^2 u - \alpha^2\partial_t u = f(x,t)\,. \tag{20.50}$$

Here, $u = u(x,t)$ denotes the concentration of a substance which spreads diffusively and is being pumped into the environment in a way described by the source term $f(x,t)$. We take our boundary conditions to be

$$u(0,t) = 0,\ u(L,t) = 0,\ u(x,0) = \phi(x)\,. \tag{20.51}$$

To solve Eq. (20.50), we first consider the simpler problem of the unforced case, described by the homogeneous PDE

$$\partial_x^2 u - \alpha^2 \partial_t u = 0\,. \tag{20.52}$$

We attack Eq. (20.52) in a standard way by trying separation of variables:

$$u(x,t) = X(x)T(t)\,. \tag{20.53}$$

Substituting this expression for u into Eq. (20.52) yields

$$X''(x)T(t) - \alpha^2 X(x)\dot{T}(t)\,. \tag{20.54}$$

We note that by dimensional analysis, the units of α^2 must be time over length squared. We can rearrange the previous equation to

$$\frac{X''}{X} = \frac{\alpha^2 \dot{T}}{T}\,. \tag{20.55}$$

In this equation, each term depends on one and only one variable, x on the left and t on the right. Changing the value of the time t would cause the right-hand side of the equation to change all by itself, which is nonsense. Therefore, both sides of this equation must equal a constant, which for notational convenience we write $-\lambda^2$ (implying that λ has units of inverse length). Then

$$\begin{aligned} X'' &= -\lambda^2 X, \\ \dot{T} &= -\frac{\lambda^2}{\alpha^2} T\,. \end{aligned} \tag{20.56}$$

These equations readily admit the solutions

$$\begin{aligned} X(x) &= A\sin(\lambda x) + B\cos(\lambda x)\,, \\ T(t) &= C\exp\left(-\frac{\lambda^2}{\alpha^2}t\right)\,. \end{aligned} \tag{20.57}$$

Combining these, and absorbing the constant C by redefining A and B appropriately, gives

$$u(x,t;\lambda) = \exp\left(-\frac{\lambda^2}{\alpha^2}t\right)\left[A\sin(\lambda x) + B\cos(\lambda x)\right]. \tag{20.58}$$

To make further progress, we consider the boundary conditions. Because $u(0,t) = 0$ for all times t, we can say that $B = 0$. Then, because $u(L,t) = 0$, we know that $\sin(\lambda L) = 0$, which we can achieve by fixing λ to be one of the set

$$\lambda_n = \frac{n\pi}{L}\,. \tag{20.59}$$

Putting these deductions together, we can say that

$$u(x,t;n) = A_n \exp\left[-\left(\frac{n\pi}{L\alpha}\right)^2 t\right]\sin\left(\frac{n\pi}{L}x\right)\,. \tag{20.60}$$

We calculate the coefficients A_n from the initial condition that the concentration as a function of position is given by $\phi(x)$.

$$u(x,0) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{L}x\right) = \phi(x)\,. \tag{20.61}$$

This is nothing else but a Fourier series. Thus,

$$A_n = \frac{2}{L}\int_0^L dx'\,\phi(x')\sin\left(\frac{n\pi}{L}x'\right)\,. \tag{20.62}$$

With the unforced diffusion problem basically squared away, we tackle the case with the forcing term $f(x,t)$. Using our experience with the unforced scenario as a guide, we try an Ansatz for $u(x,t)$ of the form

$$u(x,t) = \sum_{n=1}^{\infty} T_n(t)\sin\left(\frac{n\pi}{L}x\right)\,, \tag{20.63}$$

and we expand the forcing term as a Fourier series,

$$f(x,t) = \sum_{n=1}^{\infty} f_n(t)\sin\left(\frac{n\pi}{L}x\right)\,. \tag{20.64}$$

The diffusion equation (20.50) must hold for each term in the sums over n. Taking derivatives with respect to position pulls out factors of $(n\pi/L)$, and the omnipresent sine factors cancel out, leaving

$$\alpha^2\dot{T}_n(t) + \left(\frac{n\pi}{L}\right)^2 T_n(t) = f_n(t)\,. \tag{20.65}$$

This is a first-order, ordinary differential equation, albeit inhomogeneous. We can solve it by multiplying throughout by the integrating function $\exp\left[\left(\frac{n\pi}{\alpha L}\right)^2 t\right]$. Then,

$$\alpha^2\dot{T}_n(t)\exp\left[\left(\frac{n\pi}{\alpha L}\right)^2 t\right] + \left(\frac{n\pi}{L}\right)^2 T_n(t)\exp\left[\left(\frac{n\pi}{\alpha L}\right)^2 t\right] = f_n(t)\exp\left[\left(\frac{n\pi}{\alpha L}\right)^2 t\right]\,. \tag{20.66}$$

The left-hand side of this equation can be rewritten using the product rule for derivatives, so

$$\alpha^2\partial_t\left[\exp\left[\left(\frac{n\pi}{\alpha L}\right)^2 t\right] T_n(t)\right] = f_n(t)\exp\left[\left(\frac{n\pi}{\alpha L}\right)^2 t\right]\,. \tag{20.67}$$

The next step is just the Fundamental Theorem of Calculus and a bit of algebraic rearrangement:

$$T_n(t) = \frac{\exp\left[-\left(\frac{n\pi}{\alpha L}\right)^2 t\right]}{\alpha^2}\int_0^t dt'\,f_n(t')\exp\left[\left(\frac{n\pi}{\alpha L}\right)^2 t'\right] + A_n\exp\left[-\left(\frac{n\pi}{\alpha L}\right)^2 t\right]\,. \tag{20.68}$$

We incorporate the initial condition by noting that $A_n = \phi_n$, where the Fourier coefficients ϕ_n are defined by

$$\phi(x) = \sum_{n=1}^{\infty} \phi_n \sin\left(\frac{n\pi}{L}x\right) . \tag{20.69}$$

Let us neglect the A_1 term on the presumption that the initial conditions are just too boring when compared with the input signal. Suppose that the input is well-described by its first Fourier coefficient,

$$f_1(t') = \frac{2}{L}\int_0^L dx'\, f(x',t') \sin\left(\frac{\pi x'}{L}\right) . \tag{20.70}$$

Discarding everything we have discarded, then,

$$u(x,t) \approx \frac{2\exp\left[-\left(\frac{\pi}{\alpha L}\right)^2 t\right]}{\alpha^2 L} \sin\left(\frac{\pi x}{L}\right) \int_0^t dt'\, \exp\left[\left(\frac{\pi}{\alpha L}\right)^2 t'\right] \int_0^L dx'\, f(x',t') \sin\left(\frac{\pi x'}{L}\right) . \tag{20.71}$$

Units check: By integrating over x', we pick up a dimension of length, which cancels with the units of the L^{-1} in the prefactor. By integrating over t', we pick up a dimension of time, which cancels against the α^{-2} in the prefactor. Overall, dimensional analysis indicates that $u(x,t)$ must have units of length squared times the units of $f(x,t)$, which checks with the convention established in Eq. (20.50).

We now turn to the case where the input is *stochastic*. That is, we consider a space- and time-dependent concentration $u(\vec{x},t)$ which varies in accordance with the equation

$$\partial_t u(\vec{x},t) = \frac{1}{\alpha^2}\nabla^2 u + s(\vec{x},t) . \tag{20.72}$$

Here, the "rainfall" function $s(\vec{x},t)$ is taken to vary from one realization of the experiment to another; we characterize it by ensemble averages. Because we've already talked about what happens with a constant input, we assume that the noise has zero mean:

$$\langle s(\vec{x},t)\rangle = 0 . \tag{20.73}$$

"White noise" implies a flat power spectrum, which by Wiener–Khinchin in turn tells us that the autocorrelation of $s(\vec{x},t)$ is a delta function, or more precisely, a product of delta functions in the time and space variables:

$$\langle s(\vec{x},t)s(\vec{x}',t')\rangle = 2D\delta^d(\vec{x}-\vec{x}')\delta(t-t') . \tag{20.74}$$

We repeat the steps we followed before, but in d dimensions. The result is the analogue of Eq. (20.68). Starting from flat initial conditions,

$$u(\vec{x},t) = \int \frac{d^d\vec{q}}{(2\pi)^d} e^{-i\vec{q}\cdot\vec{x}} \int_0^t dt'\, \exp\left[-\frac{1}{\alpha^2}q^2(t-t')\right] s(\vec{q},t) . \tag{20.75}$$

This implies we need to know about the behavior of the noise $s(\vec{x},t)$ in the Fourier-transformed basis. The ensemble mean has a familiar form,

$$\langle s(\vec{q},t)\rangle = 0\,, \tag{20.76}$$

as does the ensemble correlation,

$$\langle s(\vec{q},t)s(\vec{q}',t')\rangle = 2D(2\pi)^d\delta(t-t')\delta^d(\vec{q}+\vec{q}')\,. \tag{20.77}$$

The change in sign within the delta function merits a brief comment. We find the correlation for the Fourier-transformed noise function in terms of the original correlator:

$$\begin{aligned}\langle s(\vec{q},t)s(\vec{q}',t')\rangle &= \int d^d\vec{x}d^d\vec{x}'\,e^{i\vec{q}\cdot\vec{x}+i\vec{q}'\cdot\vec{x}'}\langle s(\vec{x},t)s(\vec{x}',t')\rangle \\ &= \int d^d\vec{x}d^d\vec{x}'\,e^{i\vec{q}\cdot\vec{x}+i\vec{q}'\cdot\vec{x}'}2D\delta^d(\vec{x}-\vec{x}')\delta(t-t') \\ &= 2D\delta(t-t')\int d^d\vec{x}\,e^{i\vec{x}\cdot(\vec{q}+\vec{q}')} \\ &= 2D\delta(t-t')(2\pi)^d\delta^d(\vec{q}+\vec{q}')\,. \end{aligned} \tag{20.78}$$

One important quantity of interest is the expected width of the concentration distribution, which we can define by

$$w^2(t,L) \equiv \frac{1}{L^d}\int d^d\vec{x}\,\big\langle u(\vec{x},t)^2\big\rangle\,, \tag{20.79}$$

or in Fourier space,

$$w^2(t,L) = \frac{1}{L^d}\int \frac{d^d\vec{q}}{(2\pi)^d}\big\langle |u(\vec{q},t)|^2\big\rangle\,. \tag{20.80}$$

If we introduce the decay rate

$$\gamma(\vec{q}) \equiv \frac{1}{\alpha^2}q^2\,, \tag{20.81}$$

then we can rewrite our expression for $w^2(t,L)$ as

$$w^2(t,L) = \int \frac{d^d\vec{q}}{(2\pi)^d}\frac{D}{\gamma(\vec{q})}\left(1-e^{-2\gamma(\vec{q})t}\right)\,. \tag{20.82}$$

The behavior observed in this system depends on the timescale of observation. Two timescales are significant: $t_{\rm min}$, the characteristic time for transport across the shortest distance at which the diffusion equation is a viable approximation, and $t_{\rm max}$, the timescale for transport across the system as a whole. The former depends on the distance a, the scale below which a coarse-grained model like the diffusion equation breaks down; the latter depends upon L, the overall system size. For $t \ll t_{\rm min}$,

$$w^2(t,L) = \int \frac{d^d\vec{q}}{(2\pi)^d}\frac{D}{\gamma(\vec{q})}2\gamma(\vec{q})t = \frac{2Dt}{a^d}\,. \tag{20.83}$$

In the opposite extreme, when $t \gg t_{\max}$,

$$w^2(t,L) = \int \frac{d^d\vec{q}}{(2\pi)^d} \frac{D\alpha^2}{q^2} \,. \tag{20.84}$$

We find that

$$w^2(t,L) \propto D\alpha^2 \begin{cases} a^{2-d} & \text{for } d > 2 \quad (\chi = 0) \\ \log(L/a) & \text{for } d = 2 \quad (\chi = 0^+) \\ L^{2-d} & \text{for } d < 2 \quad (\chi = \frac{2-d}{2}) \,. \end{cases} \tag{20.85}$$

We have here defined the *roughness exponent* χ which gives the power-law dependence of the divergence width on the system size L:

$$\lim_{t\to\infty} w(t,L) \propto L^\chi \,. \tag{20.86}$$

In the intermediate regime, $t_{\min} \ll t \ll t_{\max}$,

$$w^2(t,L) \propto D\alpha^2 \int dq\, q^{d-3} \left(1 - e^{-2q^2 t/\alpha^2}\right) \,, \tag{20.87}$$

or

$$w^2(t,L) \propto D\alpha^2 \left(\frac{\alpha^2}{t}\right)^{\frac{d-2}{2}} \int_{t/t_{\max}}^{t/t_{\min}} dy\, y^{\frac{d-4}{2}} \left(1 - e^{-2y}\right) \,. \tag{20.88}$$

When $d < 2$, this integral is convergent, and when $d \geq 2$, it is dominated by its upper limit.

At short times, we have

$$\lim_{t\to 0} w(t,L) \propto t^\beta \,, \tag{20.89}$$

and

$$w^2(t,L) \propto \begin{cases} D\alpha^2 a^{2-d} & \text{for } d > 2 \quad (\beta = 0) \\ D\alpha^2 \log(t/t_{\min}) & \text{for } d = 2 \quad (\beta = 0^+) \\ D\alpha^d t^{(2-d)/2} & \text{for } d < 2 \quad (\beta = \frac{2-d}{4}) \,. \end{cases} \tag{20.90}$$

20.5 Directed Percolation

Another stochastic process that lives outside the thermal-equilibrium mindset is *directed percolation.* As we discussed back in §4.3, percolation is the study of how channels or connections form in randomized media. Previously, we covered isotropic percolation, in which the structure formation happens without any preferred direction. In directed percolation, we fix a "downhill" direction in which flow can occur. We can also think of this as demanding that the process develops over time, rather than considering only a frozen end result.

Directed percolation is an idealized model of fluid flow through a porous medium. We can begin with the mental image of a regular lattice of channels, some of whose junction points were blocked. If the fraction of blocked points was too large, fluid

flowing through the lattice from top to bottom would always be stymied, but if the blockages were sparse, the fluid could percolate downwards through the channels. We can treat this model in any number of dimensions; the crucial point is that there is a preferred direction. And because the fluid only spreads downhill, there is another picture available. We can also think of the tip of each stream as a corpuscle executing a *random walk*. These walkers can merge, if two fluid streams come together at a common point, and they can reproduce, which happens when a fluid stream splits and its offshoots progress along two different channels. The walkers can also perish: This corresponds to a stream which encounters a point where no further propagation is possible.

In short, directed percolation in a $(d+1)$-dimensional lattice, with one dimension singled out as the axis of gravity, can equally well be thought of as *merging, reproducing and perishing random walks* in a d-dimensional space. Given a value of the dimension d, there exists a directed percolation universality class. The hallmarks of the DP universality class in a chosen dimension are the *critical exponents* that describe how the dynamics evolve over time and how they depend upon the "knobs" of the model. As with isotropic percolation, one can do a renormalization-group analysis of DP. The task is considerably more arduous than the scaling study we did back in §4.3; it involves bringing over ideas from field theory and requires considerable effort to obtain even approximate answers. The happier news is that because the critical exponents are the same for all systems belonging to a universality class, a discovery about one such system can be translated over to all the others. Experience in nonequilibrium statistical mechanics, codified in a conjecture of Grassberger and Janssen [152, 153], suggests that the main requirement for a model to belong to a DP universality class is the existence of an absorbing state, a configuration (like an empty lattice) that fluctuations can take the system into but not out of.

It is interesting to follow the history of the directed-percolation concept. It was first proposed in 1957 [154]. The mathematics necessary to treat it cleverly was invented (or, rather, adapted from a different area of physics) in the 1970s, and then forgotten, and then rediscovered by somebody else [155]. Connections with other subjects were made. Experiments were carried out on systems which *almost* behaved like the idealization, but always turned out to differ in some way... until 2007, when the behavior was finally caught in the wild [156]. And this experiment, which at last observed a DP-class phase transition with quantitative exactness, used a liquid crystal substance (*N*-(4-Methoxybenzylidene)-4-butylaniline) which wasn't synthesized until 1969 [157].

This rather puts the lie to the notion that a scientific hypothesis must be amenable to immediate falsification by experiment. The model here is not an esoteric proposal for quantum gravity, but an idealization of water flowing through coffee grounds. And not only did it take half a century to go from a mathematician's mind to a laboratory bench, but that journey depended on tools which did not exist when the model was first conceived.

One place we find DP behavior is in *spatial predator–prey models*. Consider a two-dimensional lattice. Each lattice site can be empty, occupied by a prey organism or occupied by a predator. With each tick of a clock, a prey organism can reproduce into an adjacent empty site with some probability g, a predator can take over an adjacent

prey site with some probability τ, and a predator can die with probability v. For each choice of (g, v), there will be a critical consumption rate τ_c below which the predator population will die out. Predator extinction is an absorbing state, and the extinction transition turns out to be DP-class. A predator population eating just fast enough to sustain itself grows as a power law, with the exponent predicted by DP theory. We can also observe DP-class transitions in *evolutionary games,* where players following different strategies accumulate benefits and costs depending on their interactions with their neighbors which affect their ability to reproduce [158].

21 Identical Quantum Particles

21.1 Introduction

In this chapter, we will explore what happens we impose a requirement of symmetry or interchangeability upon an aggregate of quantum particles. The first step in doing so is giving an adequately precise mathematical meaning to everyday words like "interchangeable", in the context of quantum physics. We do this by putting constraints on the kinds of quantum states that we can ascribe to multiparticle systems. To begin with, we will consider states that are extremal in state space, that is, states having the form $\rho = |\psi\rangle\langle\psi|$. We recall that we can write a joint state for two particles by taking the tensor product of single-particle states, which for brevity we will write by combining the symbols we put inside the kets. For example, if we characterize our preparation of particle 1 by saying that we expect a position measurement will yield a result $\vec{x}_1$, we can call that preparation $|\vec{x}_1\rangle$ for short, and to express an uncorrelated joint preparation of particles 1 and 2, we write

$$|\vec{x}_1, \vec{x}_2\rangle := |\vec{x}_1\rangle \otimes |\vec{x}_2\rangle. \tag{21.1}$$

This might express the assertion that we expect to find particle 1 in the kitchen and particle 2 in the living room. All fine so far, but it does run into trouble with the idea that elementary particles — electrons, photons and their ilk — ought to be *featureless.* Anywhere one electron will do, another electron should do equally well. If Alice wants to write a quantum state for a pair of particles that are *in principle indistinguishable,* then she should not write a quantum state that gives them distinguished roles. A state like $|\vec{x}_1, \vec{x}_2\rangle$ is no good, because if Alice swaps the labels of the two particles, what was in the left slot goes into the right and vice versa, yielding $|\vec{x}_2, \vec{x}_1\rangle$, which is mathematically a different object than she started with. Alice should instead pick a ket with the property that swapping the left and right slots leaves the ket unaffected. One option is *symmetrize* the tensor-product ket by taking a balanced linear combination of the original and the swapped versions:

$$|\psi_+\rangle = \frac{1}{\sqrt{2}}(|\vec{x}_1, \vec{x}_2\rangle + |\vec{x}_2, \vec{x}_1\rangle). \tag{21.2}$$

Applying the label-exchange operation to this ket simply exchanges one term in the sum for the other, leaving the total unchanged. Another course of action is to symmetrize but with a minus sign:

$$|\psi_-\rangle = \frac{1}{\sqrt{2}}(|\vec{x}_1, \vec{x}_2\rangle - |\vec{x}_2, \vec{x}_1\rangle). \tag{21.3}$$

Now, the label-exchange operation turns $|\psi_-\rangle$ into $-|\psi_-\rangle$, but this leaves $|\psi_-\rangle\langle\psi_-|$ unchanged, and thus the predictions for all possible experiments upon the particle pair are unaffected.

We can discuss these two symmetrization schemes together and extend them to incorporate more particles by bringing in the idea of *permutations.* Given a list with a recognizable order, like $(1, 2, 3, 4, 5)$, we can rearrange its elements, say by successively swapping pairs. Flipping 1 with 3 and then 4 with 5 yields a new list, $(3, 2, 1, 5, 4)$. Any rearrangement can be decomposed into pairwise swaps, and moreover, any rearrangement can be given a sign factor by raising -1 to the power of the number of pairwise swaps needed to implement it. This way of defining a sign factor has the appealing property that when we apply the permutations P and P' to obtain a new permutation P'', the sign of P'' is the product of the signs of P and P'. Accordingly, let us denote the sign of a permutation P by $(-1)^P$. Introducing the variable $\eta = \pm 1$, we combine and generalize our two previous examples above:

$$|\psi_\eta\rangle \propto \sum_P \eta^P P|\vec{x}_1, \ldots, \vec{x}_N\rangle. \tag{21.4}$$

The prefactor is chosen to ensure proper normalization, $\langle\psi_\eta|\psi_\eta\rangle = 1$.

Crucially, if we take the $\eta = -1$ option and two of the $\vec{x}_i$ are *equal,* then $|\psi_-\rangle$ will implode to the zero vector! This is easily seen in the two-particle example.

These are the two options for building symmetrized multiparticle states that see use in physics. The $\eta = +1$ option describes *bosons* and the $\eta = -1$ choice describes *fermions.* It is worthwhile to ask what other mathematical possibilities might exist, in what circumstances they can be excluded, and how the classification of particles into bosons and fermions relates to other things we know about them. This rapidly leads into topics that are mathematically and physically quite interesting and involved, but which do not directly pertain to statistical physics at our level today. We will, therefore, take the division of particles into bosons and fermions as an empirically driven addition to quantum mechanics, much as quantum mechanics is itself an empirically motivated addition to probability theory.

In fact, much of what we will have to know about bosons and fermions can be encapsulated by the principle of *Pauli exclusion* that applies to the latter.

21.2 Partition Functions for Noninteracting Identical Particles

The upshot of these considerations is that when particles are indistinguishable, we only have to keep track of *the number of particles* associated with each "bin". Moreover, if we declare that our particles are bosons, this *occupation number* can be arbitrarily large, whereas if we are working with fermions, at most one can be associated with any given "bin". In the previous section, we thought of the "bins" as labeled by position vectors, but we will more often be labeling them by momentum vectors instead. A typical scenario will involve enumerating the possible values of momenta (i.e., the

possible outcomes of a momentum measurement), using some extra given information to see how many "bins" share a momentum label (e.g., two polarization values per momentum value), and then calculating expected occupation numbers using Boltzmann weighting.

For the purposes of this chapter, we will work with *ideal gases*. That is, we will not include extra interactions like electrical forces between the particles. The consequences of the *statistical* implications of working with bosons or fermions will already be quite enough to handle! Everything will follow from the grand partition function. We can write this either for fermions or for bosons by taking a sum of the N-particle canonical partition functions, weighting each by the fugacity:

$$\mathcal{Q} = \sum_{N=0}^{\infty} e^{\beta\mu N} Z_N. \tag{21.5}$$

Let us call the fermionic grand partition function $\mathcal{Q}_-$ and the bosonic counterpart $\mathcal{Q}_+$. Then, using s to denote the "bin" labels,

$$\mathcal{Q}_\eta(\beta,\mu) = \sum_{N=0}^{\infty} e^{\beta\mu N} \sum_{\{n_s\}} \exp\left(-\beta \sum_s E(s) n_s\right) \delta\left(\sum_t n_t - N\right). \tag{21.6}$$

where the possible values of n_s are restricted appropriately for fermions or for bosons. The sum inside the exponential becomes a product of exponentials, and the sum over N hits the delta function to give

$$\mathcal{Q}_\eta(\beta,\mu) = \sum_{\{n_s\}} \prod_s \exp\left(-\beta(E(s)-\mu)n_s\right). \tag{21.7}$$

From these, we can compute the average particle numbers:

$$N_- = \sum_s \frac{1}{e^{\beta(E(s)-\mu)}+1}, \tag{21.8}$$

$$N_+ = \sum_s \frac{1}{e^{\beta(E(s)-\mu)}-1}. \tag{21.9}$$

So, first we calculate all quantities like $\mathcal{G}$, E, P and so forth as functions of μ. We then calculate N as a function of μ, so we can express $\mathcal{G}$, E and P in terms of N. (This last step is usually only possible in certain limits, $T \to 0$ and $T \to \infty$.)

A recurring technique in this area is to separate the physical constants (like the degeneracy factor and the mass) from the sums and integrals, as much as possible. This will result in expressions for thermodynamically meaningful quantities, like the pressure P, in terms of constant-filled prefactors and "special functions" defined by integrals over unitless variables. In particular, we will frequently see functions of the following form:

$$f_m^\eta(z) = \frac{1}{(m-1)!} \int_0^\infty \frac{dx\, x^{m-1}}{z^{-1}e^x - \eta}, \text{ with } \eta = \pm 1. \tag{21.10}$$

Here, the number m is a parameter that will take positive integer or half-integer values. We can define factorials of half-integers by way of analytic continuation; the one fact to remember is that

$$\left(\frac{1}{2}\right)! = \frac{\sqrt{\pi}}{2}. \tag{21.11}$$

It is useful to know the small-z limit, which we can evaluate as follows:

$$\begin{aligned} f_m^\eta(z) &= \frac{1}{(m-1)!}\int_0^\infty dx\, x^{m-1}(ze^{-x})(1-\eta ze^{-x})^{-1} \\ &= \frac{1}{(m-1)!}\int_0^\infty dx\, x^{m-1}\sum_{\alpha=1}^{\infty}(ze^{-x})^\alpha \eta^{\alpha+1} \\ &= \sum_{\alpha=1}^{\infty}\frac{z^\alpha \eta^{\alpha+1}}{(m-1)!}\int_0^\infty dx\, x^{m-1}e^{-\alpha x} \\ &= \sum_{\alpha=1}^{\infty}\eta^{\alpha+1}\frac{z^\alpha}{\alpha^m} \\ &= z + \eta\frac{z^2}{2^m} + \frac{z^3}{3^m} + \eta\frac{z^4}{4^m} + \cdots . \end{aligned} \tag{21.12}$$

In the last step, we have used the fact that η to an even power is always unity. We can hazard a guess that the math books will have relevant information about these functions, because if we take $\eta = +1$ and $z = 1$, our expression becomes the *Riemann zeta function* of the parameter m:

$$f_m^*(1) = \sum_{k=1}^{\infty}\frac{1}{m^k} = \zeta(m). \tag{21.13}$$

To fix as much physical intuition into place as possible, we will start in the limit of vanishingly small temperature and see how fermions will behave.

21.3 Ideal Fermi Gases

We begin by taking an empty box that is length L on each side and trying to fill it up, obeying the constraint that we can associate at most some small number g of fermions with the same quantum-state label. The eigenvalues for the Hamiltonian, and thus the energy levels, for one particle in a cubical box are

$$E(n_x, n_y, n_z) = \frac{\pi^2\hbar^2}{2mL^2}(n_x^2 + n_y^2 + n_z^2), \tag{21.14}$$

where each index n_x, n_y and n_z is a positive integer. The number of modes available below an energy E is, approximately, the volume of a sphere in "n-space" of radius

$$\xi = \sqrt{\frac{2mL^2}{\pi^2\hbar^2}}, \tag{21.15}$$

divided by 8 because we only need the positive octant of it:

$$\frac{1}{8}\frac{4\pi}{3}\left(\frac{2mL^2}{\pi^2\hbar^2}\right)^{3/2}. \tag{21.16}$$

Differentiating this with respect to the energy gives a "density" of available modes:

$$\nu(E) = \frac{V}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} E^{1/2}. \tag{21.17}$$

Alternatively, we can write this as a function of the wave number k, which is related to the momentum p by de Broglie's formula:

$$k = \frac{p}{\hbar}. \tag{21.18}$$

The density of available modes as a function of k is

$$\nu(k) = \frac{V}{(2\pi)^3}4\pi k^2, \tag{21.19}$$

where we have chosen a suggestive way of writing the constants in order to hint at the *volume element* $d^3\vec{k}$ in 3-dimensional mode-label space. The total number of fermions is the number we get by integrating up to the highest occupied wave number, or equivalently, the highest occupied momentum. Introducing a constant factor g to account for the possibility that multiple fermions can be assigned the same momentum (if, for example, the states also have spin-up or spin-down labels), we obtain

$$N = gV\int_0^{p_F}\frac{4\pi p^2 dp}{(2\pi\hbar)^3}. \tag{21.20}$$

Dividing both sides by the volume gives the number density:

$$n = \frac{N}{V} = \frac{gp_F^3}{6\pi^2\hbar^3} = \frac{gk_F^3}{6\pi^2}. \tag{21.21}$$

The Fermi energy is the energy value corresponding to the top of the Fermi sea, i.e., the kinetic energy for a particle in a box corresponding to momentum p_F:

$$\mathcal{E}_F = \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2}{2m}\left(\frac{6\pi^2 n}{g}\right)^{2/3}. \tag{21.22}$$

We can likewise define a Fermi temperature as

$$T_F = \frac{\mathcal{E}_F}{k_B}; \tag{21.23}$$

this will set the comparison point for deciding whether temperatures are "large" or "small".

The total energy is found just like the total number of particles. We can tell that it must have a form close to our expression for N, but with two extra powers of k_F, because our integrand will include the kinetic energy, which goes as k^2. The energy is also extensive and is conveniently expressed as an energy per unit volume,

$$\varepsilon = \frac{E}{V} = \frac{g\hbar^2 k_F^5}{20\pi^2 m}. \tag{21.24}$$

With this, we can apply thermodynamics. A small change in E can be broken down into a change due to varying the volume and a change due to varying the particle number:

$$dE = \left.\frac{\partial E}{\partial V}\right|_N dV + \left.\frac{\partial E}{\partial N}\right|_V dN. \tag{21.25}$$

The coefficient of dV is $-P$, and the coefficient of dN is μ.

We can also write the infinitesimal change dE as

$$dE = d(\varepsilon V). \tag{21.26}$$

The d operator acts on εV by the product rule, and from the above, the only input to ε that can vary is the Fermi wave number k_F. We therefore apply the chain rule to $d\varepsilon$:

$$dE = \varepsilon dV + V\frac{d\varepsilon}{dk_F}\left[\left.\frac{\partial k_F}{\partial N}\right|_V dN + \left.\frac{\partial k_F}{\partial V}\right|_N dV\right]. \tag{21.27}$$

This suggests that we can find P and μ by matching coefficients of dV and of dN between our two expressions for dE. To do so, we evaluate

$$dn = d\left(\frac{N}{V}\right) = \frac{dN}{V} - \frac{N}{V^2}dV, \tag{21.28}$$

and we note that this must also equal $(dn/dk_F)dk_F$. We deduce that

$$\frac{\partial k_F}{\partial N} = \frac{1}{V}\frac{1}{dn/dk_F}, \tag{21.29}$$

$$\frac{\partial k_F}{\partial V} = -\frac{1}{V}\frac{n}{dn/dk_F}. \tag{21.30}$$

Now, we can put it all together in our formula for dE:

$$dE = \frac{d\varepsilon}{dk_F}\left(\frac{1}{dn/dk_F}dN - \frac{n}{dn/dk_F}dV\right) + \varepsilon dV. \tag{21.31}$$

Now, we match coefficients of dN and dV and simplify:

$$\mu = \frac{d\varepsilon/dk_F}{dn/dk_F} = \frac{p_F^2}{2m}, \tag{21.32}$$

$$P = \frac{(d\varepsilon/dk_F)n - \varepsilon(dn/dk_F)}{dn/dk_F} = \frac{2}{3}\varepsilon. \tag{21.33}$$

We see that the ideal gas exerts pressure even at vanishing temperature! This nonclassical effect, born of Pauli exclusion, is known as *degeneracy pressure.*

To understand the ideal Fermi gas at nonvanishing temperature, we bring in the grand canonical distribution. We recall that the "grand free energy" is given by

$$G = E - TS - \mu N = -k_B T \log Q. \tag{21.34}$$

This equation bridges the statistical ($\log Q$) with the thermodynamic (E, S and their friends). From the expression for an infinitesimal change in the free energy,

$$dG = -SdT - Nd\mu - PdV, \tag{21.35}$$

we deduce a way to calculate the pressure:

$$P = -\left.\frac{\partial G}{\partial V}\right|_{\mu,T}. \tag{21.36}$$

Consider a two-dimensional gas of noninteracting fermions. First, we'd like to find out about the system at $T = 0$; at absolute zero, what is the relationship between the Fermi energy ϵ_F and the density n? Second, we'd like to find the chemical potential in terms of density and temperature, $\mu = \mu(n, T)$. (Usually, it's much easier to do the inverse, but in 2D we can do this.) After we've done these calculations, we will be able to use these results to obtain the specific heat at low temperatures.

At vanishingly small temperatures, the fermions will occupy the lowest possible energy state, consistent with the requirements of the exclusion principle. The fermions occupy all states up to the cutoff ϵ_F. Therefore, the number of particles is given by the density of states multiplied by the area in k-space which is accessible:

$$\frac{Ag}{(2\pi)^2} \cdot \pi k_F^2 = N. \tag{21.37}$$

The energy spectrum is just a kinetic term,

$$\epsilon(k) = \frac{\hbar^2 k^2}{2m} \tag{21.38}$$

which gives a Fermi energy

$$\begin{aligned} \epsilon_F &= \frac{\hbar^2 2\pi}{m} \cdot \frac{N}{Ag} \\ &= \frac{2\pi\hbar^2 n}{mg}. \end{aligned} \tag{21.39}$$

In the $T \to 0$ limit, the chemical potential is just the Fermi energy:

$$\mu(0) = \frac{2\pi\hbar^2 n}{mg}. \tag{21.40}$$

We expect that as we increase the temperature, the fermions will be excited into higher-energy states. The deformation of the occupation curve cannot be symmetric, however, because it is restricted on the left. To first order, μ will stay at ϵ_F for small T. The chemical potential tells us *where the occupation curve is centered*; it must move to the left to keep the total particle number constant. (To see this, examine Eq. (21.8): at $\epsilon = \mu$, the occupation number is half its maximum.)

Using Eq. (21.8), we calculate

$$\begin{aligned} N &= \sum_k \frac{1}{e^{\beta(\hbar^2 k^2/2m - \mu)} + 1} \\ &= \int \frac{Agd^2k}{(2\pi)^2} \frac{1}{e^{\beta(\hbar^2 k^2/2m - \mu)} + 1} \\ &= \int_0^\infty \frac{Ag}{(2\pi)^2} 2\pi k dk \frac{1}{e^{\beta(\hbar^2 k^2/2m - \mu)} + 1}. \end{aligned}$$

As usual, we change variables. Define

$$x := \beta \frac{\hbar^2 k^2}{2m} \tag{21.41}$$

so that the integral simplifies to

$$\begin{aligned} N &= \int_0^\infty \frac{Ag}{2\pi} \frac{m}{\beta\hbar^2} \frac{1}{z^{-1}e^x + 1} dx \\ &= \frac{Ag}{2\pi} \frac{m}{\beta\hbar^2} f_1^-(z). \end{aligned} \tag{21.42}$$

Normally, this is where we would stop. Because we know the properties of the $f(z)$ functions, however, and because we have dropped down to two dimensions, we can press it a little further. In the small-z limit, we obtain from Eq. (21.12) that

$$\begin{aligned} f_1^-(z) &= \sum_{\alpha=1}^{\infty} (-1)^{\alpha-1} \frac{z^\alpha}{\alpha} \\ &= \log(1+z). \end{aligned} \tag{21.43}$$

Substituting this result yields

$$N = \frac{A}{2\pi} \frac{m}{\beta\hbar^2} \log\left(1 + e^{\beta\mu}\right). \tag{21.44}$$

Solving for μ gives the result that

$$\mu = k_B T \log\left(e^{\frac{2\pi\beta\hbar^2 n}{m}} - 1\right). \tag{21.45}$$

Finally, we want to calculate the low-temperature specific heat. Naturally, we first calculate the energy.

$$E(T) = g \sum_k \frac{\hbar^2 k^2/2m}{e^{\beta(\hbar^2 k^2/2m - \mu)} + 1} \tag{21.46}$$

Again, we can change this sum into an integral and swap variables for the dimensionless x. The only difference is that the numerator gives us an extra factor of x:

$$\begin{aligned} E &= \int_0^\infty \frac{Ag}{2\pi}\frac{m}{\beta\hbar^2}\frac{x/\beta}{z^{-1}e^x+1}dx \\ &= \frac{Ag}{2\pi}\frac{m}{\beta\hbar^2}f_2^-(z). \end{aligned} \tag{21.47}$$

We have the energy as a function of μ, but we'd much rather think about its dependence upon N. Since we're in the low-T limit, we can use a large-z expansion due to Sommerfeld:

$$f_2^-(z) \approx \frac{(\log z)^2}{2}\left[1+\frac{\pi^2}{3(\log z)^2}\right]. \tag{21.48}$$

From the definition of z,

$$\begin{aligned} E &= \frac{Ag}{2\pi}m\beta^2\hbar^2\left[\frac{(\beta\mu)^2}{2}+\frac{\pi^2}{6}\right] \\ &= \frac{Ag}{2\pi}m\hbar^2\left[\frac{\mu^2}{2}+\frac{\pi^2}{6}(k_BT)^2\right] \end{aligned}$$

where we have assumed that μ does not vary significantly at low T. Differentiating with respect to temperature now gives

$$C_V = k_B\frac{\pi Agmk_BT}{6\hbar^2}. \tag{21.49}$$

One can also derive C_V up to constants using dimensional and heuristic arguments.

21.4 Electrons in a Lattice and the Band-Gap Model

Consider a 1D series of square-well potentials arranged along the x-axis. The width of each well is a, and the distance between left-edges of neighboring wells is b. Because of periodicity, we write the wave function

$$\phi_k(x) = e^{ikx}U_k(x), \tag{21.50}$$

where the function $U_k(x)$ is given by

$$U_k(x) = \frac{1}{\sqrt{N}}\sum_n e^{ik(nb-x)}\phi_n(x) \tag{21.51}$$

$$= U_k(x+mb). \tag{21.52}$$

The wavefunction $\phi_k(x)$ has a form incorporating a plane wave with the periodicity of the potential. If we restrict ourselves to measuring at discrete values x_n following the periodicity of the potential, our wavefunction will be given by e^{ikx_n} times a constant, *i.e.,* a plane wave.

We can define a reciprocal space for our network of connected sites. This reciprocal space is mapped by the coördinate k; all the information of the system is contained within an interval $2\pi/a$ along the k-axis, where a is the spacing between network sites. We shall define our zone of interest to be

$$k \in \left[-\frac{\pi}{a}, \frac{\pi}{a}\right]. \tag{21.53}$$

Within this zone, which contains all the information of the system, we have a spectrum of distinct k-values (eigenmodes) separated by

$$\delta k = \frac{2\pi}{L} \tag{21.54}$$

where L is the end-to-end length of the system. These values of k fall into the *first Brillouin zone.*

The energy spectrum for an electron placed on a lattice is a standard result. Given a 1D lattice of sites,

$$E_k = E_0' - A\cos(bk). \tag{21.55}$$

In 3D, the analagous result works out to be

$$E_{\vec{k}} = E_0' - A(\cos k_x b + \cos k_y b + \cos k_z b), \tag{21.56}$$

but we shall restrict our attention to the 1D case. Define the new quantity

$$\mathcal{E}_k \equiv E_k - E_{\min} = A(1 - \cos kb), \tag{21.57}$$

and examine $\mathcal{E}_k$ for $|k| = \frac{2\pi}{\lambda} \ll \frac{2\pi}{a}$, that is, for long wavelengths, $\lambda \gg a$. In this case,

$$\mathcal{E}_k \approx \frac{Ak^2b^2}{2} = \frac{\hbar^2k^2}{2m_{\text{eff}}} = \frac{p^2}{2m}. \tag{21.58}$$

In the long-wavelength limit, we can take a and b to be interchangeable.

$$A = \frac{4E_0}{\gamma},\ E_0 = \frac{\pi^2\hbar^2}{2ma^2} \Rightarrow b \approx a. \tag{21.59}$$

We can solve for the effective mass, as follows:

$$m_{\text{eff}} = \gamma\hbar^2\frac{2ma^2}{4\pi^2\hbar^2a^2}e^{\frac{\gamma\Delta}{a}} \approx \frac{\gamma m}{2\pi^2}e^{\frac{\gamma\Delta}{a}} \approx me^{\frac{\gamma\Delta}{a}} \gg m. \tag{21.60}$$

Therefore, the particles on the lattice behave as free particles with an effective mass m_{eff} much greater than their rest mass m.

One can perform the same analysis for arbitrary k, in which case one obtains an effective mass with a k dependence: $m_{\text{eff}} = m_{\text{eff}}(k)$.

We now turn to a situation where we can apply these one-particle concepts to a gas of many particles. *Neglecting interactions among these particles,* we can do this in

a relatively straightforward fashion. According to Pauli exclusion, we can place two electrons in each Bloch solution, with the two electrons having opposite spins.

In the case that each atom contributes one electron ($N' = N$), we will have a half-filled band of solutions whose upper edge is the Fermi energy

$$\mathcal{E}_F = \mathcal{E}_{\min} + \frac{\Delta E}{2}. \tag{21.61}$$

Here, it is legitimate to model the electrons as a free gas of particles with effective mass m_{eff}. In the second case, where each atom contributes two electrons ($N' = 2N$), the band will be filled completely, and the Fermi energy will be

$$\mathcal{E}_F = \mathcal{E}_{\min} + \Delta E = \mathcal{E}_{\max}. \tag{21.62}$$

Consider a half-filled band. The Fermi energy corresponds to some Fermi wavenumber k_F. If no electric field is applied, the current is strictly zero, because there are equally many left-moving and right-moving modes in the Brillouin zone. In the presence of a field,

$$\tilde{\mathcal{E}}_k = \mathcal{E} \pm eEL, \tag{21.63}$$

where the sign depends upon the direction of motion, *i.e.,* upon the sign of k.

$$\tilde{\mathcal{E}}_k = \begin{cases} \mathcal{E}_k - eEL & \text{left} \\ \mathcal{E}_k + eEL & \text{right} \end{cases} \tag{21.64}$$

We wil have more electrons moving to the right than to the left, implying a current. Because we will have an electric current for an arbitrarily small field at $T = 0$, we can say that this situation describes a metal.

Now, take the case of a completely filled band. There are no states available arbitrarily close to the Fermi energy. Therefore, at $T = 0$, the material is a perfect insulator. In other words, half the particles will still be in left-moving modes and the other half in right-moving modes, for sufficiently small applied field E. In order to have conduction in a material whose band is completely filled, we must excite electrons to empty states in the *second* band, the set of states due to the next-higher energy level of the individual atom. The number of electrons excited into these states at finite temperature is governed by the Boltzmann weight $e^{-\beta\tilde{\Delta}E}$. A change in temperature will have an exponential effect upon the number of potential charge carriers. In fact, we can observe resistivities varying over twenty orders of magnitude, thanks to this Boltzmann factor.

Unfortunately, the model developed in the previous section predicts that the elements in the second column of the Periodic Table — beryllium, magnesium, calcium, strontium, barium and radium — would be electrical insulators. This is not the case: in fact, these elements all have clearly metallic properties.

In our very simple model, if one has the band width ΔE much greater than the band gap $\tilde{\Delta}E$, one will recover the metallic band structure, since the two neighboring bands will *overlap.* One can sometimes effect this change by applying pressure.

Real elements can have considerable overlap of atomic orbitals. The model developed above, which treats this overlap as a perturbation, is not a good departure point for quantitative comparisons.

21.5 Collapsed Stars

We model a dead star as a sphere of matter containing electrons, protons and neutrons, the latter two of which we can lump together under the label of "baryons". If there are N electrons within a radius R, the number density is thus of order N/R^3. Looking back at Eq. (21.21), we see that the number density also goes as the cube of the Fermi momentum, so we can set a typical scale of electron momentum within the star as $\hbar n^{1/3}$. Using the relativistic energy-momentum relation, we have that

$$\varepsilon_F \sim \hbar n^{1/3} c \sim \frac{\hbar c N^{1/3}}{R}. \tag{21.65}$$

Most of the *mass* of the star comes not from electrons but from the baryons, and so the gravitational potential energy of the star will be on the order $-GM^2/R$. The number of baryons is comparable to the number of electrons, so we can divide this by N to get a potential energy per electron that is roughly

$$\varepsilon_G \sim -\frac{GMm_B}{R}, \tag{21.66}$$

where m_B is the baryon mass. (The mass difference between protons and neutrons is irrelevant at this level of precision.) To find an equilibrium between gravitation and degeneracy pressure, we minimize the total energy per electron:

$$\varepsilon = \varepsilon_F + \varepsilon_G = \frac{\hbar c N^{1/3}}{R} - \frac{GNm_B^2}{R}. \tag{21.67}$$

The bigger N is, the more the negative term will dominate, because N grows more quickly than $N^{1/3}$. So, when N is small, ε is positive and can be decreased by increasing the radius R. This will decrease the Fermi energy ε_F and eventually the electrons will become nonrelativistic, with ε_F scaling with p_F^2, which is proportional to $1/R^2$. In this regime, the gravitational term ε_G must eventually dominate, because $1/R$ is larger than $1/R^2$, and so ε must eventually go negative. For very large R, the energy ε will tend to zero from below. Overall, then, there will be a stable equilibrium at some value of R.

$$\varepsilon(R) = \frac{a}{R^2} - \frac{b}{R} \tag{21.68}$$

The derivative with respect to the radius is

$$\frac{d\varepsilon(R)}{dR} = -\frac{2a}{R^3} + \frac{b}{R^2}, \tag{21.69}$$

which equals zero when $R = 2a/b$.

The behavior is very different when N is large. In this case, we never transition to a regime where one term has a $1/R^2$ dependence, so we can make ε go arbitrarily far negative by decreasing R. If N is too big, then *degeneracy pressure cannot stand against gravity.* We determine what counts as "too big" by setting $\varepsilon = 0$ and solving for N. This gives

$$N_{\max} \sim \left(\frac{\hbar c}{Gm_B^2}\right)^{3/2}, \tag{21.70}$$

which works out to be about 2×10^{57}. Multiplying this by the mass of a proton gives a maximum mass of about 1.5 times that of the Sun, which is a reasonable figure for an astrophysical quantity.

21.6 Bose–Einstein Condensation

Let us recall how far we have come. We started our treatment of quantum physics with the concept of quantum states, saying that a state ρ is pure if $\rho = \rho^2$, and that we can just as well specify a pure state by giving the ray it projects onto: $\rho = |\psi\rangle\langle\psi|$. Then we took it as a further constraint that our states have to be symmetrized. This meant that our basis states are characterized fully by occupation numbers:

$$|\{n_s\}\rangle := |n_1, n_2, n_3, \ldots\rangle \tag{21.71}$$

Ascribing this pure state to an aggregate of quantum particles means that when we go to measure bin s, we are absolutely confident that we will find n_s particles there. For fermions, n_s is always 0 or 1, while for bosons, n_s can be any nonnegative integer.

We saw that we can handle ideal gases either of bosons or fermions by introducing a parameter $\eta = \pm 1$ and writing a grand canonical distribution,

$$p_\eta(\{n_s\}) \propto \prod_s \exp\left(-\beta(E(s) - \mu)n_s\right). \tag{21.72}$$

Normalizing this gives a grand canonical partition function $\mathcal{Q}_\eta$, which is conveniently expressed as

$$\log \mathcal{Q}_\eta = -\eta \sum_s \log\left(1 - \eta \exp(\beta\mu - \beta E(s))\right). \tag{21.73}$$

And we get an expectation value for the number of particles found in bin s:

$$\langle n_s \rangle_\eta = -\frac{\partial \log \mathcal{Q}_\eta}{\partial(\beta E(s))} = \frac{1}{z^{-1}e^{\beta E(s)} - \eta}. \tag{21.74}$$

In practice, we found that these conceptual bins are labeled by a momentum or wavenumber vector $\vec{k}$ and a polarization index α, with the latter just contributing a factor of g to our formulae. For a box of side length L, a step of $(2\pi)/L$ in k-space picks up a mode, so the "density" in k-space is $gV/(2\pi)^3$. Nonrelativistically, the energy dependence for free particles is $E(\vec{k}) = \hbar^2 k^2/2m$. Equating $-k_B T \log \mathcal{Q}_\eta$ with the free energy gives

$$\beta P_\eta = \frac{\log \mathcal{Q}_\eta}{V} = \frac{g}{\lambda^3} f^\eta_{5/2}(z), \tag{21.75}$$

where

$$\lambda = \sqrt{\frac{2\pi\hbar^2}{m k_B T}} \tag{21.76}$$

and

$$f^\eta_m(z) = \frac{1}{(m-1)!} \int_0^\infty \frac{dx\, x^{m-1}}{z^{-1}e^x - \eta}. \tag{21.77}$$

This gives the pressure and the density in terms of $z = e^{\beta\mu}$:

$$\beta P_\eta = \frac{g}{\lambda^3} f^\eta_{5/2}(z), \tag{21.78}$$

$$n_\eta = \frac{g}{\lambda^3} f^\eta_{3/2}(z). \tag{21.79}$$

But it is easier to think in terms of the density, so we combine and eliminate:

$$P_\eta = n_\eta k_B T \left[1 - \frac{\eta}{2^{5/2}} \left(\frac{n_\eta \lambda^3}{g}\right) + \cdots\right]. \tag{21.80}$$

Quantum mechanics will seriously kick in when $n_\eta \lambda^3 \geq g$, the *degenerate* limit.

We now focus on the case of bosons, $\eta = +1$. The expectation value for the number of particles at momentum $\vec{p}$ is

$$\langle n_{\vec{p}} \rangle = \frac{1}{z^{-1} e^{\beta E(p)} - 1}. \tag{21.81}$$

Something weird happens when $z \to 1$. If $E(0) = 0$, then the expectation value $\langle n_0 \rangle$ will go screwy!

It is easiest to see what's going on by examining the density, $n = (g/\lambda^3) f^+_{3/2}(z)$. We recall the series expansion

$$f^+_m(z) = \sum_j \frac{z^j}{j^m}. \tag{21.82}$$

For bosons, we must have $z < 1$, so

$$f^+_m(z) < \sum_j \frac{1}{j^m} = \zeta(m). \tag{21.83}$$

This implies a bound on n:

$$n < \frac{g}{\lambda^3} \zeta\left(\frac{3}{2}\right) \approx 2.612 \frac{g}{\lambda^3}. \tag{21.84}$$

What happens if we push the density beyond this bound?

When we moved from sums to integrals, we assumed there was no singular business! This means there can be a spike at $p = 0$ that we did not account for.

By setting $n\lambda^3/g$ equal to $\zeta(3/2)$ and solving for the temperature, we can find the critical value

$$T_c(n) = \frac{2\pi\hbar^2}{m k_B} \left(\frac{n}{g\zeta(\frac{3}{2})}\right)^{2/3}. \tag{21.85}$$

If our system is dense and cold enough, then we will obtain *Bose–Einstein condensation.* The density of *excited* states goes as $T^{3/2}$, and all the rest fall into the $\vec{k} = 0$ mode.

One intriguing consequence is that only the *excited fraction* contributes to the pressure! The condensed modes, so to speak, do not come out and play. Consequently,

$$\beta P = \frac{g}{\lambda^3}\zeta\left(\frac{5}{2}\right) \approx 1.341\frac{g}{\lambda^3}. \tag{21.86}$$

This value is *independent of density.*

Also, only excited modes contribute to the specific heat. We can obtain a good sense of how the specific heat varies with temperature by the usual heuristic of counting the number of modes that are brought out to play. The radius k_e in $\vec{k}$-space of excited modes is set by the thermal energy:

$$\frac{\hbar^2 k_e^2}{2m} = k_B T. \tag{21.87}$$

So, the total excitation energy will go as

$$E_e \propto V k_e^3 k_B T. \tag{21.88}$$

Differentiating by T, we obtain the specific heat at constant volume,

$$C_V \propto V k_B T^{3/2}. \tag{21.89}$$

This dependence holds up until the condensate is dispersed at T_c.

21.7 Superconductivity

Superconductivity was discovered by Heike Kamerlingh Onnes in 1911, using elemental mercury cooled to cryogenic temperatures. Kamerlingh Onnes observed that the electrical resistivity fell to zero at a temperature of roughly 4 K. One important aspect of superconductivity is a phenomenon known as the *Meissner effect*: Superconductors expel magnetic flux. A superconductor contains no magnetic field lines (except in a surface layer typically tens of nanometers thick). Reasoning from Maxwell's Equations, a perfect conductor may still have $\vec{B}$-field lines inside; it is the *change* of flux that is forbidden. Thus, it was seen that superconductors were not merely materials with zero resistivity. Superconductors are examples of quantum mechanics having measurable macroscopic effects, and they are used in many applications where large electric currents must be sustained for long durations of time.

Perfect conductivity, the Meissner effect and other properties can all be explained by the Bardeen–Cooper–Schreiffer (BCS) theory. In 1950, it was observed that the transition temperature of different mercury isotopes depends upon the mass of the postive mercury ions among which the electrons circulated. This was the first indication that vibrations in the superconductor's crystal lattice played a role in the phenomenon. According to BCS, vibrations of the crystal lattice are quantized as *phonons,* effective particles which are analagous to the photons of electromagnetism. An electron passing through a superconductor's crystal lattice distorts the crystal by virtue of its

electric field; these distortions propagate as phonons and can interact with other electrons. If the temperature is sufficiently low, the electron-phonon interaction can cause electrons to form bound configurations called *Cooper pairs,* another type of effective particle. Because individual electrons are "spin-1/2" fermions, the Cooper pairs act as integral-spin bosons: Exchanging both electrons in a pair changes the wavefunction's sign twice, leaving the overall state unchanged. Because Cooper pairs obey Bose–Einstein statistics, they can congregate in a many-particle state of low energy. The superconductor's energy spectrum contains a gap between this ground state and the next lowest permissable state; without this amount of energy, an electron cannot be separated from this Bose–Einstein condensate. Since interactions do not transfer enough energy to cross this energy gap, electrons scattering off crystal imperfections do not contribute to electrical resistance. This, in qualitative terms, is the origin of a superconductor's perfect conductivity.

The BCS theory is counterintuitive in many ways. The coherence length of a Cooper pair — i.e., the separation between two electrons bound by phonon exchange — is on the order of $10^{-4}\,\mathrm{cm}$. Compared to typical atomic dimensions, which are angstrom-scale or less, this is a macroscopic distance. What's more, many Cooper pairs can exist in the same volume, each individual electron only interacting with the other in its pair, and not with those bound in other pairs.

We can explain the Meissner effect and get a quantitative expression for it using what we know so far and a little "received wisdom" about the BCS theory. To start with, we need to know how to talk about momentum in electromagnetism, and so we begin with the Lorentz force law:

$$\vec{F} = q\vec{E} + q\vec{v} \times \vec{B}. \tag{21.90}$$

A charge feels no magnetic force unless it is already moving, and the force it does feel is perpendicular both to the applied field and its current direction of motion.

We recall from E&M class how to write the magnetic and electric fields in terms of the vector and scalar potentials, $\vec{A}$ and V:

$$\vec{B} = \vec{\nabla} \times \vec{A}, \tag{21.91}$$

$$\vec{E} = -\vec{\nabla}V - \frac{\partial \vec{A}}{\partial t}. \tag{21.92}$$

Substituting these into the Lorentz force law, we obtain

$$m\frac{d\vec{v}}{dt} = -q\vec{\nabla}V - q\frac{\partial \vec{A}}{\partial t} + q\vec{v} \times (\vec{\nabla} \times \vec{A}). \tag{21.93}$$

At this juncture, we can invoke a vector-calculus identity,

$$\vec{v} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\vec{v} \times \vec{A}) - (\vec{v} \cdot \vec{\nabla})\vec{A}. \tag{21.94}$$

Substituting in and rearranging, we find,

$$\frac{d}{dt}(m\vec{v}) + q\frac{\partial \vec{A}}{\partial t} + q(\vec{v} \cdot \vec{\nabla})\vec{A} = -q\vec{\nabla}(V - \vec{v} \cdot \vec{A}). \tag{21.95}$$

The combination of derivatives

$$\frac{\partial \vec{A}}{\partial t} + (\vec{v} \cdot \vec{\nabla})\vec{A} \tag{21.96}$$

measures how fast we would find the vector potential changing if we were riding along with the moving particle. This *comoving* derivative has a contribution due to the external field itself changing, and a contribution due to the particle moving into a region where the vector potential may be stronger or weaker. If we decide to call this combination $d\vec{A}/dt$, then

$$\frac{d}{dt}(m\vec{v} + q\vec{A}) = -q\vec{\nabla}(V - \vec{v} \cdot \vec{A}). \tag{21.97}$$

On the left, we have the derivative of a momentum-like quantity, while on the right, we have minus the gradient of something like a potential energy. So, we have recovered the form of Newton's second law, but incorporating the vector potential $\vec{A}$. The quantity $m\vec{v}+q\vec{A}$ reduces to the familiar, "kinematic" momentum $m\vec{v}$ when the vector potential vanishes, and it is known as the *canonical* momentum.

The BCS theory of superconductors tells us that if we measure a Cooper pair, we will typically find its total momentum to be $\vec{p} = 0$. This tells us to zero out the canonical momentum:

$$m\vec{v} + q\vec{A} = 0. \tag{21.98}$$

The current density is given by the density of moving charges, the total charge carried by each particle, and the velocity:

$$\vec{J} = nq\vec{v}. \tag{21.99}$$

Combining this with the previous equation, we get that

$$\vec{J} = -\frac{nq^2}{m}\vec{A}. \tag{21.100}$$

This is conventionally written as the *London equation*

$$\mu_0 \lambda^2 \vec{J} = -\vec{A}, \tag{21.101}$$

where we have introduced the vacuum permeability constant μ_0 and defined the length scale

$$\lambda = \sqrt{\frac{m}{n\mu_0 q^2}}. \tag{21.102}$$

Maxwell tells us that magnetic fields curl around electric currents. Vector calculus then tells us that the curl of the current density is given by the Laplacian of the magnetic field:

$$\vec{\nabla} \times \vec{J} = -\frac{1}{\mu_0}\vec{\nabla}^2\vec{B}. \tag{21.103}$$

But the curl of $\vec{J}$ gives us back the curl of the vector potential, which is the magnetic field, and so

$$\vec{\nabla}^2 \vec{B} = \frac{1}{\lambda^2} \vec{B}. \tag{21.104}$$

We can appreciate the meaning of this by imagining a big block of superconductor, filling the region $x > 0$, while in the volume $x \leq 0$, we establish a constant magnetic field $\vec{B} = B_0 \hat{z}$. What is the magnetic field within the superconductor? By symmetry, everything drops out except

$$\frac{d^2 \vec{B}}{dx^2} = \frac{1}{\lambda} \vec{B}, \tag{21.105}$$

which has the solution

$$\vec{B} = B_0 \hat{z} e^{-x/\lambda}. \tag{21.106}$$

Our length scale λ establishes the *penetration depth,* beyond which the magnetic field is *expelled* from the superconductor.

Let's look at some actual data! First, we need to demonstrate that we can produce a superconductor in the laboratory. So, we rig up a probe containing a cylindrical arrangement of two concentric solenoids, and we lower it into a dewar of liquid helium. We feed current into the outer solenoid, producing a field along its axis. When a sample of metal placed along the axis is cooled below its transition temperature T_c, the Meissner effect will eject the flux lines inside the sample, changing the flux linked by the inner solenoid. The voltage between the inner solenoid's terminals will therefore change, indicating passage through the superconductor transition. To gauge the temperature at which this transition happens, we can use a silicon diode, which reads out increasing voltage with decreasing temperature.

We plot the results in Figure 21.1. We did not do so terribly: The reference books say that the transition temperature T_c for vanadium is 5.4 kelvins, whereas we put it at about 6.2 kelvins. But in our experiment, the temperature sensor was a little *above* the vanadium sample, so the helium vapor being pumped up from below probably cooled it less!

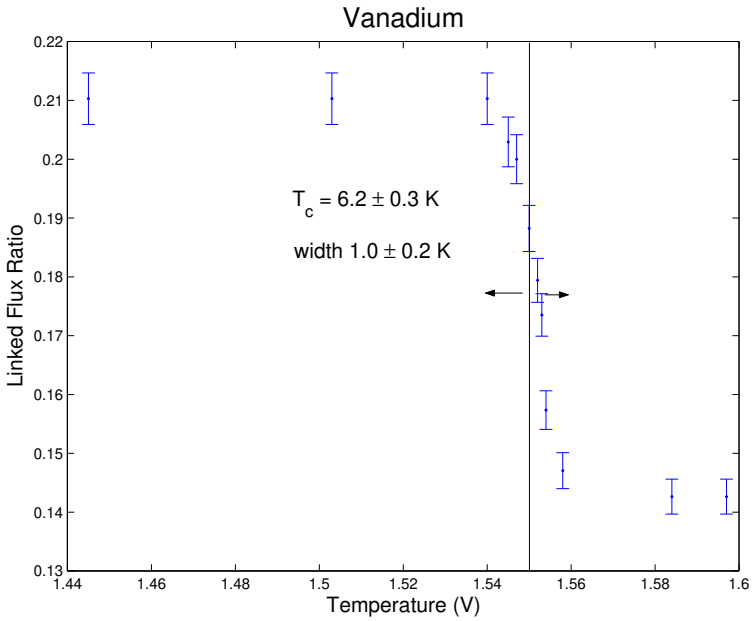


Figure 21.1: Meissner-effect temperature dependence for a vanadium sample. The horizontal axis is voltage read from our silicon thermosensor; lower temperatures are on the right. The vertical axis is the ratio of voltages across the inner and outer solenoids, a measure of the flux linked through the inner cylinder.

22 Useful Mathematical Background: Linear Algebra

One reason that a physics course can be more painful than strictly necessary is that the student might have seen the mathematical prerequisites without having grown comfortable with them. It is very easy for the professor to think, "Well, they've taken three semesters of calculus by now, so obviously they know how to do such-and-such." Consequently, the professor will make a step from one equation to the next that the student might not follow, perhaps just because their memory was rusty.

It is also often the case that the mathematical background we assume for a physics course might be scattered across three or four different math textbooks, with significant fractions of each textbook irrelevant for physics purposes. Accordingly, in this chapter we will review as much as feasible of the mathematics that we will end up using. Parts of this review may be very "elementary" compared to the reader's current standing, in which case feel proud of how far you have come!

22.1 Numbers

The first thing we need to understand is the *types of numbers* that we will meet while doing quantum physics. We develop progressively more expansive numerical environments by starting with a simpler idea of what a "number" is and does, and then asking questions that we cannot answer within that setting — questions that *feel like* they ought to have answers. Perhaps the most friendly place to start is with the idea of counting.

The "Ishango bone" is an artifact that is roughly 20,000 years old. Carved into it are marks that sure *look* like somebody was taking tallies, but of course what exactly they were counting is unknown to us. (If mathematics ever feels difficult, just remind yourself that humanity has been developing it for 20,000 years, so having to take a little while to work at understanding it is only to be expected.) The modern approach is to say that when we count, we use the *natural numbers,* which indicate the sizes of sets of things. It is a matter of personal taste whether we start the natural numbers with "zero" (the size of the empty set) or with "one". For myself, I like including zero, because it seems right that we should be able to say we have none of something. Of course, the discovery that we can treat zero as a number took a long time. If I have no apples, do I have a number of apples? (For that matter, if I have one apple, do I have a number of apples?) One Egyptian papyrus that survives from the Thirteenth Dynasty, around 3,800 years ago, is a balance sheet of sorts, recording foodstuffs collected and redistributed by the pharaoh's court. A hieroglyph most familiar from the word for

"beautiful" appears in cases where the income exactly met the expenses. It also occurs in architectural contexts, as a label for a reference level; parts of buildings were marked as being so-and-so many cubits above or below the ground or the pavement [159]. This sign, which we in retrospect would call an early kind of zero, might have had the sense of being satisfyingly balanced.

The natural numbers have a notion of *addition,* because we can take a set of a objects and a set of b objects and combine them to form a set of size $a+b$. We should keep in mind, however, that this abstract mathematical manipulation may or may not faithfully reflect physical reality. As Eugenia Cheng points out, if you give a child two cookies and then give them two more cookies, they might not have four cookies [160]. Sticking with the abstract manipulations, we note that addition of natural numbers is *commutative*: $a + b = b + a$. It is also *associative*: $a + (b + c) = (a + b) + c$. We teach these big words to small students, but in so doing we often lose track of the important point that the reason we have the big words is so that we can talk about the concept, and we talk about the concept because it *isn't always true.* There are mathematical operations which are not commutative — where the order matters. We will meet important examples soon. In this course, we will be spending more time with failures of commutativity than failures of associativity, but the latter exist and are significant too.

We can also *multiply* natural numbers. We can think of multiplication as iterated addition ("five times three" is five laid down three times). We can also think of the product of two sets as making a new set from all the possible pairings of an element from the first and an element from the second. This gives us a visualization for why multiplication is also commutative. Five sets of three pandas apiece gives the same product as three sets of five pandas apiece; we can just demarcate the same rectangular grid of pandas either horizontally or vertically.

An *inverse* is an action that undoes another. Subtraction is the inverse of addition: Subtracting 2 from a natural number undoes the effect of adding 2. And here we confront the issue that the natural numbers do not provide answers to all the inverse actions that we can imagine. $4-2$ is a natural number, but $2-4$ is not. Yet shouldn't we be able to treat $2-4$ as a number of some kind? Even the scribes who wrote that Thirteenth Dynasty papyrus recognized that on some days, the expenditures of Pharaoh's court exceeded the income. The extra amount that needs to be collected to achieve balance should be a quantity that we can manipulate numerically. We declare, therefore, that for every natural number a there is a new number $0-a$. Or, to say it another way, every natural number a has an *additive inverse*, $-a$, which enjoys the property that $a + (-a) = 0$. The natural numbers, together with all their additive inverses, furnish the *integers.*

Just as subtraction undoes addition, division undoes multiplication. And just as with subtraction, division also takes us out of the natural numbers. We can divide 6 by 2 or by 3 and get a natural number, but we cannot divide 6 by 4. And yet, shouldn't we be able to? In ancient Egypt, wages were paid in grain.[1] Suppose we have to divide

[1] And when the workers didn't get their grain, they went on strike, as happened at the village now called Deir el-Medina, circa 1158 BCE.

6 jars of wheat among 4 workers. Surely that must be *some* kind of number, something that we can handle numerically in a legitimate way. It's the number with the property that, if we multiply it by 4, we get back to the original 6. Developing this idea, we arrive at the *rational numbers.* These include all ratios a/b of two integers a and b, except when b is zero. The rational number a/b, when multiplied by b, yields a. Two ratios that might look different, say a/b and c/d, are declared to be the same number if $ad = bc$.

(Why can't we divide by zero? Well, think of modifying our last example. We can share 6 jars of wheat among 2 workers, or among 4 workers, but what does it mean to share 6 jars of wheat among 0 workers? The question "How much wheat do each of the 0 workers receive?" is nonsensical.)

Adding two rational numbers always gives a rational number, and so does subtracting or multiplying, as well as dividing (whenever division is defined, i.e., not by zero). We say that the rationals are *closed* under these operations. But they are not closed under *every* kind of manipulation that we can imagine; there are still questions we can ask for which we should expect numerical answers, even though the rationals can't provide them. To take the next step, we recall how we got from adding to multiplying, and we repeat: *Exponentiation,* or *taking to a power,* is repeated multiplication. Hopefully, you can recognize some typical examples with smallish numbers instinctively:

$$2^2 = 4,\ 3^2 = 9,\ 2^3 = 8\,, \tag{22.1}$$

and so forth. These examples illustrate that the order of the numbers matters — three squared is not the same as two cubed. So, there are two different ways we can talk about inverting the act of exponentiation. We could ask, "What number to the second power equals 9?" Or, we could ask, "To what power must we raise 3 to get 9?" The first is the *square root* of 9, and the second is the *logarithm* to the base 3 of 9. In this case, both kinds of inverse give rational numbers (indeed, integers). But this is not always so!

Traditionally, one demonstrates that taking square roots can leave the world of rationals behind using the square root of 2. For variety, let's do it with the *golden ratio* instead. This number solves the quadratic equation

$$x^2 - x - 1 = 0\,. \tag{22.2}$$

The quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{22.3}$$

tells us that there are two solutions. One of them is negative and the other positive; the positive solution gets a special symbol,

$$\phi = \frac{1 + \sqrt{5}}{2}\,. \tag{22.4}$$

This is approximately 1.618, or 809/500, but it is not *exactly* that, or exactly *any* rational number. We can prove this using *continued fractions.* Any ordinary fraction,

like 19/7, can be expanded out into an alternate form of nested reciprocals. We start by finding the integer part. In this case, 7 goes into 19 twice, with 5 left over:

$$\frac{19}{7} = 2 + \frac{5}{7}\,. \tag{22.5}$$

Now, we flip the fractional part. It is the inverse of its inverse, like so:

$$\frac{19}{7} = 2 + \frac{1}{\frac{7}{5}}\,. \tag{22.6}$$

Next, we repeat. 5 goes into 7 once with 2 left over, and so:

$$\frac{19}{7} = 2 + \frac{1}{1 + \frac{2}{7}}\,. \tag{22.7}$$

We get another fraction, which we again write as the inverse of its inverse:

$$\frac{19}{7} = 2 + \frac{1}{1 + \frac{1}{7/2}} = 2 + \frac{1}{1 + \frac{1}{3 + \frac{1}{2}}}\,. \tag{22.8}$$

And now we notice that the flipping trick just gives us a sum of integers:

$$\frac{1}{2} = \frac{1}{\frac{2}{1}} = \frac{1}{1+1}\,, \tag{22.9}$$

and so

$$\frac{19}{7} = 2 + \frac{1}{1 + \frac{1}{7/2}} = 2 + \frac{1}{1 + \frac{1}{3 + \frac{1}{1+1}}}\,. \tag{22.10}$$

Any rational number can be "unspooled" in this way to give a chain of integers, in this case the list $2, 1, 3, 1, 1$. Going the other way, given a continued fraction, we can always spool it back up by starting at the end and doing the arithmetic. But what if there *is* no end? Then that continued fraction cannot have come from a rational number, because the "unspooling" procedure for rational numbers always stops.

We know that ϕ satisfies the quadratic equation we started with, so

$$\phi^2 - \phi - 1 = 0\,. \tag{22.11}$$

Let's try to write a continued fraction for ϕ. We divide through everything in this equation by ϕ, and we get

$$\phi - 1 - \frac{1}{\phi} = 0\,. \tag{22.12}$$

Solve this for ϕ, and we find

$$\phi = 1 + \frac{1}{\phi}\,. \tag{22.13}$$

This is starting to look right. Let's try to iterate. We have an equation for ϕ, so we plug it into where ϕ appears on the right-hand side.

$$\phi = 1 + \frac{1}{1 + \frac{1}{\phi}}\,. \tag{22.14}$$

Uh-oh. We can keep doing this, and it'll never stop!

$$\phi = 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \cdots}}} \,. \tag{22.15}$$

"Unspooling" the golden ratio gives an endless chain: $1, 1, 1, 1, \ldots$. There is no way to spool this back and get a rational number, and no rational number can produce a chain like this. So, the golden ratio is irrational. This immediately tells us that $\sqrt{5}$ is also irrational.

Extending the world of rationals to encompass all of the irrationals gives us the *real numbers.* This is perhaps the most technically demanding extension of all those that we are discussing today. If one tries to make sense of the real numbers by considering all decimal expansions, for example, one has to be very careful when giving a meaning to the "dot dot dot" in decimal expansions that never terminate. It can be done, but the devil is very much in the details, and the result of filling in the gaps between the rationals to obtain an unbroken number line has some peculiar properties. For example, between any two rational numbers there is an irrational, and between any two irrationals there is a rational. Moreover, there are *more* irrationals than rationals, even though the set of all rationals is infinitely big. Pursing these topics leads to deep mathematics, easily more than we could fill a course with. However, the mathematics one develops in that pursuit has not, so far, been relevant for physics. Perhaps this is a truth of nature, or perhaps physicists have missed opportunities by not trying hard enough to find applications of it. Because this course is nearly all about things that have been solidly worked out, rather than speculative matters at or beyond the frontier, we will resist the temptation to traipse off and instead take the real numbers as established.

We are still not done with asking questions that our number system cannot answer. We can take square roots of positive numbers, and we can see that $\sqrt{0} = 0$, but what about negative numbers? We recall that the additive inverse of a number a is that number which, when we add it to a, gives zero:

$$a + (-a) = 0 \,. \tag{22.16}$$

This means that the additive inverse of $-a$ must be a again, because that's what we add to $-a$ in order to get zero. In other words, "the negative of a negative is a positive":

$$-(-a) = 0 \,. \tag{22.17}$$

It also follows from our basic notions that taking the additive inverse is the same as multiplying by -1:

$$-a = (-1) \cdot a \,. \tag{22.18}$$

Why? Well, let's try it out:

$$a + (-1) \cdot a = (1) \cdot a + (-1) \cdot a = (1 + (-1)) \cdot a = 0 \cdot a = 0 \,. \tag{22.19}$$

Therefore, -1 times -1 is just 1, and indeed a negative times a negative always gives a positive.

We can think of this geometrically by imagining a number line and considering what the arithmetic operations do to that number line. Addition shifts the line one way or the other. For example, the act of adding 2 is that shift which takes the point 0 and shifts it to lie at the original location of the point 2. Multiplying by a positive number stretches or squeezes the number line as though it were made of elastic. For example, multiplying by 3 is that stretch which pulls out the line until the point 1 is located at the original location of the point 3. Multiplying by a number smaller than 1 squeezes the line. The identity

$$a \cdot \frac{1}{a} = 1 \tag{22.20}$$

says, in this picture, that corresponding stretches and squeezes undo each other.

Multiplying by -1 flips the number line around the origin. And so, quite naturally, two flips brings it back to where it started! Multiplying by any other negative number is going to do a flip and a stretch or squeeze. Consequently, there is no negative number that can square to a positive. If b is a square root of a, then

$$1 \cdot b \cdot b = a\,. \tag{22.21}$$

That is, starting at 1, and applying the transformation "multiply by b" twice, we must arrive at a. But multiplying by b twice means taking either *no* flips or *two* flips. There's no way to get *one single flip,* and so no way to reach a if a is negative.

What geometrical transformation, when applied twice in succession, is the same as flipping? *Rotating by a quarter turn.* To take square roots of negative numbers, we have to move off the number line and into the number *plane.* When we went from integers to rationals, we made new numbers by taking pairs of the old ones and then declaring how those pairs should behave. We do that again, except now we take pairs of real numbers and make them into *complex numbers*:

$$z = (a; b)\,, \tag{22.22}$$

with a and b both real. The real numbers live within the *complex plane*; they are the line $(a; 0)$. We insist that our rules of arithmetic still work, and we say that multiplying by the complex number $(0; 1)$ effects a quarter turn of the complex plane. Multiplying by $(0; 1)$ twice in succession is the same as multiplying by $(-1; 0)$. We traditionally abbreviate $(0; 1)$ with the letter i and write the complex number z as the sum of real and "imaginary" parts:

$$z = a + bi\,. \tag{22.23}$$

The sum of two complex numbers $z = a + bi$ and $w = c + di$ is

$$z + w = (a + c) + (b + d)i\,, \tag{22.24}$$

and their product is

$$zw = (a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i\,. \tag{22.25}$$

Reflecting the complex plane through the real number line flips the sign of the imaginary part, giving the *complex conjugate* of z:

$$z^* = (a; -b) = a - bi\,. \tag{22.26}$$

The conjugate of i is also its negative, and it is also a square root of -1. Multiplying a complex number by its conjugate yields

$$zz^* = (a + bi)(a - bi) = a^2 + b^2\,, \tag{22.27}$$

which is the square of the distance from the origin $(0; 0)$ to the point z. This will be particularly important in quantum mechanics. We denote the distance from z to the origin by the absolute value notation, since it generalizes that idea:

$$zz^* = |z|^2\,. \tag{22.28}$$

Because the complex numbers live in a plane, we can label them in two different ways. We can use rectangular coordinates, like we have been, or we can alternatively label them in a *polar* system by giving a distance from the origin and an angle from some convenient axis, like the real line. The relation between these two labelings is provided by the *exponential* function. You perhaps saw this function first in calculus, defining it by the infinite sum

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\,. \tag{22.29}$$

This function has the nice property that it is its own derivative. (In fact, it is uniquely specified by asking for a function that is its own derivative and whose value at $x = 0$ is 1.) It might not be immediately obvious that this is the same as raising a special number e to the power x, but it's true:

$$e^x = (\exp(1))^x = \exp(x)\,. \tag{22.30}$$

We can equivalently define the exponential function as the limit

$$\exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n\,. \tag{22.31}$$

Let us use this expression to raise e to an imaginary power, it. If n is big enough, then t/n is going to be small, and the complex number $1 + it/n$ will be a point near 1, a little above the real line if t is positive and a little below if t is negative. Multiplying by $1 + it/n$ is a tiny turn and a tiny stretch: We have to rotate the real line by a little angle and then pull it out just enough so that the number 1 ends up at $1 + it/n$. Raising this to the nth power means repeating this turn and stretch n times. We are stacking up n thin triangles to make a sort of spiral. The bigger n is, the thinner the triangles, and the more tightly the spiral is wound — the more it resembles an arc of a circle.

This picture leads us to *Euler's formula,*

$$e^{it} = \cos t + i \sin t \,, \tag{22.32}$$

and thence to the polar scheme for writing complex numbers:

$$re^{it} = r\cos t + ir\sin t = (r\cos t; r\sin t) \,. \tag{22.33}$$

It is easier to add complex numbers in rectangular coordinates, but it is easier to multiply them in polar coordinates: To implement two successive turn-and-stretches, we just multiply the radii and add the angles. If $z = re^{it}$ and $w = se^{iu}$, then

$$zw = rse^{it}e^{iu} = rse^{i(t+u)} \,. \tag{22.34}$$

In one sense, the complex numbers are "the end of the line". The *fundamental theorem of algebra* says that the solutions to any polynomial equation we can write with complex coefficients will again be complex numbers. More specifically, if

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n \,, \tag{22.35}$$

with all of the a_k complex, then $f(z)$ has at most n distinct zeroes, all of them in the complex plane. Thus, there is no way that the problem which led us to the complex numbers, trying to solve an equation like $z^2 = -1$, will recur. One *can* keep going in a different way, by doubling the dimension of the number world: from 1 for the real numbers, up to 2 for the complex plane, and then 4 for the *quaternions*, 8 for the *octonions* and so on. The resulting number systems look "curiouser and curiouser" the further you go. Multiplication of quaternions is no longer commutative, and multiplication of octonions is not even associative. Some topics in quantum physics can be addressed using quaternions, but this is nonstandard (physicists have other tools with which they are much more comfortable for those problems). Octonions show up in rather esoteric circumstances, and anything beyond that is sufficiently pathological from a quantum-physics perspective that we will have no need of it.

We follow tradition and denote the natural numbers by $\mathbb{N}$, the rational numbers by $\mathbb{Q}$, the integers by $\mathbb{Z}$ (for the German *Zahlen*) and the complex numbers by $\mathbb{C}$.

22.2 Combinatorics

To start, let's update an example from a handy little classic, which I first found an old physics book on a colleague's "miscellaneous" shelf: *University of Chicago Graduate Problems in Physics,* by Cronin, Greenberg and Telegdi (Addison–Wesley, 1967). It looked like fun, so I started working through some of it. The opening problem in the "statistical physics" chapter turns out to be a good place to begin our review. Editing slightly to adjust for shifts in the cultural milieu, the problem runs as follows:

> Asami is meeting Korra for lunch downtown. Korra is E blocks east and N blocks north of Asami, on the rectangular street grid of downtown

> Republic City. Because Asami is eager to meet Korra, her path never doubles back. That is, each move Asami takes must bring her closer to Korra on the street grid. How many different routes can Asami take to meet Korra?

Asami must traverse a total of $E + N$ block edges. At each intersection, she must choose whether to move east or to move north, unless she is on the same street as Korra, at which point she keeps going in a straight line along that street. Her path can be represented as a string of ns and es, of total length $E + N$ characters. The question is how many different ways such a string can be written. That is, given a total of $E + N$ slots, how many ways can we pick a subset of E of them to fill with es?

Phrasing the problem this way, we see that it's a job for *binomial coefficients.* The number of ways to pick K things out of a set of M items is

$$\binom{M}{K} = \frac{M!}{K!(M-K)!}. \tag{22.36}$$

This is one of the more enthusiastic formulas that one encounters in everyday mathematics. The factorial $M!$ of a positive integer M is the product of M with $M - 1$ and $M - 2$ and so on, down to 1:

$$M! = M(M-1)\cdots 1. \tag{22.37}$$

If we ever find ourselves needing to take the factorial of 0, we use the special rule that $0! = 1$. The factorial operation counts the number of ways to arrange M items in a line. The special rule $0! = 1$ amounts to saying that if we have *no* items, there's only one possible way to order them, *i.e.,* do nothing, because it's all we can do!

Here, we need the number of ways to pick E items out of a set of size $E + N$, which is

$$\binom{E+N}{E} = \frac{(E+N)!}{E!\,N!}. \tag{22.38}$$

This is what the problem asked us to find. Note that we would have gotten the same answer if we decided to pick N items out of $E + N$.

It's good practice to check the "limiting cases" of a solution. If we derive a complicated formula which applies across a wide swath of scenarios, and if we have an answer we trust for a special case, then we should check that our complicated calculation gives the right answer for that special case. Here, we know that if Korra and Asami start off on the same street, there's only one path which meets the "never doubling back" criterion. In such a situation, either $N = 0$ or $E = 0$. Let's plug that into our boxed equation:

$$\frac{(E+0)!}{E!\cdot 0!} = \frac{E!}{E!\cdot 1} = 1. \tag{22.39}$$

Simple enough!

Likewise, if the initial configuration has Korra one block north and one block east of Asami, then it's pretty plain that Asami has two possible paths that she can take: she

can go north-and-then-east, or she can travel east-and-then-north. Here, $E = N = 1$, and our formula says

$$\frac{(E+N)!}{E!\,N!} = \frac{2!}{1!\,1!} = 2. \tag{22.40}$$

The figure in the Chicago book gives for specific numbers $E = 3$ and $N = 4$. Plugging these in, we find

$$\frac{(E+N)!}{E!\,N!} = \frac{7!}{3!\,4!} = \frac{7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2}{3\cdot 2\cdot 4\cdot 3\cdot 2} = 7\cdot 5 = 35. \tag{22.41}$$

22.3 Trigonometry

Let's start with the most useful formula I learned later than I should have:

$$e^{i\theta} = \cos\theta + i\sin\theta. \tag{22.42}$$

This equation, due to Leonhard Euler, relates two ways of labelling points in the complex plane. The left-hand side is in the spirit of polar coordinates: it locates a point on the unit circle by turning an angle θ from the horizontal. The right-hand side is rectangular in character: it tells us to reach a destination by taking a step of length $\cos\theta$ in the horizontal direction and then one of length $\sin\theta$ in the vertical direction.

First trick: take the derivative of Eq. (22.42) with respect to θ.

$$\frac{d}{d\theta}e^{i\theta} = ie^{i\theta} \tag{22.43}$$

$$= i(\cos\theta + i\sin\theta) \tag{22.44}$$

$$= -\sin\theta + i\cos\theta. \tag{22.45}$$

The real part of Eq. (22.42), which was a cosine, has become minus a sine; the imaginary part, which was a sine, has become a cosine. This meshes with our basic calculus lessons.

Now, multiply two numbers of the form $e^{i\theta}$:

$$e^{iA}e^{iB} = e^{i(A+B)}, \tag{22.46}$$

by the rules of exponentials. But using Eq. (22.42),

$$e^{i(A+B)} = \cos(A+B) + i\sin(A+B). \tag{22.47}$$

This must be equal to the product

$$e^{iA}e^{iB} = (\cos A + i\sin A)(\cos B + i\sin B) \tag{22.48}$$

$$= \cos A\cos B - \sin A\sin B + i\sin A\cos B + i\sin B\cos A. \tag{22.49}$$

Matching the real parts of the two expressions, we find the addition formula for cosines:

$$\cos(A+B) = \cos A\cos B - \sin A\sin B. \tag{22.50}$$

Likewise, by matching the imaginary parts, we find the addition formula for sines:

$$\sin(A+B) = \sin A \cos B + \sin B \cos A. \tag{22.51}$$

Setting $A = B$ in these expressions yields the double-angle formulae,

$$\cos(2A) = \cos^2 A - \sin^2 A, \tag{22.52}$$
$$\sin(2A) = 2\sin A \cos A. \tag{22.53}$$

Combining the double-angle formula for cosines with the Pythagorean identity,

$$\sin^2 A + \cos^2 A = 1, \tag{22.54}$$

we can equivalently say that

$$\cos(2A) = 2\cos^2 A - 1 = 1 - 2\sin^2 A. \tag{22.55}$$

It can also be useful to know that

$$\tan(2A) = \frac{\sin(2A)}{\cos(2A)} \tag{22.56}$$
$$= \frac{2\sin A \cos A}{\cos^2 A - \sin^2 A} \tag{22.57}$$
$$= \frac{2\sin A \cos A}{\cos^2 A - \sin^2 A} \cdot \frac{1/\cos^2 A}{1/\cos^2 A} \tag{22.58}$$
$$= \frac{2\tan A}{1 - \tan^2 A}. \tag{22.59}$$

If we flip the sign of the angle in Eq. (22.42), we get the same thing as if we took the complex conjugate:

$$e^{-i\theta} = \cos(-\theta) + i\sin(-\theta) = \cos\theta - i\sin\theta. \tag{22.60}$$

Adding this to Eq. (22.42), the imaginary parts cancel out:

$$e^{i\theta} + e^{-i\theta} = 2\cos\theta. \tag{22.61}$$

We can, therefore, write the cosine function as the sum of two complex exponentials:

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}. \tag{22.62}$$

This can be handy when one has to manipulate expressions involving combinations of exponential and trigonometric functions. We also have that

$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}. \tag{22.63}$$

Well, that's high-school trigonometry sorted.

Seriously! Life would have been so much easier then — well, that part of life *inside* the classroom, anyway — if I'd only been properly introduced to the idea that a complex number stands for a stretch and a twist. (See http://math.ucr.edu/home/baez/trig.html for a concurring opinion.)

22.4 Vectors

Vectors are one of the topics that we hope you've met before. A prototypical physicist definition of a vector goes something like this: "A vector is a quantity with both magnitude and direction." This means that the mathematical manipulations of vectors can affect their size and their orientation, and these affects must play nicely with each other. The first example of a vector one meets during a physics education is probably *displacement,* e.g., the vector pointing from our classroom to the nearest place to buy coffee. A significant portion of introductory classical mechanics is piling more examples of vectors on top of this idea: velocity, acceleration, momentum and force are all vector quantities. Keeping these old friends in mind, we can think of a few properties that vectors satisfy in general.

First, we can change the size of a vector by rescaling it. In symbols, if $\vec{v}$ is a vector and α is a number, then $\alpha\vec{v}$ is also a valid vector. Next, we can change the direction of a vector by adding a step in a different orientation: If $\vec{v}$ and $\vec{w}$ are both displacement vectors, for example, then we can combine them to form a new vector $\vec{v}+\vec{w}$. Turning these ideas around, we can use them to define what we mean by a vector — we can take them as axioms for a *vector space.* Axiomatically, then, a vector space is a set V equipped with *addition* and *scalar multiplication* operations that meet the following conditions:

- The sum $\vec{v}+\vec{w}$ of any two elements $\vec{v},\vec{w}$ within V also belongs to V.
- Vector addition is commutative and associative:

$$\vec{v}+\vec{w}=\vec{w}+\vec{v}\,, \tag{22.64}$$
$$(\vec{v}+\vec{w})+\vec{u}=\vec{v}+(\vec{w}+\vec{u})\,. \tag{22.65}$$

- There is a *zero vector* within V, i.e., an element $\vec{0}$ that satisfies

$$\vec{0}+\vec{v}=\vec{v} \tag{22.66}$$

  for every $\vec{v}\in V$.

- Every element $\vec{v}\in V$ has an additive inverse, an element $-\vec{v}$ such that

$$(-\vec{v})+\vec{v}=\vec{0}. \tag{22.67}$$

- Addition and scalar multiplication interact properly:

$$\alpha(\vec{v}+\vec{u})=\alpha\vec{v}+\alpha\vec{u}\,, \tag{22.68}$$
$$(\alpha_1+\alpha_2)\vec{v}=\alpha_1\vec{v}+\alpha_2\vec{v}\,, \tag{22.69}$$
$$(\alpha_1\alpha_2)\vec{v}=\alpha_1(\alpha_2\vec{v})\,, \tag{22.70}$$
$$0\vec{v}=\vec{0}\,, \tag{22.71}$$
$$1\vec{v}=\vec{v}\,. \tag{22.72}$$

All of these properties ought to feel familiar enough for vectors like displacement and velocity. But one of the great themes of quantum physics is that *functions* can be vectors too. For example, consider the set of all polynomial functions of some fixed degree:

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \,. \tag{22.73}$$

The sum of any two such functions is again a function we can write in the same form — adding these polynomials can't up the degree above n. We can multiply through all the coefficients by a constant α, and this gives us another polynomial of the same degree. Taking all the coefficients equal to 0 gives us a zero polynomial. All the properties we listed for a vector space are satisfied.

Getting a little more fancy, the subject of *Fourier series* is all about treating functions like $\sin\theta$ and $\cos\theta$ as vectors.

You may have noticed some manipulations familiar from introductory classical mechanics that we have not mentioned yet. When you learned about displacement, velocity and acceleration, you probably took a lot of dot products; the dot product is an operation $\vec{u}\cdot\vec{w}$ that "eats" two vectors and produces a plain old number, or in our new lingo, a scalar. You also must have taken your share of cross products — an operation that "eats" two vectors and gives back, not a scalar, but another quantity that also has both magnitude and direction. Our axiomatic definition of a vector space is deliberately rather bare-bones. Including a well-behaved operation that turns pairs of vectors into scalars makes a vector space into an *inner product space.* The idea of an inner product generalizes the dot product we met when we were younger. Many notations exist, and so we might as well pick one that looks a little exotic compared to the dot symbol. We'll write the inner product of vectors $\vec{v}$ and $\vec{u}$ as $(\vec{v},\vec{u})$, for now. An inner product must satisfy some conditions:

$$(\vec{v},\vec{v}) \geq 0 \,, \tag{22.74}$$

with equality if and only if $\vec{v}$ is the zero vector. Also, the inner product needs to distribute over the other vector operations in a nice way:

$$(\vec{v},\alpha\vec{u}+\beta\vec{w}) = \alpha(\vec{v},\vec{u}) + \beta(\vec{v},\vec{w}) \,. \tag{22.75}$$

Finally, we will be including complex numbers in our vectors, and we will require that

$$(\vec{v},\vec{u})^* = (\vec{u},\vec{v}) \,. \tag{22.76}$$

Where did this come from? Why should it be that the order matters, but only up to a complex conjugation? The intuition is that we want the *length of a vector* to be a *real number* like it was back in first-year mechanics. Let's say that $\vec{u}$ is a two-dimensional vector, and we try to take its dot product with itself to get the square of its length: Can we just do

$$(\vec{u},\vec{u}) = \vec{u}\cdot\vec{u} = u_1^2 + u_2^2 \,? \tag{22.77}$$

This worked fine and didn't give us any problems when u_1 and u_2 were real numbers, but if they are complex, then we're not guaranteed that the sum of the squares will

be real. If we instead complex-conjugate one of the inputs, then the sum becomes

$$u_1^* u_1 + u_2^* u_2 = |u_1|^2 + |u_2|^2 \,, \tag{22.78}$$

which is guaranteed real. Making the choice to complex-conjugate one of the arguments means that the order of the arguments will matter.

Generalizing the cross product in a meaningful way is harder.

When we get to quantum mechanics, we will be using vectors in "$\mathbb{C}^d$", the vector space made of vectors d components long, where each component is a complex number. And we will be writing them in *Dirac notation.* A "ket" is a column vector:

$$|\psi\rangle = \begin{pmatrix} z_1 \\ \vdots \\ z_d \end{pmatrix} , \tag{22.79}$$

and its corresponding "bra" is the row vector made by transposing and complex-conjugating a ket:

$$\langle\psi| = (z_1^* \ \cdots \ z_d^*). \tag{22.80}$$

Putting a bra and a ket together makes a bra[c]ket, which by the rules of the notation is just a number — the inner product:

$$\langle\psi|\psi\rangle = |z_1|^2 + \cdots + |z_d|^2. \tag{22.81}$$

This implies that $\langle\psi|\psi\rangle \geq 0$, and it is only equal to 0 if the vector $|\psi\rangle$ is the zero vector itself. Moreover, for any pair of vectors $|\psi\rangle$ and $|\phi\rangle$,

$$(\langle\psi|\phi\rangle)^* = \langle\phi|\psi\rangle. \tag{22.82}$$

We will call

$$||\psi|| = \sqrt{\langle\psi|\psi\rangle} \tag{22.83}$$

the *norm* of the vector $|\psi\rangle$, and we'll say that a vector is *normalized* if its norm is 1. Any vector can be made normalized by dividing each element by the norm.

A *spanning set* for a vector space is a set of vectors such that any vector in the space can be written as a linear combination of those in the spanning set:

$$|\psi\rangle = \sum_{i=1}^{d} a_i |v_i\rangle. \tag{22.84}$$

For example, we can make a spanning set for $\mathbb{C}^2$ by taking the vectors

$$|v_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \ |v_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} , \tag{22.85}$$

in which case

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = a_1|v_1\rangle + a_2|v_2\rangle. \tag{22.86}$$

But we could equally well use the spanning set

$$|\tilde{v}_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix},\ |\tilde{v}_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}, \tag{22.87}$$

because we can say

$$\begin{pmatrix}a_1\\a_2\end{pmatrix} = \frac{a_1+a_2}{\sqrt{2}}|v_1\rangle + \frac{a_1-a_2}{\sqrt{2}}|v_2\rangle. \tag{22.88}$$

These are both examples of a kind of spanning set we will often use, the *orthonormal bases.* An orthonormal basis for a vector space is a spanning set $\{|v_i\rangle\}$ for that space that enjoys the property

$$\langle v_i|v_j\rangle = \delta_{ij}\,. \tag{22.89}$$

A set of vectors $\{|v_i\rangle : i = 1,\ldots,n\}$ is *linearly dependent* if there is a way to write one of them as a linear combination of the others. This is equivalent to saying that there is some set of coefficients, not all of which are zero, such that

$$\sum_{i=1}^{n} a_i|v_i\rangle = 0\,. \tag{22.90}$$

All linearly independent sets that span the same space must have the same size, a number that is called the *dimension* of the space.

Linear operators are transformations of a vector space back to itself that obey the rules

$$L(|v_1\rangle + |v_2\rangle) = L|v_1\rangle + L|v_2\rangle,\ L(\alpha|v_1\rangle) = \alpha L|v_1\rangle\,, \tag{22.91}$$

for arbitrary vectors $|v_1\rangle, |v_2\rangle$ and scalar α. This means that we can specify what a linear operator is in terms of its action upon a spanning set:

$$L\left(\sum_i a_i|v_i\rangle\right) = \sum_i a_i L|v_i\rangle\,. \tag{22.92}$$

Moreover, if $\{|v_i\rangle\}$ is an orthonormal basis, then $\langle v_i|L|v_j\rangle$ are the *elements* of L in that basis. If we have some linear operator L defined in a higher-level, more abstract way, we can "compile it down to machine language" by introducing an orthonormal basis and writing the operator as a *matrix*:

$$[L]_{ij} = \langle v_i|L|v_j\rangle\,. \tag{22.93}$$

We write matrices as rectangular grids of numbers. It follows from this expression that the *columns* of the matrix L give the results of applying L to the *basis vectors.*

22.5 Matrix Manipulations

Matrices act rather like numbers themselves, in that under the right conditions we can add and multiply them. Sometimes we can even divide! The crucial insight is that

applying two transformations in succession is a transformation too, and so if L and M are matrices that represent linear transformations which can be done in sequence, then there must be a matrix LM that is their product. We can figure out how this *matrix multiplication* works by first applying M to a basis vector and then applying L to the result of that. The result we tease out by thinking carefully on this is that we must sum over the repeated index:

$$[LM]_{ij} = \sum_k [L]_{ik}[M]_{kj} \,. \tag{22.94}$$

This tells us that the *identity matrix,* which acts like the number 1, is a square matrix with 1's along the diagonal and 0's everywhere else. We will denote identity matrices by I; the size of the matrix will nearly always be clear from context. This also implies that matrix multiplication, unlike the multiplication of real or complex numbers, is not always commutative. The *commutator* of two matrices captures how much the order of their multiplication matters:

$$[L, M] = LM - ML \,. \tag{22.95}$$

If matrix arithmetic is to be like ordinary numerical arithmetic, matrix multiplication should include *inverses.* If we take a number a which is not 0, then we know there is a number a^{-1} which when multiplied by a gives 1. Do matrices have inverses? Sometimes, they do.

We can understand the problem better by working through the idea explicitly for a 2×2 matrix. Let us deduce the inverse of the matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} . \tag{22.96}$$

We want to fill in the question marks in this equation:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} ? & ?? \\ ??? & ???? \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{22.97}$$

First, let's try to get the bottom left-hand entry to work out.

$$\begin{pmatrix} c & d \end{pmatrix} \begin{pmatrix} ? \\ ??? \end{pmatrix} = 0. \tag{22.98}$$

We can satisfy this if we choose the column vector to have d in the top slot and $-c$ in the bottom:

$$\begin{pmatrix} c & d \end{pmatrix} \begin{pmatrix} d \\ -c \end{pmatrix} = 0. \tag{22.99}$$

Now, we try for the top right corner:

$$\begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} ?? \\ ???? \end{pmatrix} = 0. \tag{22.100}$$

We fill this in with

$$\begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} -b \\ a \end{pmatrix} = 0. \tag{22.101}$$

Combining these two column vectors into a 2×2 matrix, we get

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & 0 \\ 0 & -bc+ad \end{pmatrix}. \tag{22.102}$$

We're almost there! We need to rescale our inverse matrix by the quantity $1/(ad-bc)$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \left[\frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \right] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{22.103}$$

We called the original matrix M, so we can call the inverse matrix M^{-1}. We have shown that

$$M^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{22.104}$$

To find the inverse of a 2×2 matrix, we flip the elements on the diagonal, take the negatives of the off-diagonal entries and then divide the whole matrix by the quantity $ad - bc$. If $ad - bc = 0$, then we cannot divide by it, so the matrix has no inverse.

This quantity is important enough that it deserves a name. We call it the *determinant* of the matrix:

$$\det(M) = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc. \tag{22.105}$$

Note that the determinant of the identity matrix is 1, and the determinant of M^{-1} is the inverse of the determinant of M. This suggests a general rule, which turns out to be true: for two matrices M_1 and M_2,

$$\det(M_1)\det(M_2) = \det(M_1 M_2). \tag{22.106}$$

Consider the matrix

$$M = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \tag{22.107}$$

and what it does to the unit vectors in the x and y directions:

$$M \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \; M \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}. \tag{22.108}$$

The tips of these vectors are two corners of a square with area 1. Multiplying by M gives us a rectangle with area 6, which is the determinant of M. Next, consider the matrix

$$T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{22.109}$$

which interchanges the x and y directions. The matrix T has a determinant of -1. Its action is to leave the area of the unit square unchanged, but to flip it about its diagonal.

These two examples illustrate a general property of the determinant. Its value tells us the area of the parallelogram that the matrix transforms the unit square into; and if there is a flip involved in the transformation, the determinant picks up a minus sign. This is not too difficult to establish. We just follow what matrix multiplication does to the corners of the unit square, find the area of the big rectangle that encloses the resulting parallelogram, and subtract away the triangles surrounding it.

The determinant is one of the two most important mappings which turn matrices into numbers. The other is the *trace,* which is the sum of the diagonal elements:

$$\mathrm{tr}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + d. \tag{22.110}$$

The trace operation also interacts with matrix multiplication. An important property is that if we have three matrices M_1, M_2 and M_3, then

$$\mathrm{tr}(M_1 M_2 M_3) = \mathrm{tr}(M_2 M_3 M_1) = \mathrm{tr}(M_3 M_1 M_2). \tag{22.111}$$

If we wrote the labels M_1, M_2 and M_3 clockwise around a circle, we could get these three orderings by picking a starting label and reading around, clockwise. Such rearrangements are called *cyclic permutations.*

So far, we have only defined the determinant and the trace for 2×2 matrices, but we can also define them for bigger matrices, in such a way that their nice properties still hold true.

22.6 2D Rotation Matrices

In this section, we'll cover an important example of 2×2 matrices, the *rotation matrices* that describe turns of the two-dimensional flat plane.

We start with a point in the (x, y) plane, and we break out the trig functions to relate Cartesian and polar coordinates:

$$x = r\cos\theta,\ y = r\sin\theta. \tag{22.112}$$

Now, what happens when we take the point (x, y) and rotate it around the origin by some fixed amount? The *distance* to the origin doesn't change, but the *angle* certainly does: the angle shifts by the amount we turn. If the original angle was θ and we call the size of the turning ϕ, then the *new* point will be located at the position labeled by r and $\theta + \phi$.

Where is this in Cartesian coordinates? Well, adopting the common practice of using apostrophes or "primes" to denote transformed variables, then the previous formula gives us the values of "x prime" and "y prime":

$$x' = r\cos(\theta + \phi),\ y' = r\sin(\theta + \phi). \tag{22.113}$$

We can easily work out what these expressions are in terms of the sines and cosines of θ and ϕ, just using the addition formulas.

$$x' = x\cos\phi - y\sin\phi, \tag{22.114}$$

$$y' = x\sin\phi + y\cos\phi. \tag{22.115}$$

There's a certain pleasing regularity to the appearance of the cosines and sines. We can get at this more neatly if we *separate* the coordinate variables x and y from the functions which transform them. First, let's consider x and y joined together: they're both labels for marking the same point in space, so we should by rights consider them something of a "married couple". We'll write this combined quantity as a column vector:

$$\begin{pmatrix} x \\ y \end{pmatrix} . \tag{22.116}$$

After a transformation like a rotation, our vector will have primes on its components:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} . \tag{22.117}$$

The vector before the transformation and the new vector produced by the transformation are related, and in writing the vectors in this fashion, we can "pull out" the part of the equations which is the transformation proper, separating it from the variables on which the transformation acts:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} . \tag{22.118}$$

Here, we see matrix multiplication at work:

$$\begin{pmatrix} x\cos\phi - y\sin\phi \\ x\sin\phi + y\cos\phi \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} . \tag{22.119}$$

To get the entry in the first row of the product vector, we multiply the first row of the *rotation matrix* by the first (and only) column of the original vector, item-by-item, and add them up. To get the second row of the result, we multiply the *second* row of the matrix by the column entries, item-by-item, and add *them* up. Here, we're multiplying a 2×2 matrix by a 2×1 matrix to get another 2×1 matrix, but we could multiply bigger matrices. If we were multiplying two 40×40 matrices, the entry in "cellblock 11, 38" would be the product of the entries in the 11^{th} row of the first times those in the 38^{th} column of the second, added together.

Our notion of "married couple" admits a natural generalization. What about transformations in three-dimensional space? We can get a handle on this by identifying a property of two-dimensional rotations, and thus of 2×2 rotation matrices, which holds true in higher dimensions, and then use that to tell us about bigger rotation matrices. The salient feature is that rotations do not change angles between vectors. Imagine two rays coming out of the origin of the 2D plane, at angles θ_1 and θ_2 from the horizontal axis. Crank those rays around all you want, and the angle between them will always be $\theta_2 - \theta_1$. In particular, two *orthogonal* rays will rotate into another pair of orthogonal rays. Recall the helpful fact that the columns of a matrix give the results of transforming the basis vectors by that matrix. So, the columns of a "generalized rotation" matrix must be orthogonal to each other! We can verify this holds true for the rotation matrix we deduced above:

$$\cos\phi \cdot (-\sin\phi) + \sin\phi \cdot \cos\phi = 0 \,, \tag{22.120}$$

as desired. In fact, the columns are not just orthogonal, but orthonormal, because the norm of each column is 1:

$$\cos^2 \phi + \sin^2 \phi = 1 \,. \tag{22.121}$$

If the columns of a matrix M are an orthonormal basis, then taking the transpose of M makes a new matrix M^{T} whose *rows* are orthonormal. Matrix multiplication means taking inner products of rows against columns, so

$$M^{\mathrm{T}} M = I \,. \tag{22.122}$$

In other words, the inverse of M is its transpose. Matrices which enjoy this property are called *orthogonal matrices.*

22.7 Changes of Basis

We developed the rotation matrices by grabbing a vector and moving it through an angle ϕ. We can also think of this transformation as changing our *viewpoint* on the vector, as if we tilted our heads by the opposite angle, $-\phi$. One thing matrices are very good for is changes of perspective, or in the language of linear algebra, changes of basis.

Suppose we pick a basis and then put a lot of work into deducing the matrix that represents a transformation. Then we realize that a different basis is also interesting. Can we find the matrix that encodes our transformation from the perspective of our new basis? The straightforward thing to do is to turn the new basis vectors into the old, apply the transformation, and then convert the result back into the new basis. Let M be the matrix representing the transformation using the old basis, and let T be the matrix that turns the new basis vectors into the old ones. Then our recipe for our new matrix M' is just multiplication:

$$M = T^{-1} M T \,. \tag{22.123}$$

The columns of the matrix T tell us what the new basis vectors look like in the old basis.

22.8 Eigenvalues and Eigenvectors

It is often convenient to characterize a transformation by what it leaves unchanged. For example, we talk of a rotation around a given axis, which moves everything except the points lying along that axis. A "characteristic vector", or *eigenvector,* is a vector which a transformation leaves unchanged, or changes only in a simple way. Specifically, if M is a matrix, and $\vec{v}$ is an eigenvector of M, then

$$M\vec{v} = \lambda \vec{v}, \tag{22.124}$$

for some number λ. This number is called the *eigenvalue* of the matrix M associated with the eigenvector $\vec{v}$.

It is helpful to work out the eigenvalues of a 2×2 matrix, not only because the exercise illustrates the general idea, but also because it's a calculation which comes up over and over again in practice. We start by defining the matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{22.125}$$

Let $\vec{v}$ be an eigenvector of M, with λ its corresponding eigenvalue. Then

$$(M - \lambda I)\vec{v} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{22.126}$$

This will be satisfied if

$$\det(M - \lambda I) = 0. \tag{22.127}$$

The determinant is easy to work out:

$$\begin{aligned} \det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} &= (a-\lambda)(d-\lambda) - bc \\ &= ad - \lambda(a+d) + \lambda^2 - bc \\ &= \lambda^2 - \lambda \mathrm{tr} M + \det M. \end{aligned} \tag{22.128}$$

We solve the quadratic equation for the eigenvalue λ, obtaining the pair of solutions

$$\lambda_\pm = \frac{\mathrm{tr} M \pm \sqrt{(\mathrm{tr} M)^2 - 4 \det M}}{2}. \tag{22.129}$$

You can check from this equation that

$$\lambda_+ + \lambda_- = \mathrm{tr} M, \tag{22.130}$$

and with just a little more work,

$$\lambda_+ \lambda_- = \det M. \tag{22.131}$$

These properties will hold true for larger matrices: The sum of the eigenvalues will be the trace, and the product of the eigenvalues will be the determinant.

The *adjoint* $L^\dagger$ of a matrix L is found by flipping it across its main diagonal and complex-conjugating all of the entries:

$$[L^\dagger]_{ij} = L^*_{ji}. \tag{22.132}$$

One helpful fact about the adjoint is that

$$(LM)^\dagger = M^\dagger L^\dagger. \tag{22.133}$$

If a matrix is equal to its own adjoint, we call it *self-adjoint* or *Hermitian.* If the adjoint of a matrix is the *inverse* of that matrix, then we call that matrix *unitary.* (These are the souped-up, complex counterparts of the orthogonal matrices we introduced earlier.)

The eigenvalues of a self-adjoint matrix are real, and those of a unitary matrix lie on the unit circle in the complex plane $\mathbb{C}$. A matrix L that commutes with its adjoint $L^\dagger$ is called *normal*. Both the self-adjoint and the unitary matrices are examples of normal matrices. Moreover, normal matrices have the property that they can be *diagonalized*: There exists some orthonormal basis $\{v_i\}$ such that

$$L = \sum_i \lambda_i |v_i\rangle\langle v_i| \,, \tag{22.134}$$

where $\{\lambda_i\}$ are the eigenvalues of L. The set of eigenvalues is often called the *spectrum,* and writing a matrix in this way is called its *spectral decomposition.* If two different matrices use the same basis of eigenvectors in their spectral decompositions, they are *simultaneously diagonalizable.* Two self-adjoint matrices are simultaneously diagonalizable when they commute. The ket-bra combinations in this expression are an example of a construction that turns up all the time, wherein we turn a vector $|v\rangle$ into a special kind of linear operator we call a *projector.* A linear operator P is a projector if it satisfies $P^2 = P$. Projectors are self-adjoint by construction.

Suppose that P is the projector defined by the ket-bra combination (or "outer product") of $|v\rangle$ with itself:

$$P = |v\rangle\langle v| \,. \tag{22.135}$$

Now, let $|u\rangle$ be any other vector, and "sandwich" P on both sides:

$$\langle u|P|u\rangle = \langle u|v\rangle\langle v|u\rangle = \langle u|v\rangle(\langle u|v\rangle)^* = |\langle u|v\rangle|^2 \,. \tag{22.136}$$

This is necessarily a real number, and more than that, it must be zero or larger. Consequently, all projectors are what we call *positive semidefinite.* A linear operator is "p.s.d." if it satisfies $\langle u|L|u\rangle \geq 0$ for all vectors $|u\rangle$. If we pick a basis and write a p.s.d. operator as a matrix, that matrix will be self-adjoint, and all its eigenvalues will be nonnegative.

Let $\{|v_i\rangle : i = 1, \ldots, n\}$ be an orthonormal basis. Then

$$\sum_{i=1}^{n} |v_i\rangle\langle v_i| = I \,. \tag{22.137}$$

This is called "providing a *resolution of the identity.*"

Suppose that $|v\rangle$ and $|u\rangle$ are two eigenvectors of the same self-adjoint operator L. Let's say,

$$L|v\rangle = \lambda_v|v\rangle, \;\; L|u\rangle = \lambda_u|u\rangle \,. \tag{22.138}$$

Then we can take the adjoint of either of these equations:

$$\langle u|L = \langle u|\lambda_u \,. \tag{22.139}$$

And combining this with the eigenvector equation for $|v\rangle$, we find

$$\langle u|L|v\rangle = \lambda_u\langle u|v\rangle = \lambda_v\langle u|v\rangle \,. \tag{22.140}$$

The only way to satisfy this equality of $\lambda_u \neq \lambda_v$ is if the other factor is zero. Eigenvectors of a self-adjoint operator that are associated with distinct eigenvalues must be orthogonal.

If functions are secretly vectors, can there be an inner product for them, too? We take inspiration from the complex-conjugated dot product. Suppose that f and g are two functions defined on the real line $\mathbb{R}$. Then we conjugate one of them and multiply it by the other at each point, adding as we go:

$$(f, g) = \int dx\, f(x)^* g(x)\,. \tag{22.141}$$

The norm of a function is its inner product with itself:

$$(f, f) = \int dx\, |f(x)|^2\,. \tag{22.142}$$

In quantum physics, we will care most about functions that are well-behaved in that they have a finite "length", so most of our functions will be *square-integrable.* I.e., we will ensure that the inner product of a function with itself will not diverge.

Another question raised by regarding functions as vectors is whether we can have linear operators that act on them. Sure we can! For example, the act of multiplying by x is a linear operator, because

$$x(\alpha f) = \alpha(xf)\,, \tag{22.143}$$

and

$$x(f + g) = xf + xg\,. \tag{22.144}$$

Taking the derivative of a differentiable function is also a linear operator, since

$$\frac{d}{dx}(\alpha f) = \alpha \frac{df}{dx}\,, \tag{22.145}$$

and also,

$$\frac{d}{dx}(f + g) = \frac{df}{dx} + \frac{dg}{dx}\,. \tag{22.146}$$

We can even take the commutator of these operators, just like we took the commutator of matrices. All we need is a test function f to apply them against:

$$x\frac{d}{dx}f - \frac{d}{dx}xf = x\frac{d}{dx}f - f - x\frac{d}{dx}f = -f\,. \tag{22.147}$$

Because this is true for any function f where these operations make sense, we can abstract the f away and say

$$\left[x, \frac{d}{dx}\right] = -1\,. \tag{22.148}$$

Is there a sensible notion of an *adjoint* for operators on functions? When we dealt with matrices, we defined the adjoint by transposing and complex-conjugating. The

counterpart of that for functions is not immediately clear. But observe how the matrix adjoint plays with the inner product:

$$(|v\rangle, M|u\rangle) = \sum_j v_j^* \sum_k [M]_{jk} u_k \,. \tag{22.149}$$

This is the same thing we'd get if we moved M to the other side of the inner product and took its adjoint:

$$(M^\dagger|v\rangle, |u\rangle) = \sum_k \left(\sum_j [M^\dagger]_{kj} v_j \right)^* u_k = \sum_k \left(\sum_j [M]_{jk}^* v_j \right)^* u_k = \sum_{jk} v_j^* M_{jk} u_k \,. \tag{22.150}$$

We can carry this property over to the general setting and say that the adjoint is the operator for which

$$(f, Lg) = (L^\dagger f, g) \tag{22.151}$$

is satisfied. For example, the adjoint of the operator "multiply by the real number x" is that same operator:

$$(f, xg) = \int dx\, f^* x g = \int dx\, (fx)^* g = (xf, g) \,. \tag{22.152}$$

What about the operator $\frac{d}{dx}$? Here, we get a little sneaky. Recall the product rule:

$$\frac{d}{dx}(fg) = \frac{df}{dx} g + f \frac{dg}{dx} \,. \tag{22.153}$$

By the fundamental theorem of calculus, integrating a derivative gives us the original function, plus an unknown constant. If we integrate both sides of the previous equation, we can absorb those unknown constants into the other integrals:

$$fg = \int dx\, \frac{df}{dx} g + \int dx\, f \frac{dg}{dx} \,. \tag{22.154}$$

This is just the rigmarole for integrating by parts, if we isolate one of those integrals:

$$\int dx\, \frac{df}{dx} g = fg - \int dx\, f \frac{dg}{dx} \,. \tag{22.155}$$

The left-hand side now has the form of an inner product, though in order to get a specific number out of it we'd need to specify a range of integration. Let us say that both f and g go to zero as x goes to $\pm\infty$; this is ensured by their both being square-integrable. Then we integrate over all $\mathbb{R}$ and the fg term vanishes. We get

$$\left(\frac{df}{dx}, g \right) = - \left(f, \frac{dg}{dx} \right) \,. \tag{22.156}$$

Therfore, the adjoint of the operator $\frac{d}{dx}$ is the operator $-\frac{d}{dx}$.

Talking about self-adjoint operators is a little more complicated than self-adjoint matrices, because one runs into the kind of issue that physicists don't like to worry about. To have true self-adjointness, we must ensure that an operator and its adjoint are defined on the same domain. We will avoid issues like this as much as we can; for the bulk of a course like this, operators on functions can be imagined to work like "infinity × infinity" matrices.

22.9 3D Rotations

While rotations in two dimensions commute, rotations about *different axes in three dimensions* don't: In 3D, the order of rotation operations matters. Rotation matrices provide a way of teasing out the essence of this odd behavior and presenting the geometrical phenomenon in a useful form.

First, we need to bulk up our machinery from two dimensions to three. This just means making the matrices bigger.

In 3D, we can rotate around the axis of our choice. To get started, we'll pick three axes such that x denotes the distance to the right, y denotes the distance up and z denotes the distance forward, "out of the chalkboard." Negative numbers imply a placement in the opposite direction.

Up until now, we've been discussing rotations which mix together x and y. In this picture, we can see that such an operation is a turn around the z-axis. We'll denote it by $R_z(\phi)$, where ϕ is the angle through which we turn. In matrix form,

$$R_z(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{22.157}$$

We've just added an extra row and column to go from 2D to 3D; this extra realm is filled with zeros except for the lower right corner, which is 1. This is the matrix way of saying that a rotation around the z-axis leaves z unchanged.

We can also write a matrix for rotation around the x-axis. It will *also* have sines and cosines of the angle ϕ, and a row and column containing zeros except for a solitary 1. Why? Because a rotation around the x-axis leaves the value of x unchanged. We can figure out what goes in the other spaces by observing that a rotation of $\pi/2$ radians (90 degrees) around the x-axis will take a point on the y-axis onto the z-axis. This means that the entry in the second column, third row must be the sine of ϕ (remember that the sine of $\pi/2$ is 1). The same rotation will take a point on the z-axis onto the negative y-axis, so the entry in the third column, second row must be $-\sin\phi$.

$$R_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}. \tag{22.158}$$

We can figure out a matrix for rotations about the y-axis in much the same way.

Skipping ahead to the punchline,

$$R_y(\phi) = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix}. \tag{22.159}$$

Note what happens to each of these matrices when we rotate by *zero* radians. Because the sine of zero is zero and the cosine of zero is one, each of the rotation matrices reduce to the same thing, a 3×3 matrix whose entries are all zero, except along the *diagonal,* where they are all identically 1. This "do nothing" operation is called the *identity,* and the *identity matrix* is

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{22.160}$$

Multiplying any matrix by the identity gives back the original matrix; it's the equivalent of the number 1 on the familiar number line. (Identity matrices exist for all sizes $n \times n$ but here we'll only need the 3×3 version.)

Now that we've graduated high school, it's time to go to college!

We're going to take "baby steps" and consider *infinitesimal* rotations. That is, we're going to look at the form these matrices take when the angle ϕ is very small indeed. Recall that for small angles, the sine of the angle is just the angle (when you're working in radians). Cosines of such small angles are approximately 1. This means that each of the rotation matrices becomes a matrix containing only ones, zeros and two instances of the small angle. Furthermore, each entry on the *diagonal* of a rotation matrix is either 1 or a cosine, but in the small-angle regime, the cosines become 1, so the diagonal is all ones, just like the identity matrix.

We're going to take advantage of this fact and write the rotation matrices for small angles as the identity matrix *plus* another matrix. We can be even sneakier if we realize that the entries in this other matrix are all either 0 or the small angle, which by convention we'll call epsilon, ϵ. We can *factor out* the number ϵ to give a matrix which contains just zeros and ones. (Multiplying a number times a matrix just multiplies each element in the matrix by that number.) Thus we say,

$$R_j(\epsilon) = I + \epsilon J_j. \tag{22.161}$$

Here, the letter j denotes any one of the axes x, y or z — we mean that this equation is true for all of them, and we don't want to write the same thing three times. I get the following for the matrix J_x:

$$J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{22.162}$$

We can turn around and substitute this into the previous equation: multiply each entry by ϵ, add 1 to each diagonal entry, and we get back the small-angle version of R_x.

We can play the same game with the y-axis rotation matrix to get J_y:

$$J_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \tag{22.163}$$

Again, multiply by ϵ and add the identity matrix to check. Finally,

$$J_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{22.164}$$

The three matrices J_j are called the *generators* of the rotation transformations.

Remember, we're trying to study what happens when we perform successive rotations around different axes. So, let's multiply the generators of x- and y-axis rotations. The matrix multiplication isn't so bad, because everything is either 0 or 1, and the result is the following:

$$J_x J_y = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{22.165}$$

But, oh la, what happens when we multiply the same matrices *in the other order*? This is an example of matrix multiplication failing to be commutative.

$$J_y J_x = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{22.166}$$

The two products, taken in opposite orders, are *different!* By how much do they differ? Well, we can just subtract one from the other:

$$J_x J_y - J_y J_x = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = J_z. \tag{22.167}$$

Surprise! The difference is just the *third* generator, J_z.

That's an odd enough coincidence that we should be tempted to press further. What about the generators of x- and z-axis rotations? First, try one order:

$$J_x J_z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{22.168}$$

Then, try the other:

$$J_z J_x = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{22.169}$$

Again, the answers differ, and again, we subtract to find how much they differ:

$$J_z J_x - J_x J_z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} = J_y. \tag{22.170}$$

Double whammy! The "mismatch" between the z and x generators is just the y generator.

Well, now there's nothing stopping us from trying the next pair:

$$J_y J_z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{22.171}$$

Again, we check the multiplication in the other order:

$$J_z J_y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \tag{22.172}$$

The matrices differ by the amount...

$$J_y J_z - J_z J_y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = J_x. \tag{22.173}$$

All of the generators are tied up, somehow, such that combinations of two give back the third.

Here, we introduce a more compact notation, since we're going to be dealing with a good many quantities which fail to commute. For any two thingamajigs A and B,

$$[A, B] = AB - BA. \tag{22.174}$$

In this notation, the results we worked out a moment ago take on an interesting form:

$$[J_x, J_y] = J_z, \ [J_z, J_x] = J_y, \ [J_y, J_z] = J_x. \tag{22.175}$$

Notice the pattern of the subscripts. Each equation lists the generators in the order xyz, but sometimes "wrapped around" with a cyclic permutation.

This is, almost, the form of the rotation operations used in quantum mechanics. The next step, which should be a brief one to cover, will be to roll these three equations into one, taking advantage of their cyclic character.

22.10 Einstein Summation and Levi-Civita Symbols

Earlier, we took a look at rotations and found a way to summarize their behavior using commutator relations. Recall that the commutator of A and B is defined to be

$$[A, B] = AB - BA. \tag{22.176}$$

For real or complex numbers, the commutator vanishes, but as we saw, the commutators of *matrices* can be non-zero and quite interesting. We recognized that this would have to be the case, since we used matrices to describe rotations in three-dimensional space, and rotations about different axes in 3D do not commute. Looking at very small rotations, we also found that the commutators of *rotation generators* were tied up together in a way which involved cyclic permutations. Today, we'll express this discovery more neatly, using the *Einstein summation convention* and a mathematical object called the *Levi-Civita tensor*.

First, let's ground ourselves by noting a few basic properties of commutators. We observe that the commutator of anything with itself vanishes:

$$[A, A] = 0, \tag{22.177}$$

and that swapping the order of the bracketed symbols introduces a minus sign:

$$[A, B] = -[B, A]. \tag{22.178}$$

In the previous section, we looked at *infinitesimal* rotations, where we twisted through a very small angle ϵ, and we found that the matrix for rotating around axis j took the following form:

$$R_j(\epsilon) = I + \epsilon J_j. \tag{22.179}$$

The matrices J_j were comprised of ones and zeros, and they satisfied the following commutator relationships:

$$[J_x, J_y] = J_z,\ [J_z, J_x] = J_y,\ [J_y, J_z] = J_x. \tag{22.180}$$

Alternatively, we saw, these relations could be written this way:

$$[J_1, J_2] = J_3,\ [J_3, J_1] = J_2,\ [J_2, J_3] = J_1. \tag{22.181}$$

Now, we'd like to combine these three equations into one "package."

Let's think for a moment about *vectors,* which we used before as a way of combining several numbers into a "marriage" (with possibly many more than two spouses). We said that a vector has both a magnitude (say, 10 kilometers per hour) and a direction (*e.g.,* 35 degrees west of due north). We can write a vector as a list of numbers, organized if you'd like into a one-column or one-row matrix; the first number might give an easterly speed in kph, the second a northerly speed and the third a vertical speed, for example. Going between the magnitude-direction description and the list of "components" just requires a little trigonometry, which for the 2D case we explored earlier. (Well, you *could* go off and invent yourself a non-orthonormal basis in some abstract vector space, but patience! We're not quite there yet.)

We noted earlier that the dot, or inner, product of two vectors means multiplying their components, one by one, and adding the total up. The dot product of two velocity

vectors in the 2D plane would be, for example, the product of vector 1's easterly component and vector 2's easterly component, *plus* vector 1's northerly component times vector 2's northerly component. That's enough of a mouthful to warrant expressing in symbols instead:

$$\vec{v} \cdot \vec{u} = v_1 u_1 + v_2 u_2 + \cdots + v_N u_N. \tag{22.182}$$

Here, we recognize that a vector could in principle have an arbitrarily large number of components. We can squeeze still harder by introducing a *summation* symbol, a big version of the Greek letter Σ:

$$\vec{v} \cdot \vec{u} = \sum_{i=1}^{N} v_i u_i. \tag{22.183}$$

This says exactly what we said before, but more stylishly. If we trust ourselves to remember where the "index" i starts and stops, we might omit the 1 and N for brevity:

$$\vec{v} \cdot \vec{u} = \sum_{i} v_i u_i. \tag{22.184}$$

The final step in streamlining the notation is due to Einstein, who realized that most of the time when you see indices repeated across variables — like the i we have on two variables here — you're going to be summing over them. So, why keep the summation sign around at all? Einstein found that he didn't have to, and in his honor, we call the idea that *summation is implied whenever you have repeated indices* the "Einstein summation convention." In our example, this means we can write the dot product like this:

$$\vec{v} \cdot \vec{u} = v_i u_i. \tag{22.185}$$

At last, our equations are starting to look like *physics!*

OK, it's time to stare down the commutators. Here is, again, what we're trying to study:

$$[J_1, J_2] = J_3, \ [J_3, J_1] = J_2, \ [J_2, J_3] = J_1. \tag{22.186}$$

Notice that each of these three equations has the form of a commutator between two generators yielding the third generator. We might try to summarize them by saying that in general, the commutator of generator i with generator j is the third generator, k. If we're more careful, we'll note that if i equals j, for example, the result is *zero.* It's also possible to get *minus* a generator on the right-hand side (how?), so we're really talking about getting *something times* that third generator:

$$[J_i, J_j] = (??) J_k. \tag{22.187}$$

Let's give that question mark a name. Whatever it is, its value will depend upon i, j and k. By convention, this object is identified using the Greek letter ϵ — it's not an infinitesimal angle; in fact, it's not an infinitesimal anything, but that's the letter everybody uses. (The Greek alphabet is only so big.) To make sure we don't forget which epsilon we're talking about, and to indicate what it depends upon, we attach i, j and k as subscripts:

$$[J_i, J_j] = \epsilon_{ijk} J_k. \tag{22.188}$$

Aha! We've caught ourselves repeating indices. Since each of i, j and k can be 1, 2 or 3, this equation really means

$$[J_i, J_j] = \epsilon_{ij1}J_1 + \epsilon_{ij2}J_2 + \epsilon_{ij3}J_3. \tag{22.189}$$

Hmmm. Have we made life any easier? That is, we've managed to squeeze three commutator equations into one, but we have to find out what this "epsilon thing" — or *Levi-Civita tensor* — is all about before we can claim success. To begin with, let's look at the commutator of J_1 with J_2. By our formula,

$$[J_1, J_2] = \epsilon_{12k}J_k, \tag{22.190}$$

which by the Einstein summation convention expands out to

$$[J_1, J_2] = \epsilon_{121}J_1 + \epsilon_{122}J_2 + \epsilon_{123}J_3, \tag{22.191}$$

but *since we already know* that J_1 with J_2 gives J_3, we can say that

$$\epsilon_{123} = 1, \tag{22.192}$$

and furthermore that

$$\epsilon_{121} = \epsilon_{122} = 0. \tag{22.193}$$

Following the same procedure with the other two commutators tells us that

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1. \tag{22.194}$$

Here we see the cyclic nature of the commutators making itself manifest: If the subscripts are a *cyclic permutation* of the sequence 123, the value of the Levi-Civita symbol is 1.

Next, we note what happens when we *swap* the variables i and j. We know that in general, swapping the things inside the commutator brackets gives a minus sign:

$$[J_j, J_i] = -\epsilon_{ijk}J_k, \tag{22.195}$$

but looking at the formula where we introduced the Levi-Civita tensor, we see that

$$[J_j, J_i] = \epsilon_{jik}J_k. \tag{22.196}$$

Comparing these two, we observe a neat property:

$$\epsilon_{ijk} = -\epsilon_{jik}. \tag{22.197}$$

The Levi-Civita symbol is what we call *antisymmetric*. A "symmetry" means that we can transform an object and have it look the same as it did before; rotating a symmetrical vase around its axis, for example, yields a configuration which looks the same as the vase before the transformation. Here, we perform an operation — swapping the indices — and we find that the result is almost, but not quite the same as the initial quantity: it's got an extra minus sign.

As we mentioned earlier, for any value of i,

$$[J_i, J_i] = 0, \tag{22.198}$$

which implies that

$$\epsilon_{iik} = 0. \tag{22.199}$$

Note that this is consistent with the antisymmetry: If we *swap* the two instances of the index i, we get that

$$\epsilon_{iik} = -\epsilon_{iik}, \tag{22.200}$$

and the only way for a thing to equal *minus itself* is for that thing to be *zero.* Furthermore, looking at the original three commutators, we see that when a particular index is on the left-hand side, it *can't be on the right.* If J_1 is used on the left, J_1 won't be on the other side. This lets us write that

$$\epsilon_{iji} = \epsilon_{jii} = 0, \tag{22.201}$$

which is just the cyclic permutation of the statement we had two equations ago.

The fact that swapping indices introduces a minus sign means that, for example,

$$\epsilon_{213} = -\epsilon_{123} = -1. \tag{22.202}$$

We can equally well say that

$$\epsilon_{132} = -\epsilon_{312} = -1, \tag{22.203}$$

or that

$$\epsilon_{321} = -\epsilon_{231} = -1. \tag{22.204}$$

Summarizing what we've found,

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1, \tag{22.205}$$

$$\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1. \tag{22.206}$$

This exhausts the six possible ways of writing the numbers 1, 2 and 3 in a chosen order without repeating a digit. We observe that the cyclic permutation property holds whether the result is $+1$ (when the subscripts are "the right way around") or the result is -1 (when two indices have been swapped). Also, if any two indices have the same value, the result is always 0.

Our new toy, the Levi-Civita tensor (also known as the Levi-Civita symbol and the permutation tensor) is a *completely antisymmetric* object. We could define the analogous contraption for any number of dimensions, but we'll work with three most of the time. One finds this object in calculations of matrix determinants, vector cross products and other places, and while all of that will probably come up sooner or later, the use for which we will employ it now is in commutator relations:

$$[J_i, J_j] = \epsilon_{ijk} J_k. \tag{22.207}$$

Incidentally, physicists typically define the rotation generators with a factor of i, the square root of -1, worked in:

$$R_j(\theta) = I - i\theta J_j \text{ (for } \theta \text{ small)}. \tag{22.208}$$

This brings a factor of i into the commutator relation:

$$[J_i, J_j] = i\epsilon_{ijk} J_k. \tag{22.209}$$

Somehow, people live with themselves even after they use the letter i to represent two different things in the same equation.

22.11 Tensor Products and Partial Traces

The inner product of two vectors yields a scalar. We can also define a product that takes in two vectors and produces another vector. One particularly useful way of doing so is the *tensor* or *Kronecker* product. Basically, we use one vector as a template for stamping out copies of the other. For example, start with the two vectors

$$\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}, \ \vec{u} = \begin{pmatrix} c \\ d \end{pmatrix}. \tag{22.210}$$

Then the tensor product of these two vectors is

$$\vec{v} \otimes \vec{u} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}. \tag{22.211}$$

We can also take the tensor product of matrices. For example, let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \ B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}. \tag{22.212}$$

The tensor product of these two matrices is

$$A \otimes B = \begin{pmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & de & df \\ cg & ch & dg & dh \end{pmatrix}. \tag{22.213}$$

So to speak, we slot copies of the second matrix into the shape made by the first.

The tensor product is noncommutative: Generally, $A \otimes B$ will not equal $B \otimes A$. However, it is associative:

$$(A \otimes B) \otimes C = A \otimes (B \otimes C). \tag{22.214}$$

Moreover, the tensor product has a nice interplay with ordinary matrix multiplication:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD). \tag{22.215}$$

The conceptual inverse of taking tensor products is taking *partial traces.* These are written by the trace operator tr with an extra subscript saying what part is being "traced over". We can define

$$\mathrm{tr}_2(A \otimes B) = A\mathrm{tr}B. \tag{22.216}$$

By requiring the action of tr_2 to be *linear* on its input, this actually establishes what tr_2 does on *all* matrices, even those that can't be written as a tensor product. To be more explicit about this, let's work out what the matrix $A\mathrm{tr}B$ is, in terms of the entries of A and B:

$$A\mathrm{tr}B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} (e+h) = \begin{pmatrix} ae+ah & be+bh \\ ce+ch & de+dh \end{pmatrix}. \tag{22.217}$$

Each term in these sums is an element of the matrix $A \otimes B$:

$$A\mathrm{tr}B = \begin{pmatrix} [A\otimes B]_{00} + [A\otimes B]_{11} & [A\otimes B]_{02} + [A\otimes B]_{13} \\ [A\otimes B]_{20} + [A\otimes B]_{31} & [A\otimes B]_{22} + [A\otimes B]_{33} \end{pmatrix}. \tag{22.218}$$

But we can slot any 4×4 matrix into this formula, whether or not it came from a tensor product! If M is an arbitrary 4×4 matrix, then

$$\mathrm{tr}_2 M = \begin{pmatrix} [M]_{00} + [M]_{11} & [M]_{02} + [M]_{13} \\ [M]_{20} + [M]_{31} & [M]_{22} + [M]_{33} \end{pmatrix}. \tag{22.219}$$

If we take the trace of $\mathrm{tr}_2 M$, we get the trace of M, which makes sense. We can define $\mathrm{tr}_1 M$ analogously, by working out and then extending the validity of $(\mathrm{tr}A)B$.

23 Useful Mathematical Background: Differential Equations

23.1 Common Examples

Differential equations are expressions that characterize a function in terms of its derivatives. Finding a function that satisfies a given differential equation is known as "integrating" that differential equation. Sometimes, we can do this on sight. For example, if we are told that

$$\frac{dy}{dx} = m, \tag{23.1}$$

then we know that

$$y = mx + b, \tag{23.2}$$

because any function of the latter form will satisfy the former. As is almost ubiquitous in the land of differential equations, an extra constant has appeared, because differentiating any constant b by x yields zero. Asking that the slope of y is m says nothing about its intercept.

If we are told that

$$\frac{dy}{dx} = y, \tag{23.3}$$

then we can fumble our way through our memories of calculus and realize eventually that any function of exponential form will do:

$$y = be^x. \tag{23.4}$$

Here, the undetermined constant is a multiplicative factor. Very often in physics, we meet differential equations very much like this, only with an extra constant already included:

$$\frac{dy}{dx} = my. \tag{23.5}$$

Integrating this differential equation merely requires combining the previous solution with the chain rule:

$$y = be^{mx}. \tag{23.6}$$

This describes exponential growth if m is positive and exponential decay if m is negative. Often, the input variable is *time,* and we deduce the value of b from the value of y at time zero.

Another type of differential equation that is frequently encountered in physics involves a second derivative with respect to time. If we imagine a particle of mass M moving in one dimension while subject to an force, Newton's second law tells us that

$$M\frac{d^2x}{dt^2} = F(x) = -\frac{dV}{dx}, \tag{23.7}$$

where we have written the force as minus the derivative of the potential energy. A position x_0 is a point of *stable equilibrium* if the force exactly at x_0 vanishes, and if small steps away from x_0 in either direction cause the force to push back towards x_0. This is equivalent to saying that x_0 must be a *minimum* of the potential energy curve $V(x)$. Time and time again in physics, we study a system by finding its points of stable equilibrium and then seeing what happens when we push it slightly from those points. In the bulk of cases, this means approximating the potential energy $V(x)$ near x_0 by a Taylor series, i.e., drawing the parabola that best mimics $V(x)$ when x is close to x_0:

$$V(x) \approx V(x_0) + \frac{1}{2}\left.\frac{d^2V}{dx^2}\right|_{x_0}(x - x_0)^2. \tag{23.8}$$

The second derivative of V, which is also minus the first derivative of the force, measures the steepness of the potential well. We can call it K for brevity. Because x_0 is just a constant value,

$$M\frac{d^2}{dt^2}(x - x_0) = M\frac{d^2x}{dt^2} = -K(x - x_0). \tag{23.9}$$

Introducing the shifted variable $y = x - x_0$,

$$\frac{d^2y}{dt^2} = -\frac{K}{M}y. \tag{23.10}$$

Here, we are being told that the function $y(t)$ reproduces itself, up to a negative constant factor, when differentiated twice. This brings us into the realm of trigonometric functions (§22.3). Recalling that the second derivative of the sine is minus the sine, and likewise for the cosine, we discover

$$y = A\sin(\omega t) + B\cos(\omega t). \tag{23.11}$$

where we have defined $\omega := \sqrt{K/M}$. If we know the position y at time $t = 0$, then we can find B, and if we know the velocity at time zero, then we know the quantity A. We can also use the addition formulae from §22.3 to write this as

$$y = C\sin(\omega t + \phi). \tag{23.12}$$

These have all been examples of *ordinary* differential equations, because the unknown functions have each been functions of a single variable, and we have only taken ordinary derivatives with respect to that variable. A *partial* differential equation involves partial derivatives. Perhaps the first topic where a physics student encounters

these, in the ordinary way of the curriculum, is in the topic of *wave motion.* Waves are disturbances that over time progress through space, and to a good first approximation they often maintain their shape while doing so. This leads us to consider functions of position and time that "slide along" without changing shape. How do we write a function with this property? Let us say that if we take a snapshot of a wave at a particular time, its contours follow a function $f(u)$. The middle of the wave pattern is at $u = 0$. As time flows on, the only thing that really changes is *where the center point of the pattern has moved.* If it started at position 0, then after t seconds, it will be at position vt. So, if we want to know what the wave is like at position x, then we just need to compare that value of x with the center position vt, because only the distance to the center point matters. So, we describe a wave pattern by a function $f(u)$, and to see what the wave is doing at position x and time t, we evaluate $f(x - vt)$.

If we differentiate $f(u)$ with respect to x while holding t constant, then the function f changes value only through how changing x changes u. But

$$\frac{\partial u}{\partial x} = \frac{\partial (x - vt)}{\partial x} = 1. \tag{23.13}$$

Likewise, if we hold x fixed and vary the time t, then $f(u)$ will change based on how much changing t changes u:

$$\frac{\partial u}{\partial t} = -v. \tag{23.14}$$

Consequently, the *second* derivative of f with respect to position is related to its second derivative with respect to time:

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 f}{\partial t^2}. \tag{23.15}$$

To check the units here, note that v is a measure of length per time, so $1/v^2$ has units of time squared per length squared. Differentiating twice by time brings in units of inverse time squared, so both sides have units of inverse length squared (times whatever the units of f might be).

Eq. (23.15) is the *wave equation.* Much of the theory of differential equations in undergraduate physics is the care and feeding of the wave equation, along with the equation for exponential change,

$$\frac{df}{dt} = \lambda f, \tag{23.16}$$

and the *simple harmonic oscillator* equation

$$\frac{d^2 f}{dt^2} = -\omega^2 f. \tag{23.17}$$

We will encounter others as we go.

What if we make the growth (or decay) rate of the exponential dependent upon t? Can we integrate the differential equation

$$\frac{df}{dt} = \lambda(t) f \ ? \tag{23.18}$$

It seems like the solution here should not be too dissimilar to what we had before. We ought to try an exponential again, but it needs to give us a $\lambda(t)$ when we differentiate. Thinking back to our calculus homework, we recall that differentiating e^u with respect to x pulls down, not u itself, but its derivative du/dx. So, we let $\Lambda(t)$ be an integral of $\lambda(t)$:

$$\Lambda(t) = \int^t dt' \, \lambda(t). \tag{23.19}$$

The derivative of this at t is just $\lambda(t)$, by the fundamental theorem of calculus. Now, we evaluate

$$\frac{d}{dt} e^{\Lambda(t)} = \lambda(t) e^{\Lambda(t)}. \tag{23.20}$$

So, any function of the form

$$f(t) = C \exp\left(\int^t dt' \, \lambda(t) \right) \tag{23.21}$$

will solve our differential equation. If $\lambda(t)$ is constant, we simply recover the special case that we had before. Note that we can pick any antiderivative of $\lambda(t)$: A constant shift can be absorbed into the choice of constant factor C out in front of the exponential.

It may so happen that our system of interest is being affected by an external influence whilst it is also doing its own thing. For example, if we charge up a capacitor and connect the two sides of it through a resistor, the voltage across the capacitor will decay exponentially. The process of periodically re-charging the capacitor would require an additional term in our differential equation, representing the extra input as a function of time.

So, let us consider differential equations of the form

$$\frac{df}{dt} - \lambda(t) f = \mu(t). \tag{23.22}$$

The left-hand side represents the system's own internal dynamics, and the right-hand side captures the outside influence. Let $\Lambda(t)$ be an antiderivative of $\lambda(t)$, as before, and consider the derivative of the product $e^{-\Lambda(t)} f$.

$$\begin{aligned} \frac{d}{dt}(e^{-\Lambda(t)} f) &= -\lambda(t) e^{-\Lambda(t)} f + e^{-\Lambda(t)} \frac{df}{dt} \\ &= e^{-\Lambda(t)} \left(\frac{df}{dt} - \lambda(t) f \right) \\ &= e^{-\Lambda(t)} \mu(t). \end{aligned} \tag{23.23}$$

We can integrate both sides:

$$e^{-\Lambda(t)} f = \int^t dt' \, e^{-\Lambda(t')} \mu(t'). \tag{23.24}$$

With one final turn of the algebra crank, we obtain

$$f = f_0 e^{\Lambda(t)} + e^{\Lambda(t)} \int_{t_0}^{t} dt'\, e^{-\Lambda(t')} \mu(t'). \tag{23.25}$$

As usual, we have some undetermined constants, which we have here written f_0 and t_0. A special case worth understanding is if $\lambda(t)$ and $\mu(t)$ are both constant. Then $\Lambda(t)$ can be taken to be λt, and

$$f = f_0 e^{\lambda t} + e^{\lambda t} \int_{t_0}^{t} dt'\, \mu e^{-\lambda t'}. \tag{23.26}$$

Pulling the constant μ out of the integral and evaluating, we get

$$f = f_0 e^{\lambda t} + \left(-\frac{\mu}{\lambda}\right) e^{\lambda t}\, e^{-\lambda t}\Big|_{t_0}^{t} = f_0 e^{\lambda t} - \frac{\mu}{\lambda}(1 - e^{\lambda(t-t_0)}). \tag{23.27}$$

If λ is negative, then the first term is an exponential decay, and the second term in parentheses will be so as well. After an interval of time has passed, then f will look like $-\mu/\lambda$. This is an example of a differential equation displaying *transient* behavior that dies out over time, leaving the long-term behavior governed by the external influence.

There is a certain "rabbit out of a hat" aspect to how we solved Eq. (23.22). Where exactly did the notion come from to take the derivative of that particular product of functions? It works, but without a way to see how one might guess it, the approach feels unsatisfying. This is unhappily emblematic of many lessons in differential equations: The motivation for a technique is not clear, and one is left wondering whether each new method being taught is a trick that only works for the one problem where it was introduced.

The mathematician Gian-Carlo Rota wrote a textbook about differential equations. Later [161], he reflected on the experience:

> One of several unpleasant consequences of writing such a textbook is my being called upon to teach the sophomore differential equations course at MIT. This course is justly viewed as the most unpleasant undergraduate course in mathematics, by both teachers and students. Some of my colleagues have publicly announced that they would rather resign from MIT than lecture in sophomore differential equations. No such threat is available to me, since I am incorrectly labeled as the one member of the department who is supposed to have some expertise in the subject, guilty of writing an elementary textbook still in print.

23.2 Driven, Damped Harmonic Oscillator

If there is any big theme or overarching lesson to learn when the physics department ships you over to the mathematics building to learn "Diff. Eq.," perhaps it is this. *Functions are secretly vectors,* and so manipulations of functions that *respect the*

structure of vector space make for easier problems. By the first part, we mean that given any function f, we can scale it by a (real or complex) number α and get a new, equally legitimate function. We can take any two functions f and g and add their values at each point to obtain a new function $f + g$. The basic manipulations we do with arrows when we learn vectors in high-school physics apply to functions just as well. A linear operator is a map from the world of functions back to itself such that

$$L(\alpha f) = \alpha L(f) \tag{23.28}$$

and

$$L(f + g) = L(f) + L(g)\,. \tag{23.29}$$

For example, consider a body of mass m attached to a Hookian spring, with an external driving force $F(t)$ applied as well. The position $x(t)$ satisfies the differential equation

$$m\frac{d^2x}{dt^2} = F(t) - kx\,, \tag{23.30}$$

which we can rearrange into

$$\ddot{x} + \frac{k}{m}x = \frac{F(t)}{m}\,. \tag{23.31}$$

Multiplying a function by a constant is a linear operation, as is taking the derivative, so this is all saying that a certain linear operator applied to $x(t)$ must give the function $F(t)/m$:

$$L(x) = \frac{F(t)}{m}\,. \tag{23.32}$$

Even though position is a real-number quantity, we can see how L behaves upon complex-valued inputs. All we do is write a complex function $z(t)$ as the sum of a real-valued $x(t)$ and i times a real-valued $y(t)$:

$$L(z) = L(x + iy) = L(x) + iL(y)\,. \tag{23.33}$$

For two complex numbers to be equal, their real parts must match. And for this particular choice of L, a real-valued input will have a real-valued output. Therefore,

$$L(x) = \mathrm{Re}\, L(z)\,. \tag{23.34}$$

If we didn't notice this property of L, or if we had a more difficult problem where that property did not hold true, we could still pull off a similar trick. Suppose that $f(t)$ is a complex-valued function, and $z_1(t)$ and $z_2(t)$ are two functions that satisfy

$$L(z_1) = f,\ L(z_2) = f^*\,. \tag{23.35}$$

Then by linearity we have

$$L(z_1 + z_2) = L(z_1) + L(z_2) = f + f^* = 2\mathrm{Re}\, f\,. \tag{23.36}$$

So, we can pick $f(t)$ such that its *real part* is the physical function we care about, and then a *linear combination* of the two solutions $z_1(t)$ and $z_2(t)$ will give us the physical answer that we want.

This really pays off when the driving force is sinusoidal. Let its amplitude be F and its frequency be ω. Then we can write

$$L(z) = \ddot{z} + \frac{k}{m}z = \frac{F}{m}e^{i\omega t}\,. \tag{23.37}$$

We know that differentiating has to give an exponential out, so we put an exponential in, and we guess a solution of the form

$$z(t) = z_0 e^{i\omega t}\,. \tag{23.38}$$

Differentiating with respect to time gives

$$\dot{z} = i\omega z,\ \ddot{z} = -\omega^2 z\,. \tag{23.39}$$

Our differential equation becomes merely an algebraic equation

$$-\omega^2 z + \frac{k}{m}z = \frac{F}{m}e^{i\omega t}\,, \tag{23.40}$$

in which we can cancel the time dependence from both sides:

$$-\omega^2 z_0 + \frac{k}{m}z_0 = \frac{F}{m}\,. \tag{23.41}$$

The constant we left unspecified is then

$$z_0 = \frac{F/m}{(k/m) - \omega^2}\,. \tag{23.42}$$

We can incorporate the influence of *damping* in addition to driving by supposing that a force proportional to the speed counteracts the motion:

$$m\ddot{x} = F(t) - m\gamma\dot{x} - kx\,, \tag{23.43}$$

or, introducing $\omega_0^2 = k/m$ as the square of the natural frequency,

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = \frac{F(t)}{m}\,. \tag{23.44}$$

The complex generalization of this for a cosine-shaped $F(t)$ is

$$\ddot{z} + \gamma\dot{z} + \omega_0^2 z = \frac{F}{m}e^{i\omega t}\,. \tag{23.45}$$

As before, we will want to cancel the time dependence, so we try a $z(t)$ of the same form. The derivatives again become multiplications, and

$$z_0 = \frac{F/m}{\omega_0^2 - \omega^2 + i\gamma\omega}\,. \tag{23.46}$$

This $z(t)$ is a periodic function: It describes an oscillator that keeps doing the same thing for as long as $F(t)$ does. But with damping comes the possibility of *transients,* wobbles that die out as the friction bleeds their energy away. To understand this aspect of the oscillator's behavior, let us set the driving force to zero.

$$\ddot{z} + \gamma\dot{z} + \omega_0^2 z = 0\,. \tag{23.47}$$

A decaying oscillation ought to look something like a trig function (an exponential of an imaginary number) curtailed by a factor that shrinks over time. We guess that the latter will also be exponential, and so we have the exponential of a pure imaginary times the exponential of a real value, which comes together to make

$$z(t) = z_0 e^{i\alpha t} \tag{23.48}$$

for some α we have yet to figure out. Plugging this into our differential equation, we obtain

$$(-\alpha^2 + i\gamma\alpha + \omega_0^2) z_0 e^{i\alpha t} = 0\,. \tag{23.49}$$

This is satisfied if $z = 0$, which means that nothing is happening at all, or if

$$\alpha^2 - i\gamma\alpha - \omega_0^2 = 0\,, \tag{23.50}$$

which is more interesting. Solving the quadratic, we find

$$\alpha = \frac{i\gamma \pm \sqrt{-\gamma^2 + 4\omega_0^2}}{2} = \frac{i\gamma}{2} \pm \sqrt{\omega_0^2 - \frac{1}{4}\gamma^2}\,. \tag{23.51}$$

Our solution $z(t)$ involves the exponential of this times i, so the real part, the exponential decay, is governed by $\gamma/2$, while the imaginary part, which tells us the frequency of oscillation, is the natural frequency ω_0 modified by a damping contribution. All quite reasonable! We'll call the modified frequency ω_γ for short:

$$\omega_\gamma = \sqrt{\omega_0^2 - \frac{1}{4}\gamma^2}\,. \tag{23.52}$$

We have two solutions $z_\pm(t)$ corresponding to the two solutions of the quadratic formula. This is just what we want, because at the end we want a real-valued $x(t)$ for our physical solution. This means taking a linear combination in which the imaginary parts cancel each other out:

$$x(t) = e^{-\gamma t/2}(A e^{i\omega_\gamma t} + A^* e^{i\omega_\gamma t})\,, \tag{23.53}$$

where A is a number we have to get from other information, like the initial conditions of a particular scenario.

Now for the real fun. Suppose that z_1 is a solution of the damped harmonic oscillator equation with a driving term $(F/m)e^{i\omega t}$, and z_2 is a solution of the same equation but without driving. In symbols,

$$L(z_1) = \frac{F}{m} e^{i\omega t}\,, \tag{23.54}$$

while

$$L(z_2) = 0\,. \tag{23.55}$$

Then, by linearity,

$$L(z_1 + z_2) = L(z_1) + L(z_2) = \frac{F}{m}e^{i\omega t} + 0 = \frac{F}{m}e^{i\omega t}\,. \tag{23.56}$$

The *superposition* of a periodic solution and a transient solution will again solve the equation with driving included!

23.3 Fourier Transforms

Another thing we do with vectors is to take their dot product. This is closely related to yet another very important thing, writing a vector in a basis. If $\vec{v}$ lives in 3-dimensional space, for example, we know that

$$\vec{v} = v_x\hat{x} + v_y\hat{y} + v_z\hat{z}\,, \tag{23.57}$$

and

$$\vec{v}\cdot\hat{x} = v_x\,, \tag{23.58}$$

because the dot product of perpendicular vectors is zero. So,

$$\vec{v} = (\vec{v}\cdot\hat{x})\hat{x} + (\vec{v}\cdot\hat{y})\hat{y} + (\vec{v}\cdot\hat{z})\hat{z}\,. \tag{23.59}$$

A function on the line $\mathbb{R}$ works in much the same way! The dot product of two vectors $\vec{v}$ and $\vec{w}$ in three-dimensional space is obtained by multiplying their corresponding components and adding up the products. Let's invent a dot product for functions that does the same thing:

$$f\cdot g = \int dx\, f(x)g(x)\,. \tag{23.60}$$

This is fine provided that our functions f and g are both real-valued. If they can be complex-valued, it makes sense to take the complex conjugate of one of them, so that the dot product of a function with itself (which should be the square of a "length") always works out to be a real number.

$$f\cdot g = \int dx\, f(x)^* g(x)\,. \tag{23.61}$$

What functions can serve the role of basis vectors, like $\hat{x}$ and company? These functions ought to be *orthogonal* to one another, with respect to the dot product we have invented.

Consider the integral

$$\int_{-\pi}^{\pi} d\theta\, \cos(\theta)\sin(\theta)\,. \tag{23.62}$$

Referring back to the unit circle, we know that $\cos(-\theta) = \cos(\theta)$, whereas $\sin(-\theta) = -\sin(\theta)$. Consequently, for each point in the region over which we are integrating, there

is another point on the opposite side of the origin at which the value of the integrand is exactly negated. Adding up over the whole interval, then, all the contributions must cancel out completely:

$$\int_{-\pi}^{\pi} d\theta \, \cos(\theta)\sin(\theta) = 0\,. \tag{23.63}$$

This integrand will vanish when integrated over many other regions, too. For example,

$$\sin(\pi+\theta) = -\sin(\theta),\ \cos(\pi+\theta) = -\cos(\theta)\,, \tag{23.64}$$

and so

$$\cos(\pi+\theta)\sin(\pi+\theta) = \cos(\theta)\sin(\theta)\,, \tag{23.65}$$

which in turn implies that

$$\int_{-\pi}^{\pi} d\theta \, \cos(\pi+\theta)\sin(\pi+\theta) = 0\,. \tag{23.66}$$

Changing variables to $\theta' = \pi + \theta$, we obtain

$$\int_{0}^{2\pi} d\theta' \, \cos(\theta')\sin(\theta') = 0\,. \tag{23.67}$$

Symmetry also tells us that sines and cosines of different frequencies are still orthogonal to each other:

$$\int_{-\pi}^{\pi} d\theta \, \cos(n\theta)\sin(m\theta) = 0\,, \tag{23.68}$$

for integers n and m. It is more difficult to show that sine waves of different frequencies are also orthogonal to each other, but this can be done with some trig-identity manipulation, and likewise for cosine waves of different frequencies.

The essential idea of the whole Fourier business is using trig functions as basis vectors! To be more sophisticated, we can also work with complex exponentials instead. We also have the freedom to pick the interval over which we integrate to suit the problem at hand. For example, taking the interval $0 \le t \le T$, we can expand a function $f(t)$ as

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos\left(\frac{2\pi n t}{T}\right) + b_n \sin\left(\frac{2\pi n t}{T}\right)\right]\,, \tag{23.69}$$

where

$$a_n = \frac{2}{T}\int_0^T f(t)\cos\left(\frac{2\pi n t}{T}\right)\,, \tag{23.70}$$

$$b_n = \frac{2}{T}\int_0^T f(t)\sin\left(\frac{2\pi n t}{T}\right)\,. \tag{23.71}$$

We can also center our interval upon the origin, $-T/2 \le t \le T/2$, and use the same formulae except for changing the limits of integration. And converting over to complex numbers, we can expand $f(t)$ as

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{in\omega_0 t} \,, \tag{23.72}$$

where $\omega_0 = 2\pi/T$ and

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-in\omega_0 t} dt \,. \tag{23.73}$$

23.4 Green's Functions

The classic example is the electric potential due to a point charge in otherwise empty space. This means solving Poisson's equation for a source that is a delta-function spike in the void. Let's say that the spike is at $\vec{r}$, and so

$$\vec{\nabla}^2 G(\vec{r}, \vec{s}) = -4\pi\delta(\vec{r} - \vec{s}) \,, \tag{23.74}$$

where the derivatives act upon $\vec{s}$. The delta function depends only upon the magnitude of its input, and the boundary conditions are rotationally symmetric. So, the Green's function can only depend upon the distance between $\vec{r}$ and $\vec{s}$:

$$G(\vec{r}, \vec{s}) = G(|\vec{r} - \vec{s}|) \,. \tag{23.75}$$

Without loss of generality, we move the origin to $\vec{r}$ and go over to spherical coordinates, in which

$$\frac{1}{s^2} \frac{d}{ds} \left(s^2 \frac{dG}{ds} \right) = -4\pi\delta(\vec{s}) \,. \tag{23.76}$$

Next, we integrate both sides over a spherical volume of radius R. The sifting property of the delta function gives the integral of the right-hand side, and integrating the left-hand side means summing over spherical shells of area $4\pi s^2$. After cancellations, we obtain

$$R^2 \frac{dG}{dR} = -1 \,. \tag{23.77}$$

Consequently, the radial dependence of the Green's function G is an elementary integral of a polynomial. Moving the spike to wherever we want, we have

$$G(\vec{r} - \vec{s}) = \frac{1}{|\vec{r} - \vec{s}|} \,. \tag{23.78}$$

23.5 Legendre Transformation

Let $f(x)$ be a smooth, convex function, i.e., a function whose second derivative is everywhere positive. Think "parabola", perhaps with some wobbles. For convenience,

suppose that $f(x) > 0$ everywhere. The slope of $f(x)$,

$$s(x) = \frac{df}{dx}\,, \tag{23.79}$$

is always increasing. The tangent to the curve f at the point x will have slope $s(x)$, of course. But where will that tangent line intersect the y-axis? The tangent line provides the hypotenuse to a right triangle whose horizontal leg has length x and whose vertical leg has length $xs(x)$. The segment of the vertical leg that is above the x-axis is $f(x)$. Call the length of the other segment, the part below the axis, $g(x)$. We then have

$$f(x) + g(x) = xs(x)\,. \tag{23.80}$$

The function g tells us the y-intercept of the tangent line:

$$g(x) = x\frac{df}{dx} - f(x)\,. \tag{23.81}$$

But if the slope s is always increasing as x increases, then we can solve for x in terms of s; each quantity tracks the other. Therefore, we can also say

$$g(x(s)) = x(s)s - f(x(s))\,. \tag{23.82}$$

From our original function of x, we have constructed a new function of s. This is the essence of the Legendre transformation: trading off between coordinate and slope [162].

24 Supplement on Classical Dynamics

24.1 Introductory Lagrangian Mechanics

In our youth, we encounter momentum generally by being told that it is equal to mass times velocity. We then learn, or notice, that the time derivative of mv is the force; momentum is that which is changing when a nonzero net force is present. At a still later stage, we encounter the radical thought that momentum is that which is conserved when the dynamics are invariant under spatial translation.

To do the classical point particle in the Lagrangian style, we associate a number to each possible path between an initial position and time (x_i, t_i) and a final position and time (x_f, t_f). This number, the *action,* is the integral of the *Lagrangian,* which is the difference between the particle's kinetic and potential energies:

$$S = \int L(t)dt = \int (K - V)dt. \tag{24.1}$$

Using the nonrelativistic expression for kinetic energy,

$$S = \int_{t_i}^{t_f} \left[\frac{1}{2} m(\dot{x}(t))^2 - V(x(t)) \right] dt\,. \tag{24.2}$$

We now seek the path for which the action is *stationary.* That is, we want the path with the property that, under slight perturbations, the action remains unchanged. (This is an elaboration of the idea from basic calculus that at an extremum of a function, a small shift of the input causes negligible change to the output.) If we perturb the trajectory from $x(t)$ to a new function $x(t) + \delta x(t)$, then to first order, the action changes by an amount

$$\delta S = \int_{t_i}^{t_f} \left[m\dot{x}(t) \frac{d}{dt} \delta x(t) - V'(x(t)) \delta x(t) \right] dt\,. \tag{24.3}$$

We can pull out the $\delta x(t)$ using integration by parts. This creates a pair of boundary terms that vanish thanks to the premise that $\delta x(t_i) = \delta x(t_f) = 0$. What's left over is

$$\delta S = \int_{t_i}^{t_f} \delta x(t) \left[-m\ddot{x}(t) - V'(x(t)) \right] dt\,. \tag{24.4}$$

This will vanish as desired for an arbitrary perturbation if the quantity in brackets equals zero at all times. But that just means we have Newton's second law: Mass times acceleration is minus the derivative of the potential energy, which is the force.

In terms of the Lagrangian $L = K - V$, the $-V'$ term in the above is the derivative of L with respect to x, and the $-m\ddot{x}$ term is minus the time derivative of the derivative of L with respect to $\dot{x}$. Phrased this way, Newton's second law is an example of an *Euler–Lagrange equation,*

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = 0\,. \tag{24.5}$$

This form is amenable to generalization to more complicated Lagrangians.

If the variable x does not appear in our Lagrangian, then the Euler–Lagrange equation tells us that

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = 0\,. \tag{24.6}$$

Or, in other words, the quantity

$$p_x = \frac{\partial L}{\partial \dot{x}} \tag{24.7}$$

is constant in time. This is a very easy special case of *Noether's theorem,* which tells us that symmetries of a Lagrangian imply conservation laws. Momentum in the x direction is the "Noether charge" whose conservation is deduced from a Lagrangian's invariance under spatial translations.

24.2 From Lagrangians to Hamiltonians

Let us follow a particle along its physical path and see how the Lagrangian changes over time. The Lagrangian can change value if the position x changes, or if the velocity $\dot{x}$ changes:

$$\frac{dL}{dt} = \frac{\partial L}{\partial x}\frac{dx}{dt} + \frac{\partial L}{\partial \dot{x}}\frac{d\dot{x}}{dt}\,. \tag{24.8}$$

From the Euler–Lagrange equation, we recognize the partial derivative in the first term as the time derivative of the momentum. And the second term is the momentum times the second derivative of the position:

$$\frac{dL}{dt} = \frac{dp_x}{dt}\frac{dx}{dt} + p_x\frac{d^2x}{dt^2}\,. \tag{24.9}$$

The right-hand side here has product-rule form.

$$\frac{dL}{dt} = \frac{d}{dt}\left(p_x\frac{dx}{dt}\right)\,. \tag{24.10}$$

Therefore, along the physical path, as long as L has no intrinsic time dependence, the quantity

$$H = p_x\frac{dx}{dt} - L \tag{24.11}$$

is constant over time. This is the *Hamiltonian* for the point particle. Instead of being a function of x and $\dot{x}$ like the Lagrangian, it is regarded as a function of x

and p. Mathematically speaking, the Hamiltonian is the Legendre transform of the Lagrangian. *Hamilton's equations* specify the dynamics:

$$\frac{dx}{dt} = \frac{\partial H}{\partial p}\,, \quad \frac{dp}{dt} = -\frac{\partial H}{\partial x}\,. \tag{24.12}$$

If H is independent of x, then p is constant. The Hamiltonian language for classical physics treats position and momentum on fairly even footing, regarding them both as inputs for the Hamiltonian function. On the other hand, in the Lagrangian language, the trajectory is paramount, and momentum is a derived quantity.

It is often, but not always, the case that the Hamiltonian is the sum of kinetic and potential energies.

24.3 Tryptich of Examples: Bead on a Rotating Hoop

In this section, we will solve a problem using Newtonian language, followed by Lagrangian methods and then again with Hamiltonian techniques. This will illustrate how to interconvert between these ways of thinking and avoid at least one common pitfall. Our problem of interest is a bead on a rotating hoop. Suppose that we have a circular wire hoop of radius R that rotates around the vertical axis, and call its angular speed about that axis Ω. A small bead of mass m slides along the hoop. Denote by θ the angle from the bottom of the hoop. We want to understand how the bead moves: What differential equation can we write for θ as a function of time? Does the bead have any equilibrium points? If it does, and they are stable, how does the bead oscillate around them?

This will involve a lot of equations. Before calculating anything, what are the answers?

For simplicitly, begin by imagining for a moment that the hoop is not rotating. Then, intuitively, there should be two equilibrium points, one at the top ($\theta = \pi$) and one at the bottom ($\theta = 0$). Displace the bead from the lowest point, and it will slide back under gravity, so the $\theta = 0$ equilibrium point will be stable. Put the bead at the top of the hoop and give it a tiny push, and gravity will slide it further down, so the $\theta = \pi$ equilibrium point will be unstable. Now, if the hoop is spinning about its vertical axis fast enough, the bead will be pinned to the side, so the rotating hoop can have two more equilibrium points. The $\theta = \pi$ fixed point is still unstable: centrifugal force and gravity both act to slide the bead away from it. The fixed points on the sides will be stable, and the fixed point at $\theta = 0$ will become unstable.

What is "fast enough"? By dimensional analysis, the only way to combine the information we're given and get a quantity with units of s^{-1} is to take $\sqrt{g/R}$. So, up to numerical prefactors, this should be our first guess for a frequency of oscillation or an angular speed of rotation.

To do a Newtonian analysis, we need to add up the forces. First of all, gravity will always be pulling the bead down, with a force of magnitude mg. Contact forces from the hoop will keep the bead on its track. It is most convenient to work in a reference frame which is rotating with the hoop: this means we only have to keep track of the

one dynamical variable, $\theta(t)$. Like all forces which appear when we transform into noninertial frames, the centrifugal force will be proportional to the mass m. It also depends on the angular speed, Ω, and the bead's current radius, which by trigonometry is $R\sin\theta$. So, we have mg downward, $m\Omega^2 R\sin\theta$ outward and the contact force from the hoop, which ensures that the net force is tangential to the hoop. Calling $\hat{y}$ the upward unit vector and $\hat{x}$ the outward unit vector (in the rotating frame), we need the radial and tangential components of $-\hat{y}mg + \hat{x}m\Omega^2 R\sin\theta$.

The vector from the center of a unit circle to the point θ radians from its lowest point is $\hat{x}\sin\theta - \hat{y}\cos\theta$. The radial vector pointing from the circle back to its center is the negative, $-\hat{x}\sin\theta + \hat{y}\cos\theta$. The tangent line at that point is perpendicular to the radius vector, so the vector of the tangent line is $\hat{x}\cos\theta + \hat{y}\sin\theta$. To check, note that at $\theta = 0$, the tangent vector is horizontal, and at $\theta = \frac{\pi}{2}$ and $\theta = \frac{3\pi}{2}$, it is vertical. Dotting the radial vector with the forces deduced earlier gives

$$F_T = -mg\sin\theta + m\Omega^2 R\sin\theta\cos\theta\,. \tag{24.13}$$

Equating the tangential force with $mR\ddot{\theta}$,

$$\ddot{\theta} = -\frac{g}{R}\sin\theta + \Omega^2\sin\theta\cos\theta\,. \tag{24.14}$$

For $\Omega = 0$ and θ small, this reduces to the simple harmonic oscillator with frequency $\sqrt{g/R}$. When $\Omega = 0$, $\ddot{\theta}$ vanishes if $\theta = 0$ or $\theta = \pi$, and if $\Omega = \sqrt{\frac{g}{R}}$,

$$\ddot{\theta} = -\Omega^2\sin\theta(1-\cos\theta)\,. \tag{24.15}$$

A third equilibrium appears if Ω is sufficiently large. Take

$$\frac{g}{R\Omega^2}\sin\theta = \sin\theta\cos\theta\,, \tag{24.16}$$

and note that this can be satisfied if

$$\cos\theta = \frac{g}{R\Omega^2}\,, \tag{24.17}$$

which has a solution if

$$\frac{g}{R\Omega^2} \le 1\,,\ \text{or}\ \Omega \ge \sqrt{\frac{g}{R}}\,. \tag{24.18}$$

The critical rotation rate emerges as expected.

To solve this problem with a Lagrangian or Hamiltonian approach, we consider the energy of the system. Angular kinetic energy depends on the moment of inertia and the angular speed, as $\frac{1}{2}I\omega^2$. We have two types of rotation in this problem: the hoop spinning on its axis and the bead moving radially about the hoop. If the bead were held at a fixed angle θ, then the hoop spin alone would give it a kinetic energy of $\frac{1}{2}mR^2\Omega^2\sin^2\theta$. Its gravitational potential energy would be $mgR(1-\cos\theta)$. The kinetic energy due to changing θ is the kinetic energy of a point mass rotating in a circle of radius R at angular speed $\dot{\theta}$, so the total kinetic energy is

$$K = \frac{1}{2}mR^2(\dot{\theta}^2 + \Omega^2\sin^2\theta)\,. \tag{24.19}$$

To find the Lagrangian for a mechanical system, we take the difference between kinetic and potential energies:

$$L(\theta, \dot{\theta}) = \frac{1}{2} mR^2(\dot{\theta}^2 + \Omega^2 \sin^2\theta) - mgR(1 - \cos\theta)\,. \tag{24.20}$$

The key quantities are the partial derivatives of the Lagrangian:

$$\frac{\partial L}{\partial \theta} = mR^2\Omega^2 \sin\theta\cos\theta - mgR\sin\theta, \tag{24.21}$$

$$\frac{\partial L}{\partial \dot{\theta}} = mR^2\dot{\theta}\,. \tag{24.22}$$

The Euler–Lagrange equation,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) = \frac{\partial L}{\partial \theta}\,, \tag{24.23}$$

becomes

$$mR^2\ddot{\theta} = mR^2\Omega^2 \sin\theta\cos\theta - mgR\sin\theta\,. \tag{24.24}$$

Dividing through by mR^2, we get

$$\ddot{\theta} = -\frac{g}{R}\sin\theta + \Omega^2 \sin\theta\cos\theta\,, \tag{24.25}$$

as before.

Our first stab at writing a Hamiltonian is to sum of the kinetic and potential energies, written in terms of θ and the canonical momentum conjugate to it, p_θ. We find this canonical momentum by taking the partial derivative of the Lagrangian with respect to $\dot{\theta}$:

$$p_\theta = mR^2\dot{\theta}\,. \tag{24.26}$$

We attempt to form the Hamiltonian:

$$H(\theta, p_\theta) = \frac{p_\theta^2}{2mR^2} + \frac{mR^2\Omega^2}{2}\sin^2\theta + mgR(1 - \cos\theta)\,. \tag{24.27}$$

Hamilton's equations of motion are

$$\dot{\theta} = \frac{\partial H}{\partial p_\theta}\,, \tag{24.28}$$

$$\dot{p}_\theta = -\frac{\partial H}{\partial \theta}\,. \tag{24.29}$$

Consequently,

$$\dot{\theta} = \frac{p_\theta}{mR^2}\,, \tag{24.30}$$

$$\dot{p}_\theta = -mgR\sin\theta - mR^2\Omega^2\sin\theta\cos\theta\,. \tag{24.31}$$

Taking another time derivative of the first equation,

$$\ddot{\theta} = \frac{\dot{p}_\theta}{mR^2} \,. \tag{24.32}$$

So, dividing our expression for $\dot{p}_\theta$ by mR^2, we get

$$\ddot{\theta} = -\frac{g}{R}\sin\theta - \Omega^2 \sin\theta\cos\theta \,. \tag{24.33}$$

This doesn't look right — a sign flipped!

Let's go back to the definition of the Hamiltonian as the Legendre transform of the Lagrangian:

$$H = p_\theta \dot{\theta} - L \,. \tag{24.34}$$

Working this out yields

$$H = mR^2\dot{\theta}^2 - \frac{1}{2}mR^2(\dot{\theta}^2 + \Omega^2\sin^2\theta) + mgR(1-\cos\theta) \tag{24.35}$$

$$= \frac{1}{2}mR^2\dot{\theta}^2 - \frac{1}{2}mR^2\Omega^2\sin^2\theta + mgR(1-\cos\theta) \,. \tag{24.36}$$

In terms of the canonical variables,

$$H(\theta, p_\theta) = \frac{p_\theta^2}{2mR^2} - \frac{1}{2}mR^2\Omega^2\sin^2\theta + mgR(1-\cos\theta) \,. \tag{24.37}$$

Using *this* Hamiltonian,

$$\dot{\theta} = \frac{p_\theta}{mR^2} \,, \tag{24.38}$$

$$\dot{p}_\theta = mR^2\Omega^2\sin\theta\cos\theta - mgR\sin\theta \,. \tag{24.39}$$

Now, taking a time derivative and dividing through gives

$$\ddot{\theta} = -\frac{g}{R}\sin\theta + \Omega^2\sin\theta\cos\theta \,, \tag{24.40}$$

once again. The crucial point is that the kinetic energy depends on θ as well as on $\dot{\theta}$.

This is a good time to check the stability of the fixed points. Stability requires

$$\frac{\partial^2 H}{\partial\theta^2} > 0 \,. \tag{24.41}$$

We've already taken one derivative with respect to θ, so

$$\frac{\partial^2 H}{\partial\theta^2} = mR^2\Omega^2(\sin^2\theta - \cos^2\theta) + mgR\cos\theta \,. \tag{24.42}$$

When $\theta = 0$,

$$\frac{\partial^2 H}{\partial\theta^2} = -mR^2\Omega^2 + mgR \,, \tag{24.43}$$

which is positive if $\Omega < \sqrt{g/R}$. When $\theta = \pi$,

$$\frac{\partial^2 H}{\partial \theta^2} = -mR^2\Omega^2 - mgR\,, \tag{24.44}$$

which is always negative, so the upper fixed point is always unstable. The third and fourth fixed points satisfy $\cos\theta = \frac{g}{R\Omega^2}$, so it's easiest to rewrite the stability condition in terms of cosines alone, using the Pythagorean identity:

$$\frac{\partial^2 H}{\partial \theta^2} = mR^2\Omega^2(1 - 2\cos^2\theta) + mgR\cos\theta \tag{24.45}$$

$$= mR^2\Omega^2\left(1 - 2\frac{g^2}{R^2\Omega^4}\right) + \frac{mg^2}{\Omega^2} \tag{24.46}$$

$$= mR^2\Omega^2 - \frac{mg^2}{\Omega^2}\,. \tag{24.47}$$

This is positive if

$$\Omega^4 > \frac{g^2}{R^2}\,, \tag{24.48}$$

which is just the condition for the fixed point to exist in the first place. So, all the fixed points have the stability properties which we guessed in the beginning.

Finally, consider small displacements near $\theta = 0$ when oscillations near there can be expected to be stable. The equation of motion can be written

$$\ddot{\theta} = \left(\Omega^2\cos\theta - \frac{g}{R}\right)\sin\theta\,. \tag{24.49}$$

If we replace $\sin\theta$ with θ and $\cos\theta$ with 1, then

$$\ddot{\theta} = \left(\Omega^2 - \frac{g}{R}\right)\theta\,, \tag{24.50}$$

which has the form of the simple harmonic oscillator with a frequency determined by how far the hoop rotation speed differs from the critical value. The critical value separates sinusoidal from exponential behavior. But what happens at the critical value itself?

The equation of motion is

$$\ddot{\theta} = -\Omega^2\sin\theta(1 - \cos\theta)\,. \tag{24.51}$$

We can Taylor expand this around $\theta = 0$ by finding its derivative there:

$$\frac{d}{d\theta}\left[-\Omega^2\sin\theta(1-\cos\theta)\right] = -\Omega^2\left[\cos\theta(1-\cos\theta) + \sin\theta(1+\sin\theta)\right] \tag{24.52}$$

$$= -\Omega^2\left[\cos\theta + \sin\theta + \sin^2\theta - \cos^2\theta\right]. \tag{24.53}$$

Evaluated at $\theta = 0$, this vanishes. A second derivative yields

$$\frac{d^2}{d\theta^2}\left[-\Omega^2\sin\theta(1-\cos\theta)\right] = -\Omega^2\left[-\sin\theta + \cos\theta + 4\sin\theta\cos\theta\right]. \tag{24.54}$$

Evaluating this at $\theta = 0$ gives $-\Omega^2$. Having to go beyond the first derivative is an indication that the oscillations are in fact anharmonic.

If Ω becomes very big, then the stable equilibria approach the equator. Gravity becomes irrelevant, and the motion is that due to centrifugal force only, so we get harmonic oscillations with frequency Ω.

24.4 The Maxwell Equations and Electromagnetic Waves

Now and then, we need some results from classical electromagnetism in our pocket, so let's work a bit with the Maxwell equations. In cgs units, where the dimensions of the electric and magnetic fields are the same, these equations are the following:

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho \tag{24.55}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{24.56}$$

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t} \tag{24.57}$$

$$\vec{\nabla} \times \vec{B} = \frac{4\pi}{c}\vec{\jmath} + \frac{1}{c}\frac{\partial \vec{E}}{\partial t}\,. \tag{24.58}$$

The second equation, saying that $\vec{B}$ is divergenceless, incorporates the fact that no magnetic monopole has ever been observed. An exam problem once tricked me quite completely by asking about a situation where there was an infinite flat sheet of material in the (x, y) plane; above the plane, the magnetic field was $B\hat{z}$, and below the plane, it was $-B\hat{z}$. What is the flat sheet made out of? Under the clock, I didn't have the wit to see that this configuration of the magnetic field is just like the *electric* field due to a uniform flat sheet of electric charge, so the flat sheet in this problem is actually a uniform density of magnetic monopole material!

Moving right along from my embarrassing memories, let's take the divergence of the curl of $\vec{B}$:

$$\vec{\nabla} \cdot (\vec{\nabla} \times \vec{B}) = \frac{4\pi}{c}\vec{\nabla} \cdot \vec{\jmath} + \frac{1}{c}\frac{\partial}{\partial t}(\vec{\nabla} \cdot \vec{E}) = 0\,. \tag{24.59}$$

Using Gauss' law,

$$\frac{4\pi}{c}\vec{\nabla} \cdot \vec{\jmath} + \frac{1}{c}\frac{\partial}{\partial t}(4\pi\rho) = 0\,. \tag{24.60}$$

Canceling the common factors yields the statement that charge is continuously conserved:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{\jmath} = 0\,. \tag{24.61}$$

To find wave equations for the electric and magnetic fields, we take the curl of the

curl:

$$\begin{aligned}\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) &= \vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E} \\ &= -\frac{1}{c}\frac{\partial}{\partial t}(\vec{\nabla} \times \vec{B})\,. \end{aligned} \tag{24.62}$$

Using the first and fourth Maxwell equations, we get

$$\begin{aligned}4\pi\vec{\nabla}\rho - \vec{\nabla}^2\vec{E} &= -\frac{1}{c}\frac{\partial}{\partial t}\left(\frac{4\pi}{c}\vec{\jmath} + \frac{1}{c}\frac{\partial \vec{E}}{\partial t}\right) \\ &= -\frac{4\pi}{c^2}\frac{\partial \vec{\jmath}}{\partial t} - \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2}\,. \end{aligned} \tag{24.63}$$

Rearranging terms provides a three-dimensional wave equation for $\vec{E}$ with sources given by the charges and currents:

$$\frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} - \vec{\nabla}^2\vec{E} = -4\pi\vec{\nabla}\rho - \frac{4\pi}{c^2}\frac{\partial \vec{\jmath}}{\partial t}\,. \tag{24.64}$$

We do the analogous thing with the magnetic field:

$$\begin{aligned}\vec{\nabla} \times (\vec{\nabla} \times \vec{B}) &= \vec{\nabla}(\vec{\nabla} \cdot \vec{B}) - \nabla^2 \vec{B} \\ &= \frac{4\pi}{c}\vec{\nabla} \times \vec{\jmath} + \frac{1}{c}\frac{\partial}{\partial t}(\vec{\nabla} \times \vec{E})\,. \end{aligned} \tag{24.65}$$

Referring again to the Maxwell equations,

$$\frac{1}{c^2}\frac{\partial^2 \vec{B}}{\partial t^2} - \vec{\nabla}^2\vec{B} = \frac{4\pi}{c}\vec{\nabla} \times \vec{\jmath}. \tag{24.66}$$

It's often easier to solve problems with potentials than with fields. The magnetic field $\vec{B}$ is purely transverse, so it has no scalar potential associated with it. The electric field $\vec{E}$ can have a scalar and a vector potential. Because the Maxwell equations tie the dynamics of these fields together, we'll need one scalar and one vector potential in all.

The no-monopoles condition implies that the magnetic field can be written as the curl of a vector potential:

$$\vec{B} = \vec{\nabla} \times \vec{A}\,. \tag{24.67}$$

If we were doing electrostatics, we could write the electric field as the gradient of a scalar potential:

$$\vec{E} = -\vec{\nabla}\phi\,. \tag{24.68}$$

The expression for $\vec{E}$ in terms of both ϕ and $\vec{A}$ should reduce to this when nothing is changing. So, we expect we should use the time derivative of the vector potential. Dimensional analysis tells us to include a factor of $1/c$. The correct formula is

$$\vec{E} = -\vec{\nabla}\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}\,. \tag{24.69}$$

Plugging these expressions into the wave equations for $\vec{E}$ and $\vec{B}$ yields, after some algebra,

$$-\vec{\nabla}^2\phi - \frac{1}{c}\frac{\partial}{\partial t}\vec{\nabla}\cdot\vec{A} = 4\pi\rho\,, \tag{24.70}$$

$$\vec{\nabla}^2\vec{A} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = \vec{\nabla}(\vec{\nabla}\cdot\vec{A}) - \frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{\partial}{\partial t}\vec{\nabla}\phi\,. \tag{24.71}$$

We have a wave equation for the vector potential $\vec{A}$, and we've a relationship between its longitudinal component and two scalar fields. The next step is to choose a gauge. We have two options: First, the *Coulomb gauge* makes radiation problems look like electrostatic ones; and second, the *Lorentz gauge* makes relativistic invariance manifest. We shall investigate each option in turn.

The Coulomb gauge is one in which we define away the longitudinal part of $\vec{A}$. This implies

$$-\vec{\nabla}^2\phi = 4\pi\rho, \tag{24.72}$$

which looks like Coulomb's law but is time-dependent. In addition,

$$\vec{\nabla}^2\vec{A} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = -\frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{\partial}{\partial t}\vec{\nabla}\phi\,. \tag{24.73}$$

The left-hand side of this equation is purely transverse. Therefore, the longitudinal components of the two terms on the right-hand side must cancel each other. The gradient of a scalar is a longitudinal vector field, so

$$\vec{j}_L = \frac{1}{4\pi}\frac{\partial}{\partial t}\vec{\nabla}\phi\,. \tag{24.74}$$

From the Coulomb-like equation, we can apply our knowledge of Green's functions and deduce that

$$\phi(\vec{r},t) = \frac{1}{4\pi}\int\frac{d^3\vec{r}'\rho(\vec{r}',t)}{|\vec{r}-\vec{r}'|}\,. \tag{24.75}$$

The impossibility of superluminal signaling implies that in the Coulomb gauge, $\phi(\vec{r},t)$ cannot itself be a physical quantity. Returning to the definition of the electric field $\vec{E}$ in terms of $\vec{A}$, the longitudinal part of $\vec{E}$ must fall as $1/r^2$ as r goes to infinity. The source which survives is the transverse component of the current, $\vec{j}_T$.

In the Lorentz gauge, the story begins with a different definition of the longitudinal field $\vec{A}_L$.

$$\vec{\nabla}\cdot\vec{A} + \frac{1}{c}\frac{\partial\vec{A}}{\partial t} = 0\,. \tag{24.76}$$

This is, secretly, a statement which we can write compactly in four-vector notation as $\partial^\mu A_\mu = 0$, but that's a subject for later. Using this choice of gauge, we get that

$$-\vec{\nabla}^2\vec{A} + \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = \frac{4\pi}{c}\vec{j}\,, \tag{24.77}$$

$$-\vec{\nabla}^2\phi + \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = \frac{4\pi}{c}\rho\,. \tag{24.78}$$

Working in the Lorentz gauge, we find that the dynamics are expressed by two parallel wave equations, one for the vector potential and one for the scalar potential.

We change among gauges by changing the longitudinal part of $\vec{A}$:

$$\vec{A}' = \vec{A} + \vec{\nabla}\chi \,. \tag{24.79}$$

Concurrently, we must transform the scalar potential:

$$\phi' = \phi - \frac{1}{c}\frac{\partial \chi}{\partial t} \,. \tag{24.80}$$

So far, we have been talking about electromagnetic waves in empty space. What happens when a wave passes through matter? For simplicity, we will study the case of a homogeneous medium, characterized by its dielectric constant κ and its magnetic permeability μ, and we will assume that the medium has no preferred direction.

In addition to the familiar $\vec{E}$ and $\vec{B}$ fields, it is helpful now to introduce $\vec{D} = \kappa\vec{E}$ and $\vec{B} = \mu\vec{H}$.

Writing ρ_f and $\vec{\jmath}_f$ for the free charges and currents,

$$\vec{\nabla} \cdot \vec{D} = 4\pi\rho_f \tag{24.81}$$

$$\vec{\nabla} \times \vec{H} = \frac{4\pi}{c}\vec{\jmath}_f + \frac{1}{c}\frac{\partial \vec{D}}{\partial t} \,. \tag{24.82}$$

Taking the curl of the curl of $\vec{H}$ gives

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{H}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{H}) - \nabla^2 \vec{H} \,. \tag{24.83}$$

We can use the Ampère–Maxwell law on the left-hand side:

$$\frac{4\pi}{c}\vec{\nabla} \times \vec{\jmath}_f + \frac{1}{c}\frac{\partial}{\partial t}(\vec{\nabla} \times \vec{D}) = -\nabla^2 \vec{H} \,. \tag{24.84}$$

Do we know the curl of $\vec{D}$? We do if we are entitled to say that

$$\vec{\nabla} \times \left(\frac{\vec{D}}{\kappa}\right) = -\frac{1}{c}\frac{\partial(\mu\vec{H})}{\partial t} \,. \tag{24.85}$$

Rearranging,

$$\vec{\nabla} \times \vec{D} = -\frac{\kappa\mu}{c}\frac{\partial \vec{H}}{\partial t} \,. \tag{24.86}$$

So,

$$\frac{\kappa\mu}{c^2}\frac{\partial^2 \vec{H}}{\partial t^2} - \nabla^2 \vec{H} = \frac{4\pi}{c}\vec{\nabla} \times \vec{\jmath}_f \,. \tag{24.87}$$

This is a sourced wave equation with a speed

$$v = \frac{c}{\sqrt{\kappa\mu}} \,. \tag{24.88}$$

We can call $n = \sqrt{\kappa\mu}$ the *index of refraction.*

> This velocity is so nearly that of light, that it seems we have strong reason to conclude that light itself (including radiant heat, and other radiations if any) is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws. If so, the agreement between the elasticity of the medium as calculated from the rapid alternations of luminous vibrations, and as found by the slow process of electrical experiments, shows how perfect and regular the elastic properties of the medium must be when not encumbered with any matter denser than air. If the same character of the elasticity is retained in dense transparent bodies, it appears that the square of the index of refraction is equal to the product of the specific dielectric capacity and the specific magnetic capacity.
> —James Clerk Maxwell, *A Dynamical Theory of the Electromagnetic Field,* §1.20 (1864)

To get a wave equation for $\vec{E}$ under these circumstances, start with the vector identity

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E}\,, \tag{24.89}$$

and plug in:

$$\vec{\nabla} \cdot \vec{E} = \frac{4\pi\rho_f}{\kappa}\,, \tag{24.90}$$

$$\vec{\nabla} \times \vec{B} = \frac{4\pi\mu}{c}\vec{\jmath}_f + \frac{\kappa\mu}{c}\frac{\partial \vec{E}}{\partial t}\,. \tag{24.91}$$

So,

$$-\frac{1}{c}\frac{\partial}{\partial t}\left(\frac{4\pi\mu}{c}\vec{\jmath}_f + \frac{\kappa\mu}{c}\frac{\partial \vec{E}}{\partial t}\right) = \frac{4\pi}{\kappa}\vec{\nabla}\rho_f - \nabla^2\vec{E}\,. \tag{24.92}$$

Rearranging terms,

$$-\frac{\kappa\mu}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} + \nabla^2\vec{E} = \frac{4\pi}{\kappa}\vec{\nabla}\rho_f + \frac{4\pi\mu}{c^2}\frac{\partial \vec{\jmath}_f}{\partial t}\,. \tag{24.93}$$

Again, this is a sourced wave equation with effective speed $v = c/\sqrt{\kappa\mu}$.

24.5 From Newton–Euler to Hamilton–Jacobi

Let's say that we want to reformulate "Newtonian" particle mechanics so that it looks analogous to wave motion, in order to make our knowledge of one math subject applicable to another. A light ray, minding its own business, propagates in the direction perpendicular to its wavefronts. So, the particle momentum should be perpendicular to lines of something, meaning that it should be the *gradient* of something:

$$\vec{p} = \vec{\nabla} S\,. \tag{24.94}$$

We don't know what S is, except that it's a field whose value presumably depends upon position and time, and it ought to satisfy some equation. Can we find an equation for it? Well, from mechanics we know that the time derivative of the momentum is the force. And in the absence of friction, we can write the force as minus the gradient of the potential energy, which feels relevant here:

$$\frac{d\vec{p}}{dt} = -\vec{\nabla}V\,. \tag{24.95}$$

Combining these equations, we get

$$\frac{d}{dt}\vec{\nabla}S = -\vec{\nabla}V\,. \tag{24.96}$$

But wait! A particle riding along in a field sees *two* contributions to the field value changing. If the particle is sitting still, then the field value at its current spot will change if the field itself changes. A *moving* particle can also just move over to a spot where the field value is different, even if the field is constant over time. We learn how to cover this in hydrodynamics, by splitting the derivative into two terms, one with a velocity dependence:

$$\frac{d}{dt}\vec{\nabla}S = \frac{\partial}{\partial t}(\vec{\nabla}S) + (\vec{v}\cdot\vec{\nabla})\vec{\nabla}S\,. \tag{24.97}$$

And *this* is what equals minus the gradient of the potential.

(Where does that dot product $\vec{v}\cdot\vec{\nabla}$ come from? Well, imagine we're in one dimension, moving through a field f. In a tiny interval of time dt, we move a distance $v\,dt$, and over this distance, the field value changes by an amount $(v\,dt)(\partial f/\partial x)$. Extending this to three dimensions gives the expression above.)

Now, we can write the $\vec{v}$ here as the momentum divided by the mass m, and we have declared that the momentum is the gradient of our mystery field. Therefore,

$$\frac{\partial}{\partial t}(\vec{\nabla}S) + \frac{1}{m}(\vec{\nabla}S\cdot\vec{\nabla})\vec{\nabla}S = -\vec{\nabla}V\,. \tag{24.98}$$

Let's exchange the order of differentiation on that first term:

$$\vec{\nabla}\left(\frac{\partial}{\partial t}S\right) + \frac{1}{m}(\vec{\nabla}S\cdot\vec{\nabla})\vec{\nabla}S = -\vec{\nabla}V\,. \tag{24.99}$$

It looks almost like each term can be the gradient of something. So, we stare at the one that isn't for a while, until we remember from calculus that

$$\frac{d}{dx}\left[\left(\frac{df}{dx}\right)^2\right] = 2\frac{df}{dx}\frac{d^2f}{dx^2}\,. \tag{24.100}$$

By the same logic, we find that our equation for S is what we'd get by taking the gradient of both sides of

$$\frac{\partial S}{\partial t} + \frac{1}{2m}(\vec{\nabla}S)^2 = -V\,. \tag{24.101}$$

Or, isolating the time derivative and flipping the signs to give it a more conventional presentation,

$$-\frac{\partial S}{\partial t} = \frac{1}{2m}(\vec{\nabla} S)^2 + V \,. \tag{24.102}$$

The *Hamilton–Jacobi equation* just generalizes this! We can write it as

$$-\frac{\partial S}{\partial t} = H(\vec{x}, \vec{\nabla} S, t) \,. \tag{24.103}$$

Another common notation is to write q for position instead of $\vec{x}$, suppressing the little arrow and just remembering that q can be a whole list of coordinates (for multiple particles, potentially). Then

$$-\frac{\partial S}{\partial t} = H\left(q, \frac{\partial S}{\partial q}, t\right) \,, \tag{24.104}$$

with

$$p = \frac{\partial S}{\partial q} \,. \tag{24.105}$$

The discussion in this section so far was inspired by Rosen [163]. We have actually followed Rosen's logic in reverse, starting with the familiar instead of justifying the unfamiliar by transforming it until it looks recognizable.

We can relate our "something" field S to quantities discussed earlier by differentating it with respect to time, with the understanding that it can depend upon t explicitly as well as implicitly through the coordinates q:

$$\frac{dS}{dt} = \frac{\partial S}{\partial q}\dot{q} + \frac{\partial S}{\partial t} = p\dot{q} - H \,. \tag{24.106}$$

From Eq. (24.11), we recognize this as the Lagrangian. The mysterious quantity S is the action for the special case of the physical path!

We discover something else interesting by taking the spatial gradient of H and then interchanging the partials:

$$\frac{\partial H}{\partial q} = -\frac{\partial}{\partial q}\frac{\partial S}{\partial t} = -\frac{\partial}{\partial t}\frac{\partial S}{\partial q} \,. \tag{24.107}$$

Or, in short,

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} \,, \tag{24.108}$$

meaning that *Hamilton's equation of motion has emerged as a Maxwell relation.* But we only have the time derivative of the momentum; what about the time derivative of the position? Can we also get Hamilton's equation for that? Well, we have treated the Hamilton–Jacobi quantity S like an energy in thermodynamics, differentating it twice in two different orders, so by analogy, let's *subtract* something from it like we do when we define a free energy:

$$X = S - pq \,. \tag{24.109}$$

The equality of second partial derivatives says that

$$\frac{\partial^2 X}{\partial t \partial p} = \frac{\partial^2 X}{\partial p \partial t} \,. \tag{24.110}$$

Recall that we defined S to have no dependence upon p, so $\partial X/\partial p$ only picks up the term we subtracted:

$$\frac{\partial}{\partial t}\frac{\partial X}{\partial p} = \frac{\partial}{\partial t}(-q) \,. \tag{24.111}$$

And in the other order,

$$\frac{\partial}{\partial p}\frac{\partial X}{\partial t} = \frac{\partial}{\partial p}\frac{\partial S}{\partial t} = \frac{\partial}{\partial p}(-H) \,. \tag{24.112}$$

So, our Maxwell relation is the other Hamilton equation:

$$\frac{dq}{dt} = \frac{\partial H}{\partial p} \,. \tag{24.113}$$

This treatment follows that of Baez [164, 165], who then goes on to discuss symplectic geometry.

Back in §6.2, we were able to understand Maxwell relations by expressing the conservation of energy in terms of a certain determinant being equal to 1. Can we do the same for Hamilton's equations of motion? We used the first law of thermodynamics to say that

$$dP\, dV = dT\, dS \,, \tag{24.114}$$

and from this we deduced that

$$\frac{\partial(T,S)}{\partial(P,V)} = 1 \,. \tag{24.115}$$

By analogy, let us say formally that

$$dx\, dp = dt\, dH \,, \tag{24.116}$$

and so

$$\frac{\partial(t,H)}{\partial(x,p)} = 1 \,. \tag{24.117}$$

By the chain rule,

$$\frac{\partial(t,H)}{\partial(x,p)}\frac{\partial(x,p)}{\partial(A,B)} = \frac{\partial(t,H)}{\partial(A,B)} \,. \tag{24.118}$$

Therefore,

$$\frac{\partial(x,p)}{\partial(A,B)} = \frac{\partial(t,H)}{\partial(A,B)} \,. \tag{24.119}$$

Let us take $A = x$ and $B = t$. Then

$$\frac{\partial(x,p)}{\partial(x,t)} = \frac{\partial(t,H)}{\partial(x,t)} = -\frac{\partial(t,H)}{\partial(t,x)} \,. \tag{24.120}$$

This reduces to

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x}\,. \tag{24.121}$$

If instead we take $A = p$ and $B = t$, then we find

$$\frac{\partial(x,p)}{\partial(p,t)} = \frac{\partial(t,H)}{\partial(p,t)}\,. \tag{24.122}$$

Applying the antisymmetry property of the Jacobian to both sides, the signs cancel:

$$\frac{\partial(p,x)}{\partial(p,t)} = \frac{\partial(t,H)}{\partial(t,p)}\,. \tag{24.123}$$

This reduces to the other of Hamilton's equations,

$$\frac{dx}{dt} = \frac{\partial H}{\partial p}\,. \tag{24.124}$$

So, the analogy seems to hold together! This approach, unlike what we did earlier by swapping the order of the partial derivatives, avoids having to invent a new quantity X to find the second of Hamilton's equations. Instead, we just take the two most obvious choices for A and B in the chain rule. But it is not so easy to interpret the equality $dx\,dp = dt\,dH$ physically, at least by appealing back to thermodynamics; what would be a circuit in the (t, H) plane? It seems that what is happening here, on some level, is that we are treating the possible dynamical trajectories of a particle as families of tuples (x, p, t, H) in $\mathbb{R}^4$, and only some families of tuples are dynamically legitimate.

Another motivation is more statistical. For specificity, let's consider the simple harmonic oscillator. Suppose that we have a point particle in a 1D harmonic potential, and because we only know its preparation imprecisely, we characterize the situation by a Liouville density $\rho(x, p)$. Essentially, we express the available information about the particle by drawing a blob on its phase space. We can locate points in the blob by the rectangular coordinates (x, p), or we can foliate the phase space into ellipses of constant energy and specify a point by saying which ellipse it lies on and how far around the point is rotated from some convenient reference line. In other words, uncertainty about position and momentum is equivalent to uncertainty about *energy* and *elapsed time* since the particle was let loose. If we say that neither description has an advantage upon the other — that an overlap between two Liouville densities ρ_1 and ρ_2 looks the same size whichever perspective we take — then we can justify setting the Jacobian to unity and thus obtaining Hamilton's equations of motion as Maxwell relations.

There's an old puzzle about just how much can be derived from the premise that time evolution preserves the overlaps between probability densities [166, pp. 186, 317–19]. Let us assume that holds in the (x, p) plane, and let us further postulate that there exists a quantity H that is constant along dynamical trajectories and whose dependence upon x and p is continuous. Then a probability blob on the plane of H versus elapsed time will remain constant in area: the values of H do not change, and elapsed time grows evenly for all trajectories. We can define $dt\,dH$ as the area element

into which $dx\,dp$ maps at one time, and by our premise of overlap preservation, the map between planes must be area-preserving at all times. We get Newton's second law, that the time derivative of the momentum is something we can call the force, once we require that the momentum is proportional to the velocity.

It is worth figuring out what the Hamilton–Jacobi S function might look like for a free particle. Suppose that the energy is E and the momentum is p. We need the partial derivative $\partial S/\partial q$ to equal the momentum, and the partial derivative $\partial S/\partial t$ should equal minus the energy. Accordingly, we take

$$S = pq - Et\,. \tag{24.125}$$

Because the momentum is a constant, the total derivative of S with respect to time is $p\dot{q} - E$, as required. But compare this with the *quantum* description of a free particle. Neglecting issues of normalization, a "momentum eigenstate" is a plane wave:

$$\psi(q,t) = e^{i(pq-Et)/\hbar}\,. \tag{24.126}$$

The quantum phase looks just like the classical Hamilton–Jacobi formula, the action for the physical path! This suggests two directions of development. One is a decomposition of the quantum wave equation for time evolution. We can write a wavefunction as

$$\psi(q,t) = Re^{iS/\hbar}\,, \tag{24.127}$$

letting q stand for as many spatial dimensions as desired. Then

$$\frac{\partial(R^2)}{\partial t} = -\vec{\nabla}\cdot\left(\frac{R^2}{m}\vec{\nabla}S\right)\,, \tag{24.128}$$

$$\frac{\partial S}{\partial t} = -\frac{1}{2m}(\vec{\nabla}S)^2 - V + \frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}\,. \tag{24.129}$$

The latter equation is just Hamilton–Jacobi plus a new term [88]. Another step we could take from here is to say that if the classical action for the physical path is the analogue to the phase of the wavefunction, then we can associate a phase to an *arbitrary* path by integrating the classical Lagrangian along it. This is the starting point of the path integral formulation of quantum mechanics.

24.6 Relativity 1: Velocity Addition

In this section, we'll recap Mermin's argument for the Einsteinian law of adding velocities [167]. This argument relies upon a thought experiment in which Alice, riding on a train, throws a ball from the back of the train towards the front, and at the sime time fires off a light pulse aimed at the front of the train. A mirror at the front of the train bounces the light pulse back. At some point midway along the train, the ball and the light pulse collide. For the purposes of drama, we can imagine that the ball explodes. The important points to remember are as follows. First, the speed of light is the same for Alice riding on the train and for Bob watching from the platform as

the train goes by. Second, the *fraction of the train's length* traversed by the ball must *also* be the same for Alice and for Bob. If the train has 100 seats all in a row and the explosion destroys seat number 3, both Alice and Bob will agree on that much.

Alice's analysis is simpler than Bob's. She says that she is at rest and the ball's velocity is u towards the front. The ball covers some fraction of the train's length before it explodes, call it $1-f$ (so that f is the fraction remaining). The light travels the full length of the train and then an additional fraction f. When the light pulse and the ball collide, they must have been traveling for the same amount of time, so the raio of their speeds is the ratio of their respective distances traversed:

$$\frac{u}{v} = \frac{1-f}{1+f}\,. \tag{24.130}$$

By cross-multiplying and then solving for f, we obtain

$$f = \frac{c-u}{c+u}\,. \tag{24.131}$$

Now, Bob's job is trickier, because he has to introduce a few more letters. He says that the train has a velocity v and the ball has a velocity w, both in the forward direction. To get the velocity addition law, we need w in terms of u, v and c. Bob starts going through the scenario, one step at a time. The light pulse takes time T_0 to hit the front of the train, and then an additional time interval T_1 to go back and hit the ball that is coming forward to meet it. Call D the distance between the front of the train and the ball at the time T_0. At this time, the light has traversed a length cT_0, and the ball has moved some lesser length wT_0, so the spacing between them must be

$$D = cT_0 - wT_0\,. \tag{24.132}$$

After another time T_1 has passed, the ball and the light have together made up this distance exactly. So, Bob says,

$$D = cT_1 + wT_1\,. \tag{24.133}$$

Equating these two gives the ratio of the time intervals:

$$\frac{T_1}{T_0} = \frac{c-w}{c+w}\,. \tag{24.134}$$

Bob then casts about for some other way to learn about these time intervals, in the hope of eliminating this fraction. He realizes he has yet to use the velocity v of the train itself. In the first time interval, the light pulse travels cT_0, while each point on the train moves a distance vT_0. Bob now realizes that the distance cT_0 must be the length L of the train *plus* the extra distance that the front of the train moved in that same time:

$$cT_0 = L + vT_0\,. \tag{24.135}$$

Next, the reflected light pulse travels a distance cT_1, heading for the collision point that was a distance fL away, while the chair that is to be destroyed moves forward a distance vT_1. So,

$$fL = cT_1 + vT_1\,. \tag{24.136}$$

We have two expressions for L, and by combining them we get

$$\frac{T_1}{T_0} = f\frac{c-v}{c+v}\,. \tag{24.137}$$

Equating our two formulae for this ratio,

$$\frac{c-w}{c+w} = f\frac{c-v}{c+v}\,. \tag{24.138}$$

Alice and Bob agree on the fraction f, so we finally have a formula for the velocity Bob measures for the ball, w, in terms of the velocity that Alice uses, u, and the relative velocity between their frames, v:

$$\frac{c-w}{c+w} = \frac{c-u}{c+u}\frac{c-v}{c+v}\,. \tag{24.139}$$

We can massage this into a more familiar form by introducing abbreviations for the numerator and denominator:

$$a = (c-u)(c-v) = c^2 - c(u+v) + uv\,, \tag{24.140}$$

$$b = (c+u)(c+v) = c^2 + c(u+v) + uv\,. \tag{24.141}$$

Then we have

$$(c-w)b = (c+w)a\,, \tag{24.142}$$

and

$$c(b-a) = w(b+a)\,, \tag{24.143}$$

which combine to say that

$$\frac{w}{c} = \frac{b-a}{b+a}\,. \tag{24.144}$$

But this sum and difference are easy to work out:

$$b + a = 2(c^2 + uv)\,, \tag{24.145}$$

while

$$b - a = 2c(u+v)\,. \tag{24.146}$$

Therefore,

$$\frac{w}{c} = \frac{2c(u+v)}{2c^2(1+\frac{uv}{c^2})}\,, \tag{24.147}$$

and doing a little cancellation,

$$w = \frac{u+v}{1+\frac{uv}{c^2}}\,. \tag{24.148}$$

24.7 Relativity 2: The Invariant Interval

Once again, Alice rides on the train, and Bob watches her go past at velocity v. Alice and Bob can come to agree that their relative motion does not affect *distances perpendicular* to the line of motion. For example, they can set up an experiment where they both leave paint marks on the same rock wall, such that if motion affected height, Alice would predict that Bob's mark would be on top and vice versa. So, both Alice and Bob agree on the floor-to-ceiling separation of the train, for example. This means that Alice can rig up a clock that works by sending a light pulse from the floor to the ceiling, or from one wall of the train to the other. One tick of this clock is the time interval t_A that light requires to traverse the perpendicular distance. But Bob sees the clock moving past at the speed v, and for him, the light pulse moves diagonally for a time t_B. He goes Pythagorean and draws a right triangle whose legs are ct_A and vt_B and whose hypotenuse is ct_B, from which he deduces that

$$t_B = \gamma t_A,\ \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}\,. \tag{24.149}$$

This is the standard light-clock argument.

Now, let E_1 and E_2 be any two events that, for Alice, happen in the same place but different times. Alice measures the time between them to be T_A. Whatever these events are, they can't be used to establish which observer is "really" moving, so the time-dilation factor has to work just as it does with a light clock, and

$$T_B = \gamma T_A \tag{24.150}$$

is the time Bob measures between the two events. For Bob, these events must happen at different places, a distance

$$D_B = vT_B \tag{24.151}$$

apart.

If we square the time lapse observed by Alice,

$$T_A^2 = \gamma^{-2} T_B^2 = (1 - v^2/c^2) T_B^2\,, \tag{24.152}$$

we find that the result has a simple expression in terms of Bob's distance and time figures:

$$T_A^2 = T_B^2 - \frac{v^2 T_B^2}{c^2} = T_B^2 - \frac{D_B^2}{c^2}\,. \tag{24.153}$$

We can do the same analysis for anyone else who comes along in motion relative to Alice. Whatever distance D and time lapse T they measure, those values will satisfy

$$T^2 - \frac{1}{c^2} D^2 = T_A^2\,. \tag{24.154}$$

We say that two events are *timelike separated* when there exists an observer for whom they happen in the same place. Thus, for any two events that are timelike separated,

the square of the time lapse between them minus the square of the distance (adjusted by the speed of light for units' sake) is the square of the *proper time* between them.

What about pairs of events for which there is no inertial observer for whom they happen at the same place? This is where a *spacetime diagram* starts to come in handy. First, we follow Mermin and choose our standard length unit to be the distance that light travels in vacuum during 1 nanosecond. This is only a couple percent different from the historical unit known as the "foot", so for brevity we will call 1 light-nanosecond 1 foot. We will also adopt Mermin's terminology of *equilocs* and *equitemps.* An equiloc is a line on a diagram of events that indicates a subset of events which a particular observer (e.g., Alice) deems to happen in the same place. Likewise, an equitemp is a line on a diagram indicating a subset of events that a particular observer deems to happen at the same time. A convenient cartographer's convention is to have equilocs pointing mostly vertical and equitemps lying mostly horizontal. A *light line* or *photon line* indicates events marked by the passing of a light pulse.

All of Alice's equilocs are parallel, because if they intersected, one event would happen in two different places. Likewise, all of her equitemps are parallel, because otherwise, some event would happen at two different times. Let E_1 be the generation of a light pulse and E_2 its detection 1 nanosecond later (as determined by Alice). Then Alice's equilocs through E_1 and E_2 are the mostly-vertical sides of a parallelogram, and her equitemps through E_1 and E_2 are the mostly-horizontal sides of that parallelogram. Because Alice measures time in nanoseconds and distance in feet, her value for c is numerically 1, and her scale factor for converting *ruler measurements on her diagram* into *distances and times* is the same whether she puts her ruler along an equiloc to get a time or along an equitemp to get a distance. The parallelogram she draws on her page, with two vertices marked E_1 and E_2, must be a rhombus.

A fun rhombus fact: The diagonals bisect the angles at the vertices. Here, that means the angle on the page between the E_1 equiloc and the light line is the same as the angle between the light line and the E_1 equitemp. Equitemps are just equilocs reflected across a light line!

For timelike separation, we had the invariant squared interval

$$I^2 = T^2 - D^2\,, \tag{24.155}$$

where we are now taking $c = 1$. Reflecting a pair of timelike-separated events across a light line gives a pair of events that are *spacelike* separated, and for these the invariant squared interval

$$I^2 = D^2 - T^2 \tag{24.156}$$

is just the square of the distance between them as measured by the observer for whom they are simultaneous.

In a relativity course, we'd go much more into the details of many scenarios and calculations. For our purposes, though, what we need is the idea that transformations from one inertial observer's reference frame to another preserve, not the Pythagorean distance, but this thing that looks like Pythagoras with a minus sign.

Bibliography

[1] J. C. Maxwell, "Diffusion," in *Encyclopædia Britannica*. 1878.

[2] R. Jeffrey, *Subjective Probability: The Real Thing*. Cambridge University Press, 2004. http://www.princeton.edu/~bayesway/.

[3] C. M. Caves, "Probabilities as betting odds and the Dutch book," Internal report, 2005. http://info.phys.unm.edu/~caves/reports/reports.html. Somewhat complementary to the approach we take here.

[4] C. A. Fuchs and R. Schack, "Bayesian conditioning, the reflection principle, and quantum decoherence," arXiv:1103.5950 [quant-ph]. Although it is motivated by quantum physics, the particularly quantum considerations do not explicitly enter until near the end.

[5] A. Hodges, *Alan Turing: The Enigma*. Princeton University Press, 1983. http://www.turing.org.uk.

[6] S. L. Zabell, "Alan Turing and the central limit theorem," *American Mathematical Monthly* **102** (1995) no. 6, 483–94, JSTOR:2974762.

[7] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, 2009. http://pgm.stanford.edu/.

[8] P. W. Humphreys, "Notes on the origin of the term 'Dutch book'," Webpage, 2008. http://people.virginia.edu/~pwh2a/dutch%20book%20origins.doc.

[9] C. A. Fuchs, M. Schlosshauer, and B. C. Stacey, *My Struggles with the Block Universe*. Preprint, 2014. arXiv:1405.2390 [quant-ph]. Possibly the record-holder for length of an arXiv posting.

[10] T. Tao, *Topics in Random Matrix Theory*. American Mathematical Society, 2012. http://terrytao.wordpress.com/books/topics-in-random-matrix-theory/.

[11] J. H. Conway, *On Numbers and Games*. Academic Press, 1976. I've yet to find a real connection between the kind of mathematical games that Conway talks about and the type of games discussed in this thesis, but I'm still looking.

[12] B. C. van Fraassen, "Belief and the Will," *The Journal of Philosophy* **81** (1984) no. 5, 235–56, JSTOR:2026388.

[13] B. Skyrms, "Dynamic coherence and probability kinematics," *Philosophy of Science* **54** (1987) no. 1, 1–20, JSTOR:187470.

[14] P. Diaconis and S. L. Zabell, "Updating subjective probability," *Journal of the American Statistical Association* **77** (1982) no. 380, 822–30, JSTOR:2287313.

[15] P. Diaconis and S. L. Zabell, "Some alternatives to Bayes's rule," in *Proceedings of the Second University of California, Irvine, Conference on Political Economy*, pp. 25–38. 1986.

[16] K. M. Page and M. Nowak, "Unifying Evolutionary Dynamics," *Journal of Theoretical Biology* **219** (2002) 93–8, PMID:12392978.

[17] M. Harper, "The replicator equation as an inference dynamic," arXiv:0911.1763 [math.DS]. Explains the formal mapping between the Bayes rule and the replicator equation.

[18] J. C. Baez, "Information geometry," *Azimuth* (2012) n.p. http://math.ucr.edu/home/baez/information/.

[19] J. Moriarty, *A Treatise on the Binomial Theorem*. Meyer and Company, 1867.

[20] F. Eggenberger and G. Pólya, "Über die Statistik verketteter Vorgänge," *Zeitschrift für Angewandte Mathematik und Mechanik* **3** (1923) no. 4, 279–89.

[21] A. Collevecchio, C. Cotar, and M. LiCalzi, "On a preferential attachment and generalized Pólya's urn model," *The Annals of Applied Probability* **23** (2013) no. 3, 1219–53. What happens when we can add marbles with new colors to the Pólya urn as we go? The urn model maps onto a network-growth model.

[22] N. Helfand, "Polya's Urn and the Beta-Bernoulli process," *University of Chicago REU 2013* (2013) n.p. http://math.uchicago.edu/~may/REU2013/.

[23] C. M. Caves, C. A. Fuchs, and R. Schack, "Unknown quantum states: The quantum de Finetti representation," *Journal of Mathematical Physics* **43** (2002) 4357, arXiv:quant-ph/0104088.

[24] J. Aczél, B. Forte, and C. T. Ng, "Why the Shannon and Hartley entropies are 'natural'," *Advances in Applied Probability* **6** (1974) no. 1, 131–46, JSTOR:1426210.

[25] J. C. Baez, T. Fritz, and T. Leinster, "A characterization of entropy in terms of information loss," *Entropy* **13** (2011) no. 11, 1945–57, arXiv:1106.1791 [cs.IT].

[26] G. Gbur, *Mathematical Methods for Optical Physics and Engineering*. Cambridge University Press, 2011. http://skullsinthestars.com/mathematical-methods-for-optics/.

[27] M. Kardar, *Statistical Physics of Particles*. Cambridge University Press, 2007. http://www.mit.edu/~kardar/teaching/index.html. The first volume of two in a textbook series.

[28] M. Kardar, *Statistical Physics of Fields*. Cambridge University Press, 2007. http://www.mit.edu/~kardar/teaching/index.html. Second in a two-part series of textbooks.

[29] H. S. Wilf, *generatingfunctionology*. Academic Press, 1994.

[30] I. H. Dinwoodie and S. L. Zabell, "Large deviations for exchangeable random vectors," *Annals of Probability* **20** (1992) no. 3, 1147–66.

[31] E. T. Jaynes, *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.

[32] C. A. Lugo and A. J. McKane, "Quasicycles in a spatial predator-prey model," *Physical Review E* **78** (2008) 051911, arXiv:0806.1287 [q-bio.PE].

[33] T. Biancalani, D. Fanelli, and F. D. Patti, "Stochastic Turing patterns in the Brusselator model," *Physical Review E* **81** (2010) 046215, arXiv:0910.4984 [cond-mat.stat-mech].

[34] H. Risken, *The Fokker–Planck Equation: Methods of Solution and Applications*. Springer, 1996.

[35] N. Champagnat and A. Lambert, "Evolution of discrete populations and the canonical diffusion of adaptive dynamics," *The Annals of Applied Probability* **17** (2007) no. 1, 102–55. http://projecteuclid.org/euclid.aoap/1171377179.

[36] B. Allen, M. A. Nowak, and U. Dieckmann, "Adaptive dynamics with interaction structure," *The American Naturalist* **181** (2013) no. 6, E139–63. One of the papers on which I leaned greatly.

[37] J. Van Cleve, "Social evolution and genetic interactions in the short and long term," *bioRxiv* (2014) n.p.

[38] C. E. Tarnita, H. Ohtsuki, T. Antal, F. Fu, and M. A. Nowak, "Strategy selection in structured populations," *Journal of Theoretical Biology* **259** (2009) no. 3, 570–81, PMC:2710410.

[39] B. Allen and M. A. Nowak, "Games among relatives revisited," *Journal of Theoretical Biology* **378** (2015) 103–16.

[40] B. Allen *et al.*, "Nonlinear social evolution and the emergence of collective action," *PNAS Nexus* **3** (2024) no. 4, pgae131, arXiv:2302.14700 [q-bio.PE].

[41] B. Allen and C. E. Tarnita, "Measures of success in a class of evolutionary models with fixed population size and structure," *Journal of Mathematical Biology* **68** (2014) no. 1–2, 109–43. http://scholar.princeton.edu/ctarnita/files/AxiomSuccessPublished.pdf. This paper builds up in a nice way the treatment of evolution as a stochastic process. The point made near the end about the way the Price equation is often misleadingly written is an important one.

[42] M. Doebeli, C. Hauert, and T. Killingback, "The evolutionary origin of cooperators and defectors," *Science* **306** (2004) 859–62.

[43] J. P. Sethna, *Statistical Mechanics: Entropy, Order Parameters, and Complexity*. Oxford University Press, 2nd ed., 2023. http://pages.physics.cornell.edu/~sethna/StatMech/.

[44] A. B. Zamolodchikov, ""Irreversibility" of the flux of the renormalization group in a 2D field theory," *JETP Letters* **43** (1986) no. 12, 730–32.

[45] R. J. Creswick, H. A. Farach, and C. P. Poole, *Introduction to Renormalization Group Methods in Physics*. John Wiley & Sons, 1992.

[46] W. D. McComb, *Renormalization Methods: A Guide for Beginners*. Oxford University Press, 2003.

[47] A. Zee, *Quantum Field Theory in a Nutshell*. Princeton University Press, 2003.

[48] Y. Bar-Yam, *Dynamics of Complex Systems*. Westview Press, 2003. http://necsi.edu/publications/dcs/index.html.

[49] J. Cardy, *Scaling and Renormalization in Statistical Physics*. Cambridge University Press, 1996.

[50] A. C. Mowery and D. T. Jacobs, "Undergraduate experiment in critical phenomena: Light scattering in a binary fluid mixture," *American Journal of Physics* **51** (1983) no. 6, 542–45.

[51] D. T. Jacobs, "Turbidity in the binary fluid mixture methanol-cyclohexane," *Physical Review A* **33** (1986) no. 4, 2605–11.

[52] D. T. Jacobs, S. M. Y. Lau, A. Mukherjee, and C. A. Williams, "Measuring turbidity in a near-critical, liquid–liquid system: A precise, automated experiment," *International Journal of Thermophysics* **20** (1999) no. 3, 877–887.

[53] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group*. Westview Press, 1992.

[54] A. Pelissetto and E. Vicari, "Critical phenomena and renormalization-group theory," *Physics Reports* **368** (2002) 549–727, arXiv:cond-mat/0012164.

[55] L. Szilard, "über die entropieverminderung in einem thermodynamischen system bei eingriffen intelligenter wesen," *Zeitschrift für Physik* **53** (1929) 840–56.

[56] T. Sagawa and M. Ueda, "Nonequilibrium thermodynamics of feedback control," *Physical Review E* **85** (2012) 021104, arXiv:1105.3262 [cond-mat.stat-mech].

[57] C. H. Bennett, "The thermodynamics of computation — a review," *International Journal of Theoretical Physics* **21** (1982) 905–40.

[58] C. J. Adkins, *Equilibrium Thermodynamics*. Cambridge University Press, 3rd ed., 1983.

[59] D. J. Ritchie, "A simple method for deriving Maxwell's relations," *American Journal of Physics* **36** (1968) 760.

[60] V. Ambegaokar and N. D. Mermin, "Question #78. A question about the Maxwell relations in thermodynamics," *American Journal of Physics* **69** (2001) 405.

[61] D. J. Ritchie, "Answer to Question #78. A question about the Maxwell relations in thermodynamics," *American Journal of Physics* **70** (2002) 104.

[62] H. S. Leff, "Answer to Question #78. A question about the Maxwell relations in thermodynamics," *American Journal of Physics* **70** (2002) 104.

[63] D. L. Goodstein, *States of Matter*. Prentice-Hall, 1975.

[64] W. Blum and L. Rolandi, *Particle Detection with Drift Chambers*. Springer-Verlag, 1993.

[65] S. F. Biagi, "Monte Carlo simulation of electron drift and diffusion in counting gases under the influence of electric and magnetic fields," *Nucl. Instr. and Meth. A* **421** (1999) 234–240.

[66] U. Becker *et al.*, "Consistent measurements comparing the drift features of noble gas mixtures," *Nucl. Instr. and Meth. A* **421** (1999) 54–59.

[67] U. Becker *et al.*, "Gas R&D Home Page," http://cyclotron.mit.edu/drift/www, 2005.

[68] K. Kumar, "The physics of swarms and some basic questions of kinetic theory," *Phys. Rep.* **112** (1984) no. 5, 319–375.

[69] R. Nesbet, "Variational calculations of accurate e^--He cross sections below 19 eV," *Phys. Rev. A* **20** (1979) no. 1, 58–70.

[70] G. Fraser and E. Matheison *Nucl. Instr. and Meth. A* **247** (1986) 544.

[71] H. R. Skullerud, "The stochastic computer simulation of ion motion in a gas subjected to a constant electric field," *J. Phys. D* **1** (1968) 1567–1568.

[72] S. Longo *Plasma Sources Sci. Technol.* **9** (2000) 468–476.

[73] R. Crompton, M. Elford, and A. Robertson *Aust. J. Phys* **23** (1970) 667.

[74] I. Shkarofsky, T. Johnston, and M. Bachynski, *The Particle Kinetics of Plasmas*. Addison-Wesley, Reading, Massachusetts, 1966.

[75] K. Huang, *Statistical Mechanics*. Wiley, 1987.

[76] R. Balescu, *Equilibrium and Non-Equilibrium Statistical Mechanics*. Wiley, 1975.

[77] J. Binney and S. Tremaine, *Galactic Dynamics*. Princeton University Press, Princeton, New Jersey, 1987.

[78] M. Hénon *Astron. Astrophys.* **114** (1983) 211.

[79] R. White, R. Robson, and K. Ness, "On approximations involved in the theory of charged particle transport in gases in electric and magnetic fields at arbitrary angles," *IEEE Transactions on Plasma Science* **27** (1999) no. 5, 1249–1253.

[80] W. Allis and H. Allen, "Theory of the Townsend method of measuring electron diffusion and mobility," *Phys. Rev.* **52** (1937) 703–707.

[81] R. White, K. Ness, R. Robson, and B. Li, "Charged-particle transport in gases in electric and magnetic fields crossed at arbitrary angles: Multiterm solution of Boltzmann's equation," *Phys. Rev. E* **60** (1999) no. 2, 2231–49.

[82] K. Ness, "Multi-term solution of the Boltzmann equation for electron swarms in crossed electric and magnetic fields," *J. Phys. D* **27** (1994) 1848–1861.

[83] N. Bohr, "On the quantum theory of line-spectra," in *Sources of Quantum Mechanics*. Dover, 1968. Reprinted from the Mémoires de l'Académie Royale des Sciences et des Lettres de Danemark, Copenhague, Section des Sciences, 8me série t. IV, 1918; the introduction is dated November 1917.

[84] M. Born, "Quantum mechanics," in *Sources of Quantum Mechanics*. Dover, 1968. Reprinted from "Über Quantenmechanik" originally published in 1924.

[85] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[86] J. Rau, *Statistical Physics and Thermodynamics: An Introduction to Key Concepts*. Oxford University Press, 2017.

[87] N. D. Mermin, "Quantum mechanics vs local realism near the classical limit: A Bell inequality for spin s," *Physical Review D* **22** (1980) 356–61.

[88] A. Peres, *Quantum Theory: Concepts and Methods*. Kluwer, 1993.

[89] A. Garg, "Agnostic detector error, Wigner functions, and the classical limit of the high-spin Einstein–Podolsky–Rosen experiment," arXiv:1811.12247.

[90] N. D. Mermin, "Hidden variables and the two theorems of John Bell," *Reviews of Modern Physics* **65** (1993) no. 3, 803–15.

[91] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, "Noncommuting mixed states cannot be broadcast," *Physical Review Letters* **76** (1996) 2818, arXiv:quant-ph/9511010.

[92] J. Aczél, *Lectures on Functional Equations and Their Applications*. Academic Press, 1966.

[93] J. C. Baez and B. S. Pollard, "Quantropy," *Entropy* **17** (2015) 772–89, arXiv:1311.0813.

[94] J. C. Baez and T. Fritz, "A Bayesian characterization of relative entropy," *Theory and Application of Categories* **29** (2014) 421–56, arXiv:1402.3067.

[95] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*. Addison–Wesley, 1964. https://www.feynmanlectures.caltech.edu/I_39.html.

[96] C. A. Fuchs and R. Schack, "Quantum-Bayesian coherence," *Reviews of Modern Physics* **85** (2013) no. 4, 1693–1715, arXiv:1301.3274 [quant-ph].

[97] A. Brandenburger and K. Steverson, "Axioms for the Boltzmann distribution," *Foundations of Physics* **49** (2019) 444–56.

[98] N. Yunger Halpern and J. M. Renes, "Beyond heat baths: Generalized resource theories for small-scale thermodynamics," *Physical Review E* **93** (2016) 022126, arXiv:1409.3998.

[99] B. C. Stacey, "Invariant off-diagonality: SICs as equicoherent quantum states," arXiv:1906.05637.

[100] S. Alterman, J. Choi, R. Durst, S. M. Fleming, and W. K. Wootters, "The Boltzmann distribution and the quantum-classical correspondence," *Journal of Physics A* **51** (2018) 345301, arXiv:1710.06051.

[101] C. A. Fuchs, "Quantum mechanics as quantum information (and only a little more)," in *Quantum Theory: Reconsideration of Foundations*, A. Khrennikov, ed., pp. 463–543. Växjö University Press, 2002. arXiv:quant-ph/0205039.

[102] H. Barnum, C. A. Fuchs, J. M. Renes, and A. Wilce, "Influence-free states on compound quantum systems," arXiv:quant-ph/0507108.

[103] H. Barnum, S. Beigi, S. Boixo, M. B. Elliott, and S. Wehner, "Local quantum measurement and no-signaling imply quantum correlations," *Physical Review Letters* **104** (2010) 140401, arXiv:0910.3952.

[104] G. de la Torre, L. Masanes, A. J. Short, and M. P. Müller, "Deriving quantum theory from its local structure and reversibility," *Physical Review Letters* **109** (2012) 90403, arXiv:1110.5482.

[105] M. D. Kamen, *Radiant Science, Dark Politics: A Memoir of the Nuclear Age*. University of California Press, 1985. https://archive.org/details/radiantscienceda00kame/page/29.

[106] S. I. Weissmann and M. B. Weissmann, "Alan Sokal's hoax and A. Lunn's theory of quantum mechanics," *Physics Today* **50** (1997) no. 6, 15, 114.

[107] M. B. Weissmann, V. V. Iliev, and I. Gutman, "A pioneer remembered: biographical notes about Arthur Constant Lunn," *MATCH: Communications in Mathematical and in Computer Chemistry* **59** (2008) 687–708. https://match.pmf.kg.ac.rs/electronic_versions/Match59/n3/match59n3_687-708.pdf.

[108] K. Darrow *et al.*, "Oral history — session 1," American Institute of Physics, 1964. https://www.aip.org/history-programs/niels-bohr-library/oral-histories/4567-1.

[109] R. Mulliken and T. S. Kuhn, "Oral history," American Institute of Physics, 1964. https://www.aip.org/history-programs/niels-bohr-library/oral-histories/4788.

[110] L. Loeb and T. S. Kuhn, "Oral history," American Institute of Physics, 1962. https://www.aip.org/history-programs/niels-bohr-library/oral-histories/4748.

[111] H. Weyl, *Gruppentheorie und Quantenmechanik*. S. Hirzel, 2nd ed., 1931. Translated into English by H. P. Robertson; reprinted by Dover Books in 1950.

[112] J. Schwinger, *Quantum Mechanics: Symbolism of Atomic Measurements*. Springer, 2001. Edited by B.-G. Englert.

[113] B. Zwiebach, *Mastering Quantum Mechanics: Essentials, Theory, and Applications*. MIT Press, 2022.

[114] G. Baym, *Lectures on Quantum Mechanics*. Addison-Wesley, 1969.

[115] L. I. Schiff, *Quantum Mechanics*. McGraw-Hill, 1949.

[116] R. Hofstadter, "Electron scattering and nuclear structure," *Reviews of Modern Physics* **28** (1956) no. 3, 214–54.

[117] J. C. Baez, "Division algebras and quantum theory," *Foundations of Physics* **42** (2012) 819–55, arXiv:1101.5690.

[118] S. Hill and W. K. Wootters, "Entanglement of a pair of quantum bits," *Physical Review Letters* **78** (1997) 5022, arXiv:quant-ph/9703041.

[119] W. K. Wootters, "Entanglement of formation of an arbitrary state of two qubits," *Physical Review Letters* **80** (1998) 2245, arXiv:quant-ph/9709029.

[120] Z. Jiang, A. Kalev, W. Mruczkiewicz, and H. Neven, "Optimal fermion-to-qubit mapping via ternary trees with applications to reduced quantum states learning," arXiv:1910.10746.

[121] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, "Symmetric Informationally Complete measurements identify the irreducible difference between classical and quantum systems," *Physical Review Research* **2** (2020) 013074, arXiv:1805.08721.

[122] W. K. Wootters, "Statistical distance and Hilbert space," *Physical Review D* **23** (1981) 357.

[123] C. M. Caves and C. A. Fuchs, "Mathematical techniques for quantum communication theory," *Open Systems and Information Dynamics* **3** (1995) 345–56, arXiv:quant-ph/9604001.

[124] H. Zhu, Y. S. Teo, and B. G. Englert, "Two-qubit Symmetric Informationally Complete positive operator valued measures," *Physical Review A* **82** (2010) 042308, arXiv:1008.1138.

[125] I. Bengtsson, "The number behind the simplest SIC-POVM," *Foundations of Physics* **47** (2017) 1031–41, arXiv:1611.09087.

[126] B. C. Stacey, *A First Course in the Sporadic SICs*. Springer, 2021.

[127] D. J. Griffiths, *Introduction to Quantum Mechanics*. Pearson, 2nd ed., 2004.

[128] F. Cooper, A. Khare, and U. Sukhatme, "Supersymmetry and quantum mechanics," *Physics Reports* **251** (1995) no. 5/6, 267–385, arXiv:hep-th/9405029.

[129] B. Mohar, "Graph Laplacians," in *Topics in Algebraic Graph Theory*, L. W. Beineke and R. J. Wilson, eds., pp. 113–36. Cambridge University Press, 2004.

[130] L. E. Gendenshteĭn, "Derivation of exact spectra of Schrödinger equation by means of supersymmetry," *JETP Letters* **38** (1983) no. 6, 356–59. http://www.jetpletters.ac.ru/ps/1822/article_27857.shtml.

[131] J. Baez and J. Dolan, "From finite sets to Feynman diagrams," in *Mathematics Unlimited – 2001 and Beyond*, B. Engquist and W. Schmid, eds., pp. 29–50. Springer, 2001. arXiv:math/0004133.

[132] J. Morton, "Categorified algebra and quantum mechanics," *Theory and Applications of Categories* **16** (2006) no. 29, 785, arXiv:math/0601458 [math.QA].

[133] J. C. Morton and J. Vicary, "The Categorified Heisenberg Algebra I: A combinatorial representation," arXiv:1207.2054 [math.QA]. http://golem.ph.utexas.edu/category/2012/07/morton_and_vicary_on_the_categ.html.

[134] Y. Bar-Yam, *Dynamics of Complex Systems*. Westview Press, 1997. http://necsi.edu/publications/dcs/.

[135] J. Zinn-Justin, *Phase Transitions and Renormalisation Group*. Oxford University Press, 2007.

[136] A. Zee, *Quantum Field Theory in a Nutshell*. Princeton University Press, second ed., 2010.

[137] S. Coleman, *Physics 253: Quantum Field Theory*. Harvard, 1975. arXiv:1110.5013 [physics.ed-ph]. http://www.physics.harvard.edu/about/Phys253.html.

[138] C. A. Fuchs and B. C. Stacey, "Some negative remarks on operational approaches to quantum theory," in *Quantum Theory: Informational Foundations and Foils*, G. Chiribella and R. W. Spekkens, eds. Springer, 2016. arXiv:1401.7254 [quant-ph].

[139] B. C. Stacey, "Quantum theory as symmetry broken by vitality," arXiv:1907.02432 [quant-ph].

[140] C. D. Anderson, "The positive electron," *Physical Review* **43** (1933) no. 6, 491–94.

[141] C. V. Sukumar, "Supersymmetry and the Dirac equation for a central Coulomb field," *Journal of Physics A* **18** (1985) L697–701.

[142] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Physical Review* **40** (1932) 749–59.

[143] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, "Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction," *Physical Review A* **86** (2012) 012103, arXiv:1111.5057 [quant-ph].

[144] L. Catani, M. Leifer, D. Schmid, and R. W. Spekkens, "Why interference phenomena do not capture the essence of quantum theory," *Quantum* **7** (2023) 1119, arXiv:2111.13727 [quant-ph].

[145] J. B. DeBrota, C. A. Fuchs, and R. Schack, "Quantum dynamics happens only on paper: QBism's account of decoherence," arXiv:2312.14112 [quant-ph].

[146] J. C. Baez, "Rényi entropy and free energy," arXiv:1102.2098 [quant-ph].

[147] J. Cao *et al.*, "5.72 Statistical Mechanics," MIT OpenCourseWare, 2012. https://ocw.mit.edu/courses/5-72-statistical-mechanics-spring-2012/.

[148] H. K. Janssen and B. Schmittmann, "Field theory of long time behaviour in driven diffusive systems," *Zeitschrift für Physik B* **63** (1986) no. 4, 517–20.

[149] P. L. Garrido *et al.*, "Long-range correlations for conservative dynamics," *Physical Review A* **42** (1990) 1954–68.

[150] G. Grinstein *et al.*, "Conservation laws, anisotropy, and 'self-organized criticality' in noisy nonequilibrium systems," *Physical Review Letters* **64** (1990) no. 16, 1927–30.

[151] Z. Cheng *et al.*, "Long-range correlations in stationary nonequilibrium systems with conservative anisotropic dynamics," *Europhysics Letters* **14** (1991) no. 6, 507–13.

[152] H. K. Janssen, "On the nonequilibrium phase transition in reaction-diffusion systems with an absorbing stationary state," *Zeitschrift für Physik B* **42** (1981) no. 2, 151–54.

[153] P. Grassberger, "On phase transitions in Schlögl's second model," *Zeitschrift für Physik B* **47** (1982) no. 4, 365–74.

[154] S. R. Broadbent and J. M. Hammersley, "Percolation processes I. Crystals and mazes," *Mathematical Proceedings of the Cambridge Philosophical Society* **53** (1957) no. 3, 629–41.

[155] D. C. Mattis and M. L. Glasser, "The uses of quantum field theory in diffusion-limited reactions," *Reviews of Modern Physics* **70** (1998) no. 3, 979–1001.

[156] K. A. Takeuchi, M. Kuroda, H. Chaté, and M. Sano, "Experimental realization of directed percolation criticality in turbulent liquid crystals," *Physical Review E* **80** (2009) 05116, arXiv:0907.4297 [cond-mat].

[157] H. Kelker and B. Scheurle, "A liquid-crystalline (nematic) phase with a particularly low solidification point," *Angewandte Chemie* **8** (1969) no. 11, 884–85.

[158] B. C. Stacey, *Multiscale Structure in Eco-Evolutionary Dynamics*. PhD thesis, Brandeis University, 2015. arXiv:1509.02958 [q-bio.PE].

[159] B. Lumpkin, "Mathematics used in Egyptian construction and bookkeeping," *The Mathematical Intelligencer* **24** (2002) no. 2, 20–25.

[160] E. Cheng, *Is Math Real? How Simple Questions Lead Us To Mathematics' Deepest Truths*. Basic Books, 2023.

[161] G.-C. Rota, "Ten lessons I wish I had learned before I started teaching differential equations," MAA invited address, Simmons College, 1997.

[162] R. K. P. Zia, E. F. Redish, and S. R. McKay, "Making sense of the Legendre transform," *American Journal of Physics* **77** (2009) 614–22, arXiv:0806.1147 [physics.ed-ph].

[163] N. Rosen, "Mixed states in classical mechanics," *American Journal of Physics* **33** (1965) 146–50.

[164] J. C. Baez, "Classical mechanics versus thermodynamics (part 1)," *Azimuth* (2012) n.p. https://johncarlosbaez.wordpress.com/2012/01/19/classical-mechanics-versus-thermodynamics-part-1/.

[165] J. C. Baez, "Classical mechanics versus thermodynamics (part 3)," *Azimuth* (2021) n.p. https://johncarlosbaez.wordpress.com/2021/09/23/classical-mechanics-versus-thermodynamics-part-3/.

[166] C. A. Fuchs, *Coming of Age with Quantum Information*. Cambridge University Press, 2011.

[167] N. D. Mermin, *It's About Time: Understanding Einstein's Relativity*. Princeton University Press, 2005.